

**Measurement of Telomere Organization using Widefield
Microscopy**

by

Tracy Duncan

B.Eng. Engineering Physics, University of Saskatchewan, 1996

B.Sc. Biochemistry, University of Saskatchewan, 1995

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

Master of Applied Science

in

THE FACULTY OF GRADUATE STUDIES

(Department of Electrical Engineering)

We accept this thesis as conforming
to the required standard

The University of British Columbia

January 2002

© Tracy Duncan, 2002

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Electrical + Computer Engineering

The University of British Columbia
Vancouver, Canada

Date Mar 22/2002

Abstract

Widefield microscopy was used to study the location of telomeres in human interphase lymphocyte nuclei in order to determine whether telomeres are required for nuclear organization. To do this, a microscope's computer control system was programmed to automatically focus the microscope, and large numbers of images were analyzed.

Cells in cell cycle stages G1 or S/G2 with long, medium, and short telomeres were photographed. Their two-dimensional images were analyzed to calculate the telomere distribution along the radius of the nucleus and to evaluate the degree of telomere clustering in space using Delaunay graphs. Telomeres in nuclei with short telomeres were shown to be distributed closer to the center of the nucleus than those in nuclei with long telomeres. The radial distribution of telomeres did not change with cell cycle phase. No changes in the degree of telomere clustering with telomere length or cell cycle phase were observed.

Contents

| | |
|---|------------|
| Abstract | ii |
| Contents | iii |
| List of Figures | vii |
| List of Tables | ix |
| 1 Introduction | 1 |
| 1.1 Microscope System at Terry Fox Laboratory | 2 |
| 1.1.1 Microscope Hardware | 2 |
| 1.1.2 PNA Probes | 4 |
| 1.2 Widefield vs Confocal Microscopy | 6 |
| 1.3 Multiplane Microscopy | 8 |
| 1.4 Telomere Biology | 8 |
| 1.4.1 Telomere Shortening | 9 |
| 1.4.2 Telomere Binding Proteins | 10 |
| 1.4.3 Telomere Organization | 11 |
| 1.5 Thesis Goals | 13 |

| | | |
|----------|--|-----------|
| 2 | Composite Images | 15 |
| 2.1 | Motivation | 15 |
| 2.2 | Methods | 16 |
| 2.2.1 | Wu Method | 17 |
| 2.2.2 | Maximum IF Method | 19 |
| 2.2.3 | Average Image | 22 |
| 2.2.4 | Object Segmentation and Optimization Method | 22 |
| 2.3 | Hardware | 22 |
| 2.4 | Test Images | 23 |
| 2.5 | Analysis and Results | 24 |
| 2.5.1 | Appearance | 24 |
| 2.5.2 | Focusing Artifact | 24 |
| 2.5.3 | Mean Square Error Between Bead Images | 25 |
| 2.5.4 | Bead Fluorescence CV and Telomere q-p Correlation | 26 |
| 2.6 | Conclusions | 27 |
| 2.7 | Improvements | 27 |
| 3 | Automatic Focusing | 29 |
| 3.1 | Motivation | 29 |
| 3.2 | Hardware | 30 |
| 3.3 | Methods | 31 |
| 3.4 | Results | 33 |
| 3.5 | Improvements | 34 |
| 4 | Telomere Organization in Human Interphase Lymphocytes | 36 |
| 4.1 | Goals and Hypotheses | 37 |

| | | |
|-------|--|----|
| 4.2 | Measures of Telomere Organization | 38 |
| 4.2.1 | Radial Distribution | 38 |
| 4.2.2 | Minimum Spanning Tree Features | 41 |
| 4.2.3 | General Delaunay Features | 43 |
| 4.2.4 | 1p Probe Distribution | 43 |
| 4.3 | Expected Differences | 44 |
| 4.4 | Materials and Methods | 44 |
| 4.4.1 | Lymphocyte Cells and Preparation | 44 |
| 4.4.2 | Image Acquisition | 46 |
| 4.4.3 | Colour Plane Registration | 48 |
| 4.4.4 | Correction for Optical Crosstalk | 50 |
| 4.4.5 | Nucleus Detection | 53 |
| 4.4.6 | Telomere Detection and Localization | 57 |
| 4.4.7 | Details of Telomere Detection | 61 |
| 4.4.8 | 1p Probe Detection | 66 |
| 4.5 | Analysis and Results | 67 |
| 4.5.1 | Number and Intensity of Telomeres Detected | 67 |
| 4.5.2 | Kolmogorov-Smirnov Statistic | 68 |
| 4.5.3 | Radial Distribution of Telomeres | 69 |
| 4.5.4 | 1p Probe Distribution | 75 |
| 4.5.5 | Delaunay Triangulation and MST Calculation | 75 |
| 4.5.6 | MST Segment Length Features | 78 |
| 4.5.7 | Nearest-Neighbour Features | 79 |
| 4.6 | Discussion | 81 |
| 4.6.1 | Telomere Radial Distributions | 83 |

| | |
|--|-----------|
| 5 Summary | 85 |
| 5.1 Conclusions | 85 |
| 5.1.1 Microscopy | 85 |
| 5.1.2 Image Processing and Telomere Organization | 86 |
| 5.2 Future Directions | 88 |
| 5.2.1 Widefield Microscope Focusing | 88 |
| 5.2.2 Measurement of Telomere Organization | 88 |
| Bibliography | 91 |

List of Figures

| | | |
|------|---|----|
| 2.1 | Microscope OTF (x scale depends on defocus) | 18 |
| 2.2 | Gaussian High Pass Filter | 19 |
| 2.3 | Composite Images of Beads Generated by the Wu, Maximum IF, Segmentation, and Average Methods plus the Best-Focused Image . | 20 |
| 2.4 | Map Images of Beads Generated by the Wu, Maximum IF, and Seg- mentation Methods | 21 |
| 4.1 | Uniform Cumulative Probability Distributions for Telomeres | 41 |
| 4.2 | Sample Voronoi and Delaunay Graphs | 42 |
| 4.3 | Lymphocyte Images Before and After Colour Plane Alignment . . . | 49 |
| 4.4 | Green Image Before Crosstalk Correction and Under-Corrected (10%) | 51 |
| 4.5 | Optimally-(20%) and Over-(28%) Corrected Green Images | 52 |
| 4.6 | Lymphocyte Image | 53 |
| 4.7 | DAPI Counterstain | 54 |
| 4.8 | Sample DAPI Image Histogram | 55 |
| 4.9 | Sobel Image | 56 |
| 4.10 | Thinned Sobel Image | 58 |
| 4.11 | Segmented DAPI Nucleus | 58 |
| 4.12 | Spots2.exe Program | 61 |

| | |
|--|----|
| 4.13 Split Telomere Utility | 65 |
| 4.14 Light Spread from Out-of-Focus Telomere | 65 |
| 4.15 Delete Telomere Utility | 65 |
| 4.16 Cumulative Distribution of G1 Long Telomere Radial Locations . . . | 70 |
| 4.17 Cumulative Distribution of G1 Medium Telomere Radial Locations . | 70 |
| 4.18 Cumulative Distribution of G1 Short Telomere Radial Locations . . | 71 |
| 4.19 Cumulative Distribution of S/G2 Long Telomere Radial Locations . | 71 |
| 4.20 Cumulative Distribution of S/G2 Medium Telomere Radial Locations | 72 |
| 4.21 Cumulative Distribution of S/G2 Short Telomere Radial Locations . | 72 |
| 4.22 Cumulative Distributions of Telomere Radii: All Lymphocyte Popu- lations | 74 |
| 4.23 Cumulative Distributions of 1p Probe Radial Locations | 76 |
| 4.24 Telomere Locations displayed on Lymphocyte Image | 76 |
| 4.25 Delaunay Triangulation Graph of Telomere Locations | 77 |
| 4.26 Minimum Spanning Tree of Telomere Locations | 77 |
| 4.27 Comparison of MST Segment Lengths | 80 |
| 4.28 Distribution of 1 Nearest Neighbour Distances | 82 |
| 4.29 Distribution of 2 Nearest Neighbour Distances | 82 |
| 4.30 Distribution of 3 Nearest Neighbour Distances | 83 |

List of Tables

| | | |
|-----|--|----|
| 4.1 | Pixel Shifts of FITC Relative to CY3 Image | 50 |
| 4.2 | Telomere Detection in Lymphocyte Populations | 68 |
| 4.3 | General MST Features-Bin Size 0.005 | 79 |
| 4.4 | First Nearest Neighbour Distances (normalized to Nuclear Radius) . | 79 |
| 4.5 | Second Nearest Neighbour Distances (normalized to Nuclear Radius) | 80 |
| 4.6 | Third Nearest Neighbour Distances (normalized to Nuclear Radius) | 80 |

Chapter 1

Introduction

This thesis aims to use a widefield microscope system to make a statistical study of the organization in the nucleus of DNA regions called telomeres. Telomeres are repeat DNA sequences at the end of chromosomes which become shorter each time a cell divides. Telomere length is important to cellular viability, as cells with critically-short telomeres undergo genetic mutations and stop functioning. The thesis aims to determine whether, in cells with short telomere sequences, the telomere position within the nucleus becomes less organized. Such an observation would suggest that, in addition to other functions, telomeres are recognized by the proteins which organize genetic matter into compartments in the nucleus, and that one way that telomere-induced aging occurs is by loss of nuclear organization. To perform the study, it was necessary to measure the telomere positions in large numbers of cells by using multiple plane microscope information to focus the microscope and by automating image analysis using a robust telomere detection program.

1.1 Microscope System at Terry Fox Laboratory

The microscope system used in this experiment is one of several used in the Terry Fox Laboratory for measuring the length of specific DNA sequences within a cell. It is a widefield system that is not equipped for imaging in three dimensions, but it is a quality system. Its capabilities stem from a combination of microscope hardware and the set of PNA probes described in sections 1.1.1 and 1.1.2. One of the goals for the thesis is to demonstrate that such a system can be used to study nuclear architecture.

1.1.1 Microscope Hardware

The microscope is Zeiss Axioplan device with a mercury/xenon light source. It is an epifluorescence microscope in which short-wavelength light from the source passes through a bandpass filter, through the microscope's objective lens, and is focused on a slide. When the objects on the slide are autofluorescent or are labelled with fluorescent dyes, they fluoresce when illuminated with the correct excitation wavelength. The fluoresced signal, of wavelength longer than the excitation wavelength, passes back through the microscope's objective lens, through a second bandpass filter, and, if allowed by the filter, through an eyepiece lens for visual inspection or to a digital camera for imaging. By using the right combination of fluorescent dyes and bandpass filters for the excitation and fluorescent light, users can discriminate between objects labelled with different dye molecules. Typically, users combine several images of a single microscope field of view, acquired using different excitation/emission wavelengths, in software. For example, the Lansdorp lab commonly labels certain segments of DNA within a cell with red dye, then stains all DNA with a blue dye. By combining images taken using the red and blue microscope filters,

users create images of cells which show the red DNA of interest's location relative to the rest of the genetic material.

The microscope was fitted with Zeiss 20x lens for examining slides manually at low resolution and a Zeiss 63x oil immersion lens which was generally used for acquiring images. The multiple excitation and single emission filters were from Chroma and were mounted on a wheel controlled with a pushbutton switch.

Images were acquired using a Xillix MicroImager digital camera. The camera is not cooled and has a Kodak 1400 CCD chip with a pixel size of $6.8 \mu m$.

The bandwidth BW allowed by a microscope's optical transfer function is determined mainly by the numerical aperture NA of the objective lens (which determines the depth of focus of the system) and the wavelength λ of light used and is given by

$$BW = \frac{2NA}{\lambda} \quad (1.1)$$

The numerical aperture of the 63 x lens used in this equation was 1.4 and the lowest emission wavelength of the fluorophores used is 400 nm, so the maximum spatial frequency ω_{max} present in the microscope image should equal the bandwidth:

$$\omega_{max} = 2\pi \frac{1.4}{400nm} = 0.022nm^{-1} \quad (1.2)$$

Using a 63x lens on the microscope with a pixel size of $6.8 \mu m$ means the system has an effective pixel size of 108 nm and a sampling frequency of $\frac{2\pi}{108nm} = 0.058nm^{-1}$. The Nyquist theorem requires that the sampling frequency must be greater than 2.3 times the maximum frequency present, or greater than $0.051nm^{-1}$. So the sampling frequency of the system satisfies the Nyquist criterion.

The camera field size is 1280 x 1024 pixels. The CCD performs analog-to-digital conversion in 14 bits and is read out in 8 bits by a Matrox image acquisition

board with three image buffers and some on-board hardware for performing image convolution operations.

The system is controlled by a 486 personal computer using an image acquisition system written in 1991. The control functions available through the personal computer include setting the camera's exposure time, specifying the 14-bit to 8-bit conversion table for the CCD readout, CCD binning, and reading out the camera images. The acquisition system was also originally intended to control the colour filter wheel and a stepper motor for the microscope stage, but these capabilities have not been used for some time. Although the acquisition system is quite old, its source code is available in the lab, so changes can be made to the system to fix or automate it for whatever tasks are required.

The quality microscope and the customizable image acquisition program make the system a good candidate to attempt to improve its capabilities for ambitious projects. The changes I made to the system are discussed in chapters 2 and 3.

1.1.2 PNA Probes

Peptide Nucleic Acid (PNA) probes are a relatively new class of genetic probes. The lab uses a number of such probes from Boston Probes. Genetic probes are large molecules containing DNA bases complementary to sequences of interest on genetic material. They are employed by melting the DNA of interest with heat, adding probe solution, and allowing the probe molecules to hybridize with complementary base sequences on the single-stranded DNA. Fluorescent probes are labeled with a fluorescent molecule which can be detected using a microscope, fluorescence-assisted cell sorter, or fluorescence autoradiography of a Southern blot. They are commonly

used to detect the presence or absence of genes of interest or to localize sequences within cell architecture. DNA probes can be ten to hundreds of base pairs long.

The probes from Boston Probes are different from conventional probes because the bases are attached to a protein backbone instead of a deoxyribose-phosphate backbone. The consequence of this is that PNA-DNA hybrids are stable in different salt and pH environments from DNA-DNA hybrids. PNA probes can be annealed with DNA of interest under conditions which do not allow DNA-DNA hybridization and so the PNA probes are not subject to competition with the complementary DNA strand to hybridize with the DNA of interest. This has several implications:

- The SNR (signal to noise ratio) for assays using PNA probes is higher than for those using DNA probes. This means that better resolution is available and more ambitious microscope analysis of these probes can be attempted.
- The light intensity from probe molecules, which (neglecting second-order effects such as quenching) is proportional to the number of probe molecules bound, is linearly dependent on the incidence of the complementary DNA sequence to the probe [1]. This means that, in addition to detecting repetitive sequences of interest, PNA probes can be used to count the numbers of repeats in those sequences.
- Good modeling of the light-collection efficiency of the microscope or other instrument is necessary to interpret this quantitative information.

Currently the lab uses the PNA probe $(A_2TC_3)_3$ to estimate *in situ* the number of T_2AG_3 repeats on the chromosomes of mice and humans. These sequences are called telomeres and are described in section 1.4. The lab also aims to exploit

the high SNR of the PNA probes to analyze nuclear architecture, including the positions of telomeres within nuclei.

1.2 Widefield vs Confocal Microscopy

The microscope system used in this lab is a widefield system. This means that its CCD camera takes an image of the microscope's entire field of view. Using the 63x lens, an image of a $110\ \mu\text{m}$ by $110\ \mu\text{m}$ region is acquired each time the CCD is exposed. Another system that is commonly used to study cellular architecture is called a confocal microscope [2]. Such a system uses a laser spot to scan over the microscope's field of view, illuminating discrete regions one at a time. Fluoresced light from illuminated voxels on a specimen is collected using a photomultiplier tube (PMT), digitized, and incorporated into an image reconstructed by the control computer. Bandpass filters similar to those described in section 1.1.1 can be used to discriminate between objects labelled with different dye molecules. The advantage of the confocal microscope is that the laser spot can be focused on and scanned across many different focal planes through the depth of the object to be imaged. The light information collected by the PMT can be combined to produce accurate and breathtaking three-dimensional images with resolution near the optical limit in the x and y directions and slightly lower resolution along the axis of the microscope. Confocal microscope systems are expensive and require specialized slide preparation.

Widefield microscopes have several advantages, which can be exploited for high-performance microscopy, over confocal systems:

- They are fast. Widefield microscopes acquire an image of their entire field of view during a single CCD exposure which is typically under 10 seconds

in duration. CCD digitization and readout are the only image processing tasks necessary for the acquisition computer system. On the other hand, confocal systems require minutes to scan a single object of interest with the laser spot and hours to reconstruct an image of the object from the PMT information acquired. This makes a single-plane widefield microscope system the reasonable choice for studies which require large sample sizes.

- Because of the scanning laser spot, confocal microscope systems bleach the fluorescent dyes in illuminated objects in a way that varies through the entire volume being examined. Bleaching correction is a complex procedure in confocal image processing and is difficult to carry out in a way that retains quantitative information about the number of dye molecules in a given voxel. Widefield microscopes also bleach their targets, but less strongly and in a way that is constant across an entire image. Because they pose less of a bleaching problem, widefield microscope systems reduce the complexity of bleaching correction for quantitative fluorescent imaging, and they may be less likely to bleach out information from low-intensity regions of interest.
- Widefield microscope optics reject some out-of-focus fluorescent light from the objects being imaged [3], and so in-focus two-dimensional widefield images can be used for quantitative imaging without massive correction to remove out-of-focus light from other image planes.

If three-dimensional imaging is not required, widefield systems are simpler, cheaper, and less prone to errors introduced during image reconstruction than confocal ones. In addition, widefield microscopy lends itself to performing statistical experiments because many images can be taken within a reasonably short time.

1.3 Multiplane Microscopy

It is possible to get accurate information about a three-dimensional structure using widefield microscopy. This is achieved by focusing the objective lens at different depths within an object and acquiring an image at each focal plane. For this project, I installed a Physik Instrument focusing device on the microscope to image several focal planes along the microscope's z-axis, as described in chapter 3, section 2.3. Several options are available to make use of multiple plane information:

1. A single image can be constructed by summing light or other features from several focal planes.
2. The information can be used to optimally focus the microscope [4].
3. Deblurring algorithms can be used to improve the microscope image at each focal plane, then the object can be rendered in three dimensions. Deblurring can be a simple procedure using a couple of image planes at a time or can be a complex, iterative procedure. Often it is performed in the Fourier domain [5, 6, 7].

In this thesis we investigate the first two uses described above for multiple plane information. They are described in chapters 2 and 3.

1.4 Telomere Biology

Telomeres are G-rich sequences of double-stranded DNA and associated proteins located at the ends of chromosomes. In humans, these regions contain the repeat sequence T_2AG_3 and range from 2000 to 20000 base pairs in length [8]. They are the only component of chromosome ends known to be conserved throughout vertebrate

evolution and therefore must be important to cell function [9]. They are proposed to have the following functions:

- They are bound by telomere recognition proteins into structures which 'cap' each chromosome and prevent end-to-end fusion of chromosomes by DNA repair polymerases [10].
- Cells may have a mechanism to monitor telomere length to record their doubling history [11]. Telomere length may therefore determine the proliferative capacity of cells.
- In yeast, telomere-binding proteins can silence genes in the telomere region by creating local changes in chromatin structure [12].
- Telomeres may play a role in mediating cellular organization. In particular, they may be the structures which are recognized and bound by the proteins which position chromosomes at prophase and metaphase [13].

1.4.1 Telomere Shortening

Average telomere length in the nucleus of most human cells diminishes upon cell division. This is due to DNA polymerases' failure to bind to the extreme end of DNA template strands when the DNA is copied. This 'end replication problem' [14] results in an average loss of several base pairs of the telomere sequence on each chromosome at each cell division. Cells of many species replace some of the lost telomere repeats using the specialized reverse transcriptase enzyme telomerase.

In mice which have had the gene coding for telomerase removed, the telomere length decreases radically with each cell division and each generation of mice. As

telomere length reaches a critical minimum, many end-to-end fusions between chromosomes take place [15]. The likely mechanism for this is that insufficiently-long telomeres are unable to recruit proteins to protect them and to distinguish them from broken single-stranded DNA in need of annealing by repair polymerases.

1.4.2 Telomere Binding Proteins

At least two proteins bind to duplex T_2AG_3 repeats in humans. The proteins, TTAGGG Repeat Binding Factors 1 and 2 (TRF1 and TRF2), are structurally and sequentially similar and exist as homeodimers and homeotetramers. They each carry a T_2AG_3 recognition site on the carboxy end and have a dimerization site on the N-end. Other organisms have proteins which are similar in structure and function, although not necessarily in amino acid sequence. Both proteins may require sufficiently-long telomere sequences in order to recognize and bind telomeres.

TRF1 is responsible for maintaining telomere length by downregulation of telomerase [16]. The most likely mechanism for this is that TRF1 binds to sufficiently-long telomeres which do not require elongation, then the TRF1-DNA complex suppresses telomerase expression. When telomeres become critically short, TRF1 no longer makes a DNA complex. Telomerase is then expressed and extends the telomere sequences.

In vitro, TRF1 also promotes pairing of DNA structures of sequence $(T_2AG_3)_{6-12}$ [16]. This finding suggests that TRF1 protein-protein interactions may have a role in organizing telomeres into clusters or discrete nuclear regions as described in section 1.4.3.

TRF2 protects chromosome ends. It is more highly conserved among mammals than TRF1, which demonstrates TRF2's importance to the cell [17]. Cells

which express mutant TRF2 protein undergo end-to-end fusions at a very high rate [18, 11]. These fusions also take place in cells with critically short telomeres, so TRF2 may be the protein which monitors telomere length and stops proliferation of cells which have short telomeres due to a long division history [11]. In that case, TRF2 must be dependent on telomere length for binding and function.

If TRF2 and TRF1's functions depend on telomere length, then it is likely that protein-mediated telomere properties, including the organization properties described below, also depend on sufficient telomere sequence length within a nucleus and should become less evident as telomere length decreases.

1.4.3 Telomere Organization

Telomere location and degree of organization differ with species and with cell type within a species. This divergence, along with the fact that the telomere recognition factors do not recognize telomeric sequences inserted within chromosomes [19], suggest that telomere organizing factors are weak. Nevertheless, many studies describe telomere organization in the nucleus [20, 12, 9, 21, 22, 19, 23, 24, 13].

Radial Distribution

Most studies of telomere position describe telomere distribution density along the radius of the nucleus. In human lymphocytes, which are the subjects of this thesis, Ferguson and Ward [23] report that this 'radial distribution' varies with cell cycle phase. Telomeres tend to be at the center of the nucleus in G1 phase (when the cell has not yet replicated its DNA), randomly distributed in S phase (when DNA is being synthesized), and oriented at the outside of the cell in G2 phase (when the cell is preparing to enter metaphase). Also of interest is the observation in yeast that,

when the telomere binding proteins yKu70p and yKu80p are absent, the location of telomeric DNA changes from clusters around to the nuclear periphery to uniform distribution throughout the nucleus [20].

The first goal of this thesis is to duplicate Ferguson and Ward's observation of cell cycle dependent telomere movement, then to see whether the observation holds true in lymphocytes which have shortened telomeres due to repeated cloning. It is hypothesized that the telomere locations will become more random and less correlated with cell cycle phase when the telomeres become shorter. If this is true, it would suggest that length-dependent telomere recognition factor proteins play a role in nuclear organization, and that loss of nuclear organization is one mechanism in which cells with critically short telomeres lose viability.

Clustering

There is less information on the local spatial organization of telomeres - specifically, whether groups of telomeres are organized together into clusters. However, it is plausible that there are organizing elements in mammal cells which bind multiple telomeres or bind single telomeres but reside in discrete nuclear compartments. Several pieces of information support this possibility:

- In the plant species *Pisum sativum* and *Vicia faba*, telomeres were observed to localize on the nuclear membrane. The telomeres were not randomly distributed throughout the membrane, but existed in clusters which might have replicative significance [19].
- In confocal microscope studies of mouse nuclei, the number of telomere repeat probe spots detected was half the number of telomeres known to be in the cells [13]. While this observation may indicate low microscope resolution or

low detection efficiency in the image analysis program, the authors speculate that it means the telomeres are closely paired at distances below microscope resolution.

- TRF1 promotes pairing of chromosomes *in vitro* [16] and may have the same function *in vivo*.

The second goal of this thesis is to create a mathematical measure of the degree of telomere clustering within a nucleus, then to observe whether telomere clustering changes with cell cycle phase or in nuclei with short telomeres. It is hypothesized that, if clustering proteins (such as possibly TRF1) depend on sufficiently-long telomeres for action, clustering should decrease with decreasing telomere length.

1.5 Thesis Goals

This thesis aims to investigate whether long telomeres are necessary elements for the organization of DNA within the nucleus. If this were so, it would prove that loss of nuclear organization is a mechanism for telomere-related cell aging. Telomere organization is studied in nuclei with decreasing telomere length to determine whether the cells with short telomeres become less organized. The first two goals of the experiment, described above, are to examine telomere radial distribution density and clustering in populations of cells with decreasing telomere length.

The third thesis goal is to demonstrate the usefulness of widefield microscopy for in a nuclear architecture study because it permits analysis of large numbers of images. The experiment in chapter 4 is accomplished in an unusual way for a biology experiment: By setting up the microscope system for automated image capture and creating a dedicated image analysis program, we make it possible to

acquire large numbers of widefield microscope images and make a statistical analysis of the telomeres detected. Hopefully this type of statistical analysis of many images will continue in the lab in the future.

Chapter 2

Composite Images

Since specimens on a slide have thickness in the z-direction (normal to the plane of the slide), not all objects of interest lie on the same focal plane. This means that the microscope can be focused at several depths through the body of the specimen, and that an image can be acquired at each depth which contains some useful information. To incorporate such multiple plane information for analyzing interphase lymphocyte images, two approaches were evaluated. The first is to combine in-focus regions from several focal planes into a single 'best-focused' composite image. The second is to use the multiple plane information to focus the microscope. Composite images are discussed in this chapter.

2.1 Motivation

A composite image is a single image which has been constructed using in-focus pixels or regions from several images of the same object. It allows the user to have accurate, in-focus information about objects on the slide without having to reconstruct or analyze the image in three dimensions. Such two-dimensional information

is important to the lab in a couple of ways:

- The image analysis programs currently used in the lab are for single images only.
- If quantitative information is desired about objects, such as the nucleus size or the amount of fluorescent light emitted, it is necessary that the objects be fully in-focus or have equivalent defocus. This is because out-of-focus objects appear larger than in-focus ones, due to the point spread function of the microscope, and because the epifluorescence microscope illuminates in-focus objects preferentially, making them appear brighter than out-of-focus objects [3].
- In-focus objects are easier to detect than out-of-focus ones. This is an important consideration for an experiment which measures object distribution probabilities, such as in chapter 4. When measuring distributions, it is important to have the same detection efficiency for all telomeres, regardless of their location in the nucleus. Use of a composite image which incorporates information from all image planes means that objects near the top and bottom are just as likely to be detected as those in the center.

2.2 Methods

Four methods were used for constructing composite images. Each method requires acquiring a stack of microscope images at varying focal planes, evaluating the images according to some criterion, then copying pixels or regions from the stack into a composite image. The resulting composite images were compared to determine the best method.

2.2.1 Wu Method

This method was proposed by pathologists Hai-Shan Wu *et al* [25]. It requires partitioning the composite image plane into regions. For each region, an image is chosen from the stack which offers the sharpest focus, and pixels from that image are copied into the region on the composite image.

Their model of image formation assumes that the point spread function for a microscope imaging system is linear and shift invariant locally. It is a cylindrical function with radius $r(d)$ dependent on the amount of defocus d . The optical transfer function (OTF) $H(\omega_1, \omega_2)$ is a Bessel function of the form

$$H(\omega_1, \omega_2) = \frac{J_1(r(d)\sqrt{\omega_1^2 + \omega_2^2})}{r(d)\sqrt{\omega_1^2 + \omega_2^2}} \quad (2.1)$$

where ω_1 and ω_2 are the spatial frequencies in the x and y directions. It is shown in figure 2.1. This OTF is a lowpass filter with cutoff frequency

$$\omega_c = \frac{2.2}{r(d)} \quad (2.2)$$

where $r(d)$ is lowest and the cutoff frequency ω_c is highest when the image is in focus. There is also noise present at all frequencies.

Each image from the stack represents light from the original object subjected to a low pass OTF with a different cutoff frequency depending on the distance from image plane to the original object. Wu *et al* analyze the images by applying a high pass filter with a cutoff frequency near that of the microscope at a moderate degree of focus. The combined effect of the microscope OTF and software filter is a bandpass filter with a constant lower frequency and an upper cutoff frequency which is highest for in-focus images. The images which are in focus will experience the widest pass band and have the most power after processing. The noise spectrum is expected to be constant across all the images.

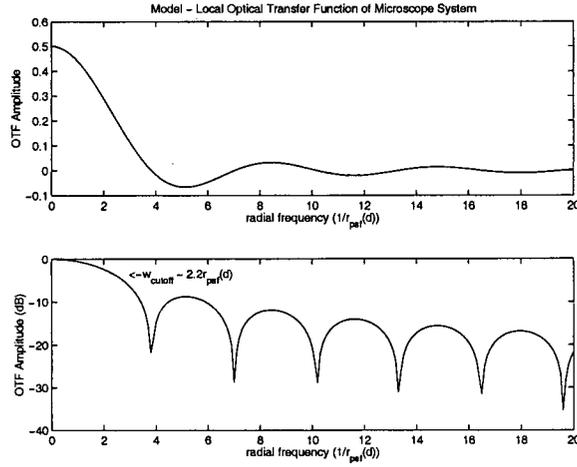


Figure 2.1: Microscope OTF (x scale depends on defocus)

Often, sharpness of focus is evaluated by applying a Laplacian or other edge-detecting gradient filter to the image, then measuring the strength of the gradient image. The Wu method is an improvement over such a method because filters like the Laplacian measure the power of the image only at the highest frequencies and are therefore very susceptible to noise. The Wu method uses a lower cutoff frequency and evaluates the image over a wider frequency band, making it more responsive to the strength of medium-frequency features in the actual image and making the noise signal a smaller portion of the output image. Also, it thoughtfully allows the user to tune the filter to incorporate *a priori* knowledge about the cutoff frequency of the microscope's OTF.

To implement this method, I convolved the images from the image stack with the Gaussian high pass filter

$$h_{g_{hp}} = C \exp\left(-\frac{n_1^2 + n_2^2}{20}\right) \text{ for } |n_1|, |n_2| \leq 7 \quad (2.3)$$

$$h_{g_{hp}} = 0 \text{ otherwise} \quad (2.4)$$

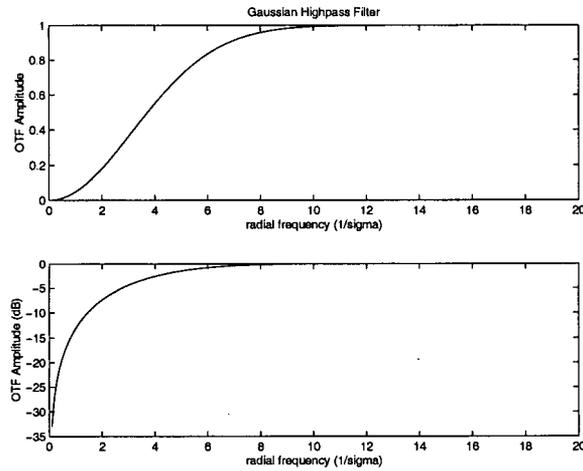


Figure 2.2: Gaussian High Pass Filter

This filter is shown in figure 2.2. For each 8×8 block in the composite image, I selected the image from the stack which had the highest power at that block's coordinates, then copied the block from the original stack image into the composite image. The resulting image of fluorescent beads is shown in the first image of figure 2.3.

2.2.2 Maximum IF Method

As outlined in [3], objects in the focal plane of an epifluorescence microscope receive more light than out-of-focus objects. Neglecting noise, the brightest regions must be the best-focused regions. To implement this method, I divided the image again into 8×8 pixel regions and chose the subimages with the highest integrated fluorescence (IF) to combine into a best-focused image.

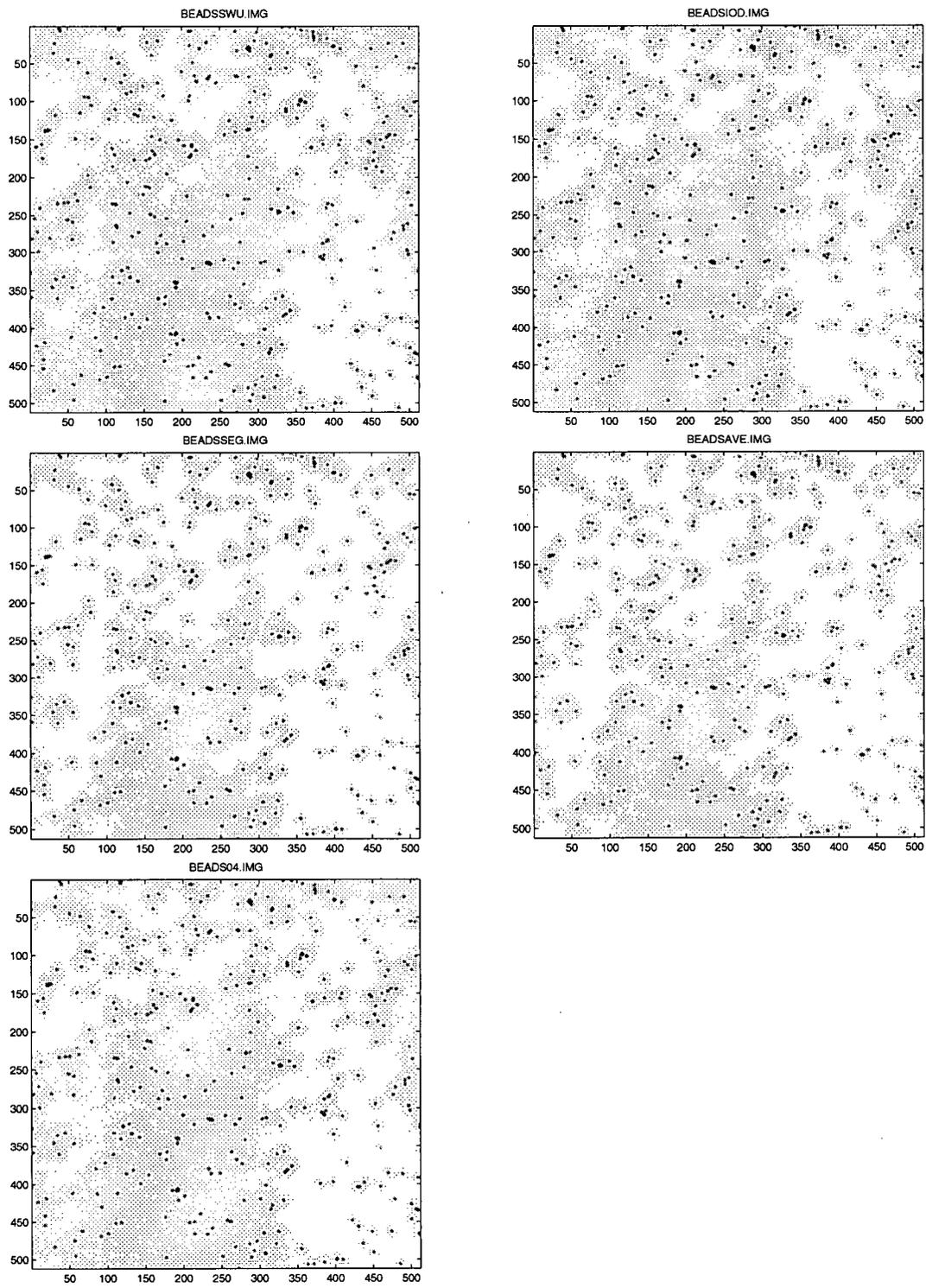


Figure 2.3: Composite Images of Beads Generated by the Wu, Maximum IF, Segmentation, and Average Methods plus the Best-Focused Image

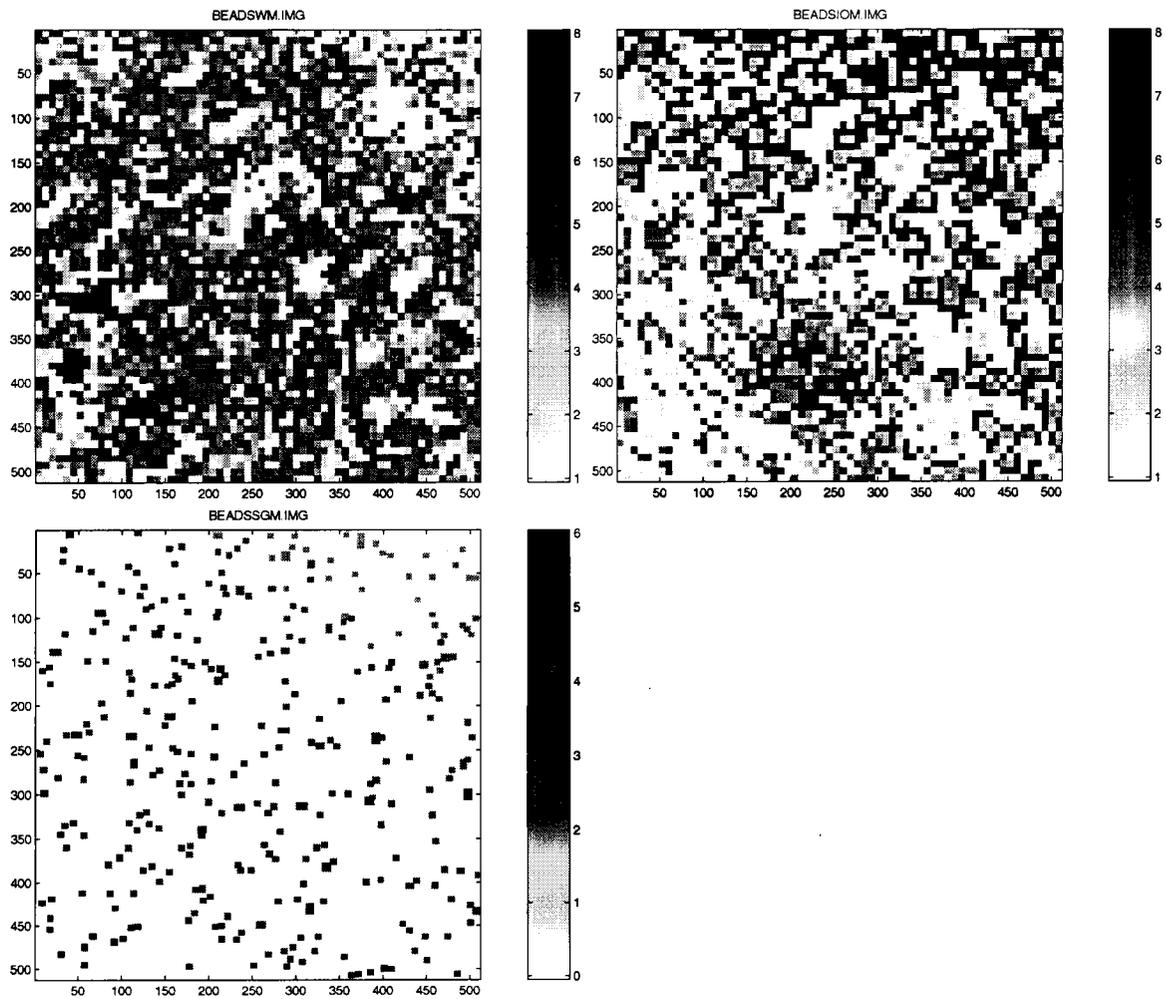


Figure 2.4: Map Images of Beads Generated by the Wu, Maximum IF, and Segmentation Methods

2.2.3 Average Image

The third method was to average all the images on the stack. An obvious advantage of this method is that it reduces background noise by a factor of \sqrt{N} , where N is the number of images averaged [26]. It is partially justified by the observation in [3] that the point spread function of the epifluorescence microscope is spatially finite in the z direction.

2.2.4 Object Segmentation and Optimization Method

The fourth method was to segment objects in the images using simple thresholding at 60% of the background histogram peak and labelling objects. For each object segmented, I chose the subimage which displayed the highest integrated fluorescence (IF) for that object, and copied the rectangular region containing the segmented object from the subimage to the composite image. To fill regions on the composite image which did not contain segmented objects, the algorithm averaged the background. This algorithm was expected to be useful for the bead and telomere images processed here, since each object is small enough that it should be entirely in focus in a single subimage.

2.3 Hardware

The piezomotive focusing device (PIFOC) which was installed to acquire multiple focal plane images is described in chapter 3, section 3.2. A fortunate hardware accident happened during this project which was useful for testing the composite image algorithms. The PIFOC was not machined to be aligned properly with the microscope optical axis (and would have to be replaced later by Physik Instrumente).

This meant that the microscope focal plane was not aligned with the plane of the slide where the objects lay, so only a linear region at the intersection of the two planes was in focus at all times. This line of intersection ran diagonally across the field of view, and meant that, as the PIFOC moved up and down while taking test images, a different diagonal strip in each image was in focus.

This focusing artifact was a good test of the algorithms because the best methods corrected it by choosing regions from the bottom through to the top of the image stack to fill diagonal strips across the composite image.

2.4 Test Images

Two sets of source images were acquired. The first set of images was of fluorescent calibration beads mounted in VectaShield. The images were 512 x 512 pixels in extent with about 300 beads in each image. The beads were well-separated and easy to resolve. The background intensity was 3 and the intensity of pixels in the bead spots was near 240. The second set of images was of human telomeres labeled with PNA probe. These images were also 512 x 512 pixels, with a mean background level of 6 and a mean spot intensity of 128.

Both sets of images were acquired in this way:

1. A routine for the PIFOC described in section 3.2 was written which acquired multiple images at interplanar spacings specified by the user.
2. An experienced user focused on the image manually.
3. The PIFOC moved 330 nm away from the slide, then was stepped back through the focal plane along the optical (z) axis of the microscope.

4. Eight images were acquired through the image focal plane at equal 110 nm spacings. The 110 nm spacing satisfies the Nyquist limit for z resolution for light microscopes.
5. The user refocused the microscope manually and took a single picture of the DAPI counterstain in the telomere images for use in later image analysis.

The mean intensity of the bead images did not decrease from the first to the last image, so photobleaching of the dyes was minimal.

2.5 Analysis and Results

The composite images were compared in four different ways: appearance, correction for the focusing artifact, mean square error (MSE) between images, and outcome of the bead coefficient of variance (CV) and telomere correlation tests.

2.5.1 Appearance

In general, the images, shown in figure 2.3, looked the same. The images formed by the Wu and by the maximum IF algorithms, in the top left and right of the figure, show discontinuities at the edges of the 8x8 regions. The background of the averaged image is smooth.

2.5.2 Focusing Artifact

The map images in figure 2.4 show which source images were selected by the Wu, the maximum IF, and the segmentation methods to be copied into each block of the composite images. White blocks in the map images correspond to blocks copied from the bottom of the image stack, while black blocks correspond to images taken

from the top. The legends at the side of each image show the index of the original image from which the blocks were copied.

The map image from the Wu algorithm shows little correlation between the indices of the images selected to fill adjacent blocks, and the choice of original image does not change diagonally across the images. In the composite images from the maximum IF method and the segmentation method, the indices of the blocks are correlated and change from low numbers at the bottom left corner to high numbers in the top left corner. This shows that the maximum IF and segmentation methods corrected the focusing artifact, while the Wu method did not. As well, the segmentation method chose the majority of the blocks from image number 4, which an experienced user judged to be the best-focused image.

2.5.3 Mean Square Error Between Bead Images

To get a quantitative idea of how the images differed, the Mean Square Error (MSE) was calculated. The MSE is given by

$$MSE = \sum_i \sum_j (u_{i,j} - v_{i,j})^2 \quad (2.5)$$

where $u_{i,j}$ and $v_{i,j}$ are the intensities of the (i, j) th pixel in the two images being compared. In general, the MSE between composite images and between the composite images and the best-focused single image (Image 4) is around 2, corresponding to an average pixel-to-pixel difference of around 1.5. This is a reasonable amount of difference to see and can be attributed to differences in the background pixels. However, the averaged image has a MSE difference of 13 to 16 with respect to the other images. Such a difference, corresponding to an average pixel-to-pixel difference of at least 4, can not be attributed to noise, since the magnitude of the background noise in the images is around 3, and there were few background pixels of magnitudes as

high as 7 or as low as 0. Therefore, the high MSE with all other images must be due to large differences in the bright objects in the image. Since these objects are the ones of quantitative interest to us, and it is reasonable to require them to be similar to the original image and all the other images, averaging multiple planes is not a good method for generating composite images.

2.5.4 Bead Fluorescence CV and Telomere q-p Correlation

The main criterion to evaluate these methods was to see how they affect the quantitative analysis of the images. This was measured in different ways for the bead and telomere images.

- The integrated fluorescence (IF) of each bead in the composite image and in the best-focused image were measured using software developed for the lab [27]. The coefficient of variance

$$CV = \frac{\sigma_{IF}}{\bar{IF}} \quad (2.6)$$

for the bead IF distributions, where σ_{IF} is the variance in the integrated fluorescences of the beads and \bar{IF} is the mean bead integrated fluorescence, was expected to be about 10 percent and was expected to be lowest for the best-focused images. In the stack of original images, the CV of the beads decreased monotonically towards the best-focused original image. The composite images of the beads failed to improve the CV of the measured bead intensity distribution. The composite image formed by the segmentation method performed best, but only performed as well as the best-focused original image.

- The telomere composite images, as well as the best-focused original telomere image, were used to measure the IF of the telomeres and calculate the cor-

relation between telomere intensities on the complementary p and q arms of the metaphase chromosomes. In metaphase chromosomes, the DNA has been fully copied. Each chromosome is coiled up and joined near its center with its coiled-up copy. These structures look like X's on the microscope image. The four arms of the X are called the p_1 , q_1 , p_2 , and q_2 arms, where p and q denote the short and long arms of the first and second copies the chromosome. Since the p_2 and q_2 arms of each chromosome are copied from the p_1 and q_1 arms respectively, it is expected that there is a high degree of correlation between the number of repeat sequences on the (p_1, q_1) and (p_2, q_2) telomeres, and that the measured correlation should increase with focus quality. The highest correlations were found in the composite images generated by the segmentation and average algorithms, but these were not significantly higher than that measured for the the best-focused original image.

2.6 Conclusions

The best algorithm for generating composite images was the segmentation algorithm. However it did not perform much better than the best-focused image from the original image stack. Therefore the best-focused image, and not any of the composite images investigated, was used for the experiment in chapter 4.

2.7 Improvements

Several modifications are possible for the composite image algorithms.

- The region size could be modified for the Wu and maximum IF algorithms.

- The cutoff point for the high pass filter used in the Wu algorithm could be modified. It is possible that it was not set sufficiently near the cutoff frequency of the low pass OTF of our microscope system.
- Most of the telomeres and beads in the slides analyzed should lie on the same plane, since they settle on their cover slips. Although the composite images algorithms evaluated here were not useful for our slides, it would be appropriate to test them on thicker objects such as interphase cells.

Chapter 3

Automatic Focusing

3.1 Motivation

Theoretically, the amount of light detected in widefield microscopy from a molecule of fluorescent material should be invariant to whether it is in or out of focus [2]. However Sedat *et al* have documented that epifluorescence microscopes illuminate in-focus objects more intensely than out-of-focus ones because the incident light path runs through the objective lens, focusing the illumination light on one plane of the specimen [3]. This effect is useful since it attenuates some out-of-focus light which would normally blur images. But also means that objects **MUST** be in focus (or an equivalent amount of defocus) for accurate, comparable counts of probe light intensity to be measured.

Therefore, to do quantitative analysis of DNA images, I implemented an automatic routine on the control computer of our microscope to focus the microscope automatically and in a user-invariant way. Other advantages of automatic focusing are that well-focused images are easier to segment and analyze than defocused images, and automatic focusing speeds up image acquisition.

3.2 Hardware

To equip the microscope for automatic focusing, a piezomotive microscope focusing device from Physik Instrumente (model P-721.10) was used. This device, made of a lead-zinc-titanium oxide (PZT) ceramic, mounts between the objective lens and the camera on the microscope. The piezomotive crystal expands and contracts linearly with voltage applied [28] to a range of 100 microns and is controlled by a -10 to 100 V power supply. The device is equipped with an inductive position sensor (linear variable differential transformer, or LVDT) which feeds back a control signal (0 to 10 V) which is proportional to the position of the objective lens. Our power supply (model P-862.50) has feedback electronics which allow the focusing device to be positioned and to know its position to an accuracy of 20 nm. This is well within the microscope's z-axis resolution. The device is novel since it moves the objective lens relative to a stationary microscope stage, rather than moving the stage.

To implement this device I created routines to perform the RS-232 asynchronous serial communications between the microscope acquisition computer and the PIFOC controller. Communications were done using ASCII strings to send commands and receive position information. I encapsulated these routines in functions of the same name as those used in the past to control the microscope's z-drive. As well, I included some new routines which use the on-board memory in the PIFOC power supply. I linked these routines with the DOS microscope acquisition software (Xillix Microimager, 1991) in place of the old stage control commands. Since the acquisition software and the compiler used to create it were very old, I wrote some of the serial communications routines in 486 assembler.

3.3 Methods

To focus on an object of interest, the microscope takes images at several predefined focal planes on the slide, scores them using one of the focusing criteria outlined below, and returns to the plane which scores the highest. The following criteria were evaluated, where $I(i, j)$ is the intensity of the (i, j) th pixel on an N by M image.

1. Absolute image Integrated Optical Density (IOD)

$$Crit_{IOD} = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (I(i, j) - thresh) \quad (3.1)$$

2. Absolute image power

$$Crit_{AP} = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (I(i, j) - thresh)^2 \quad (3.2)$$

3. Quadratic sum of x and y Laplacian Filter Result

$$L_x(i, j) = I(i + 1, j) - I(i, j) \quad (3.3)$$

$$L_y(i, j) = I(i, j + 1) - I(i, j) \quad (3.4)$$

$$Crit_{Lxy} = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (L_x^2(i, j) + L_y^2(i, j)) \quad (3.5)$$

This criterion scores the sharpness of the image edges by combining the outputs of Laplacian edge detection filters applied to the image in the x and y directions.

4. Nonlinear Laplacian Filter Result

$$nL(i, j) = Max_3(A(i, j)) + Min_3(A(i, j)) - 2I(i, j) \quad (3.6)$$

$$A(i, j) = I(i + k, j + l), \quad k = [-1, 1], \quad l = [-1, 1] \quad (3.7)$$

$$Crit_{LnL} = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (nL(i, j))^2 \quad (3.8)$$

This criterion also scores the sharpness of the image edges, but it does not use a linear filter. Instead, the Nonlinear Laplacian Filter calculates the mean of the maximum and minimum pixels from the surrounding 3x3 region $A(i, j)$, then subtracts the intensity of the pixel $I(i, j)$. In this way, it reflects the difference between the pixel at (i, j) and the mean of its surroundings.

5. Poisson-Corrected Power

$$Crit_{Poi} = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (I(i, j) - thresh)^2 - 2I(i, j)\sqrt{(I(i, j) - thresh)} \quad (3.9)$$

This method evaluates the power of the image as in method 2, but it takes into account the fact that the camera CCD obeys Poisson statistics [26] when counting photons, so the error in each pixel's intensity is equal its the square root. $Crit_{Poi}$ subtracts the second term to prevent overestimation of the power at high-intensity pixels.

For methods 1, 2, and 5, the threshold $thresh$ was set using the grayscale histogram of the image. Where $histo_{low}$ and $histo_{high}$ are the low and high shoulders of the histogram peak, the threshold is given by

$$thresh = \frac{3histo_{low} + 2histo_{high}}{5} \quad (3.10)$$

This threshold rejects 60% of the background peak of the image. It has been used in the lab for thresholding PNA images for telomere segmentation [27]. The x and y Laplacian filters were applied using the on-board convolution hardware on the Matrox image acquisition board.

3.4 Results

The best-performing criterion depends on the type of objects being examined. For metaphase chromosome images, the Laplacian filter criteria perform best. However, for beads and telomere images, the Laplacian filter criteria cause the microscope to defocus itself. This is because the Laplacian filters are edge detectors. Since the in-focus bead and telomeres have very small area, their perimeters and the power of their edge images are very small. Out-of-focus beads and telomeres have a larger area and longer perimeters. Therefore a Laplacian filter autofocus routine will always tend to move away from optimum focus of beads and telomeres. For beads and mouse telomeres, the Poisson-Corrected Power, Absolute Power, and Absolute IOD methods all converged to the same image, which also appeared to be the best-focused image by eye.

For human telomere images, which are more faint than bead or mouse telomere images, the Absolute Power method is the only method which works. It is able to accurately focus the telomere images from approximately 600 nm away. This is a sufficient distance that the computer can focus approximately 90 percent of the images that the user can quickly locate by eye.

As it stands now, the automatic focus routine uses the Absolute Power criterion. It adjusts the image acquisition time according to the image intensity, sets the camera to read out 2x2 regions of CCD pixels at a time to reduce acquisition time, takes a number of images (adjustable in a parameter file) at varying focal distances, and returns to the location whose image scores the highest on the focusing criterion. Images for the experiment in chapter 4 were acquired using automatic focusing.

3.5 Improvements

Several steps can be taken to improve the autofocus routine:

- Make both the Laplacian filter and absolute power methods available for focusing. Since both chromosome and telomere images are regularly acquired using the microscope, and since the chromosome focal plane is approximately 300 nm away from the telomere focal plane (due to differing optical paths of the two wavelengths of light in the microscope), it would be useful to add a function which
 1. focuses on the telomeres using the Absolute Power criterion
 2. takes a telomere picture
 3. moves the 300 nm to the approximate chromosome focal plane
 4. focuses using the Laplacian criterion
 5. takes the chromosome picture and stores it in a second image buffer on the Matrox board
 6. saves both images under a single filename with the appropriate 'c' or 'd' suffix.
- Make the search pattern smarter so that the microscope only searches for best-focused images in the direction of increasing focus criterion score. Firestone *et al* found that most autofocus criteria increase monotonically as the optical system approaches optimum focus [4], but this assumption should be checked, since the focus step sizes used in their study are much larger than those considered for our system.

- Fine tune the step number, step size, and number of iterations of the autofocus routine.

Chapter 4

Telomere Organization in Human Interphase Lymphocytes

The multiple plane information described in the previous two chapters was used in an experiment to measure changes in the spatial distribution of telomeres in the nucleus.

If DNA exists as long strands floating randomly through the interphase nucleus, the telomere sequences at the ends of the strands will have a uniform spatial density distribution. However, biology is not so disorganized. There are elements in the nucleus which bind regions of DNA to a nuclear matrix and hold loops of DNA in specific compartments for replication or other activities. It has been demonstrated that the ends of chromosomes are held in place within the nucleus in some species during some phases of the cell cycle [24, 9, 21, 22]. If this the case in human interphase lymphocytes, the telomeres studied in this experiment will show signs of

being spatially organized, rather than being uniformly distributed. Furthermore, if the organizing elements bind directly to the telomere repeat sequences and require sufficiently-long telomeres for function, then the level of spatial organization will decrease in the nuclei studied here which have increasingly shorter telomere repeat sequences.

If the distribution density of telomeres in the nucleus becomes more uniform with decreasing telomere length, this will suggest that loss of nuclear organization is a mechanism of telomere-regulated cell aging.

4.1 Goals and Hypotheses

It is hypothesized that telomere binding proteins such as TRF1 and TRF2, described in section 1.4.2, may be responsible for positioning telomeres within the nucleus and for organizing telomeres into clusters. These proteins have a binding affinity to telomere DNA which increases nonlinearly with telomere length, so it is hypothesized that telomeres must be sufficiently long before sufficient amounts of protein can be bound for biological function. If these hypotheses are true, nuclei should become less organized when their telomeres shorten after repeated cell division.

The first goal of this experiment is to observe how the distribution of telomeres along the radius of the nucleus changes as the telomeres become shorter. To accomplish this, we attempt to reproduce the observation that human white blood cell (lymphocyte) telomeres move radially within the nucleus as the lymphocytes progress through the cell cycle, as reported in [23]. In cells with medium-length or short telomeres, which *in vitro* have reduced ability to recruit telomere binding proteins, we attempt to measure whether the telomere radial distribution becomes more uniform, or whether any non-uniform telomere distributions become less correlated

with the cell cycle.

The second goal is to measure the degree of telomere clustering using features from the Delaunay graph of the telomere locations, then attempt to detect a change in clustering behavior during the cell cycle or in cells with short telomeres.

The third goal is to demonstrate the use of two-dimensional images to study nuclear organization. Since the human lymphocyte cells used in this experiment are roughly spherical, features such as telomere radial location and clustering should ideally be measured in three dimensions using confocal or multiple plane widefield microscopy. When they are squashed onto a slide and their three-dimensional structures are projected into a two-dimensional image, as is done in widefield microscopy, some information about their nuclear organization is lost. However, we assume that our cells retain some of the organization which was inherent in them before they were squished onto a slide and projected onto an image. By measuring changes rather than absolute levels of organization, by analyzing large numbers of nuclei, and by using Delaunay graph features to measure clustering, we attempt to detect the differences in organization levels between populations.

4.2 Measures of Telomere Organization

Three measures are used to characterize the spatial distribution of telomeres in the cell and to evaluate whether telomeres are 'organized': distribution along the radius of the nucleus, minimum spanning tree features, and general Delaunay features.

4.2.1 Radial Distribution

The first measure is to calculate the telomere distribution along the radius of the nucleus. For each of several populations of cells, a cumulative histogram of the

distances of the telomeres from the center of the nucleus is calculated. Then the histogram is compared to a uniform probability density distribution and to the other populations.

It is important to consider the effects of two-dimensional microscopy on our measurements of radial telomere distribution. When Ferguson and Ward made similar measurements in the past [23], they reconstructed three-dimensional microscope images of interphase nuclei, then projected the images in software into several planes to calculate the telomere locations. In our experiment, spherical lymphocyte cells are partially squashed onto a slide, then their features are projected by the microscope optics onto a two-dimensional image. Telomeres near the zenith of the cell in its original form will therefore appear in the squashed/projected image to be close to the center. To determine whether the telomeres are organized in some way along the nuclear radius, we compare their distribution to the following two uniform distributions:

1. If the telomeres have a uniform distribution in the nucleus, and the nucleus remains spherical when it is placed on the slide, the telomeres will have a three-dimensional probability density distribution $p(\rho)$ along the radius ρ given by

$$p(\rho) = \frac{4\pi\rho^2}{4/3\pi R_{max}^3} \quad (4.1)$$

If the uniform spherical distribution is projected down onto one image plane, the uniform probability density $p(r)$, at radius r in cylindrical coordinates, is given by the area of a cylinder shell of radius r and height $h = 2\sqrt{R_{max}^2 - r^2}$, which is twice the height of a hemisphere of radius R_{max} at cylindrical radius r :

$$p(r) = \frac{2\pi r 2\sqrt{R_{max}^2 - r^2}}{4/3\pi R_{max}^3} \quad (4.2)$$

When this '3D uniform' distribution is expressed as a cumulative probability density function, $F(r)$, it is given by

$$F(r) = \frac{\int_{\rho=0}^r 2\pi\sqrt{R_{max}^2 - \rho^2}2\rho d\rho}{4/3\pi R_{max}^3} \quad (4.3)$$

$$= 1 - \frac{4/3\pi(R_{max}^2 - r^2)^{3/2}}{4/3\pi R_{max}^3} \quad (4.4)$$

$$= 1 - \left(1 - \left(\frac{r}{R_{max}}\right)^2\right)^{3/2} \quad (4.5)$$

where r is the radial location of a telomere and R_{max} is the nuclear radius.

A distribution of telomeres similar to the one described in 4.5 would suggest that the nuclei on the slide retained their spherical shape but that the telomeres were uniformly distributed, not positioned within the cell by organizing forces.

2. If the three-dimensional conformation of the nucleus is completely destroyed by specimen preparation, the telomeres will be splattered all over the slide and will have a 'fried egg' uniform probability density distribution $p(r)$ in 2 dimensions given by

$$p(r) = \frac{2\pi r}{\pi R_{max}^2} \quad (4.6)$$

The uniform cumulative distribution $F(r)$ is given by

$$F(r) = \frac{\int_{\rho=0}^r 2\pi\rho d\rho}{\int_{\rho=0}^{R_{max}} 2\pi\rho d\rho} \quad (4.7)$$

$$= \left(\frac{r}{R_{max}}\right)^2 \quad (4.8)$$

This would be an unfortunate distribution to observe in an experiment because it would suggest that most of the nuclear structure was destroyed during cell preparation for the two-dimensional microscopy.

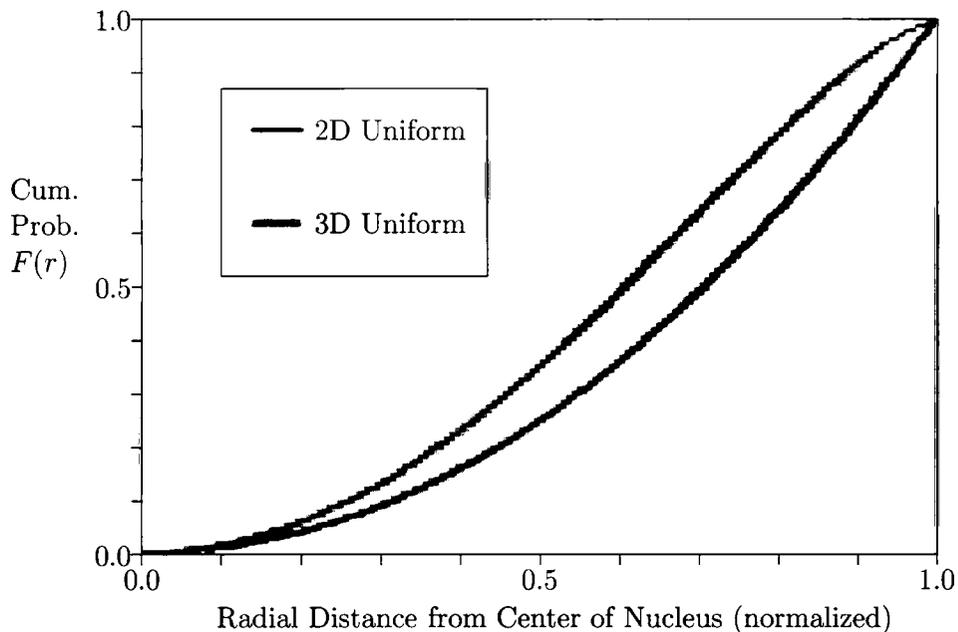


Figure 4.1: Uniform Cumulative Probability Distributions for Telomeres

The cumulative probabilities $F(r)$ reflect the fraction of all telomeres which lie within any given radius r . The uniform probability distributions from equations 4.5 and 4.8 are shown in figure 4.1. For each population of cells studied in this experiment, the radial distribution of all telomeres in the population is compared with the proposed uniform distributions and with the distributions of telomeres from other cell populations.

4.2.2 Minimum Spanning Tree Features

The second measure of organization is to measure telomere clustering using a Minimum Spanning Tree (MST) graph. To do this, a Delaunay tessellation is first constructed using the telomere centers of mass as the vertices of Delaunay triangles. The Delaunay triangulation is a graph which

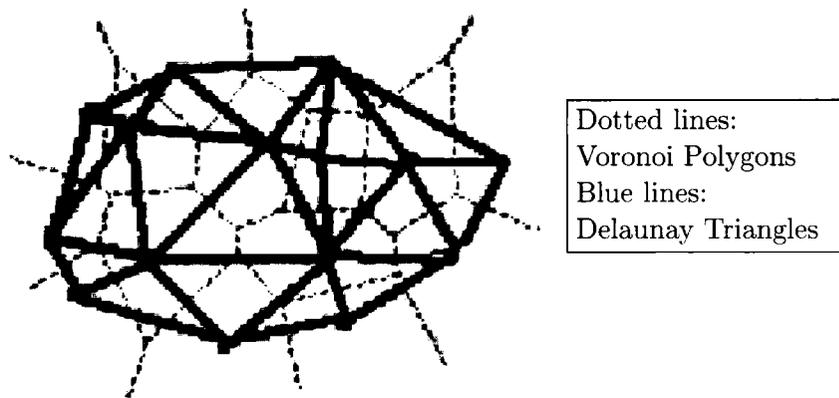


Figure 4.2: Sample Voronoi and Delaunay Graphs

- spans a set of noncollinear points
- is the dual of the Voronoi diagram, a graph which divides graph space into regions whose closest neighbour (generator) is the point at the center of a Voronoi polygon. A Voronoi diagram is shown in figure 4.2.
- whose vertices are the generators of the Voronoi polygon and whose segments are perpendicular bisectors of the Voronoi edges.

The Delaunay graph's properties have been used extensively to quantify point distributions and to compare them to the Delaunay graphs of complete spatial randomness [29]. However in small graphs the Delaunay graph's features, such as triangle internal angles and segment lengths, are very sensitive to the arrangement of the points around the edge of the graph. So the Minimum Spanning Tree (MST) subgraph was used as the second measure of telomere organization. The MST connects all the points in the graph space using segments from the Delaunay graph so that the graph length is a minimum. If the points in the graph are evenly spaced,

the segment lengths of the MST will be equal and the variance and higher-order moments of the histogram of segment lengths will be small. In highly-clustered distributions, segment lengths will differ widely and the histogram moments will be large. It is difficult to calculate a distribution corresponding to complete spatial randomness for objects in the lymphocyte nuclei, but if MST features change monotonically with shortening telomere length, that is sufficient evidence to show that telomere length influences nuclear clustering.

4.2.3 General Delaunay Features

The third measure of organization is to measure other features from the Delaunay graph and to map their change as telomere length shortens. The features chosen for use are the average distance from each point to its 1, 2, and 3 nearest neighbours. These features could reflect tightness of pairing or clustering of telomeres.

4.2.4 1p Probe Distribution

In this experiment, the lymphocyte DNA was labeled using a fluorescent probe specific to telomeres, and also with a "1p" probe which binds to a region near the telomere on the short (p) arm of chromosome 1. Since the probe is located close to the 1p telomere, changes in 1p probe distribution reflect changes in the 1p telomere's location. Since there are only 2 1p probe spots in a G1 nucleus and 4 1p spots in a G2 nucleus, detection of the 1p probe spots in each image is much easier than telomere detection. So the 1p probe spots' distribution along the radius of the nucleus was measured in order to track the distribution of the nearby telomere.

4.3 Expected Differences

Six populations of lymphocytes were analyzed. Young, medium, and old cells with long, medium, and short telomeres, as described in section 4.4.1, were sorted into populations which were in the G1 (prior to DNA replication) and S/G2 (during and post-DNA replication) cell cycle phases.

There are two possible trends in the degree of telomere organization in the nucleus which could be observed: Organization could change as cells cycle from G1 to S/G2 phases, and it could change as telomere length decreases.

4.4 Materials and Methods

4.4.1 Lymphocyte Cells and Preparation

The lymphocyte cells used in this experiment were frozen cells previously grown for a study of telomere length shortening with clone passage [30]. The cells were originally obtained from umbilical cord blood and were cloned repeatedly. The samples used for this analysis were taken from the first, 6th, and 11-12th data points from that experiment and represent cells which had undergone 20, 120, and 240 population doublings respectively. The cells were shown using flow-FISH and Southern analysis to display telomere shortening of approximately 75 base pairs per population doubling. Therefore the cells in the first data set had the longest telomeres while those in 11-12th data set had the shortest ones.

The cells had been frozen for about 3 months prior to the experiment. On thawing, they were grown in culture for a week and then sorted into G1 (prior to DNA replication) and S/G2 (during and post-DNA replications) cell cycle stages using a flow cytometer. This created the following data sets:

| Population | Data Point | Telomere Length | Expected Telomere Count |
|------------|------------|-----------------|-------------------------|
| G1-1 | 1 | long | 92 |
| S/G2-1 | 1 | long | 92-184 |
| G1-2 | 6 | medium | 92 |
| S/G2-2 | 6 | medium | 92-184 |
| G1-3 | 11-12 | short | 92 |
| S/G2-3 | 11-12 | short | 92-184 |

The lymphocytes from each population were fixed with methanol acetic acid, dropped onto slides, treated with 70% formamide to melt their DNA into single strands, and hybridized with three fluorescent molecules:

- A protein-nucleic-acid (PNA) probe specific to the telomere region, labeled with a CY3 fluorescent molecule
- A PNA probe specific to the shortest arm of chromosome 1 (band 36 on the 1p arm) near the terminus, labeled with an FITC fluorescent molecule
- DAPI fluorescent dye, which is a general DNA stain fluorescing in the blue region

The cells were not lysed using hypotonic buffer, as is commonly done during fluorescence *in situ* hybridization (FISH) preparation, in order to avoid unnecessary disturbance to the nuclear structure. One microscope slide was prepared for each lymphocyte population.

4.4.2 Image Acquisition

Images were acquired using the system described in section 1.1.1: A Zeiss Axio-plan microscope with a fluorescent mercury/xenon light source. The microscope was fitted with a Zeiss 63x oil immersion lens and a colour filter set with four excitation filters (specific to CY3, FITC, DAPI, and CY5 excitation wavelengths) and a dichroic mirror and emission filter which admit emitted light from the 4 fluorochromes. The filter set had a high band rejection of DAPI and FITC light when using the CY3 filter, but allowed approximately 20% of CY3 signal into the FITC and DAPI images. The CY5 channel was not used in this experiment.

The microscope was controlled using a personal computer and was fitted with the Physik Instrumente piezoelectric focusing device described in section 3.2 mounted between the objective lens and the body of the microscope. The algorithm described in chapter 3 was used to automatically focus on the CY3 telomere image. On instruction from the user, the computer moved the objective lens in 120 nm steps in the z-direction, acquiring images at each of the resulting focal planes, and returned to the focal plane which yielded the highest power in its image according to criterion 2 in chapter 3. The telomere image was acquired, the user switched filters, and the computer acquired green and blue images of the FITC and DAPI dyes at the same focal plane.

Images were acquired using a Xillix MicroImager digital camera with an image pixel size of 6.8 μm and a field size of 1280 x 1024 pixels. Images were saved in 8-bit grayscale in a simple "img" file format specific to BC Cancer Agency Cancer Imaging.

For each lymphocyte population, all images were taken over a single day. Since some of the slides had sufficiently large numbers of cells to have several cells in

a microscope field of view, the entire field of view was acquired. On later processing, the field of view was automatically segmented into one-cell images by

1. Thresholding the DAPI counterstain image at the peak of the image histogram. Since the full field-of-view images are dominated by background, this is an effective way to segment.
2. Eroding and dilating the images to clean up ragged edges
3. Labeling and selecting objects of size greater than 5000 pixels
4. Selecting rectangles 100 pixels wider than the object boundaries from these images
5. Combining the corresponding rectangles from the red, green, and blue images of the cell plane into 24-bit TIFF files. This file format was chosen because it was easy to read using commercial programs and because image acquisition information (exposure time, date, etc.) and intermediate experimental results (such as the location of the 1p probes) could be stored within the image at private tags.

The following numbers of images and nuclei were analyzed:

| Population | Number of Images | Number of Nuclei |
|------------|------------------|------------------|
| G1-1 | 35 | 92 |
| S/G2-1 | 33 | 46 |
| G1-2 | 25 | 92 |
| S/G2-2 | 25 | 56 |
| G1-3 | 25 | 65 |
| S/G2-3 | 25 | 51 |

Since the goal of this experiment was detection and localization of telomeres rather than measurement of light intensity, no attempt was made during image acquisition to correct for drifts in illumination intensity or for uneven illumination of the field of view.

4.4.3 Colour Plane Registration

Each of the lymphocyte images is a three-colour image with red, green, and blue planes acquired using the CY3, FITC, and DAPI microscope filters. Because of movement in the optical system caused either by vibrations in the table or by jolting of the microscope when the color filters moved, the three planes were often shifted by several pixels relative to each other. It was necessary to correct for these small shifts before doing further analysis of the images. So the green (FITC) plane was aligned with the red (CY3) plane by maximizing the correlation between the two planes.

To align, the images were thresholded, eroded, and cleaned of isolated pixels so that the images were essentially telomeres and 1p probes on a black background. The moving image was then shifted -2, -1, 0, 1, 2 pixels in the x direction. At each

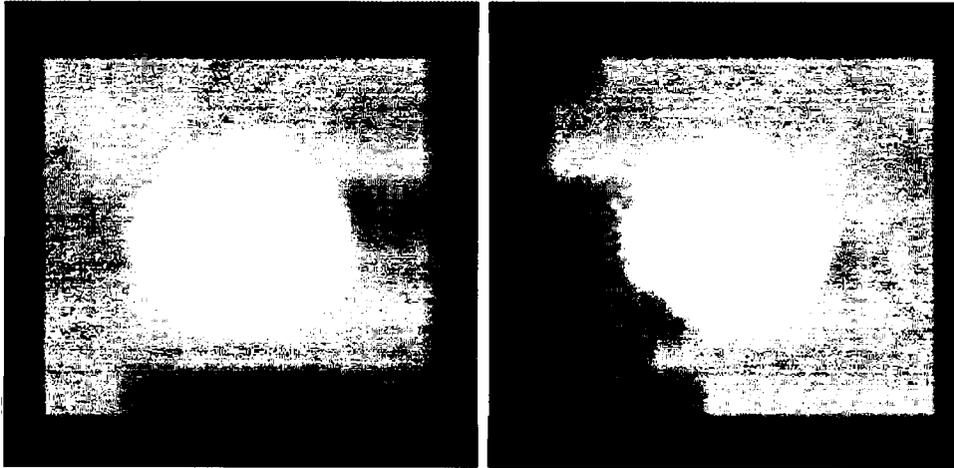


Figure 4.3: Lymphocyte Images Before and After Colour Plane Alignment

shift, the correlation between images was calculated as

$$correlation = \frac{\sum (u_i - \bar{u})(v_i - \bar{v})}{N(N - 1)\sqrt{\sigma_u} * \sqrt{\sigma_v}} \quad (4.9)$$

where u_i , \bar{u} , v_i , and \bar{v} are the i th pixel and expectation of the two images u and v , N is the number of pixels, and σ_u , σ_v are the variances of the images u and v .

The moving image was shifted relative to the reference image by the number of pixels required to maximize the correlation between the two images. Then the process was repeated using -2, -1, 0, 1, 2 shifts in the y direction. If either of the optimum x or y shifts were 2 or -2 pixels, the shift process was repeated.

Registration and alignment were easy for the FITC image relative to the CY3 telomere image because of the high level of optical crosstalk between the CY3 and FITC channels, as discussed in section 4.4.4. Successful alignment is demonstrated by the fact that small lymphocyte images cut from the same large 1024 x 1280 pixel microscope image, processed separately, all require shifts of the same number of pixels, as shown in table 4.4.3. Examples of a G1 image before and after pixel shifting are shown in figure 4.3.

| File Name | X Shift | Y Shift |
|---------------|---------|---------|
| 1123003x0.tif | 0 | 0 |
| 1123003x1.tif | 0 | 0 |
| 1123003x2.tif | -1 | 1 |
| 1123003x3.tif | -1 | 1 |
| 1123003x4.tif | -1 | 1 |
| 1123005x0.tif | 0 | 0 |
| 1123005x1.tif | 0 | 0 |
| 1123006x0.tif | 1 | 0 |
| 1123006x1.tif | 1 | 0 |
| 1123006x2.tif | 1 | 1 |
| 1123007x0.tif | 3 | 1 |
| 1123007x1.tif | 3 | 1 |
| 1123007x2.tif | 3 | 1 |
| 1123007x3.tif | 3 | 1 |
| 1123008x0.tif | 0 | 0 |
| 1123008x1.tif | 0 | 0 |
| 1123008x2.tif | 1 | 0 |
| 1123008x3.tif | 0 | 0 |
| 1123008x4.tif | 0 | 0 |
| 1123008x5.tif | 0 | 0 |
| 1123009x0.tif | 3 | 2 |
| 1123009x1.tif | 3 | 2 |
| 1123009x2.tif | 3 | 2 |
| 1123009x3.tif | 3 | 2 |
| 1123009x4.tif | 3 | 1 |
| 1123009x5.tif | 3 | 1 |

Table 4.1: Pixel Shifts of FITC Relative to CY3 Image

Registration and alignment of the DAPI images with the CY3 telomere image were not done, since there is very little crosstalk between those colour channels to use in alignment, and since it was not necessary for the images to line up exactly.

4.4.4 Correction for Optical Crosstalk

There is a high level of optical crosstalk from the CY3 (telomere) signal into the FITC (1p probe) image because their excitation wavelengths and emission wave-

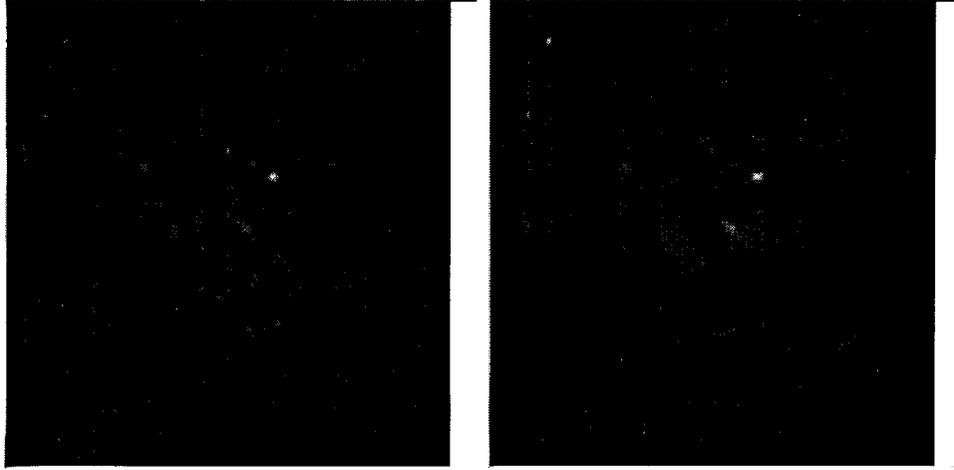


Figure 4.4: Green Image Before Crosstalk Correction and Under-Corrected (10%)

lengths are close ([512 nm, 565 nm] for CY3 and [490 nm, 525 nm] for FITC) and the bandpass filter used to capture the green FITC image has a low attenuation in the CY3 region of the spectrum. The crosstalk from the telomere image into the 1p probe image, therefore, is about 20 percent. The 1p probe image is significantly contaminated with images of telomeres, so the 1p probe signals are difficult to discriminate from the similar-sized telomeres, as shown in figure 4.4. Ideally, figure 4.4 should show only two fluorescent 1p probe spots.

The corrected FITC image was calculated using

$$u_{FITC,corr} = u_{FITC,observed} - k(time, LUT) * v_{CY3,observed} \quad (4.10)$$

$$k(time, LUT) = k_{crosstalk} \frac{time(FITC)(\Delta LUT(CY3) + 1)}{time(CY3)(\Delta LUT(FITC) + 1)} \quad (4.11)$$

where $u_{FITC,corr}$, $u_{FITC,observed}$, and $v_{CY3,observed}$ are the corrected and observed FITC and CY3 images; $time(CY3)$ and $time(FITC)$ are the exposure times; and $\Delta LUT(CY3)$ and $\Delta LUT(FITC)$ are the difference in the maximum and minimum values of the lookup table used for each image when converting the digital camera's

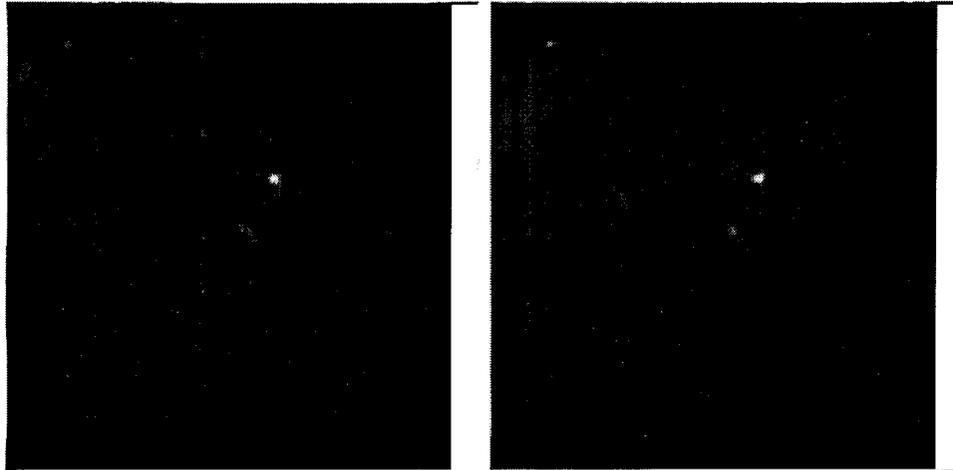


Figure 4.5: Optimally-(20%) and Over-(28%) Corrected Green Images

14-bit output into eight bits for storage.

The constant $k_{crosstalk}$ was determined by doing a series of corrections on FITC images using different constant values, then inspecting the images to see which correction factor yielded FITC images showing two 1p probe spots on a uniform background. Under-correction resulted in telomere images on the FITC image, while over-correction resulted in holes in the image where the telomeres were located, as shown in figures 4.4 and 4.5. The correct factor was found to be 0.20. Correction results were dramatic, and the coefficient was found to be consistent throughout the experiment. An FITC image which was corrected by subtracting 20 percent of the corresponding telomere image is shown in figure 4.5. Occasionally when a telomere image was overexposed and some pixels had values which had been truncated at 255, the FITC image was undercorrected in the overexposed region.

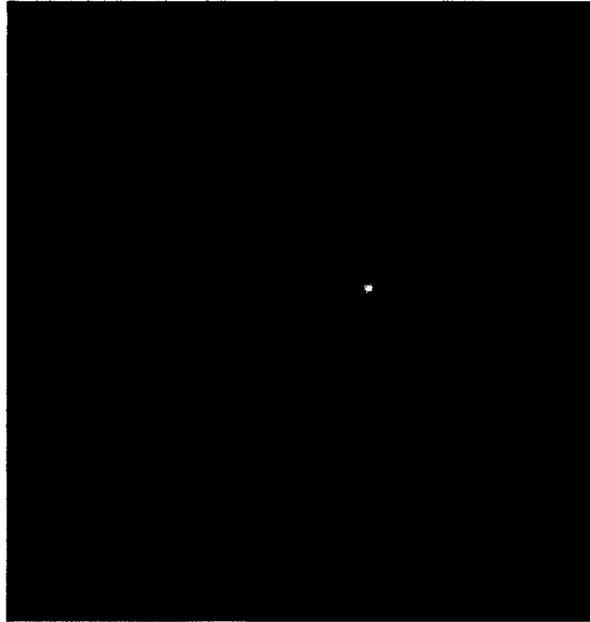


Figure 4.6: Lymphocyte Image

4.4.5 Nucleus Detection

Locating the edges of the nucleus was difficult and was done using gradient-weighted thresholding. As described in section 4.4.1, the lymphocyte nuclei were counterstained with DAPI, a nonspecific DNA stain whose intensity roughly corresponds to the DNA density in the nucleus and which is often used for identifying banding patterns on metaphase chromosomes. Since, in interphase, DNA is uncoiled and distributed throughout the nucleus, the nucleus appears as a blue blob without a definite edge, as shown in figure 4.7. For measuring telomere distribution, it was necessary to have a reproducible way to locate the lymphocyte's border and to calculate its center and radius, both for calculating the radial location of the telomeres and for normalizing inter-telomere distances to nucleus size.

Initially, the nuclear border was determined from the DAPI image histogram,

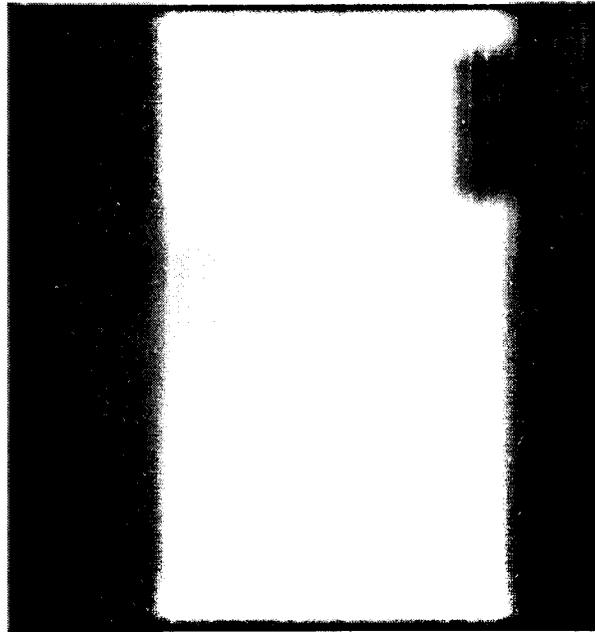


Figure 4.7: DAPI Counterstain

as shown in figure 4.8. The top and bottom edges of the histogram noise peak were located, and the background threshold was set at 60% of the difference between those gray levels, resulting in a threshold that rejected 60% of the noise peak. This threshold is used extensively in the lab to threshold DAPI metaphase chromosome images [27]. However, histogram-based thresholding was not reproducible enough to determine nuclear radii in this experiment for three reasons. First, the nuclei do not have well-defined boundaries. Second, the exposure times, degree of slide bleaching, and signal to noise ratios of the images used in this experiment varied widely. Third, the images had often been segmented out of larger full-field microscope images, and their DAPI histograms were often dominated by the image itself, rather than a background peak.

Therefore a gradient-weighted threshold was used to determine cell radius

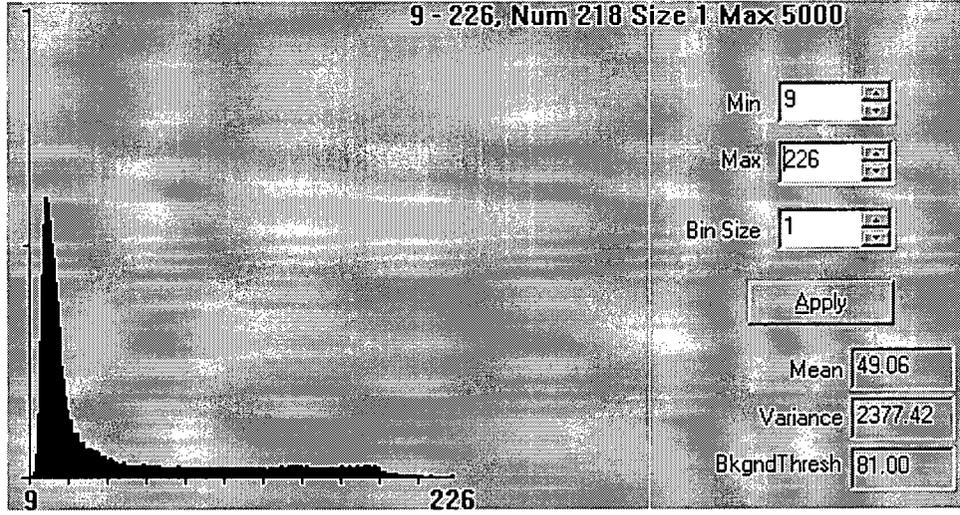


Figure 4.8: Sample DAPI Image Histogram

and location. In this scheme, the approximate edges of the DAPI image were determined by finding the pixels of largest gradient on the image. Then the mean pixel intensity at those edges was used as a threshold to segment the nucleus image. The following steps were taken:

- The DAPI image was median filtered in the x-direction to remove noise, then Sobel filtered by convolving the image with the following kernels:

$$\mathbf{H}_1 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{pmatrix} \quad \mathbf{H}_2 = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{pmatrix} \quad \mathbf{H}_3 = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{pmatrix}$$

The kernels H_1 , H_2 , and H_3 are the Sobel filters to detect edges in the horizontal, diagonal, and vertical directions. The Sobel image intensity $g(m, n)$ at the (m, n) th pixel was taken as

$$g(m, n) = \max(\text{abs}(g_1(m, n)), \text{abs}(g_2(m, n)), \text{abs}(g_3(m, n))) \quad (4.12)$$

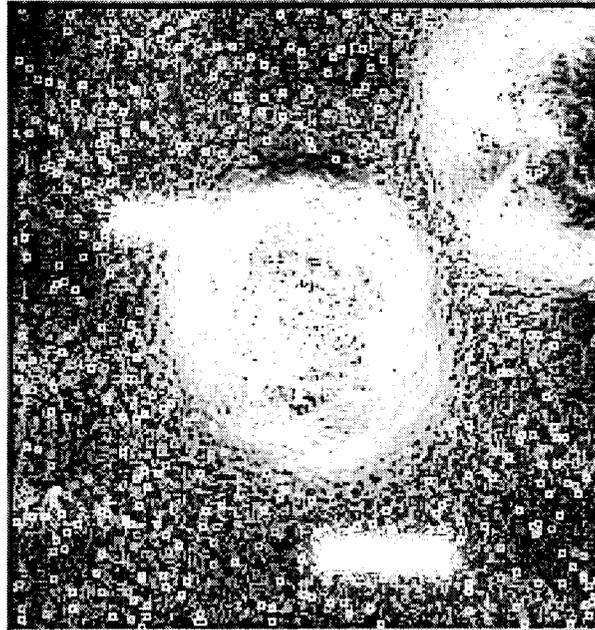


Figure 4.9: Sobel Image

where $g_1(m, n)$, $g_2(m, n)$, and $g_3(m, n)$ are the pixel intensities at (m, n) of the convolution of the DAPI image with the kernels H_1 , H_2 , and H_3 . The Sobel image is shown in figure 4.9.

- The Sobel image was thresholded.
- To exclude cells or other objects at the boundaries of the image, a mask was created from the original DAPI image. To create the mask, the DAPI image was thresholded using 60% of the background peak, the regions on the image were labeled, and the largest region was selected. To ensure that the mask had smooth edges and was sufficiently large, it was dilated into a convex object

using the operator

$$\mathbf{H}_1 = \begin{pmatrix} - & 1 & 1 \\ - & - & 1 \\ - & - & 1 \end{pmatrix}$$

and the mask was applied to the Sobel image.

- The outside edge of the Sobel image was found, then the edge was thinned and pruned. Such a thinned Sobel image is shown in figure 4.10.
- The thinned Sobel image was applied as a mask to the original DAPI image. The mean pixel intensity under the Sobel image outline was calculated and used as a threshold to find the DAPI nucleus.

After the nucleus was labeled on the DAPI image, a minimum bounding circle was fitted to the labeled region, as shown in figure 4.11. The circle's parameters were saved in the image file and were used to calculate of the radial location of the telomeres in the nucleus, and to normalize the graphs of the telomere clusters.

4.4.6 Telomere Detection and Localization

The telomere detection algorithm was the most important one written for this application. The following methods were tried for telomere detection. Eventually method 4 was used.

1. Initially, spots were counted by thresholding the telomere image based on the image histogram. The telomeres were repeatedly eroded, dilated and labeled using a pixel label algorithm. Touching spots could sometimes be resolved using extra erosions on regions with areas or integrated fluorescence significantly larger than the mean values. This method worked well on images of

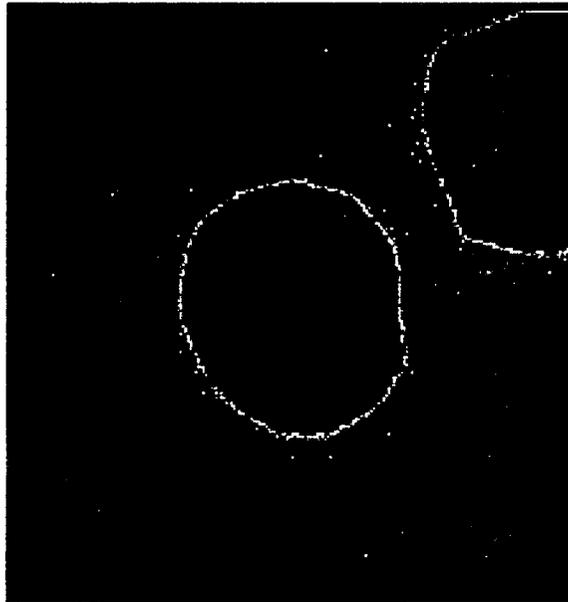


Figure 4.10: Thinned Sobel Image

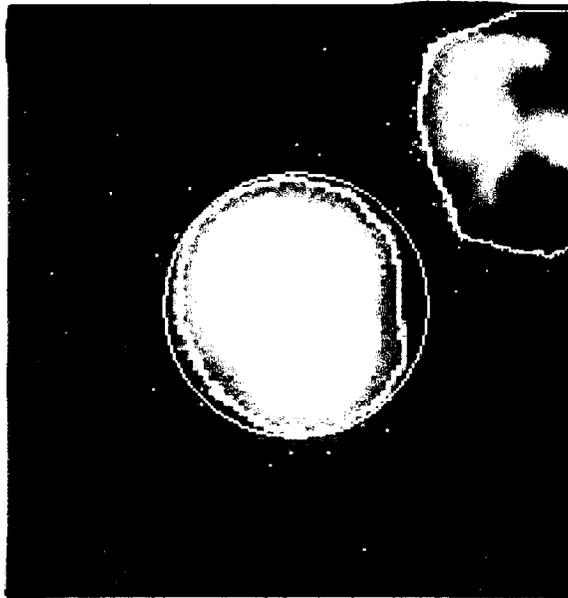


Figure 4.11: Segmented DAPI Nucleus

same-sized fluorescent beads. However it was very sensitive to original threshold selection, had difficulties resolving closely-spaced spots, and did not work well on telomere images or spots of varying intensities.

2. The tophat transform in equation 4.13 was applied to the image, then thresholded to detect spots. The tophat transform $Tophat(i, j)$ subtracts the local background from an image $I(i, j)$ with small objects on it. The filter uses a small kernel, B , of approximately the size of the spots to be detected. It first excludes the bright spots (such as telomeres) by applying a minimum filter Min_B to the image. It then finds the maximum local background by applying a maximum filter Max_B to the recently-obtained minimum image. The transform is given at the $(i, j)th$ pixel by

$$Tophat(i, j) = I(i, j) - Max_B(Min_B(i, j)) \quad (4.13)$$

Where $I(i, j)$ is the pixel intensity at (i, j) and $Max_B()$ and $Min_B()$ are the maximum and minimum pixel intensities over the region B centered at i, j . Square kernels of size 5x5 and 7x7 were used for the maximum and minimum filters. The resulting tophat image was thresholded to detect dots. This algorithm detected more of the dots than the previous one, but it did not resolve closely-spaced dots.

3. A nonlinear Laplacian filter to find the edges of dots worked far better on the telomere images. This filter $L(i, j)$ at the $(i, j)th$ pixel is given by

$$L(i, j) = Max_3(i, j) + Min_3(i, j) - 2I(i, j) \quad (4.14)$$

Where $I(i, j)$ is the pixel intensity at (i, j) and $Max_3(i, j)$ and $Min_3(i, j)$ are the maximum and minimum pixel intensities over the 3x3 region centered at

(i, j) . This algorithm detects changes in image intensity (edges) and is able to detect the dips in intensity between nearby spots. The nonlinear Laplacian transform image was thresholded to obtain a dot mask. The nonlinear Laplacian filter was effective for detecting small spots and sometimes counted between 90 and 100 spots on high signal-to-noise images, where 92 spots per image would be expected. However, the number of dots counted was threshold-dependent and sensitive to noise, so the algorithm counted too many spots in noisy images. As well, the algorithm often detected arcs along the perimeter of large spots and counted them as several separate spots.

4. The most effective telomere detection method was the Dot Label algorithm described by Netten [31]. This region growing algorithm involves selecting a high 'seed threshold', identifying all pixels above this threshold, and labeling disconnected regions as dot 'seeds'. The dot mask is expanded by reducing the threshold to a 'connecting threshold' which is three camera standard deviations below the seed threshold. The camera standard deviation is approximately the square root of the pixel intensity [26]. All pixels which are above the connecting threshold and which are connected to an existing dot are labeled as part of that dot. The algorithm iterates by selecting lower and lower seed thresholds, with connecting thresholds three camera standard deviations below. It labels dots at all seed thresholds down to just above background intensity. This method is less sensitive to noise than the nonlinear Laplacian filter and deals with large dots effectively.

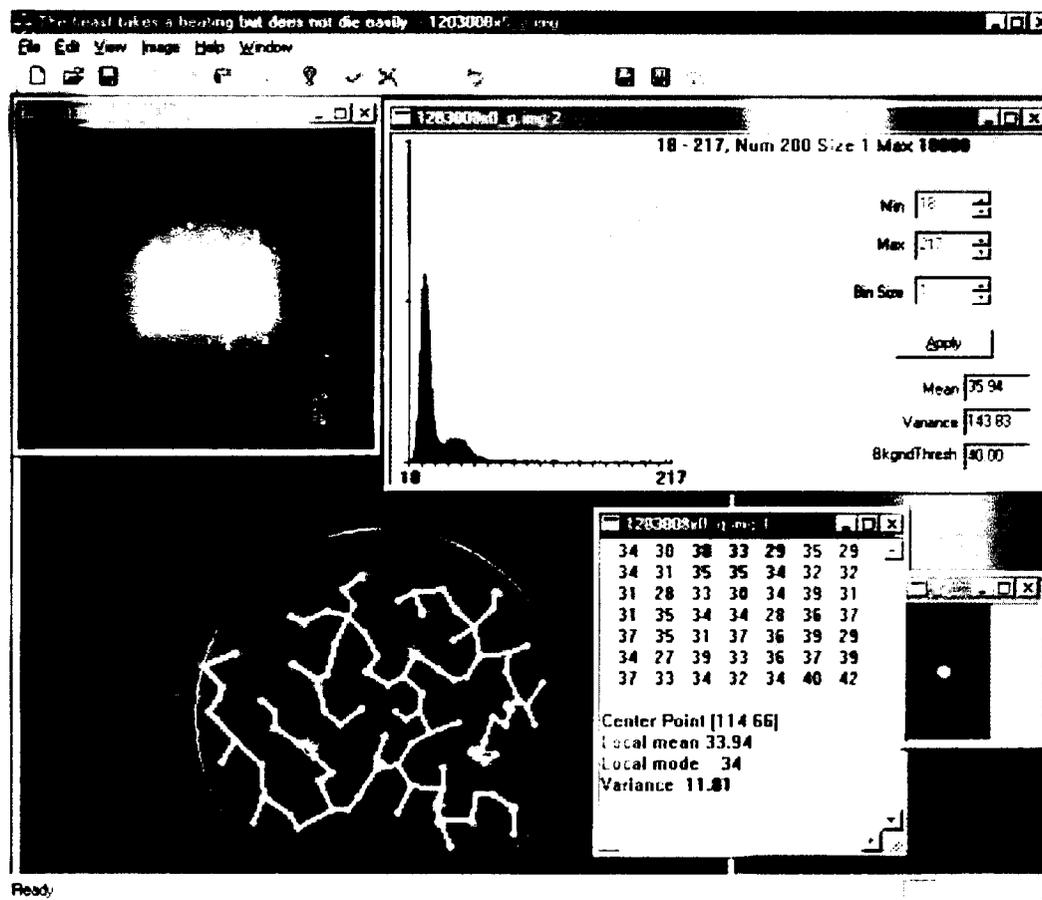


Figure 4.12: Spots2.exe Program

4.4.7 Details of Telomere Detection

Telomeres were detected using a PC program named “spots2.exe”. It was written to automate most aspects of the image analysis, including combination of the three colour planes; segmenting and saving nuclei out of the wide microscope image plane; image display, enhancement, and filtering; Delaunay graph calculation; and telomere detection. A screen shot of the application is shown in figure 4.12.

The Dot Label algorithm was used to detect telomeres in each CY3 image. The parameters for the algorithm were

| | |
|------------------------|-----------------------------------|
| Initial seed threshold | Image maximum- σ |
| Step size σ | Variance in image background peak |
| Connection threshold | Seed threshold- σ |
| New seed threshold | Old seed threshold- σ |
| Lowest seed threshold | 60% of background peak |

The variance σ in the image background peak was found from the image histogram. It was taken to be the difference in gray levels between the image mode (the gray level with the largest number of counts in the histogram) and the nearest gray level with 20% as many pixels as the mode. This difference was generally 3 to 4 gray levels and would correspond to two standard deviations in the background pixel intensity if the background noise had a Gaussian distribution. This is a much smaller threshold increment than that used in Netten [31] or suggested by Bevington [26]. However, it was necessary to use the small number in order to resolve close-together telomere spots.

At each iteration, a mask was generated containing all pixels above the seed threshold which had not been previously labeled. Objects from the mask were labeled using the region labeling described by Jain [32]. A second mask was generated containing all pixels above the connection threshold, and labeled objects were dilated within the connection image. Then the seed and connection thresholds were reduced by σ and the process was repeated while the seed threshold was still greater than the image background peak.

The small decrease in threshold σ required a large number of iterations of the algorithm, and it generally required two to three seconds to analyze each image on a Pentium 133 computer with 32 MB of memory.

Labeling dots by dilating objects repeatedly meant that some dots were dilated into the surrounding background. To correct for this, after the iterations were complete, each dot's pixels were examined, and pixels with less than 30% of the dot's maximum intensity were discarded. This requirement worked very well for preserving the round nucleus of each dot, although it caused errors in calculating the total integrated fluorescence of each probe spot.

Statistics were then gathered for each dot including its roundness, area, integrated fluorescence (IF), center of area, center of mass, and maximum pixel intensity. The roundness of a dot is given by

$$Roundness = \frac{4\pi Area}{Perimeter^2} \quad (4.15)$$

A perfect circle has roundness 1.0. Dots of roundness less than 0.60 were discarded. Dots of area less than 5 pixels were also discarded.

The Dot Label algorithm effectively found almost all the spots detectable by eye on the telomere images and in the histogram-equalized enhanced telomere images. It was limited in three ways, however, and the program was modified to allow the user to manually split or delete spots.

1. While the algorithm was able to split nearby telomeres with two similar-sized peaks, it was unable to resolve some small telomeres next to a large telomere when the small telomere did not have a peak to act as a seed for the region growing algorithm. These telomere pairs appear on the image as a bright spot with a tail. To split such a spot in the spots2.exe Dot Label program, the user clicks on the spot in the telomere image to bring up the Windows dialog box shown in figure 4.13 with a histogram-enhanced view of the local region. The user can examine the enhanced image and click on the center of a

smaller missed spot, if necessary. The program then splits the original spot's area into newly-specified objects by thresholding the local region, using the newly-specified centers of the objects as seeds, and dilating the new regions within the mask.

2. The lymphocytes were quite thick along the axis normal to the slide, and some telomeres were very out-of-focus. They appear on the images as rings, as shown in figure 4.14. The point spread function of a typical microscope's optics is a Bessel function, and the rings are likely the first maxima of such a point spread, as described in equation 2.1 [3]. The Dot Label algorithm sometimes labeled several spots on these rings without labeling the dim center of the telomere. In these instances, the user had to manually move one of the spots from the first maximum ring to the center of the ring and delete the other spots.
3. On very noisy images, such as those collected from the long-telomere S/G2 lymphocyte slide, the Dot Label algorithm labeled noise pixels as telomeres. To remove noise telomeres, the user generally clicked on possible noise telomere candidates to bring up a "delete telomere" Windows dialog box (figure 4.15) with an enhanced image of the noise telomeres to decide whether to delete them. This delete process required the largest amount of subjective input from the user. On average, ten (about 10%) of the telomeres on images from the noisy S/G2 long-telomere slide were noise peaks and needed to be deleted. On images with a high signal to noise ratio, fewer than one telomere per image was deleted.

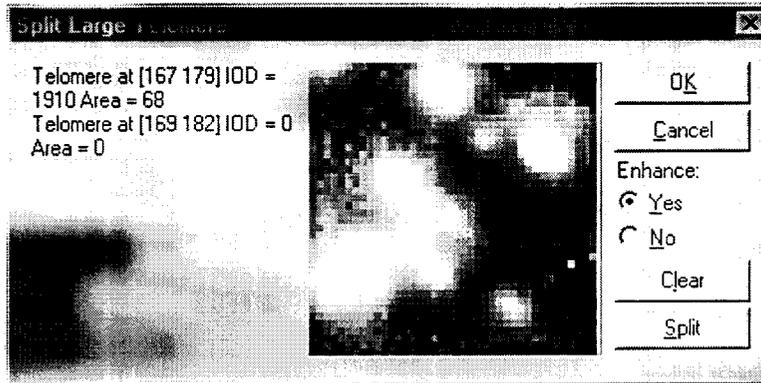


Figure 4.13: Split Telomere Utility

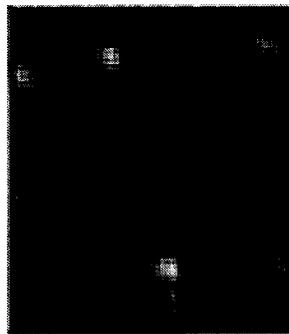


Figure 4.14: Light Spread from Out-of-Focus Telomere

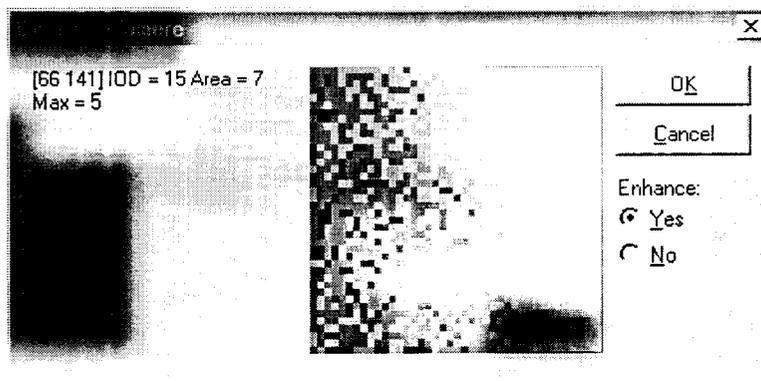


Figure 4.15: Delete Telomere Utility

4.4.8 1p Probe Detection

The fluorescent spots from the FITC probes on the short arm of chromosome 1 (1p probes) were detected in the same way as the telomeres, using the region growing algorithm described in [31]. The noise threshold was very low in the 1p images, since much background had been subtracted during the crosstalk correction described in section 4.4.4, so the algorithm was iterated until its connection threshold reached 2 or 3 gray levels. The step size σ was calculated using the full width at half maximum of the background peak of the image histogram. The morphological features used to discriminate 1p probe spot objects were an area greater than 4 pixels and a roundness greater than 0.5. In addition, it was known that there should be only two probe spots in each G1 lymphocyte image, since there are 2 alleles of chromosome 1 in G1 nuclei. In the G2 images, there should be 2 to 4 spots, since there are 2 to 4 copies of chromosome 1, depending on whether the DNA in the cell has been fully copied. Because of this, the program was modified to report the 2 or 4 brightest objects labeled in 1p images. If the next-brightest object was within 50% of the brightest objects, a window appeared in the analysis program to allow the user to manually select the desired 1p spots and number of 1p spots. This feature was useful because occasionally the 1p images had bright streaks or splotches which were obviously not probe spots, and because, if the telomere image had been overexposed, the color crosstalk correction did not remove very bright telomeres from the 1p images.

When the spots were detected, their locations were saved in the image file for later comparison with the telomere locations in the image. Their radial locations in the nucleus were calculated, a cumulative histogram of the locations was computed, and their locations were analyzed using the Kolmogorov-Smirnov test described in

section 4.5.2 to see whether the 1p spot was uniformly distributed radially within the cells.

4.5 Analysis and Results

4.5.1 Number and Intensity of Telomeres Detected

The program did not detect all the telomeres in the lymphocytes studied. In cells in G1 phase of the cell cycle, two telomeres for each of the 46 chromosomes, or 92 telomeres in total, should be present. In the S/G2 phases, the cells should have between 92 telomeres (in early S phase, when the cell has just begun to replicate DNA) and 184 telomeres (in the G2 phase, when the cells have two copies of each chromosome). In reality, the program detected between 59 and 88 percent of the telomeres in the G1 populations. Its detection efficiency for the S/G2 telomere populations is difficult to calculate, since it is unknown how many telomeres were actually present in those cells. The nuclei showed markedly different sizes and numbers of telomeres in the S/G2 populations, suggesting that some cells were in early S phase (92 telomeres) while others were in G2 phase (184 telomeres). The algorithm detected an average of 65 telomeres for the smaller G1 cells and 130 for the largest S/G2 cells, which indicate detection efficiencies of 70 percent. The average number and intensity of the telomeres in each cell are shown in table 4.2.

The reasons for the low number of telomeres detected could be several. First, the lymphocytes could be sufficiently thick on the z-direction (normal to the slide) that some of the telomeres were too far out of focus to detect. Out-of-focus telomeres with obvious Bessel rings such as the one shown in figure 4.14 were often observed in the images. Second, since the cells were not lysed with hypotonic solution before

| Population | Average Telomere Count | Average Telomere Fluorescence |
|------------|------------------------|-------------------------------|
| G1-1 | 54 | 3857 ± 38 |
| G1-2 | 72 | 7000 ± 280 |
| G1-3 | 68 | 7820 ± 130 |
| S/G2-1 | 81 | 4250 ± 130 |
| S/G2-2 | 101 | 10680 ± 220 |
| S/G2-3 | 87 | 8650 ± 170 |

Table 4.2: Telomere Detection in Lymphocyte Populations

being dyed with fluorescent telomere probe, it is possible that some of the telomeres were tightly bound into DNA-protein complexes and were inaccessible to fluorescent probe. Third, some closely-spaced telomeres could have been mislabeled by the Dot Label algorithm as a single telomere. A sample of such a telomere is shown in figure 4.13. Many of these telomeres were split by hand, but it is likely that some were not resolved. Fourth, some of the telomere spots' integrated fluorescences might have been below the noise level of the images.

It was decided that, despite the low telomere count, it was still valuable to analyze the distribution and clustering of those telomeres which could be detected.

4.5.2 Kolmogorov-Smirnov Statistic

To determine whether the distributions of telomeres in the lymphocyte populations differed between populations and from the uniform distributions which would result if the telomeres were randomly arranged, a two-tailed Kolmogorov-Smirnov test was used. Kolmogorov-Smirnov tests for agreement between two frequency distributions of data. The Kolmogorov-Smirnov statistic D is given by

$$D = \max |S_m(x) - S_n(x)| \quad (4.16)$$

where $S_m(x)$ and $S_n(x)$ are the cumulative distributions at value x of two sample distributions being tested, with m and n elements respectively.

The critical value for the test statistic for large samples and a significance $\alpha = 0.001$ is given by

$$D_{0.001} = 1.95 \sqrt{\frac{m+n}{mn}} \quad (4.17)$$

The null hypothesis of the Kolmogorov-Smirnov test is that the two samples being tested come from the same distribution. At $\alpha = 0.001$, if the test statistic is greater than the critical value, one can conclude that there is a 99.5% probability that the two sample distributions tested have different probability distributions.

4.5.3 Radial Distribution of Telomeres

After the Dot Label algorithm was used to locate telomeres, the radial locations of their centers of area were calculated relative to the bounding circle of the nucleus. Cumulative distributions of the radial location of telomeres in the G1-1, S/G2-1, G1-2, S/G2-2, G1-3, and S/G2-3 lymphocyte populations are shown in figures 4.16 to 4.21.

To determine whether the telomere distributions followed a uniform probability density function, the 1-tailed Kolmogorov-Smirnov test was used to compare the distributions to the theoretical uniform distributions proposed in section 4.2.1. The critical values to conclude that the telomere distributions are nonuniform are 0.019, 0.024, 0.029, 0.032, 0.026, and 0.029 for the distributions in figures 4.16 to 4.21.

The two-dimensional 'fried egg' distribution, which would occur if the lymphocyte contents were spilled uniformly onto a circle on the slide is shown as the green line in figures 4.16 to 4.21. All of the lymphocyte populations' telomere distri-

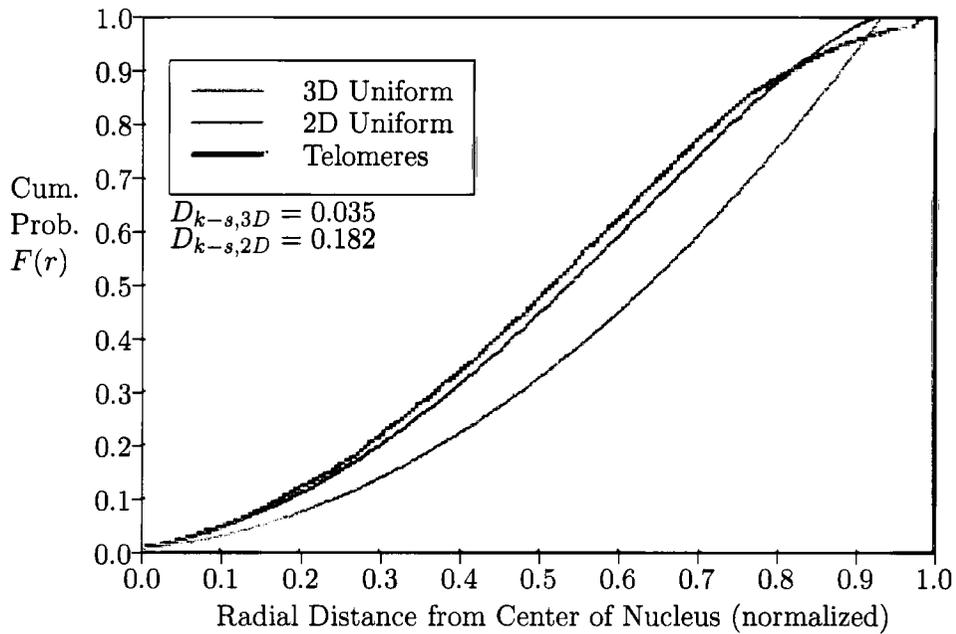


Figure 4.16: Cumulative Distribution of G1 Long Telomere Radial Locations

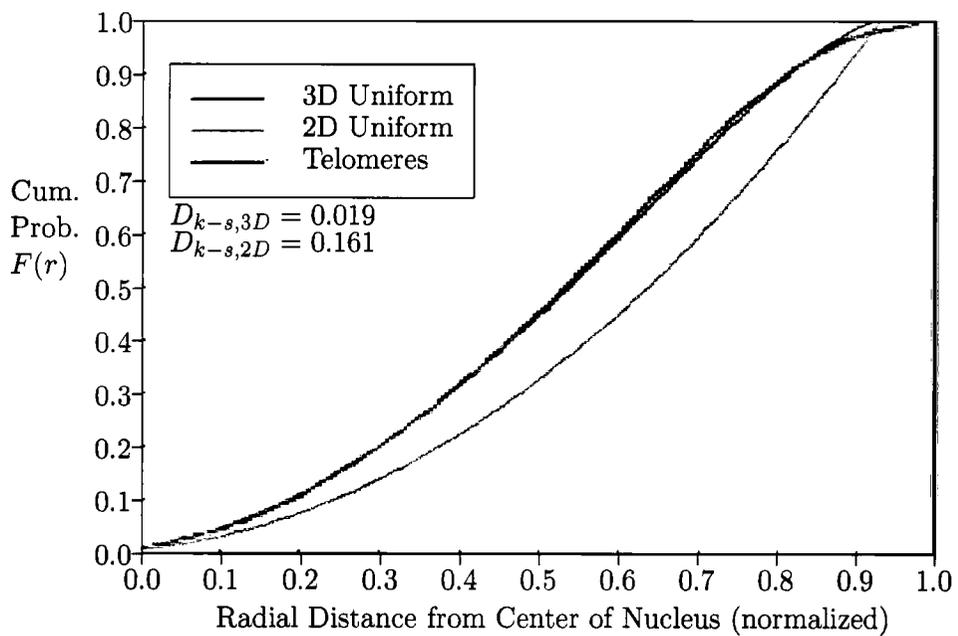


Figure 4.17: Cumulative Distribution of G1 Medium Telomere Radial Locations

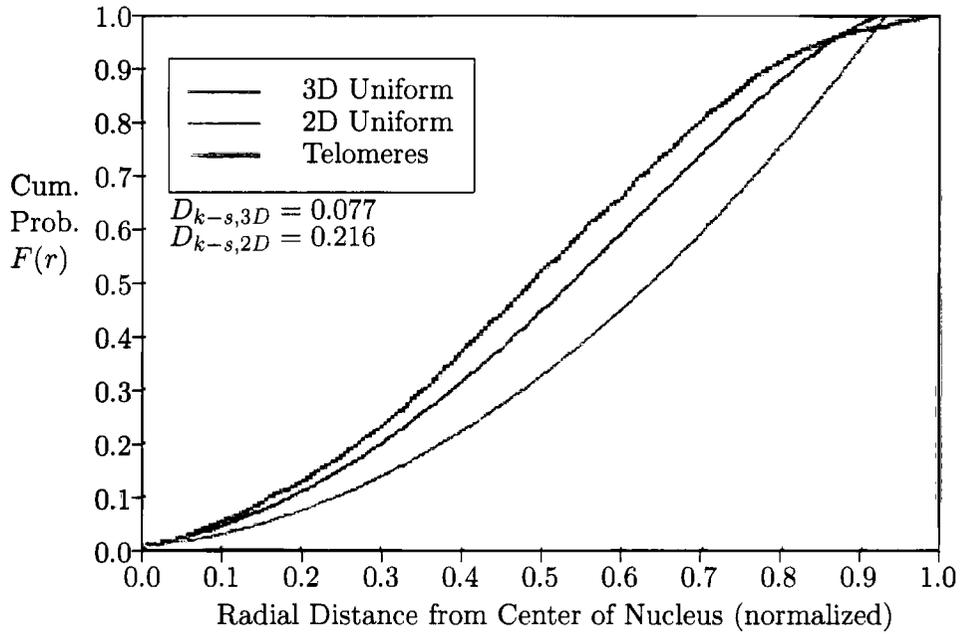


Figure 4.18: Cumulative Distribution of G1 Short Telomere Radial Locations

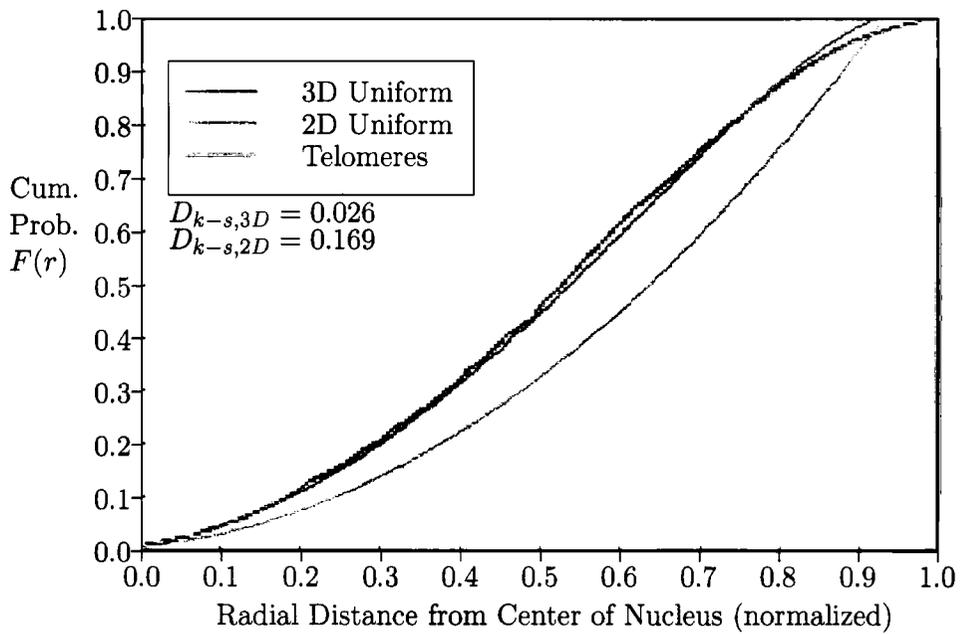


Figure 4.19: Cumulative Distribution of S/G2 Long Telomere Radial Locations

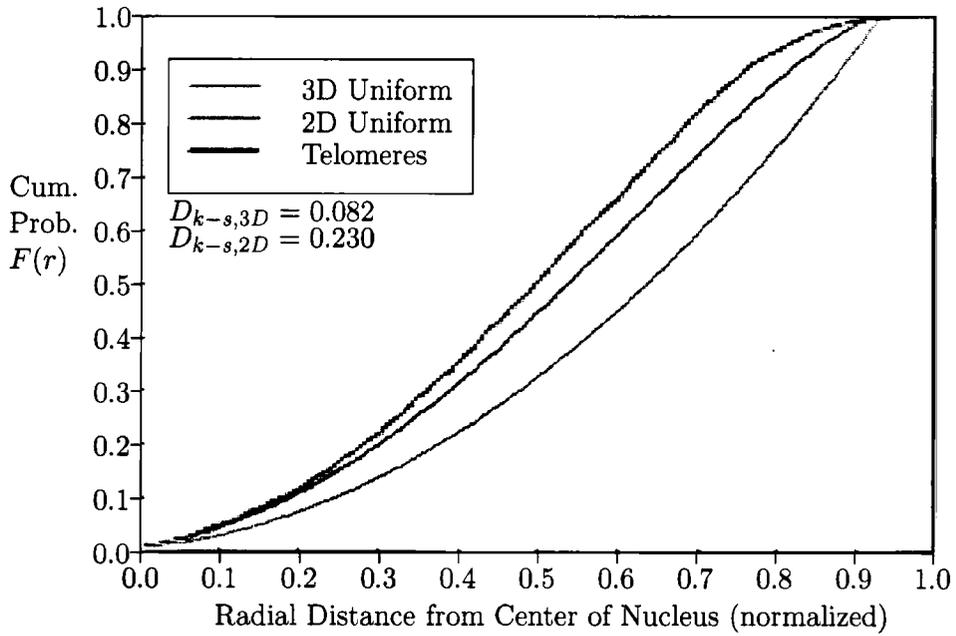


Figure 4.20: Cumulative Distribution of S/G2 Medium Telomere Radial Locations

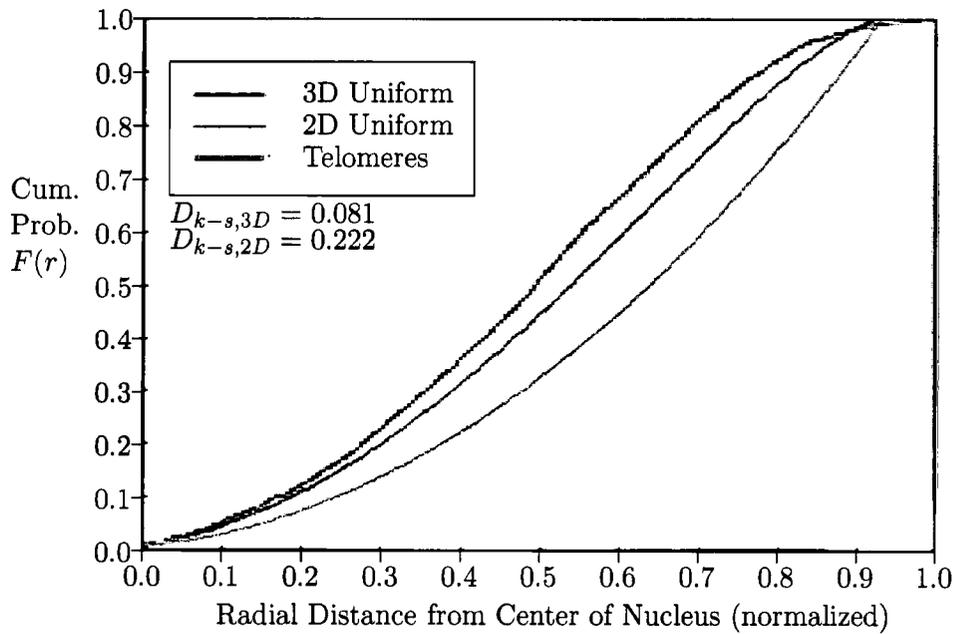


Figure 4.21: Cumulative Distribution of S/G2 Short Telomere Radial Locations

butions are significantly different from this one, so the lymphocytes have preserved some nuclear structure during slide preparation.

The three-dimensional uniform distribution which would occur if the telomeres were evenly distributed through a spherical nucleus, then viewed by projecting their locations onto the image plane, is shown as a blue-gray line in figures 4.16 to 4.21. The short-telomere populations are significantly different and more internally-distributed than uniform, while the long-telomere populations are not different from the uniform distribution. In the medium-telomere populations, the S/G2 population is more internally distributed than uniform, while the G1 population is uniformly distributed. The Kolmogorov-Smirnov statistics between the telomere distributions and the theoretical uniform distributions are shown on the graphs in figures 4.16 to 4.21. The observation that the telomeres in the non-uniformly distributed populations lie towards the inside of the nucleus suggests that organizing forces might pull telomeres (or all DNA) toward the center of the nucleus during interphase.

The radial distributions of telomeres in the nucleus for all six lymphocyte populations are compared in figure 4.22. There are two distinct groups of telomere distributions. The populations with long telomeres and the G1 population with medium-length telomeres could be statistically the same and are similar to the three-dimensional uniform distribution. The populations with short telomeres and the S/G2 population with medium telomeres could also be statistically the same, are nonuniformly distributed, and have more internally-distributed telomeres than the first group. These observations imply that, if the short telomeres in these nuclei affect nuclear organization, short telomeres are distributed towards the center of the cell. Therefore, any telomere organizing factors which are dependent on long telomeres for recognition and which affect radial distribution must pull the telomeres

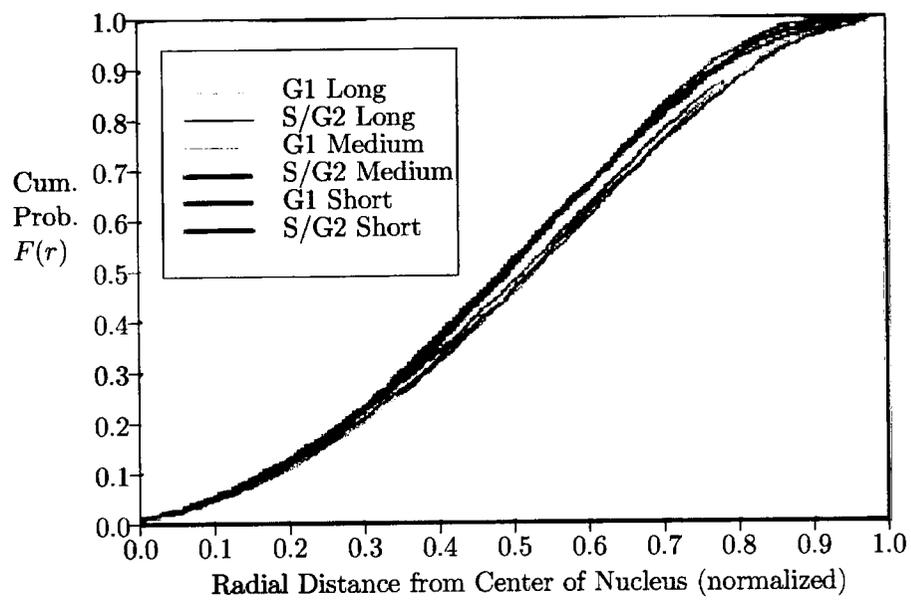


Figure 4.22: Cumulative Distributions of Telomere Radii: All Lymphocyte Populations

toward the outside of the cell in both the G1 and S/G2 cell cycle phases in human lymphocyte nuclei. No differences in the radial distribution of telomeres were observed in nuclei in different cell cycle phases. In cells with medium-length telomeres, the observation that the telomeres are distributed more internally in S/G2 than in G1 phase is counter to the observation in [23] described in chapter 1 section 1.4.3.

4.5.4 1p Probe Distribution

The 1p probe spots in all 6 lymphocyte populations were located more internally than predicted by the uniform distributions proposed in section 4.2.1. Plots of the 1p probe spot distributions for all 6 populations are shown in figure 4.23. All 1p probe spot distributions are significantly different from each other and show no trends in radial location from G1 to S/G2 cell cycle phases, or as the telomeres become shorter.

4.5.5 Delaunay Triangulation and MST Calculation

To compute the Delaunay graph of the telomere locations, the program spots2.exe first located the telomere spots on the CY3 image and displayed the spot locations in a window, as shown in figure 4.24. After the user made any manual adjustments for joined telomeres as described in section 4.4.7, the program computed the Delaunay triangulation graph of the telomere locations and its minimum spanning tree. Samples of these graphs are shown in figures 4.25 and 4.26.

The Delaunay triangulation was calculated using the Convex Hull method described in [29]. The graph of telomere locations (x, y) was transformed into a three-dimensional paraboloid $(\tilde{x}, \tilde{y}, \tilde{z})$ using the relation

$$\tilde{x} = x \tag{4.18}$$

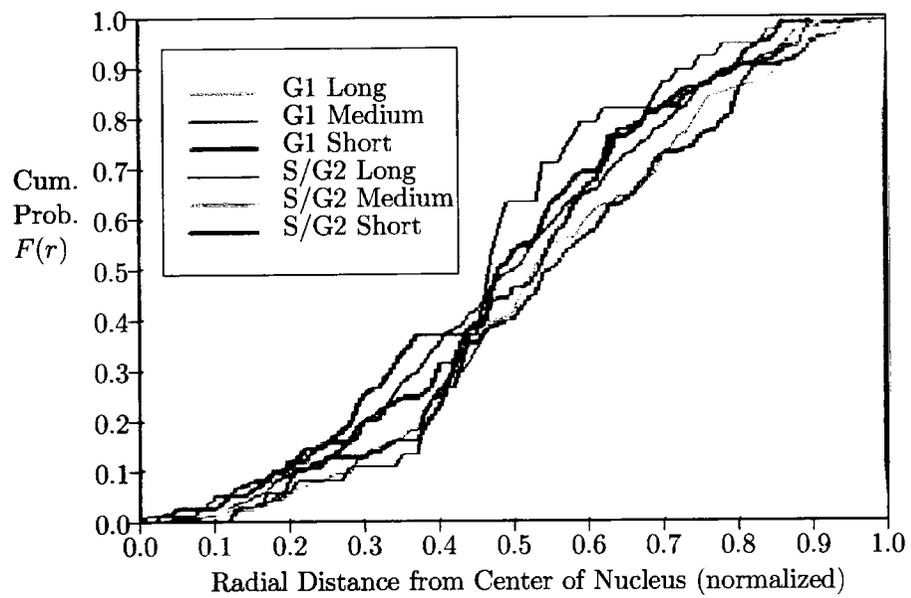


Figure 4.23: Cumulative Distributions of 1p Probe Radial Locations

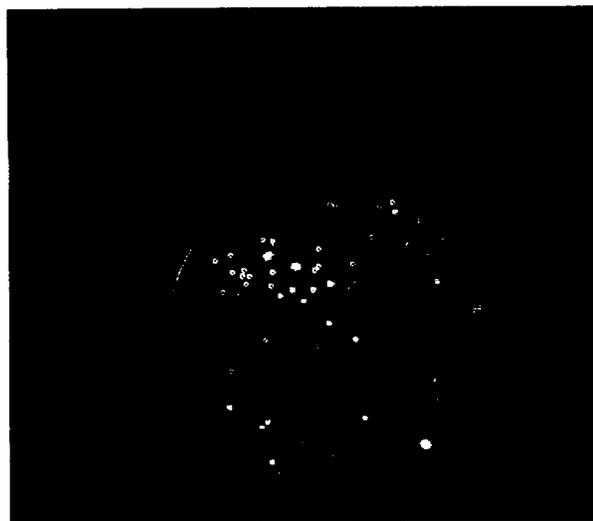


Figure 4.24: Telomere Locations displayed on Lymphocyte Image



Figure 4.25: Delaunay Triangulation Graph of Telomere Locations

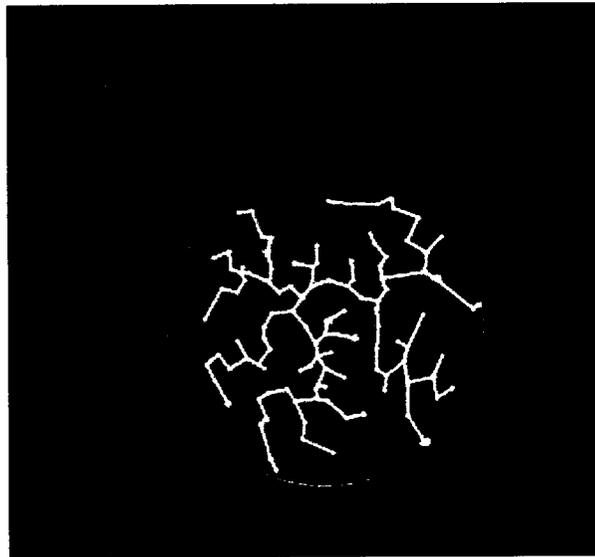


Figure 4.26: Minimum Spanning Tree of Telomere Locations

$$\tilde{y} = y \quad (4.19)$$

$$\tilde{z} = x^2 + y^2 \quad (4.20)$$

so that the height \tilde{z} of each point was equal to its distance from the origin. A property of this transformation is that triangles on the convex hull of the paraboloid correspond to Delaunay triangles on the original plane. So each triangle from the set of transformed telomere points $(\tilde{x}, \tilde{y}, \tilde{z})$ was tested to see whether it lay on the convex hull of the paraboloid by checking whether all other points lay to one side of the triangle. If so, the triangle formed by the three non-transformed points was added to the Delaunay graph.

The minimum spanning tree was calculated by sorting the segments of the Delaunay triangulation by length, selecting the shortest segment as a starting segment on the tree, then adding the shortest remaining connected segments to the tree until all the points from the Delaunay graph were spanned. The segment lengths from the minimum spanning trees for all the nuclei in each population were normalized to the nuclear radius.

4.5.6 MST Segment Length Features

The minimum spanning tree (MST) segment length histograms were calculated and shown in figure 4.27. For comparison, the histograms have been normalized to have the same area. The mean, mode, and median values of the MST segment lengths are greater in G1 than in S/G2 phase in all telomere length populations, but this is consistent with the fact that cells in G1 have fewer telomeres than cells in S/G2. Telomeres which are not clustered should, therefore, lie further apart and have greater MST segment lengths in G1 cells.

The variance, full width at half maximum (FWHM) and skew of the his-

| Population | Mean | Variance | FWHM | Mode | Skew | Median |
|-------------|------|----------|-------|------|--------|--------|
| g1 long | 0.16 | 0.0037 | 0.13 | 26 | 188582 | 0.155 |
| g1 medium | 0.14 | 0.0027 | 0.12 | 25 | 120181 | 0.135 |
| g1 short | 0.14 | 0.0033 | 0.12 | 22 | 61345 | 0.130 |
| s/g2 long | 0.12 | 0.0033 | 0.11 | 23 | 40453 | 0.115 |
| s/g2 medium | 0.11 | 0.0018 | 0.095 | 18 | 107130 | 0.110 |
| s/g2 short | 0.12 | 0.0023 | 0.10 | 20 | 66673 | 0.115 |

Table 4.3: General MST Features-Bin Size 0.005

| Population | N | Mean | Variance | FWHM | Mode | Skew | Median |
|-------------|-------|-------|----------|-------|------|--------|--------|
| g1 long | 10413 | 0.13 | 0.057 | 0.105 | 24 | 148039 | 0.125 |
| g1 medium | 5731 | 0.11 | 0.046 | 0.085 | 20 | 89134 | 0.105 |
| g1 short | 4448 | 0.11 | 0.054 | 0.090 | 15 | 48017 | 0.105 |
| s/g2 long | 3733 | 0.097 | 0.048 | 0.080 | 13 | 37618 | 0.090 |
| s/g2 medium | 5634 | 0.093 | 0.038 | 0.075 | 18 | 88972 | 0.090 |
| s/g2 short | 4425 | 0.096 | 0.043 | 0.075 | 20 | 55370 | 0.090 |

Table 4.4: First Nearest Neighbour Distances (normalized to Nuclear Radius)

tograms show no trend between G1 and S/G2 cell cycle phases and between long and short telomere populations. The general MST features are listed in table 4.3.

4.5.7 Nearest-Neighbour Features

The second measure of telomere clustering is the distribution of distances on a Delaunay graph between a given telomere point and its nearest neighbours. Changes in the moments of these distributions as the telomere length decreases would indicate changes in the strength of action of telomere-association proteins.

The segment lengths of the first, second, and third nearest neighbours to each telomere point were calculated and normalized to the nuclear radius. The first (mean), second (variance), and third (skew) moments of the segment length histograms were calculated and are shown in tables 4.4 to 4.6.

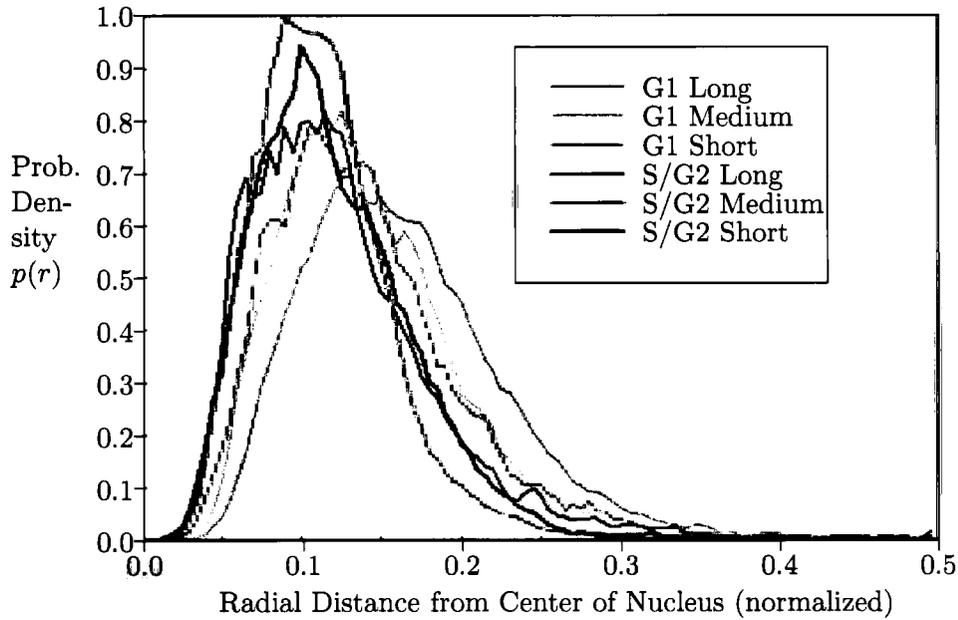


Figure 4.27: Comparison of MST Segment Lengths

| Population | N | Mean | Variance | FWHM | Mode | Skew | Median |
|-------------|-------|------|----------|-------|------|--------|--------|
| g1 long | 10413 | 0.16 | 0.058 | 0.105 | 28 | 237627 | 0.155 |
| g1 medium | 5731 | 0.13 | 0.047 | 0.080 | 23 | 142291 | 0.130 |
| g1 short | 4448 | 0.14 | 0.055 | 0.080 | 23 | 73981 | 0.125 |
| s/g2 long | 3733 | 0.12 | 0.051 | 0.075 | 18 | 56643 | 0.110 |
| s/g2 medium | 5634 | 0.11 | 0.039 | 0.070 | 20 | 135336 | 0.105 |
| s/g2 short | 4425 | 0.12 | 0.044 | 0.085 | 21 | 90886 | 0.115 |

Table 4.5: Second Nearest Neighbour Distances (normalized to Nuclear Radius)

| Population | N | Mean | Variance | FWHM | Mode | Skew | Median |
|-------------|-------|------|----------|-------|------|--------|--------|
| g1 long | 10413 | 0.19 | 0.064 | 0.105 | 33 | 274074 | 0.180 |
| g1 medium | 5731 | 0.16 | 0.051 | 0.080 | 26 | 169126 | 0.150 |
| g1 short | 4448 | 0.16 | 0.06 | 0.085 | 27 | 87633 | 0.150 |
| s/g2 long | 3733 | 0.14 | 0.057 | 0.090 | 24 | 65699 | 0.135 |
| s/g2 medium | 5634 | 0.13 | 0.043 | 0.070 | 23 | 155111 | 0.125 |
| s/g2 short | 4425 | 0.14 | 0.048 | 0.090 | 24 | 109474 | 0.130 |

Table 4.6: Third Nearest Neighbour Distances (normalized to Nuclear Radius)

Trends

- The mean nearest neighbour inter-telomere distances in the G1 populations are greater than for the S/G2 populations. However, since there are fewer telomeres in the G1 nuclei, this result can be expected in a uniform distribution.
- There is a higher variance in nearest-neighbour distances in the G1 populations than in the S/G2 populations. This is true at each clone passage and in the first, second, and third nearest neighbour cases. It is unclear whether this is significant.
- There is no trend in the mode, median, or mean of the nearest-neighbour distributions as telomere length decreases, although the G1 long-telomere and S/G2 medium- and short-telomere populations show a higher mode and median than the rest of the populations.

4.6 Discussion

This experiment failed to find evidence of telomere length-dependent changes in telomere clustering. It also did not detect the cell cycle-dependent movement of telomeres in human interphase nuclei described in Ferguson and Ward [23]. Therefore we are not able to demonstrate that telomere length-dependent proteins bind DNA into clusters or function in cell cycle-dependent DNA movement.

However, some telomere length-dependent changes in the radial distribution (section 4.6.1) were observed. Although this is weak evidence that telomere length influences nuclear organization, many improvements could be made in subsequent

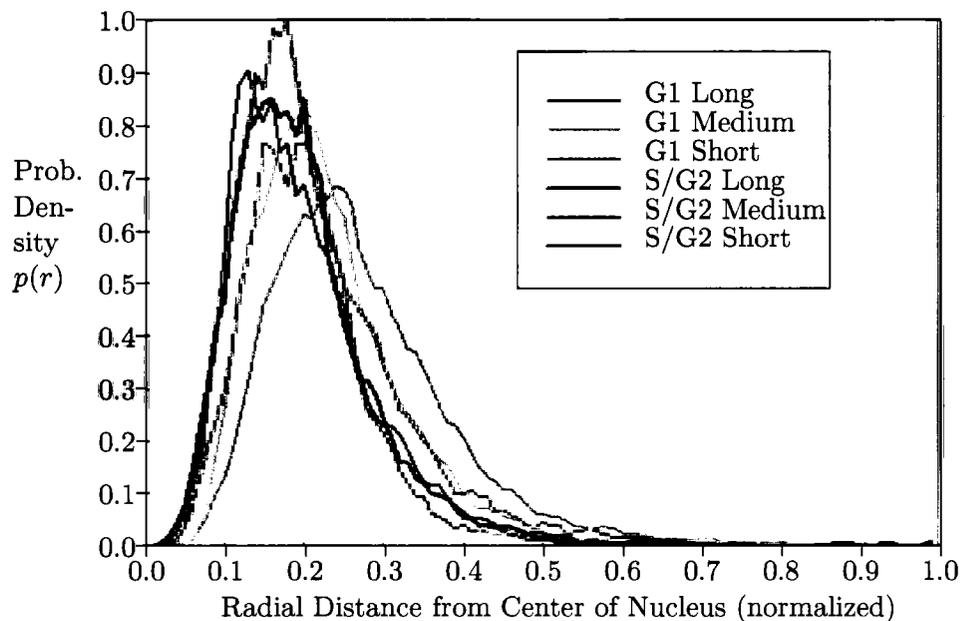


Figure 4.28: Distribution of 1 Nearest Neighbour Distances

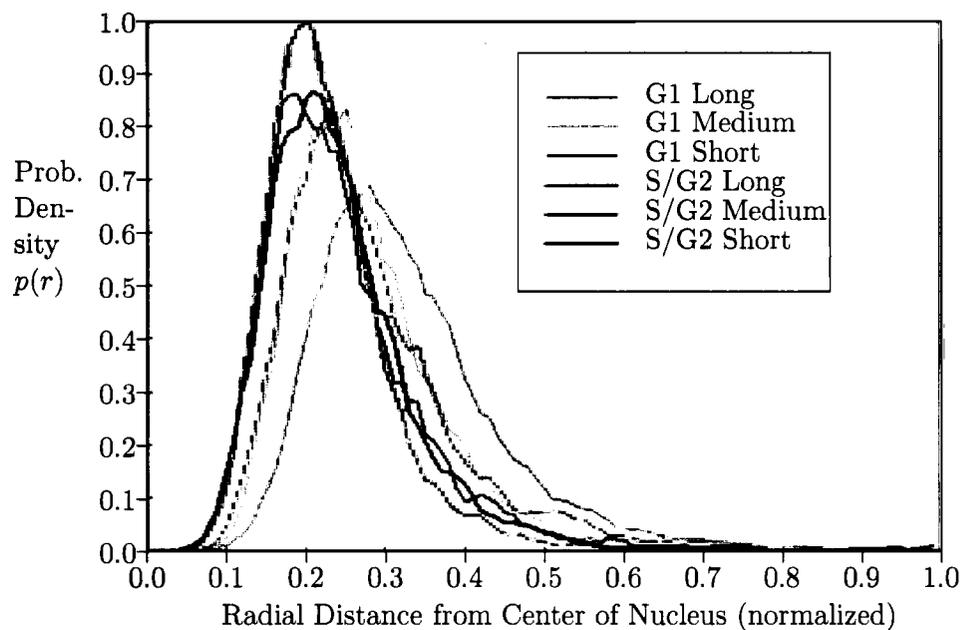


Figure 4.29: Distribution of 2 Nearest Neighbour Distances

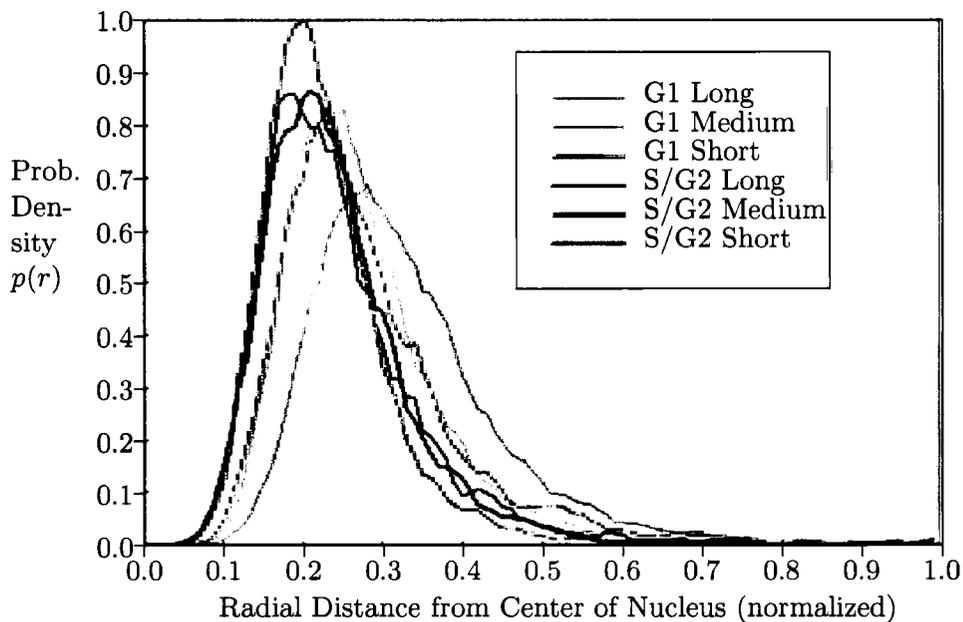


Figure 4.30: Distribution of 3 Nearest Neighbour Distances

experiments (section 5.2.2). There have been no other studies to date of Delaunay graph measurements of telomere organization, and it is possible that further studies with greater resolution might better investigate what role, if any, telomere length and the telomere ‘biological clock’ play in the structure of the nucleus.

4.6.1 Telomere Radial Distributions

In section 4.5.3 it was observed that the telomeres in the nucleus are not uniformly distributed and occur in two groups of distributions, possibly corresponding to long and short telomeres. These observations are interesting and suggest that, if telomere-length-dependent elements actually do exist and organize the telomeres in the nucleus, they pull the telomeres outward. However, the observations are not conclusive because only three telomere length data points (long, medium, and

short) were used. It is also curious that the long-telomere population's telomere distribution was consistent with the three-dimensional uniform distribution calculated, while the short-telomere population's telomere distribution was nonuniform. This contradicts the hypothesis that short telomeres should be uniformly distributed throughout the nucleus because organizing proteins are unable to bind to them.

The telomere radial distribution did not correspond to the theoretical random distributions proposed. This is an encouraging observation, since it suggests that some nuclear structure was preserved during slide preparation, and that two-dimensional widefield microscopy is useful for measuring some types of nuclear organization.

Chapter 5

Summary

5.1 Conclusions

5.1.1 Microscopy

This thesis describes several modifications which were made to a widefield microscope system in the Terry Fox Laboratory in order to study the location of telomeres in interphase nuclei. A piezomotive focusing device was installed on the microscope and control routines were written for it on the microscope's computer. The piezomotive device was used to take images of slides at many different focal planes, and information from multiple planes was used to create composite images and to focus the microscope. The composite images were not judged useful for the telomere experiment described in chapter 4, but automatic microscope focusing was successful and was used to speed up image acquisition and to ensure that all images were focused consistently.

The best criterion for microscope focusing depended on the type of objects being photographed. For small, pointlike objects, image power was the most useful criterion. Large complex objects such as DNA chromosomes were processed with

an edge-detecting filter, and the power of the image edges was used.

Images of 402 nuclei (with over 34 000 pointlike objects called telomeres) were studied using an automatic image processing application, and image statistics were compiled. It is novel to highly automate the acquisition and analysis of so many images in a microscopy experiment, and this thesis partly demonstrates the use of widefield microscopy for such a project.

5.1.2 Image Processing and Telomere Organization

Telomeres are repeat DNA sequences located at the ends of chromosomes. There are between 96 and 182 of these elements in a human cell at any time. The major experiment of this thesis studied telomere location in three populations of human lymphocytes: those with long, medium, and short telomeres in their nuclei. It attempted to measure differences in the telomere locations between these populations.

It is known that telomeres 'cap' chromosomes and protect the chromosomes from end-to-end fusion mutations. It is also known that the telomere DNA sequences must be sufficiently long in order to function this way. We hypothesized that, in addition to their protective function, telomeres might be elements which are bound by proteins which organize DNA into specific regions of the nucleus. We also hypothesized that this binding would depend on the telomeres' being sufficiently long. If these hypotheses were true, nuclei with short telomeres would be less organized than nuclei with long ones, and the telomeres would tend to be uniformly distributed in short-telomere nuclei. As well, in nuclei with short telomeres, the cell-cycle dependent movement of telomeres previously described by Ward [23] would be absent or reduced.

Two measures of telomere organization were used: telomere distribution

along the radius of the nucleus, and the Delaunay graph features of the telomere locations. Both these measures require precise determination of the nuclear radius and efficient detection of the telomeres on the images. The nuclear radius was measured using gradient-weighted thresholding, while the telomeres were detected using a region growing algorithm called Dot Label. The telomere detection efficiency was 70%.

The telomeres in nuclei with long telomeres were found to exist in a nearly-uniform distribution, while those in short-telomere nuclei were distributed closer to the center of the nucleus. This contradicts our hypothesis that the short telomeres should be unable to bind to nuclear organizing proteins and should therefore be uniformly distributed in interphase nuclei. The observation suggests that, if organizing forces act on telomeres, some forces might pull the telomeres toward the center of the nucleus, while telomere length-dependent elements pull telomeres outward.

Clustering of telomeres was analyzed using Delaunay graphs of the telomere locations. Histograms of the segment lengths of the minimum spanning trees of these graphs were compared between populations of short, medium, and long-telomere nuclei. No difference in the clustering measure was found between the populations.

We conclude that there is a slight telomere length-dependent effect on telomere organization along the radius of the nucleus, but telomere length does not affect telomere clustering. As described below, further investigation of both clustering and the radial distribution of telomeres in the nucleus could be carried out.

5.2 Future Directions

5.2.1 Widefield Microscope Focusing

The automatic focusing routine was used successfully and enthusiastically by some of the researchers in the lab. It might be useful to refine the function in the ways described in section 3.5. One suggestion is to make different focusing criteria available for different types of images. Also, automatic focus routines could be implemented on other microscopes in the lab.

5.2.2 Measurement of Telomere Organization

There are a large number of improvements which could be made to the telomere distribution experiment if further investigation were warranted:

- To examine whether telomere organization is truly dependent on telomere length, more than three telomere length populations should be used. In this experiment, even if trends in nuclear organization were observed with telomere shortening, it would have been far more conclusive if all 12 of the lymphocyte populations, with monotonically decreasing telomere lengths, from the cloning experiment described in Rufer [30] had been measured.
- Better controls or compensation in software could be made for microscope parameters such as uneven or low illumination from the xenon/mercury lamp, aberrations in the optical system, or camera noise. These factors can vary from day to day but be consistent throughout the day. Similarly, variations in slide preparation including addition of reagents or probe dyes, photobleaching, or quality of cells after thawing, can vary from slide to slide but be consistent over a slide. All the lymphocyte images for a population (G1 long, S/G2 medium,

etc) were acquired using a single day and a single slide per population (G1 long, S/G2 medium, etc.). This means that any variations in microscopy or slide preparation would show up in only one population, be consistent over a population, and would be reflected in the telomere organization statistics for that population. For example, images from the S/G2 population with long telomeres had consistently poor signal to noise ratio, which could have been due to any of the microscopy or slide preparation factors described. The telomere detection efficiency was lower for all nuclei in this population, and since more telomeres went undetected than in other slides, likely any measure of telomere organization would have different results from other slides.

- Widefield two-dimensional microscopy was used for this experiment, partly because it was much quicker than multiple plane microscopy and enabled many nuclei to be analyzed, and partly because it was the only technique available. Obviously, much information about the three-dimensional organization of the nucleus was lost when the cells were squashed onto the slide or when a projection image of the cell was acquired. Since the goal was to measure trends in organization levels, rather than absolute degrees of organization, and since some organization (such as telomere clusters or distribution near the nuclear membrane) would be preserved during squishing and projection, three-dimensional information should not have been crucial to the experiment. However, if sufficient numbers of three-dimensional images could have been acquired, three-dimensional Delaunay maps of the nuclei could be easily constructed, and the extra information would have been useful and have made organization easier to detect.

- The telomere detection efficiency varied for each nucleus. Since many telomeres went undetected, they were not included in the measures of clustering. Although the dotlabel algorithm used was the best one investigated, the experiment could be repeated if an improved algorithm were found.

Bibliography

- [1] U M Martens, M J M Zijlmans, S S S Poon, V Dragowska, J Yui, E A Chavez, R K Ward, and P M Lansdorp. Short telomeres on the p-arm of human chromosome 17. *Nature Genetics*, 18(1):76–80, Jan 1998.
- [2] J B Pawley. *Handbook of Confocal Microscopy*. Plenum Press, 1995.
- [3] Y Hiraoka, J W Sedat, and D A Agard. Determination of three-dimensional imaging properties of a light microscope system: Partial confocal behaviour in epifluorescence microscopy. *Biophys. J*, 57:325–333, Feb 1990.
- [4] L Firestone, K Cook, K Culp, N Talsania, and K Preston. Comparison of autofocus methods for automated microscopy. *Cytometry*, 12:195–206, Nov 1991.
- [5] David A Agard, Yasushi Hiraoka, Peter Shaw, and John W Sedat. Fluorescence microscopy in three dimensions. *Methods in Cell Biology*, 30:353–377, 1989.
- [6] David A Agard. Optical sectioning microscopy: Cellular architecture in three dimensions. *Ann. Rev. Biophys. Bioeng*, 13:191–219, 1984.
- [7] F S Fay, W Carrington, and K E Fogarty. Three-dimensional molecular distribution in single cells analysed using the digital imaging microscope. *Journal of Microscopy*, 2(2):133–149, Feb 1989.

- [8] D Shore. Telomeres-unsticky ends. *SCIENCE*, 281:1818–1819, Sep 1998.
- [9] T de Lange. Human telomeres are attached to the nuclear matrix. *The EMBO Journal*, 11(2):717–724, 1992.
- [10] F Ishikawa. A tale of mice without 'tails'. *Nature Genetics*, 17:257–258, Nov 1997.
- [11] B van Steensel, A Smogorzewska, and T de Lange. Trf2 protects human telomeres from end-to-end fusions. *Cell*, 92:401–413, Feb 1998.
- [12] C Nugent, G Bosco, L Ross, S Evans, A Salinger, J Moore, J Haber, and V Lundblad. Telomere maintenance is dependent on activities required for end repair of double-strand breaks. *Current Biology*, 8:657–660, May 1998.
- [13] A Dernburg, J W Sedat, W Z Cande, and H W Bass. Cytology of telomeres. In *Telomeres*, pages 295–339. Cold Spring Harbour Press, 1995.
- [14] A M Olovnikov. Molecular mechanism of morphogenesis: a theory of locational dna. *Biokhemiia*, 61(11):1948–70, 1996.
- [15] M Blasco, HW Lee, MP Hande, E Samper, PM Lansdorp, RA DePinho, and CW Greider. Telomere shortening and tumor formation by mouse cells lacking telomerase rna. *Cell*, 91:25–34, Oct 1997.
- [16] J Griffith, A Bianchi, and T de Lange. Trf1 promotes parallel pairing of telomeric tracts in vitro. *J Mol Biol*, 278(1):79–88, Apr 1998.
- [17] D Broccoli, A Smogorzewska, L Chong, and T de Lange. Human telomeres contain two distinct myb-related proteins, trf1 and trf2. *Nature Genetics*, 17:231–235, Oct 1997.

- [18] T Billaud, C Brun, K Ancelin, C E Koering, T Laroche, and E Gilson. Telomeric localization of trf2, a novel human telobox protein. *Nature Genetics*, 17:236-239, Oct 1997.
- [19] D Broccoli and H J Cooke. Effect of telomeres on the interphase location of adjacent regions of the human x chromosome. *Experimental Cell Research*, 212:308-313, 1994.
- [20] T Laroche, S Martin, M Gotta, H Gorham, F Pryde, E Louse, and S Gasser. Mutation of yeast ku genes disrupts the subnuclear organization of telomeres. *Current Biology*, 8:653-656, May 1998.
- [21] H Biessman and J M Mason. Genetics and molecular biology of telomeres. *Advances in Genetics*, 30:185-249, 1992.
- [22] A O Zalensky, M J Allen, A Kobayashi, A Zalenskaya, R Balhorn, and E M Bradbury. Well-defined genome architecture in the human sperm nucleus. *Chromosoma*, 103:577-590, 1995.
- [23] M Ferguson and D C Ward. Cell cycle dependent chromosomal movement in pre-mitotic human t-lymphocyte nuclei. *Chromosoma*, 101:557-565, 1992.
- [24] F Billia and U De Boni. Localization of centromeric satellite and telomeric dna sequences in dorsal root ganglion neurons, in vitro. *Journal of Cell Science*, 100:219-226, 1991.
- [25] Hai-Shan Wu, J Barba, and J Gil. A focusing algorithm for high magnification cell imaging. *Journal of Microscopy*, 184(2):133-142, Nov 1996.
- [26] Bevington, Phillip R, Robinson, and D Keith. *Data Reduction and Error Analysis in the Physical Sciences*. McGraw-Hill Higher Education, 1991.

- [27] Poon Steven SS. *Telomere Length Measurements using Fluorescence Microscopy*. PhD thesis, UBC, 1997.
- [28] Pzt performance test document. Quality Control Document QD2, Physic Instrumente (PI) GmbH, Dec 1996.
- [29] A Okabe, B Boots, and K Sugihara. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. John Wiley and Sons, 1992.
- [30] N Rufer, W Dragowska, G Thronbury, E Roosnek, and PM Lansdorp. Telomere length dynamics in human lymphocyte subpopulations measured by flow cytometry. *Nature Biotechnology*, 16:743-747, Aug 1998.
- [31] H Netten, L J van Vliet, Hans Vrolijk, W C R Sloos, H J Tanke, and I Young. Fluorescent dot counting in interphase cell nuclei. *Bioimaging*, 4:93-106, 1996.
- [32] Anil K Jain. *Fundamentals of Digital Image Processing*. Prentice-Hall, 1989.