Division Tracking By Live Cell Imaging and Image Processing

Jingjing Li

A Dissertation Submitted in fulfillment of the requirements
For the degree of Master by Research
University of New South Wales
Faculty of Engineering
Graduate School of Biomedical Engineering

August 1, 2011

THE UNIVERSITY OF NEW SOUTH WALES

Thesis/Dissertation Sheet

Surname or Family name: LI

First name: JINGJING Other name/s:

Abbreviation for degree as given in the University calendar: ME

School: Graduate School of Biomedical Engineering Faculty: Engineering

Title: Division Tracking By Live Cell Imaging and Image Processing

Abstract 350 words maximum:

Towards the goal of generating specific tissue from stem or progenitor cells for regenerative medicine, it will be necessary to understand the dynamics of stem and progenitor cell development and how environmental cues trigger cell migration, mitosis, apoptosis, and lineage fate. Observing the dynamic process in a continuous manner at the single-cell level will advance our knowledge of these processes. Long-term live cell imaging systems and computational methods to automatically identify and track progenitor cell migration and division will enable this study.

The aims of this thesis were to develop a live cell imaging system with semi-automated software for tracking adherent cell lines and to apply this system to study cardiac stem cell development. The imaging system was benchmarked by tracking NIH3T3 cells in vitro for 4 days. Cardiac stem cells were enriched by fluorescent activated cell sorting (FACS) from the interstitial fraction of the mouse heart. These cells form colonies (c-CFU-F) which were tracked by live cell imaging. Green fluorescent protein (GFP) transgenic mice were used to report cCFU-F cells that express β-actin or platelet-derived growth factor receptor-a (PDGFR-a). These initial studies have focused on characterizing cell motility and cell cycle dynamics of cCFU-F subpopulations.

The growth rate of NIH3T3 (482 cells) tracked by the live cell imaging system was similar to conventional culture methods. Lineage maps of PDGFR-α+ cells (164 maps) and β-actin + cells (352 maps) within passage 3 colonies were constructed by continuous cell tracking over a 5 day culture period. Two distinctive cell morphologies were indentified; large flattened-cells with low motility and highly motile spindle-shaped cells. The probabilities of mitosis of flattened- and spindle-shaped cells were estimated for each generation using Kaplan Meier statistics. There were significant differences between the cell cycle distribution and motility for these two subpopulations. Furthermore, Cox regression analysis was used to show that cell cycle progression was related to cell size and colony size. Large flattened-cells infrequently underwent asymmetric division giving birth to a small cell and large cell with a short and long cycle time respectively. These studies have illustrated the value of lineage mapping cCFU-F, leading to a deeper understanding of cCFU-F growth dynamics.

FOR OFFICE USE ONLY Date of completion of requirements for Award:

THIS SHEET IS TO BE GLUED TO THE INSIDE FRONT COVER OF THE THESIS

Declaration

I hereby declare that this submission is my own work and to the best of my knowledge. It contains no materials previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where the acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conceptions or in style, presentation and linguistic expression is acknowledged.

Jingjing Li

March, 2011

Copyright Statement

Jingjing Li

March, 2011

Authenticity Statement

I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis. No emendation of content has occurred and if there are any minor variations in formatting, they are the result of the conversion to digital format.

Jingjing Li

March, 2011

Abstract

Towards the goal of generating specific tissue from stem or progenitor cells for regenerative medicine, it will be necessary to understand the dynamics of stem and progenitor cell development and how environmental cues trigger cell migration, mitosis, apoptosis, and lineage fate. Observing the dynamic process in a continuous manner at the single-cell level will advance our knowledge of these processes. Long-term live cell imaging systems and computational methods to automatically identify and track progenitor cell migration and division will enable this study.

The aims of this thesis were to develop a live cell imaging system with semi-automated software for tracking adherent cell lines and to apply this system to study cardiac stem cell development. The imaging system was benchmarked by tracking NIH3T3 cells in vitro for 4 days. Cardiac stem cells were enriched by fluorescent activated cell sorting (FACS) from the interstitial fraction of the mouse heart. These cells form colonies (c-CFU-F) which were tracked by live cell imaging. Green fluorescent protein (GFP) transgenic mice were used to report cCFU-F cells that express β-actin or platelet-derived growth factor receptor-a (PDGFR-a). These initial studies have focused on characterizing cell motility and cell cycle dynamics of cCFU-F subpopulations.

The growth rate of NIH3T3 (482 cells) tracked by the live cell imaging system was similar to conventional culture methods. Lineage maps of PDGFR-α+ cell (164 maps) and β-actin + cell (352 maps) within passage 3 colonies were constructed by continuous cell tracking over a 5 day culture period. Two distinctive cell morphologies were indentified; large flattened-cells with low motility and highly motile spindle-shaped cells. The probabilities of mitosis of flattened- and spindle-shaped cells were estimated for each generation using Kaplan Meier statistics. There were significant differences between the cell cycle distribution and motility for these two subpopulations. Furthermore, Cox regression analysis was used to show that cell cycle progression was related to cell size and colony size. Large flattened-cells infrequently underwent asymmetric division giving birth to a small cell and large cell with a short and long cycle time

respectively. These studies have illustrated the value of lineage mapping cCFU-F, leading to a deeper understanding of cCFU-F growth dynamics.

Acknowledgements

Firstly, I would like to express a special thank you to my supervisor Dr Robert Nordon for his invaluable guidance. I am grateful for the opportunity he has given me to undertake this Masters of Engineering by Research degree, and sincerely appreciate him for always being available to provide continual help, patient tuition, supervision and his unwavering commitment to teach me what it entails to be a good research scientist.

I have been lucky to work with amazing groups of people. I am grateful to the members of the Graduate School of Biomedical Engineering (GSBmE) and the Victor Chang Cardiac Research Institute (VCCRI) at the University of New South Wales (UNSW). Thank you to all the wonderful staff in the VCCRI and GSBmE for helping me out during the study. In particular, I would like to thank Prof. Richard Harvey and Dr. Munira Xaymardan for providing the transgenetic mouse, technical training and experimental advices. I would like to acknowledge Huaying Chen for his continual help in the two years' study, particularly with the culture chamber manufacture, the imaging stage modification, and experimental designs. Thank you to Lynn Ferris, Bill Cheng, and Christine Chuang for lending a hand and being such a pleasure to work with.

It would not have been possible to produce this thesis without the assistance and support of many colleagues and friends. I would like to specially acknowledge the assistance and support from Julia Yin. Also, I would like to extend thank you to Amr Ar, Han Zhang, Hamza Toor, Walen Bai, and Paul Guo for their company and for being a wonderful friend.

Last but not least, to my parents – thank you. It would not have been possible to get this far without your understanding, love and constant support.

Jingjing Li

March 2011

Table of Content

List of Figures

List of Tables

Abbreviations and Symbols

BSA	Albumin from Bovine Serum
CFU-F	Colony-Forming Unit Fibroblast
DIC	Differential Interference Constrast
DMEM	Dulbecco's Modified Eagle Medium(D-MEM)
DT	Distance Transform
DsRed	Red Fluorescent Protein
FACS	Fluorescence Activated Cell Sorting
FBS	Fetal Bovine Serum
FCS	Fetal Calf Serum
GFP	Green fluorescent protein
IMDM	Iscove's Modified Dulbecco's Medium
MEM-α	Minimum Essential Medium-Alpha
MSC	Mesenchymal Stem Cell
NaCl	Sodium Chloride
PCM	Phase Contrast Microscope
PDGFR-α	Platelet Derived Growth Factor Alpha
QCM	Quasi-Center of Mass
RBF	Rolling Ball Filter
RGBP	Region Based Processing
SE	Structuring Element
SEM	Standard Error of the Mean
TE	Tissue Engineering
UV	Ultra Violet

Chapter 1 Introduction

1.1 Research motivation

The field of tissue engineering had been developed decades ago due to the increasing demand for tissue replacement and the lack of organ and tissue for transplantation. Diverse technologies and devices, including bioreactors, micro-fluidic devices, drug delivery systems, microinjection and microscopic imaging systems have accelerated development of this field, however tissue engineering is still far away from its stated long-term goal: 'to control and regulate the potential of natural tissue regeneration for defect repair or even organ regeneration' (Plonsey, 1999, Liao et al., 2008, Kelleher and Vacanti, 2010, Kohara and Tabata, 2010, Shokeir et al., 2010). Basically, research in tissue engineering has focused on three main themes: cells, scaffolds, and signals. However, how these elements interact is still not fully understood.

The goal of regenerative medicine is to develop specific tissues from stem and progenitor cells. Progenitor cells possess less potential to differentiate than stem cells, and lie somewhere in the differentiation pathway between a stem cell and a terminally differentiated cell. The classical view is that cells progress from undifferentiated cell types to more specialized types over multiple rounds of division. The traditional belief is that lineage commitment is an irreversible process. However recent studies have demonstrated that many cell types have a degree of cell plasticity. Plasticity is the ability of lineage-committed cells to dedifferentiate and follow a different lineage pathway (Liao et al., 2008). However, the biggest obstacle to progress is to understand how external cues direct cell migration, mitosis and lineage fate. The central question is how to create culture environments that direct differentiation of stem and progenitor cells into functional cells that can be maintained *in-vitro* and *in-vivo* (Theise, 2010).

In the future stem cell biology will consider how cell-to-cell and cell-to-matrix interactions control the formation of complex tissue architectures. Complexity and chaos theories attempt to

describe the overall behavior of a system, (colony, tissue or organ) with limited knowledge of system components (Gustavsson et al., 2003, Mullassery et al., 2008). A systems biology approach aims to describe stem cell development from a single cell and multi-cellular tissue perspective. The approach adopted in this thesis is to understand tissue development from a detailed analysis of single cell behavior and to collect enough single cell data so that one understands population dynamics as well.

The dynamics of stem and progenitor cell populations are characterized by their migration, division, change in phenotype and death. Conventional methodology for determining the cell doubling time and cell cycle is based on cell counting (Zhang et al., 2010), however important features, such as the timing of divisions and lineage choice cannot be inferred directly from this data because individual cell histories are not continuously tracked. There is a clear need to establish a quantitative and comparative framework for the analysis of cell populations by high throughput single-cell lineage tracking. This framework can also lead to a better understanding of the relationships, origins and biology of stem cells, seeking to discover ways to augment cell regeneration or stem cell therapy (Ahmed et al., 2007, Cohen et al., 2010, Eilken et al., 2008, Flaibani et al., 2009, Gustavsson et al., 2003, Khoshmanesh et al., 2008, Nordon et al., 2005, Thomas, 2010).

This research will focus on establishing methodology to track progenitor cell migration and division in vitro by live cell imaging and develop semi-automated imaging processing software. A cell line, NIH3T3, and a primary cell progenitor type, cardiac colony forming unit-fibroblasts, were tracked using the developed methodology. This approach led to a clearer understanding of the dynamics of cardiac CFU-F.

1.2 Thesis Aims

With respect to the technique of lineage mapping by live cell imaging it is hypothesized that

1. Live cell imaging is a scalable technology for high throughput study of progenitor cells.

2. Mother-daughter relationship and cell trajectories can be constructed from the time-lapse video by tracking the size and morphological changes of cells using a nearest neighbor algorithm.
3. Long-term fluorescence imaging is non-destructive to cells with suitable control of phototoxicity.

It is also hypothesized that

1. Fibroblast colony-forming assay originally developed by Friedenstein (Friedenstein, 1989a) can be adapted to quantify tissue-resident cardiac stem cells.
2. Cardiac CFU-F is a heterogeneous population composed of different cell types characterized by their cell cycle behavior and lineage potential.

To address these hypotheses the thesis has the following specific aims:

1. Customize live cell imaging platform for long term imaging of cardiac stem cells
2. Develop fluorescence imaging methods for long term tracking of GFP marked cells
3. Develop software to segment cardiac stem cell division trees from cell growth videos
4. Acquire enough data to characterize cardiac stem cell development
5. Apply Kaplan-Meier statistics and Cox regression analysis of cell lifetime data to identify the different cell types that are present in cardiac CFU-F based on cell cycle kinetics and other live cell imaging characteristics i.e., motility and cell size.
6. Quantify the transmission of cell cycle length from mother to daughter cells using lineage pedigrees.

1.3 Thesis layout

Chapter 2 A literature review covering relevant cell biological concepts; cell cycle, cell plasticity, and the biology of cardiac CFU-F. A review of live cell imaging methodologies and applications.

Chapter 3 Investigates image processing algorithms and the Matlab image analysis tool box which was applied to the problem of automated and manual tracking of cell division, cell trajectories and construction of progenitor pedigrees

Chapter 4 Evaluates the performance of live cell imaging system and image processing software by tracking NIH3T3 cells over 4 days.

Chapter 5 Characterizes development of cardiac CFU-F using the methodology developed in this thesis. Identification of two cell subtypes that comprise CFU-F with difference in cell cycle, motility and morphology. Application of Kaplan Meier and Cox regression analysis to determine the effect of cell subtype and colony size on the probability of cell mitosis.

Chapter 2 Literature Review

2.1 Introduction

In this chapter, background knowledge and terminologies for cell division and cell plasticity are reviewed. The literature of stem cell, cardiac fibroblasts, cell division tracking and image processing algorithms are then summarized and evaluated.

2.2 Background

2.2.1 Cell Division

Cell division is the process in which a parent cell divides into two or more daughter cells. It concludes the cell cycle, during which the cell replicates DNA and transcribes and expresses genes required for cell structure and function. In eukaryotes the process of separation of replicated DNA into two sets of chromosomes that are segregated into daughter cells is known as mitosis. Cell division in prokaryotes is known as binary fission. Just prior to splitting of parent cells DNA is replicated (figure 2.1 b). Eukaryotes also have germ cells which segregate each set of chromosomes into two daughter cells, a process known as meiosis. The germ cell or gamete cannot divide again until fertilization.

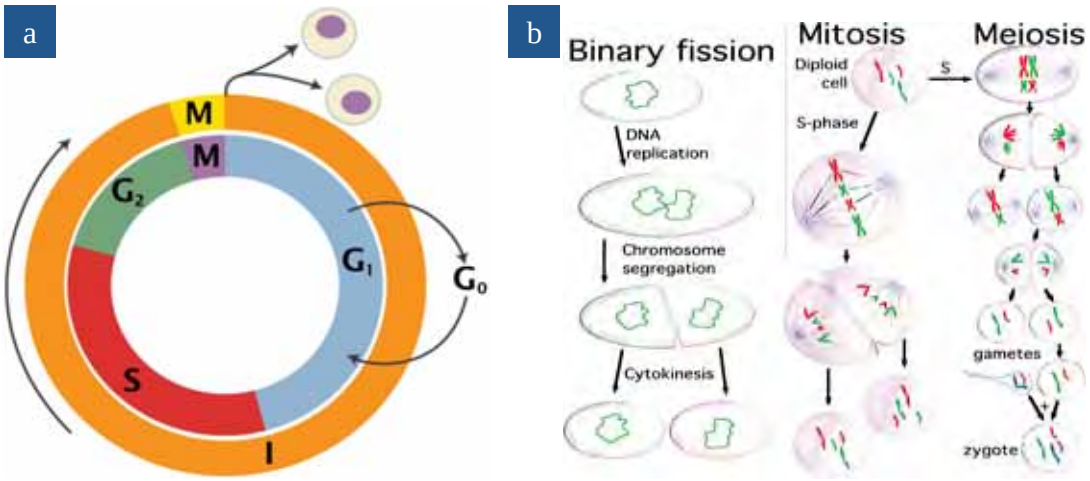


Figure 2. 1 Schematic of cell cycle and three division pattern a: Four phases of a cell cycle (Cooper, 2000) Inner ring: G0: 'resting phase', quiescent/senescent the absence of RNA synthesis, G1: Gap phase, preparation of cellular materials for S phase, S: Synthesis phase, synthesis of DNA, G2: Gap phase, assemble of materials for mitosis. Outer ring: M: Mitotic phase, distribution of duplicated DNA into two daughters; I: Interphase. b: Three types of cell division

For each cell cycle interphase refers to the time during which the cell prepares for mitosis with protein expression and cell growth referred to as gap phases (G1 and G2), duplication of centrioles and synthesis of DNA (S-phase) as depicted in figure 2.1a. Mitosis also referred to as M phase is subdivided into prophase (DNA condensation to form chromosomes), followed by prometaphase and metaphase (alignment of daughter chromatids at the central axis of the cell), anaphase (separation of chromatids) and cytokinesis (daughter cell cleavage and separation). The stages of cell cycle can be visualized using phase contrast or fluorescent microscopy or analysed by flow cytometry (Halter et al., 2009). Moreover, the rate of cell cycle progression determines the rate of cell growth and the volume of cytoplasm that is synthesized.

2.1.2 Cell Plasticity

The term plasticity refers to the capacity of tissue-derived stem cells to exhibit a phenotypic potential that includes cells comprising the tissue of origin, but also extends beyond the differentiated cell phenotypes of their resident tissue. There is an increased interest on the

plasticity of adult cells and their ability to divide or self-renew indefinitely, thus generating all the cell types of the organ from which they originate. Unlike embryonic stem cells, the use of adult stem cells in research and therapy is not considered to be controversial as they are derived from adult tissues rather than destroyed human embryos.

Theise has proposed 4 plasticity mechanisms (Theise, 2010). The canonical mechanism for cell plasticity is the classical hierarchy for cell lineage differentiation which applies to both development and adult tissue maintenance and repair. The central dogma for this mechanism is that cells become increasing restricted with respect to tissue potential of cell subtype. The second mechanism for cell plasticity is dedifferentiation followed by a re-differentiation pathway, which is commonly found in amphibians, and is also well-documented in some fetal mammals and adult mammals. Thus a differentiated cell may move up the hierarchical differentiation pathway, a process referred to as dedifferentiation. The third pathway is the change from one lineage to another in response to cues from the microenvironment, presumably without any need to dedifferentiate. In this situation there will be no evidence that stem cells are formed during a direct lineage switch. The last plasticity mechanism is nuclear reprogramming caused by cell-to-cell fusion, nuclear fusion, and post-fusion reduction division. The non-canonical pathways are most commonly referred to as ‘plasticity’, and it remains an open question the extent to which these play a prominent role in normative physiology. Thus dedifferentiation and nuclear reprogramming may only be processes that have been observed in vitro, but do not contribute in any way to normal tissue maintenance. However, from a tissue engineering perspective this is not relevant. More importantly, the big question for regenerative therapies is how stem cells can be manipulated in vitro to realize their therapeutic potential.

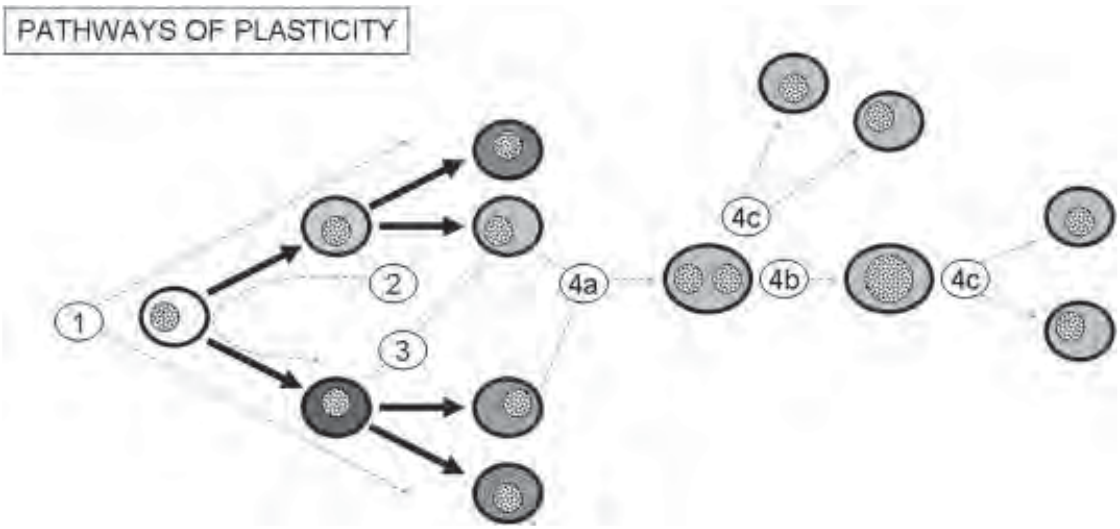


Figure 2. 2 Four pathways of cell plasticity (Theise, 2010) 1:Canonical differentiation pathway. 2: Dedifferentiation and re-differentiation. 3: lineage commitment on response to microenvironment. 4: Nuclear reprogramming by (a) cell fusion, (b) nuclear fusion, and (c) post-fusion division.

2.2 Stem Cells

2.2.2 Definitions and terminologies

Stem cells have been recognized as an importance source of cells for cell and tissue engineering applications because of their potential for self-renew through cell division and their ability to differentiate into a wide range of specific cell types. All tissues have stem cells, including the central nervous system and adult heart. However, the potential for self-renew and plasticity varies for different cell sources and at different stages of mammalian development. Recent studies demonstrate that embryonic stem cells, postnatal stem cells, and even terminally differentiated cells have a degree of cell plasticity under specific culture conditions (Theise, 2010).

This has led to some confusion as new in vitro techniques are developed to manipulate tissue-derived stem cells. Using the canonical definition of cell plasticity - cells can only differentiate

into more restricted phenotypes - stem cells are classified as totipotent, pluripotent, multipotent and unipotent. Totipotent stem cells are from the fertilized oocyte and the descendants of the first two divisions, and are able to give birth to all types of cells and organs including germ cell types. Pluripotent stem cells are the cluster of cells from the inner cell mass (ICM), also named blastocyst, which is a hollow ball of cells formed after 4 days' development of totipotent stem cells. They are able to differentiate to almost all cells derived from the three germ layers but not the placenta and supporting tissue of embryo. Pluripotency is also commonly used to indicate the potential of tissue restricted stem cells, for example pluripotent blood stem cells can give rise to all of the blood lineages. The majority of adult cells are multipotent, because their differentiating abilities are limited to a range of differentiated cell lineages that comprise the tissue or organ of origin. Unipotent stem cells are capable of generating one specific cell type, but they are less differentiated that the terminally differentiated cell type. For example a granulocyte colony grown on semisolid media will only give rise to granulocytes. The distinction between multipotent and pluripotent stem cells is somewhat subjective, depending on the field's jargon or terminology. In the field of embryonic stem cell research, pluripotency refers to the potential to generate any somatic cell type derived from the 3 layered embryo stages. Recently it has been possible to induce pluripotency in fetal fibroblast by over expression and selection of the Yamanaka (Takahashi and Yamanaka, 2006) pluripotency genes. These cells were named as induced-pluripotent stem cells (iPS) (Alison and Islam, 2009, Baraniak and McDevitt, 2010, Godara et al., 2008).

Colony forming unit-fibroblast (CFU-F) is a cell classification which is based on growth at limiting dilution of fibroblastic-like cells which form loosely associated colonies in liquid culture. These cells have a mesenchymal origin (i.e., form bone, cartilage, tendon and fat), the assay and classification originally developed by Friedenstein (FRIEDENS.AJ, 1970). The reprogramming of CFU-F from heart (cCFU-F) into functional cardiomyocytes let to the hypothesis that cCFU-F can be differentiated into a terminally differentiated lineage without first becoming a stem/progenitor cells (Tateishi et al., 2008, Wojakowski and Tendera, 2010). These observations provide the basis for a potential source of cells for cardiac regenerative therapy, though there have been very limited studies illustrating how this could be achieved. This review will endeavor to provide an overview of the biology of cCFU-F.

2.2.3 Biology of cardiac CFU-F

Fibroblasts are widely distributed connective tissue cells of mesenchymal origin that produces variety of extracellular matrix (ECM) proteins and biochemical mediators, including collagens and fibronectin (Camelliti et al., 2005, Krenning et al., 2010, Porter and Turner, 2009, Zeisberg and Kalluri, 2010). However, the identification of fibroblast is not based on its synthesis and the deposition of ECM, but generally depends on its special morphology, proliferation, and phenotypical characteristics. Morphologically, fibroblasts are flat, spindle shaped cells with multiple processes originating from the main cell body.

Cardiac fibroblasts play numerous roles in cardiac development and vascular remodeling, as well as facilitating cardiac structure and function. The lack of a basement membrane and the display of a prominent Golgi apparatus and extensive rough endoplasmic reticulum in active state are distinctive characters of cardiac fibroblasts in all cardiac tissue. Although no true specific marker are established for identifying the fibroblast phenotype in various organs, some distinct cardiac fibroblast markers are examined in human and mouse cardiac tissue, for example, discoidin domain receptor (DDR2) and fibroblast-specific protein-1(Souders et al., 2009). Other highly expressed genes in cardiac fibroblasts are cadherin-11, vimentin, beta1-integrin, fibronectin, connexins and the fasciclin gene periostin(Porter and Turner, 2009).

Myofibroblasts are a common cell type found in heart. They are characterized by a combination of fibroblast and smooth muscle markers with the exception of the myosin heavy chain (MHC) which is only expressed in mature smooth muscle cells. It was demonstrated that under appropriate conditions, for example, in response to injury, resting or quiescent fibroblasts isolated from the interstitial fraction of heart are activated with expression of a contractile phenotype including several smooth muscle markers that are not commonly observed on fibroblasts (Souders et al., 2009). Another study also reported that cardiac fibroblasts cultured in vitro at low density will differentiate into myofibroblasts as a transient phenotype (Masur et al., 1996). These cells are induced by an amniotic membrane stromal extract to return back to fibroblasts in vitro (Hansson et al., 2009, Hattori and Fukuda, 2010, Joggerst and Hatzopoulos, 2009, Krenning et al., 2010).

2.2.4 Adult cardiac CFU-F in heart repair

Heart attack, stroke, and related diseases are the leading cause of death in the Western world, outstripping deaths due to all cancers combined. Improvements to medical therapy for acute coronary artery disease (myocardial infarction, MI) have lead to an epidemic of heart failure, a condition that is expected to increase exponentially (Camelliti et al., 2005, Gonzales and Pedrazzini, 2009, Hansson et al., 2009, Porter and Turner, 2009).

Cardiovascular disease may result in loss of cardiac tissue through death of the cells by apoptosis and necrosis. The remaining myocytes are unable to reconstitute the lost tissue and the diseased heart deteriorates with time. Current therapies have limitations and are primarily focused on preventing the progression of disease, rather than repair or regeneration of healthy tissue and function. As a result, cell transplantation therapy (CTT) which has the potential to achieve cardiac repair, attracts increasing interest. It is stated in literature that suitable cells for transplantation might be cardiomyocytes, myoblasts grown from skeletal muscle, smooth muscle cells from blood vessels, or hematopoietic or mesenchymal stem cells. However, the prevalent static view of the myocardium implies that both myocyte death and regeneration have little role in cardiac homeostasis (Hansson et al., 2009, Hattori and Fukuda, 2010, Joggerst and Hatzopoulos, 2009) . Therefore, the search and investigation of resident adult cardiac stem cells had been considered futile because of the widely accepted lack of regeneration potential of this organ.

In adult heart, although cardiac myocytes occupy 75% of normal cardiac tissue volume, they only account for 30% of total cell number. The majority of the remaining cells are fibroblasts. Considering the potential diversity of functional characteristics of cardiac fibroblasts and the self-regenerating potential of adult heart, it was hypothesized that cardiac stem cells exist in the interstitial cardiac fibroblast compartment (Murry et al., 2005). Previous studies by Richard Harvey's group (Victor Chang Cardiac Research Institute, Sydney) had shown that contractile function of injured heart was recovered by injection of platelet derived growth factor with the cells that isolated from the interstitial cardiac fibroblasts compartment (Harvey et al., 2009). On this basis, it is crucial to identify, quantify, and map cardiac fibroblast lineages in vitro, and to

study how environmental cues (growth factors, etc) coordinate their dynamic behavior in cardiac development, maintenance and disease.

2.3 Cell Division Tracking

The dominant technique used by most biologists to study cellular growth kinetics is to count cells at regular intervals, and to analyse the mixture of cells using fluorescent cellular markers either by fluorescent microscopy of flow cytometry. The frequency of observations is limited by lab workers office hours, and cell growth is inferred by interpolation or a geometric cell growth model (doubling time). With this approach many cellular events such as mitosis and lineage switching are missed, and sometimes data are over interpreted. The need for continuous observation of dynamic cellular processes arose out of the need to understand the exact cellular mechanism for lineage commitment and regeneration. For example as illustrated in figure 2.3 the production of 4 progeny from a single progenitor (a) has many possible mechanisms (b) (Schroeder, 2008).

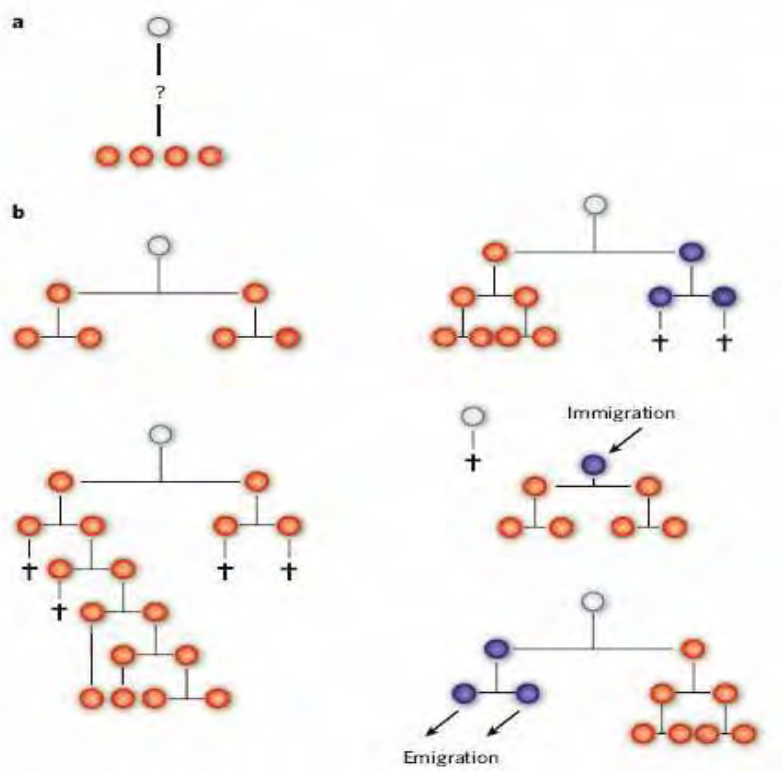


Figure 2. 3 The need for continuous observation of dynamic cellular process (Schroeder, 2008)

Nordon et al (Ko et al., 2007, Nordon et al., 2005), developed a simple sorting strategy with flow cytometry and the vital fluorescent dye CFDA-SE to track cell division in bulk culture (Nordon et al., 1997). This technique played an important role in the study of renewal and differentiation of hematopoietic stem cells. However, the assignment of cell generation number is based on the assumption of serial halving of fluorescence intensity of CFDA-SE stained cells, and is highly depend on a correct subjective gating criterion.

Kap-Hyoun et al (Ko et al., 2007), refined this method by including phenotypic analysis and a clustering and nearest neighbor algorithm for precise generation assignment, enabling analysis of multipotent cell differentiation dynamics. However this technique does not record the individual cell differentiation histories, due to the loss of relationship of individual cells because cells were not continuously tracked.

Unlike flow cytometry, which displays snapshots of the cell population distribution with respect to generation number and immunophenotype, video microscopy and image analysis completely characterise division history and phenotype for single cells with construction of single clonal family trees (De Boer and Perelson, 2005, Flaibani et al., 2009, Hasbold and Hodgkin, 2000, Ko et al., 2007, Potter and Wener, 2005, Wallace et al., 2008). Schroeder et al (Schroeder, 2008), adopted continuous long-term imaging to generate single cell pedigrees which established that endothelial cells derived from mouse mesodemal cells could (rarely) generate blood cells providing direct evidence supporting the role of haemogenic endothelium during embryonic development. In this study detailed phenotypic pedigrees were generated by in situ staining with fluorescent conjugated moAbs allowing a precise lineage map to be construct for endothelial-blood transitions.

The rapid development of cell imaging technologies has led to recent advances in live cell imaging methods. These include development of computational methods for segmentation of cellular trajectories (Cohen et al., 2010, Dufour et al., 2005, Harder et al., 2009) using phase contrast (Wang et al., 2006, White et al., 2005, Wirtz, 2009, Yang et al., 2006) or fluorescent (Levi and Gratton, 2007, Li et al., 2010, Lobutova et al., 2009, Padfield et al., 2009, Wang et al., 2005, Wang et al., 2006, Wirtz, 2009) images. The goal of this developing field is to extract biological information from time-lapse single cell observations in vivo or in vitro using automated imaging and image analysis technologies (Wang et al., 2006).

There is no single imaging method that is optimal for observation of all aspects of cell division and regeneration. To address specific biological hypotheses, an optimal combination of cellular model systems, fluorescent labeling and imaging technology is required (Schroeder, 2008). This approach raises three main technical challenges. Firstly, a suitable physiological culture system for long-term study of cell renewal and differentiation requires development. Secondly, multi-molecular markers for identification of cells or stages of cell maturation are required. Thirdly, phototoxicity resulting from cellular imaging will need to be minimized although new imaging methods such as 2 photon excitation can dramatically reduce phototoxicity. Moreover, high temporal resolution is required for highly migratory cell types. Over a period of several days or weeks, high temporal imaging yields huge volumes of data. Developing computational methods to automate identification and tracking of cells is essential, because manual analysis and visual inspection of huge image stacks is a logistic nightmare.

2.3.2 Imaging methodology

Imaging methodology requires selection of a live-cell imaging modality and techniques for enhancing cellular contrast which is related to the imaging technique. Live cell imaging modalities are transmitted-light microscopy (TLM), wide-field fluorescence microscopy (WFM), confocal laser-scanning microscopy (CLSM), multi-photon confocal laser-scanning microscopy (MP-CLSM), spinning-disk confocal microscopy (SD-CM) and total internal reflection fluorescence microscopy (TIRFM). Suggestions for selecting imaging facilities on the basis of samples are provided in figure 2.4 (Frigault et al., 2009).

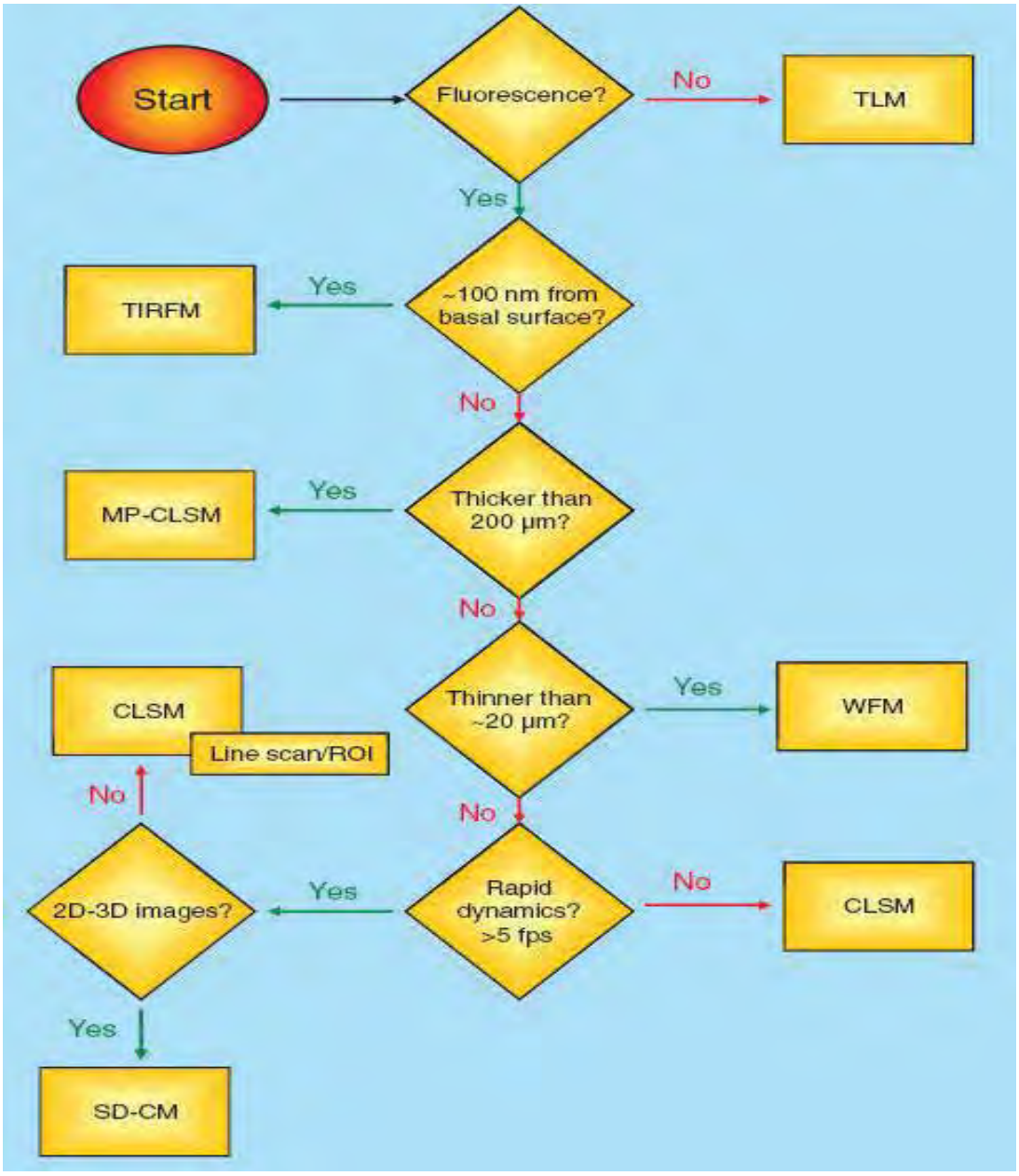


Figure 2. 4 Flow chart for selection of suitable microscope for long-term living cell imaging (Frigault et al., 2009)

Experimentation with bright field or Kohler illumination and phase contrast microscopy techniques have been attempted in live cell imaging to improve image quality. Bright field illumination illuminates a sample with an evenly dispersed source of light, and the differences in light absorption within the sample produces the contrast perceived in the image. In comparison, phase contrast microscopy (PCM) exploits the refractive properties and thickness of different

structures. PCM enhances the contrast at the cell edges in comparison to bright field illumination, but the refractive properties of the cell refract light outwards towards the cell edge creating bright halos, distorting the actual cell edge in certain areas and also at the borders of cell organelles (Gustavsson et al., 2003). Differential interference contrast (DIC) is another possible microscopy technique. Although DIC provides higher contrast at the cell edge, DIC is quite expensive and does not offer drastic improvements compared to PCM (Frigault et al., 2009). Laser scanning confocal microscopy (LCSM) has also been used to image cellular boundaries. LCSM produces extremely detailed cellular images due to the excitation of thin sample sections, but in bright field mode LCSM produces an image comparable to that of PCM. LCSM is only truly effective when samples are prepared with immuno-fluorescence techniques.

Surface staining using immuno-fluorescent staining is possible without fixation, though staining of intracellular structures requires membrane permeabilisation and fixation of cells to allow access of labeling antibodies to intracellular structures. Fixative reagents preserve many of the cellular structures, but will of course terminate the live cell imaging experiment. Hence, immuno-fluorescent staining is not a viable option for observation of internal structures within living cells.

Fluorescence is triggered by excitation of an adsorbing material at a specific wavelength. The absorption of the irradiated light energy from the excitation source causes a jump in energy of the absorbing material resulting in the radiation of fluorescence from the sample as electrons return to their ground state (Drummond and Allen, 2008, Dzyubachyk et al., 2010a, Green, 2002, Muzzey and van Oudenaarden, 2009). Photobleaching results in loss of fluorescent potential after each excitation . During real-time fluorescence imaging of time-lapse studies, continuous excitation results in rapid photobleaching and phototoxicity so excitation sources must be switched on and off to minimize these effect (Kable, 2006, Frigault et al., 2009). Ensuring a constant or replenishing supply of fluorescent dye may counteract photobleaching. This may be achieved by genetic insertion of green fluorescent protein (GFP) cDNA sequences adjacent to cellular DNA sequences coding for desired structures. Proper insertion of GFP cDNA sequences produces cellular proteins expressed with fluorescent protein tags. As a result, fluorescent protein gene reporter systems have vast potential for the field of live cell imaging.

Development of a non-destructive real-time fluorescence technique to visualize cellular membrane boundaries simplify image processing by contrast (Drummond and Allen, 2008) . However, development of Fluorescent protein probes or transgenic animals requires a significant development time for each biological application. Thus PCM is generally applicable and the most popular imaging method utilized for live cell imaging. Although the halo effect in PCM poses difficulties in cell segmentation and regular cleaning of the optical components is required to maintain high quality images, PCM is almost obligatory because of its low level of phototoxicity and high temporal and spatial resolution.

The convenience of PCM presents significant problems with interpretation and segmentation of grey scale images. Post-processing of phase contrast images is required to compensate for uneven illumination (Wang et al., 2006). Li has successfully tracked the dynamics of various cell types by PCM (Li et al., 2007). The technical challenges for segmenting live cell images are discussed in much more detail in Chapter 3.

In this study we utilized PCM in combination with fluorescent microscopy, fluorescent proteins and transgenic mouse models to collecting enough information for lineage mapping.

2.3.3 Image acquisition

There are always compromises between the quality of images, the amount of information collected and phototoxicity. Choosing suitable reagents for labeling cellular properties, keeping cells alive and minimizing the toxic effects in the imaging process are essential.

Criteria for choosing fluorescent reagents are biocompatibility, quantum efficiency, cellular uptake, photobleaching and tissue penetration. Fluorescent proteins expressed by genetic labeling certainly biocompatible. It is reported that the genes encoding eGFP, enhanced cyan fluorescent protein and enhanced yellow fluorescent protein, venus, DsRed-MST and mRFP1 (monomeric red fluorescent protein) have been expressed in stem cells or mammalian cells without an obvious change of phenotype, and considered to be highly non-toxic (Schroeder, 2008).

In addition to imaging technologies, live cell culture systems are required for long term growth of cells under the microscope objective. The live cell imaging system will require control of the following system parameters.

Temperature

The optimal temperature for cell culture is 37 ºC which mimics the human physiological environment. In order to keep a stable temperature during the cell culture, the incubator may enclose the whole microscope or just the microscope stage.

Evaporative losses

Evaporative losses during long term culture will increase media osmolarity and compromise cell growth. Thus it is necessary to humidify the samples by sparge the air/CO_2 gas mixture through a water reservoir to humidify the sample chamber, while avoiding condensation damage to the mechanical components of microscope.

pH

For mammalian cells, a bicarbonate-based buffer is used to maintain the pH of the culture medium. A 5-10% CO_2 concentration is required to maintain a physiological pH of 7.4. Frigault et al (Frigault et al., 2009), have a preference for bicarbonate/CO_2 with 25mM HEPES buffering. As shown in figure 2.5, they compare the efficiency of CO_2 dependent medium and CO_2 independent medium, and emphasized that bicarbonate contributes to many cellular process in addition to regulation of pH (Wang et al., 2006, White et al., 2005).

Light exposure

It is generally acknowledged that the damage to cells is inversely related to wavelength. Also photosensitivity to light of different wavelength differs between cells isolated from different sources. Red illumination is preferable for phase contrast imaging and 2 photon excitation for fluorescence imaging (Wang et al., 2006, White et al., 2005). In addition, light exposure can be controlled by the sampling frequency and light exposure time. Increasing the sampling frequency provides higher temporal resolution, but also may lead to a rise in photo-toxicity.

An assessment of phototoxicity is required for each kind of live cell experiment to assess the differential effects of cells and fluorescent probes, to define suitable sampling frequency and exposure time (Ford et al., 1996, Lee et al., 2002). More recently, methods have been proposed to minimize the light dose used for fluorescent excitation. Hoebe et al. designed a feedback control system to minimize the fluorescent light exposure significantly reducing phototoxicity and photobleaching (Hoebe et al., 2007, Nishigaki et al., 2006).

Focus

Most focus drift is due to the change of temperature, movement of medium, and mechanical slippage of Z-focus couplings for cheaper systems. Currently, most microscopy acquisition systems have real time software or more sophisticated external devices for automatic focusing of images. Once there are sufficient distinctive features in the field of view, autofocus software effectively maintains focus (Kable, 2006).

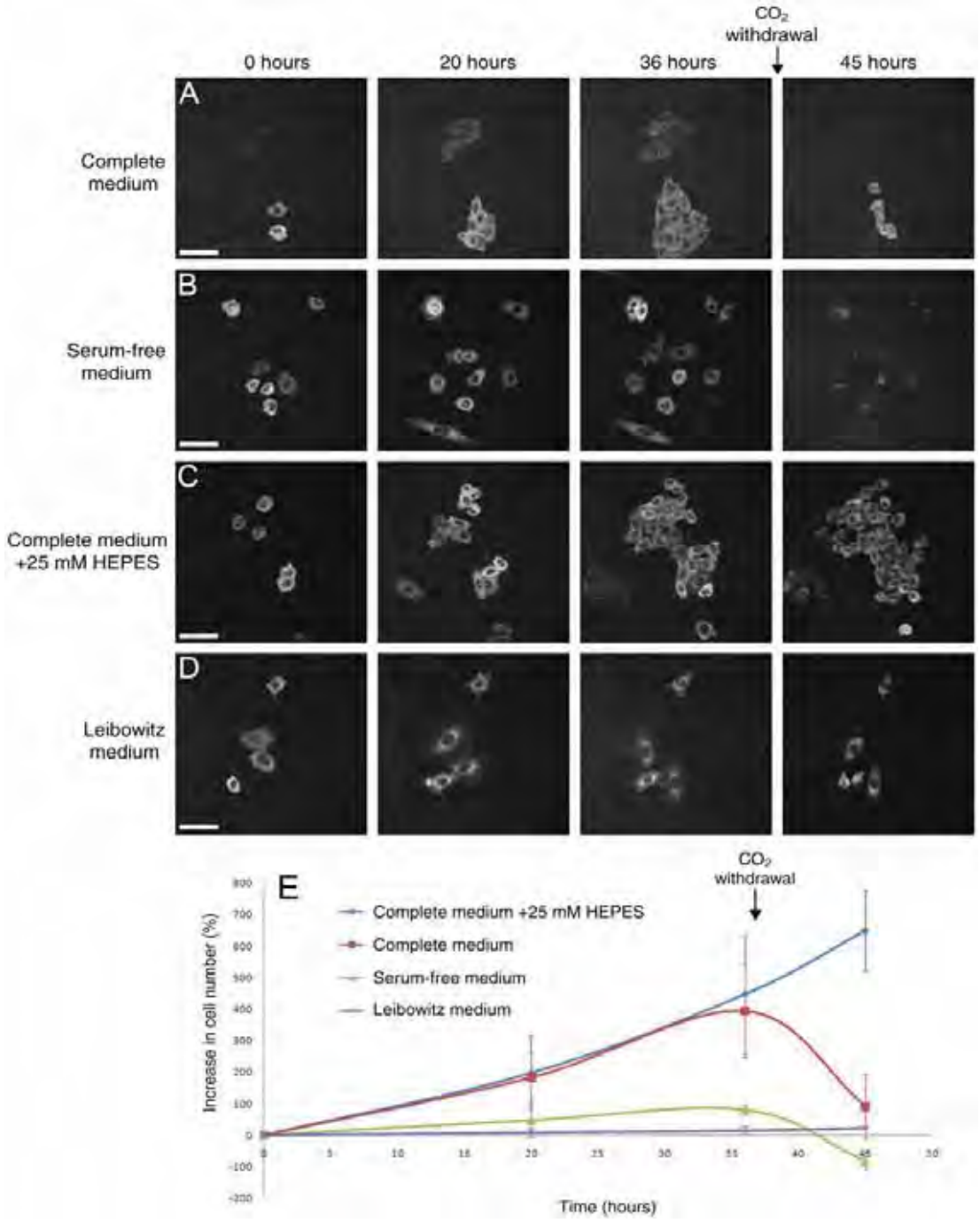


Figure 2. 5 Comparing the effect of different medium for mammalian cells (Frigault et al., 2009)

2.4 Cellular Image processing

There are several challenges to automate image processing, including the generally poor image quality (low contrast and signal-to-noise ratio SNR), the varying of density and shape of cell populations due to division and migration behaviors, as well as merging or overlapping cells (Al-Kofahi et al., 2006). Many computerized methods for cell tracking have already been proposed. Many of these are open source software (Cohen et al., 2010, Harder et al., 2009, Li et al., 2008, Swedlow and Eliceiri, 2009), however, more algorithms are required to optimize the tracking task. Several trends can be observed in the development of cell tracking methods, indicating the superiority of particular algorithms for different applications.

Generally, these methods can be classified as tracking by detection or tracking by model evolution. In the tracking by detection approach, cells are first segmented in individual frames based on intensity, texture, or gradient, then the detected cell are associated between two or more consecutive frames by optimizing certain objective functions. An example is shown in Figure 2.6 and Equation 2.1, 2.2 (Al-Kofahi et al., 2006), and the implementation of the method will be explained in full detail in chapter 3. This approach is computationally efficient and robust, however, has problems detecting mitosis or segmenting cells at high density where there are multiple touching or overlapping cells (Al-Kofahi et al., 2006).

$$p_{move}(f_i^t, f_j^{t+1}) = \frac{1}{\sqrt{(2\pi)^N |\Sigma|}} \exp\left\{-\frac{1}{2}(d_{ij}^t - \mu)^T \Sigma^{-1} (d_{ij}^t - \mu)\right\},$$

Equation 2. 1 The mathematical expression calculating likelihood for moving cells (Al-Kofahi et al., 2006)

$$p_{divide}(f_i^t, f_{jk}^{t+1}) = \frac{1}{\sqrt{(2\pi)^N |\Sigma|}} \exp\left\{-\frac{1}{2}(d_{ijk}^t - \mu)^T \Sigma^{-1} (d_{ijk}^t - \mu)\right\},($$

Equation 2. 2 The mathematical expression for likelihood of dividing cells (Al-Kofahi et al., 2006)

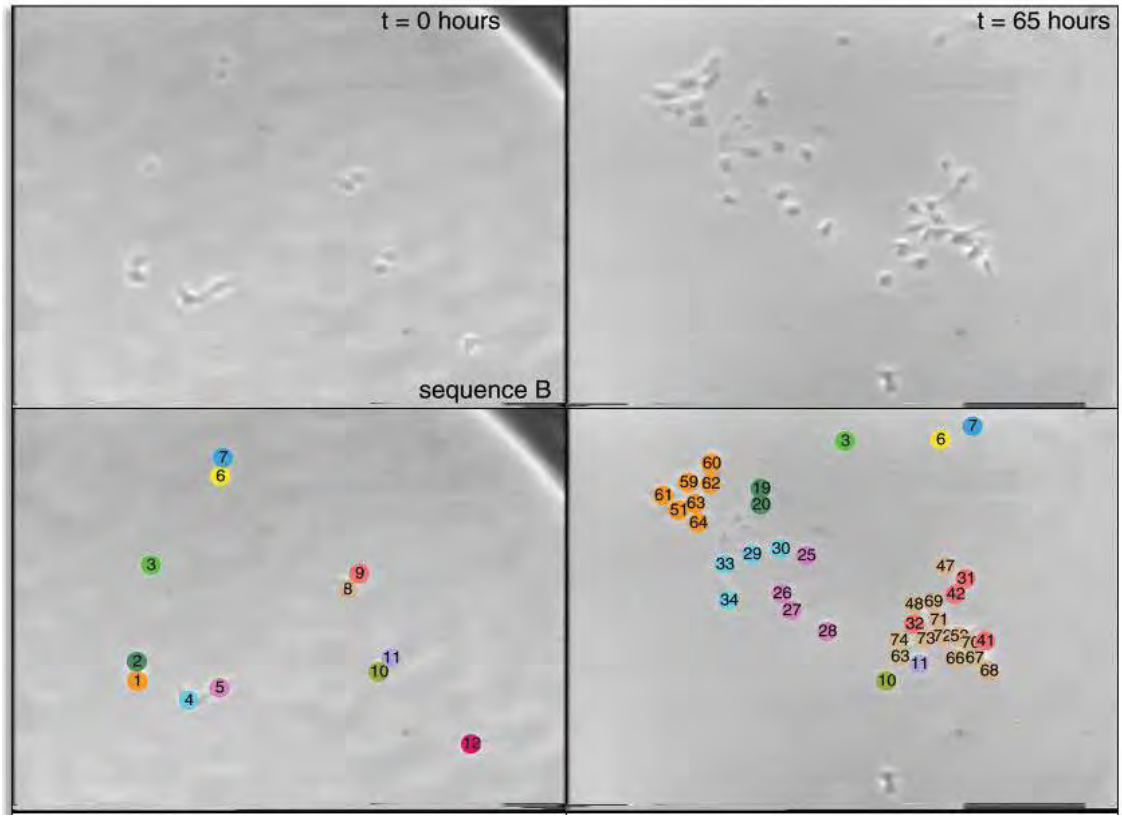


Figure 2. 6 Automatic association of segmented cells between frames(Al-Kofahi et al., 2006)

In the tracking by a model evolution approach (figure 2.7), cells are detected in the first frame, and then the detected cells and their measured features are used as a model to identify and match cells in consecutive frames. The model is also updated with each frame, and so evolves over time.

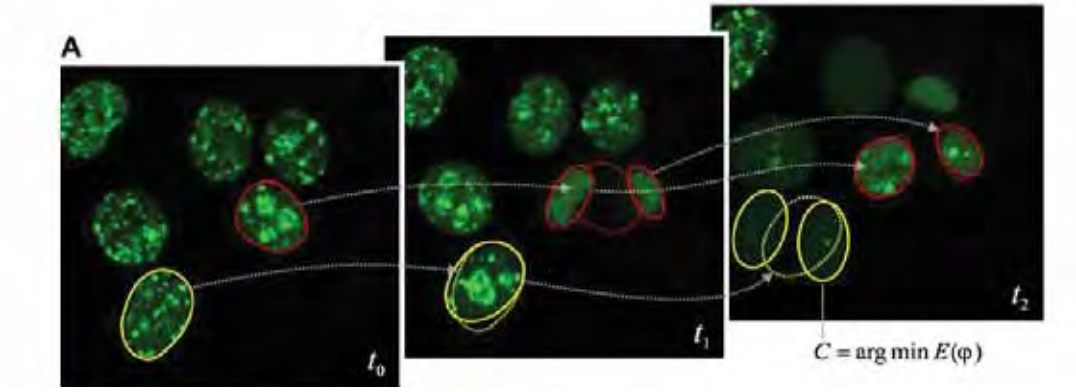


Figure 2. 7 The concept of tracking by model evolution (Meijering et al., 2009)

2.4.2 Cell segmentation

Segmentation is the process of dividing an image into parts of interests, creating a new image containing a label for each pixel indicating which segment it belongs. One approach for segmentation is to compare the value of each image pixel to a preset threshold and to score values above the threshold as foreground (Wu et al., 2010). It is the most commonly used method due to its simplicity, however, it demands high quality images in which cells are well separated and their intensities significantly and consistently differ from the background - a condition which is hard to achieve in live cell imaging process. Significant errors are inevitable when segmenting lower quality images by thresholding, which was illustrated in figure 2.8, due to image noise, autofluorescence, photobleaching, variation in illumination intensity and halos (Meijering et al., 2009)

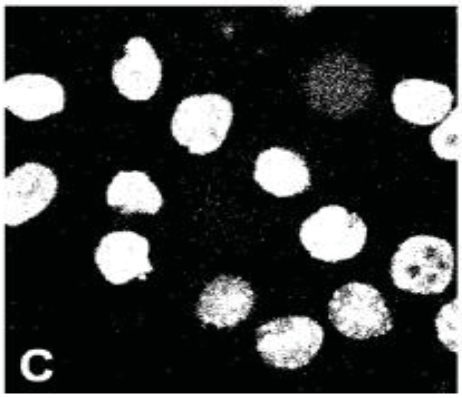


Figure 2. 8 Result of segmentation by thresholding (Meijering et al., 2009)

A more sophisticated method for cell segmentation are based fitting predetermined models to image data (figure 2.9). This template matching approach works well for images showing consistent cell shape, but fails in the case of adherent cells which have variable morphology (Zhou and Wong, 2006). A more popular approach uses the watershed transform. which considers an image as a topographic relief map, with ‘filling’ introduced from the local intensity minima and dividing the image into regions and contours, called by analogy, ‘catchment basins’ and ‘watershed’, respectively. Although this approach has a tendency to be oversensitive to noise

and to yield over-segmentation, it has been successfully applied to cell segmentation in microscopic images, by combining of carefully designed pre- and post-processing strategies, such as marking and model based segment merging. (Al-Kofahi et al., 2006, Debeir et al., 2005, Yang et al., 2006)

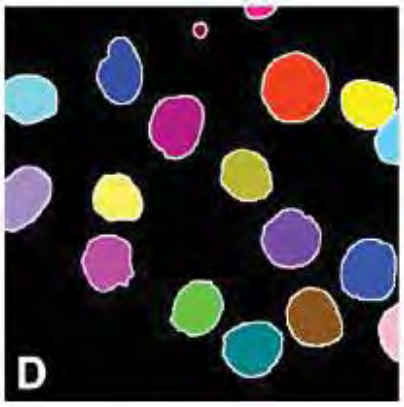


Figure 2. 9 Segmentation by a model based approach (Meijering et al., 2009)

Most recently, interest in the use of deformable models has arisen. These are classified as parametric active contours models (Debeir et al., 2004, Debeir et al., 2005, Zimmer et al., 2002a, Zimmer et al., 2002b, Zimmer and Olivo-Marin, 2005) or as geometric active contour models implemented via level-set functions (Osher and Sethian, 1988).

The latter method is deemed to be superior to the parametric models due to the ability to capture topological change spontaneously. Starting with a coarse, initial segmentation, deformable models are iteratively evolved in the image domain to minimize a predefined energy function. The specification of the energy function, which encapsulates important image features, will determine the active contour segmentation topology (Cohen et al., 2010, Dufour and Olivo-Marin, 2008, Dufour et al., 2005, Dufour et al., 2010, Dzyubachyk et al., 2008, Dzyubachyk et al., 2010b, Jun et al., 2006, Li et al., 2008, Meijering et al., 2009, Padfield et al., 2008, Padfield et al., 2009, Padfield et al., 2006, Swedlow et al., 2009, Xinyu et al., 2009). Typically the energy function is based on image features, such as intensity, gradient, texture, or objective features, such as shape, boundaries, orientation, curvature, and similarity of these features. Such a combination of image features and prior biological knowledge had been successfully applied to segmenting tasks (Zimmer and Olivo-Marin, 2005) . However, limitations still exist, such as the

dependency of the special type and behavior of cells, the requirement for special conditions, and computational intensity.

2.4.3 Cell motion tracking and trajectory mapping

After segmentation, the second step in achieving cell tracking is cell association. This process includes cell identification and linking from frame to frame in the image sequence to create the cell trajectories (figures 2.6 and 2.7). The simplest method is associating cells frame to frame using the cell's centroid using the nearest spatial distance with a constraint that prevents erroneous matching (Debeir et al., 2005). However, this method is unable to handle cells at high density or if cell motion is under sampled. To overcome the under sampling problem, additional cellular features such as intensity, area, volume, texture, surface area, curvature, orientation of axes etc were included in the determination of the closest neighbor (Al-Kofahi et al., 2006). Increasing the number of constraint features can reduce the risk of ambiguity. Similarly, Debeir, et al. applied mean-shift processes to iteratively compute cell position (Debeir et al., 2005, Yang et al., 2006) .

Several cell segmentation methods can also be naturally extended to cell association. The concept of matching the template can serve as a basis for image registration between frames. Registration refers to the process of alignment of images, using intensity or geometry based features (Sage et al., 2005, Sbalzarini and Koumoutsakos, 2005, Yang et al., 2006). Methods based on the deformable models can perform association as using the segmentation results of any frame as the initialization for minimizing the objective function in the next frame with the help of a motion filter. For example, the classic 'snake' model, adaptive snakes which also employ repulsive forces between snakes (Zimmer et al., 2002a) and shift models (Debeir et al., 2005) utilize prior frame information to initialize the next frame. These are again nearest-neighbor linking approaches, which work well when the cell density is not too high and the imaging interval is high enough to limit the cell displacement to less than an average cell diameter.

More sophisticated methods will be required to handle cell collisions, merging and division. There is a growing literature describing new image segmentation methods that may be applied to

live cell imaging. For example gradient- vector flows (Zimmer et al., 2002b) and adaptive motion filters which can handle both constant velocity movements and accelerated movements may be adapted to cell trajectory mapping (Li et al., 2007). Also probabilistic methods such as hidden Markov Chains and Bayesian Estimation may be applied to live cell motion tracking (Li et al., 2007, Li et al., 2008, Smal et al., 2006).

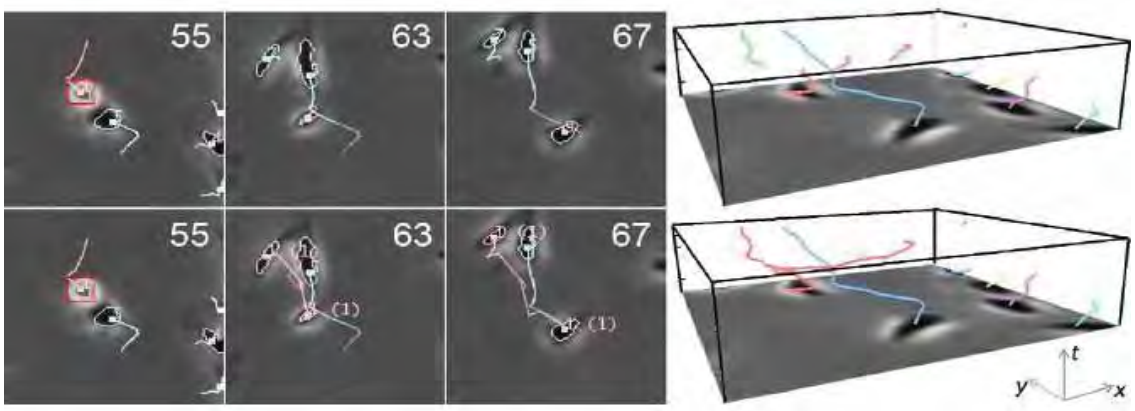


Figure 2. 10 Example of cell tracking and cell trajectories using Bayesian Estimation (Li et al., 2008)

Chapter 3 Methodologies developed for cellular image tracking

3.1 Introduction

To analyze the large amount of image data acquired by time-lapse live-cell imaging, implementation of automatic or semi-automatic image processing methods to identify and track cells, and reconstruct cell pedigrees is a powerful tool to quantify cellular behavior and cell growth. However, there are numerous difficulties automating processing of cellular images limiting progress of this technology. Common problems associated with light microscopy are poor contrast and non-homogeneous image illumination. Cell tracking is impeded by difficulties associated with cellular collisions, irregular cell shapes, and complexities associated with segmentation of phase contrast microscopic images.

Many image segmentation routines rely on homogeneous object illumination and contrast. However most imaging systems have non-uniform illumination and image acquisition hardware may not have uniform pixel sensor gain (CCD camera). Therefore it is a necessity to pre-process images prior to image segmentation and tracking. Classical algorithms for image enhancement that utilize the 2-D Fourier transform either operate in the spatial or frequency domain(Gonzalez, 2008). A preprocessing step to improve the uniformity of image intensity and contrast is described in section 3.2.1, and section 3.4.1.

The accuracy of cellular segmentation routines is limited when cells merge or touch. This is an unresolved problem for the whole field. Well-established classical algorithms for image segmentation have been implemented, namely thresholding, edge detection and morphological dilation and erosion. A maximum curvature technique was used to separate objects with circular geometry, namely fluorescently labeled nuclei. Moreover, an interactive region growing algorithm was used to define cell boundaries for fluorescently labeled cells or phase contrast images. The algorithms that were developed are described in section 3.2.2 and section 3.4.2.

The image tracking software contains six components as listed in figure 3.1; image enhancement, image segmentation, object tracking, manual confirmation, cell cycle estimation and pedigree construction. Manual confirmation is required to review and edit automatic segmentation results. Minimal user actions for editing such as mouse positioning and clicking are desirable because this will shorten analysis times.

Image enhancement consists of contrast enhancement and background non-uniformity correction. Cells were segmented using edge detection, global thresholds, curvature estimation, morphological erosion and dilation as well as employing specific constraints based on cell specific properties. Cells were tracked over consecutive frames by matching cells with almost identical classifiers using a maximum likelihood and nearest neighbour metrics. Cell cycle estimation was based on a robust statistical approach, namely estimation of the empirical probability distribution for cell cycle times using the Kaplan Meier estimator. Division histories and differentiation pathways were displayed as pedigrees.

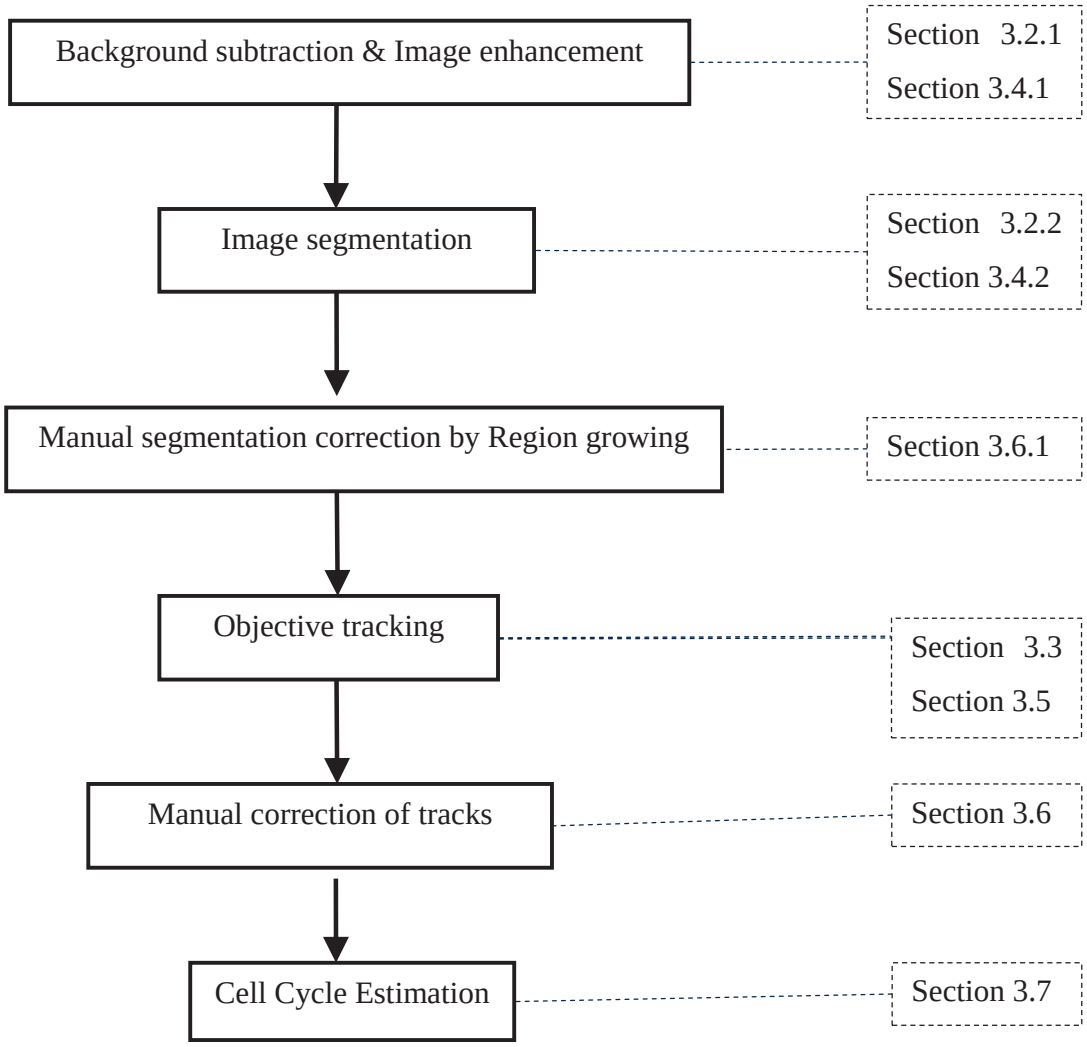


Figure 3. 1 Diagram of semi-automated cell division tracking software

3.2 Fluorescent cell image segmentation

This section described the process of how cell pedigrees are segmented from fluorescent cell image sequences.

3.2.1 Image background bias correction

The background of fluorescent images is non-uniform due to the non-uniform illumination of excitation light. Figure 3.2(a) is a background image (no objects) whose 3D intensity profile ideally should be a smooth horizontal plane. However, the 3D intensity profile of a background image (figure 3.2(c)) is typically non-uniform, the pixel intensity increases gradually towards the center of image, and decreases at the four corners of the image. Noise which results in a 'rough' surface, is also an important factor that limits contrast and edge detection.

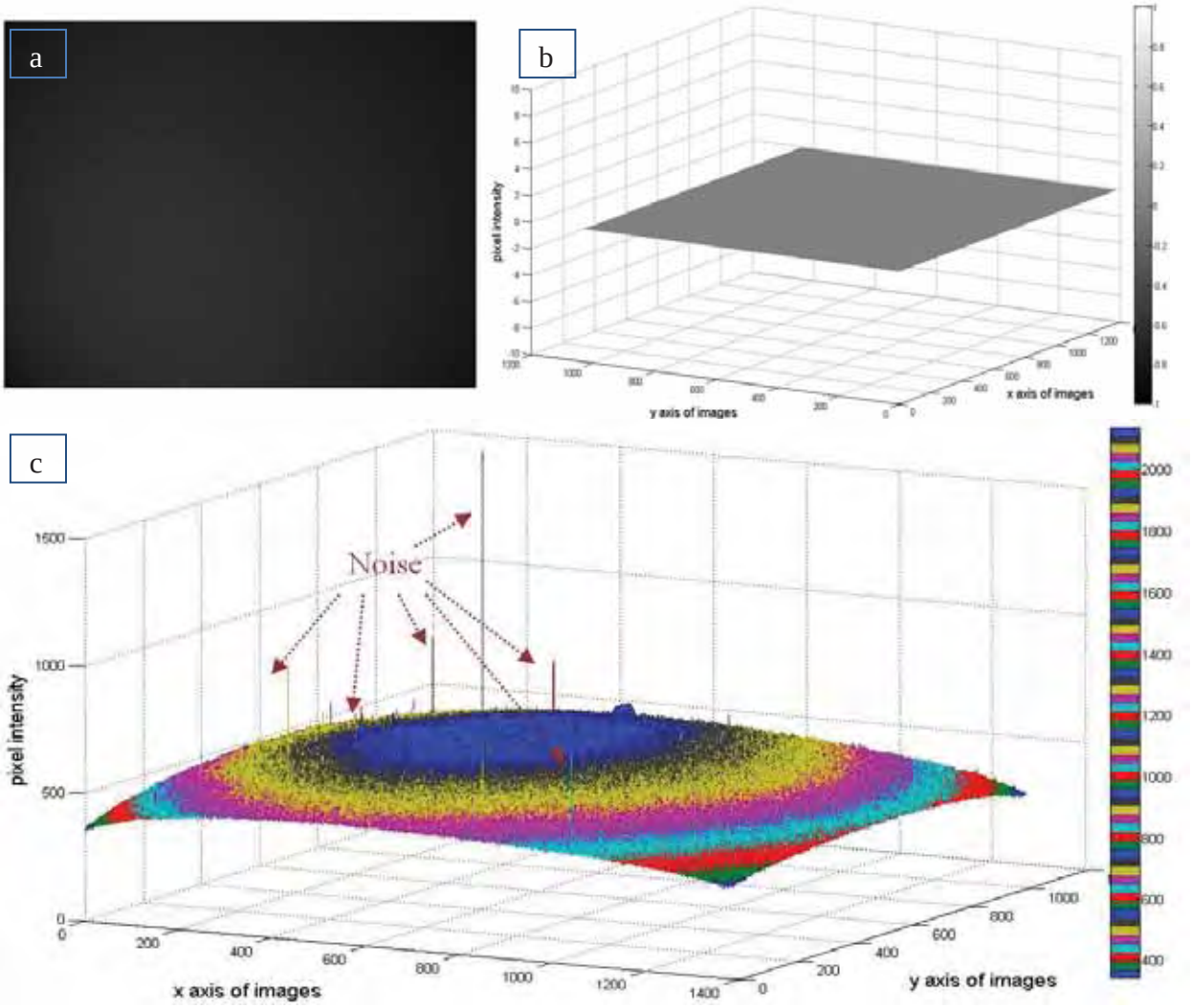


Figure 3. 2 Non-uniformity of background image intensity (a) Background image taken without objects, (b) intensity profile of ideal background, (c) intensity profile of the experimental background have background bias and noise.

3.2.1.1 Rolling-ball filter

The rolling–ball filter is a grayscale mathematical morphological filter, which was first proposed by Sternberg in 1983.(Sternberg.S, 1983) It simulates rolling a ball beneath the intensity profile of an image, removing the peaks that are untouched by the ball surface. To implement this algorithm, some backgrounds of mathematical morphological operations are essential for understanding.

Mathematical morphological image processing(Gonzalez, 2008) is based on moving a pre-defined structural element along the image, which is similar to spatially convolving the structure elements with the object image, however the process is based on set operations and therefore is nonlinear. The four important operations are dilation, erosion, opening and closing. Erosion is a 'shrinking' or 'thinning' operation, figure 3.3. Dilation is the opposite operation of erosion, which 'grows' or 'thickens' objects.

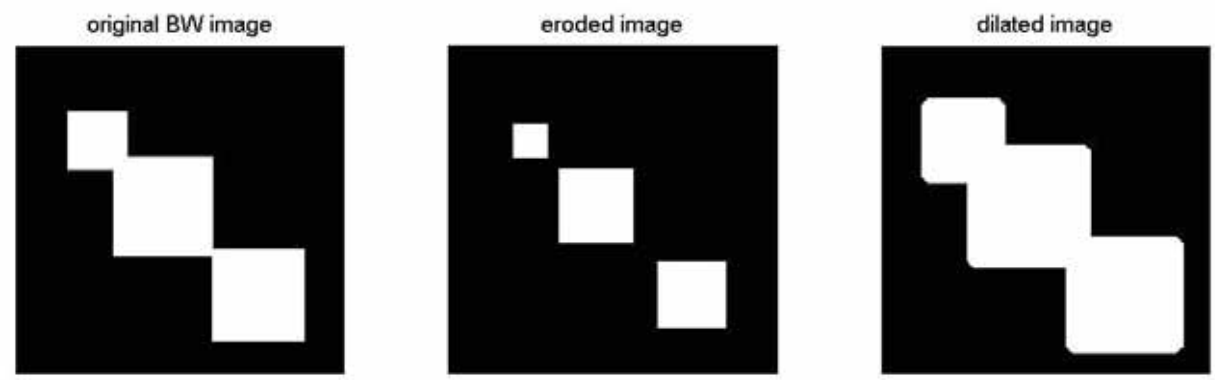


Figure 3. 3 Dilation and erosion with disk-shaped structure element sized 20 pixels

Opening(Gonzalez, 2008) performs erosion first, followed by dilation, while closing operates dilation before erosion. Generally, opening smoothes the contour of an object, breaks narrow isthmuses, and eliminates thin protrusions, while closing fuses narrow breaks and long thin gulfs, eliminates small holes and fills gaps in the contour. Definition of opening and closing operations(Gonzalez, 2008) are listed as equations below,

$$(\boldsymbol{f} \circ \boldsymbol{b})(\boldsymbol{x},\boldsymbol{y}) = \left[(f \Theta b) \oplus b\right](\boldsymbol{x},\boldsymbol{y}) \qquad \textbf{Equation 3. 1}$$

$$(\boldsymbol{f} \bullet \boldsymbol{b})(\boldsymbol{x}, \boldsymbol{y}) = \left[\left(\boldsymbol{f} \oplus \boldsymbol{b}\right) \Theta\, \boldsymbol{b}\right](\boldsymbol{x}, \boldsymbol{y}) \qquad \textbf{Equation 3. 2}$$

where *f(x,y)* and *b(i,j)* are two discrete functions on two dimensional discrete space. *f(x,y)* is the pixel value of *(x,y),* and *b(i,j)* is a structure element of disk shape.

Different shape and size of structure elements, in combination of the 4 basic morphological operations are achieve versatile aims in image processing. For binary images, 2-dimensional structure elements with different shape (square, disk, etc.) are utilized for various morphological results. While for gray scale images, 2 dimensional structure elements with a height or 3 dimensional structure elements are more useful. In this study, the rolling ball filter is implemented via the classical top-hat transform, which filters image with 3D ball-shaped structuring elements(Li et al., 2008) to perform opening operation, followed by closing operations.

$$\boldsymbol{rollball}(\boldsymbol{I}, \boldsymbol{r}) = \boldsymbol{I} - \boldsymbol{tophat}(\boldsymbol{I}, \boldsymbol{ball_r}) \qquad \textbf{Equation 3. 3}$$

where r is the radius of the rolling-ball, and $ball_r$ is a ball-shaped structural element with radius equal to r.

Cells isolated from the PDGFRα-mouse have nuclear accumulation of green fluorescent protein in cells which express the PDGFRα receptor (see Chapter 5). They have been imaged using fluorescent microscopy (figure 3.4). Comparing with the original image, figure 3.4(a), the background bias was removed from the image in figure 3.2(b). The image has a smoother more uniform background fluorescence. The estimated background bias by rolling ball filter is displayed in figure 3.2(c). However, the surface of background bias is far from smooth, which might introduce background image artifacts to the image. These variations are difficult to detect by eye, but will affect the accuracy of segmentation algorithms based on edge detection and thresholds.

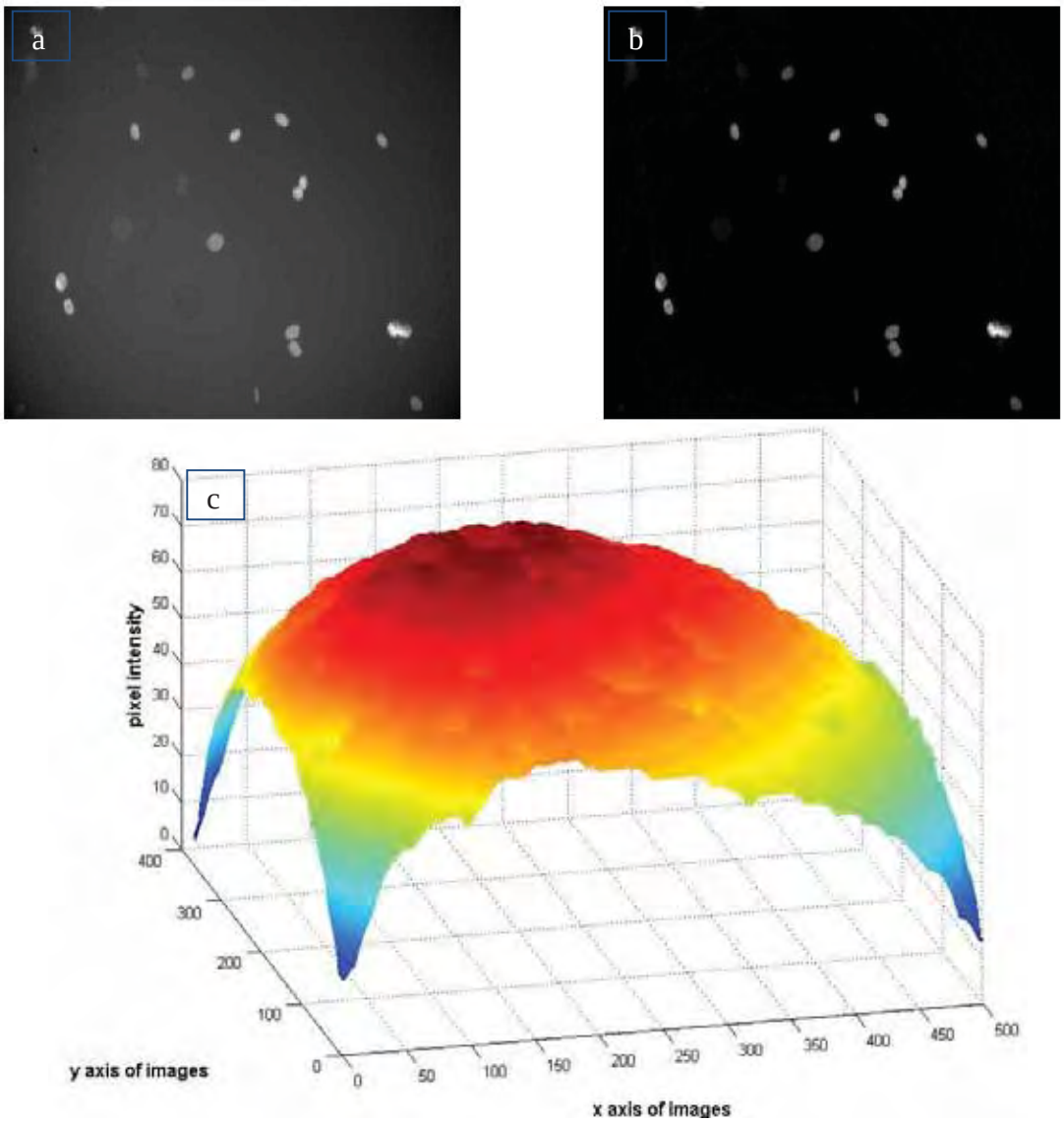


Figure 3. 4 Demo of background estimation by rolling ball filter, a: original image; b: image after correction; c: estimated bias of background

3.2.1.2 Polynomial surface fitting

The second method used to calculate a background bias was to fit a polynomial surface to the background intensity. The equation of the polynomial equation is

$$\mathbf{f(x,y) = \sum_{i=0,j=0}^{n,m} c_{ij} x^i y^j} \qquad \textbf{Equation 3. 4}$$

The surface was fitted to background image data by linear regression (Matlab Curve Fitting Toolblox). A polynomial with order $n = 3, m = 3$ was used to estimate the bias for the image shown in figure 3.5. Note that a polynomial fit is smoother than the rolling ball filter, and is does not create background image noise.

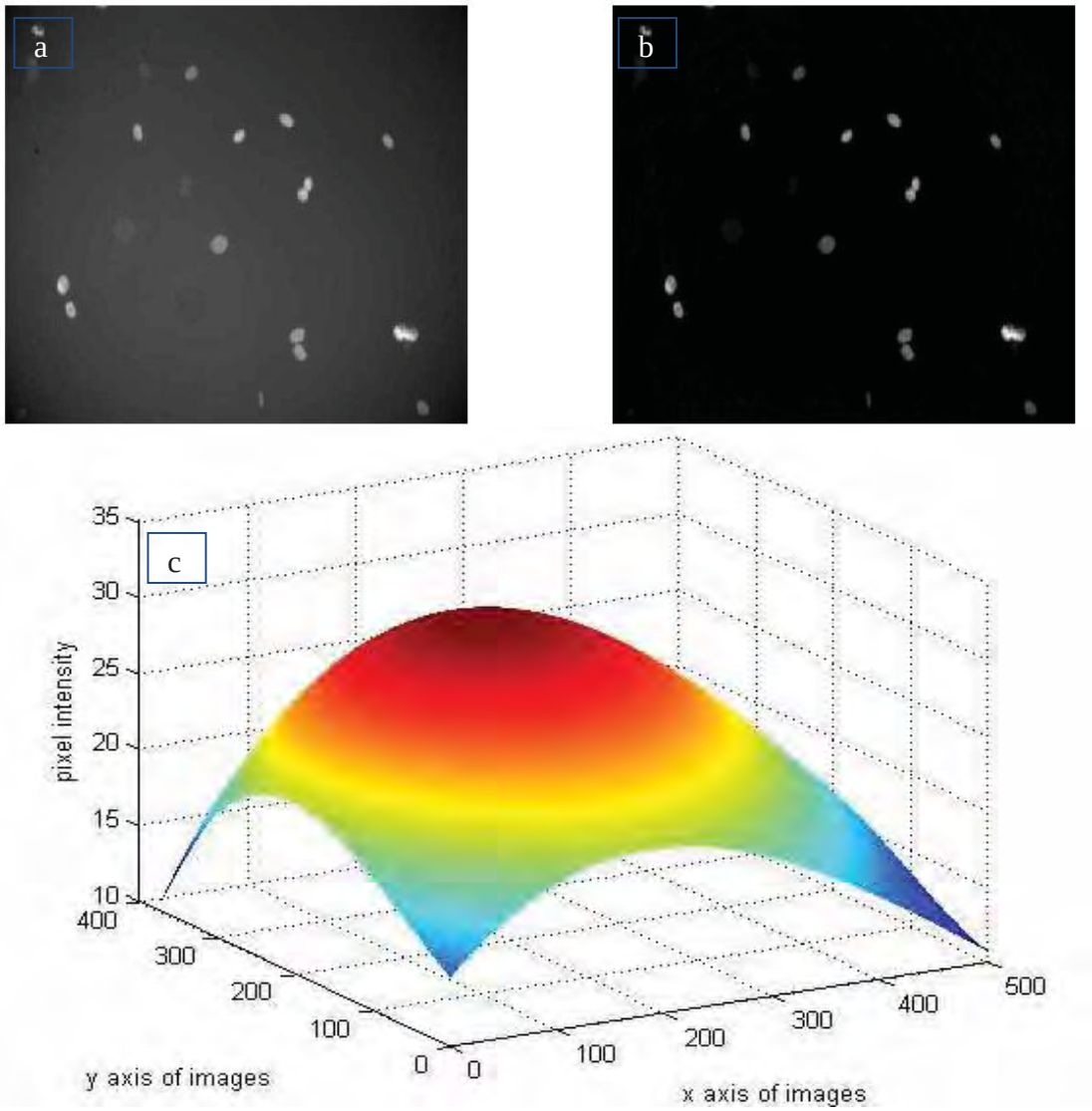


Figure 3. 5 Demo of background correction by polynomial fitting, a: original image; b: image after correction; c: background bias estimated by polynomial fitting

3.2.2 Image background subtraction

3.2.2.1 Intensity thresholding

The simplest way for thresholding is to rely on the histogram of pixel intensity distribution, shown below (figure 3.6).

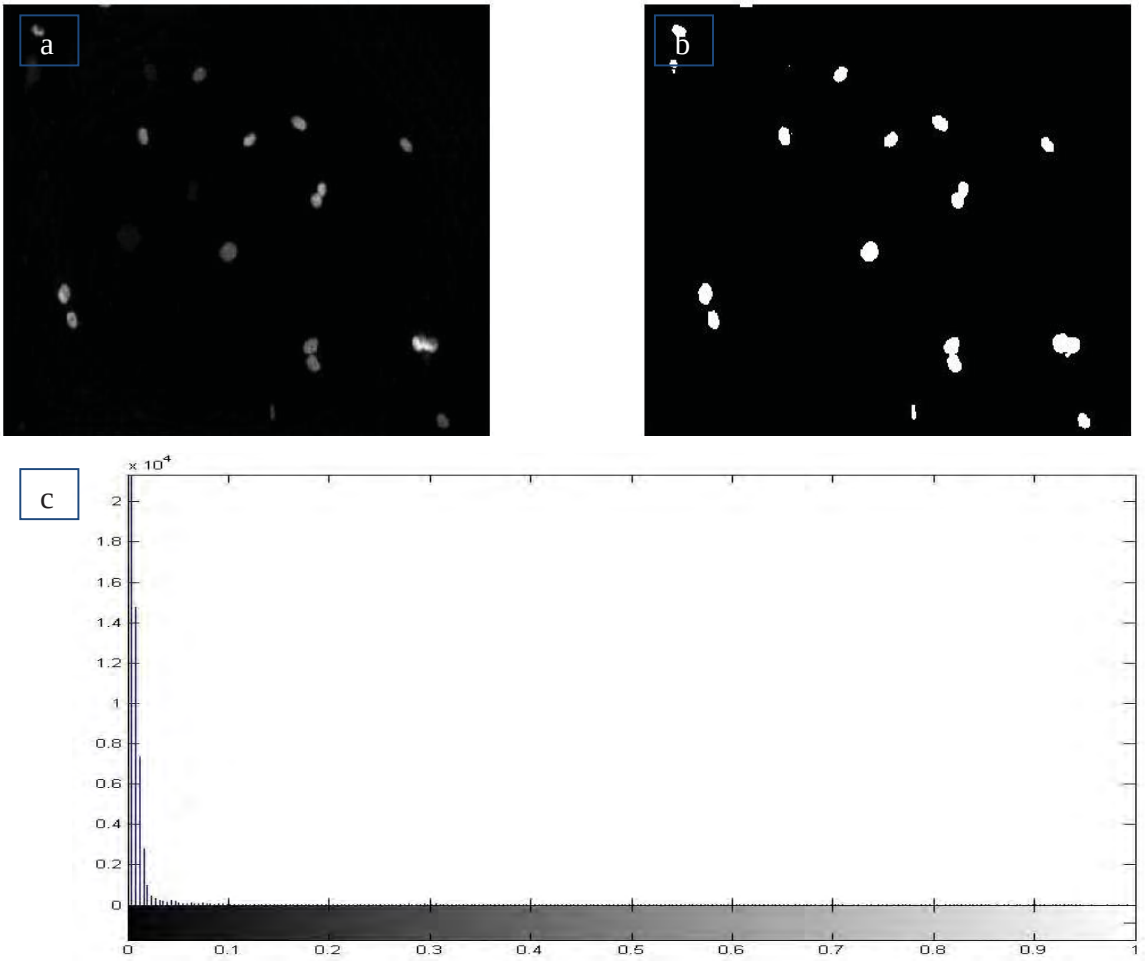


Figure 3. 6 Illustration of thresholding using the image intensity histogram (Threshold=0.1), a: original gray scale image; b: binary image after thresholding; c: histogram of pixel intensity distribution

A level for thresholding is selected using the pixel intensity distribution (figure 3.6c) and foreground information is depicted as a binary image (figure 3.6b).

Otsu's method(Otsu, 1979, Li et al., 2008, Gonzalez, 2008) for automatic threshold determination was used in this study. The Otsu one-level thresholding algorithm assumes that the image to be thresholded contains two classes of pixels (i.e.,. foreground and background), and then calculates the optimum threshold separating those two classes so that the intra-class variance (the weighted sum of variances of the two classes) is minimised(Li et al., 2008).

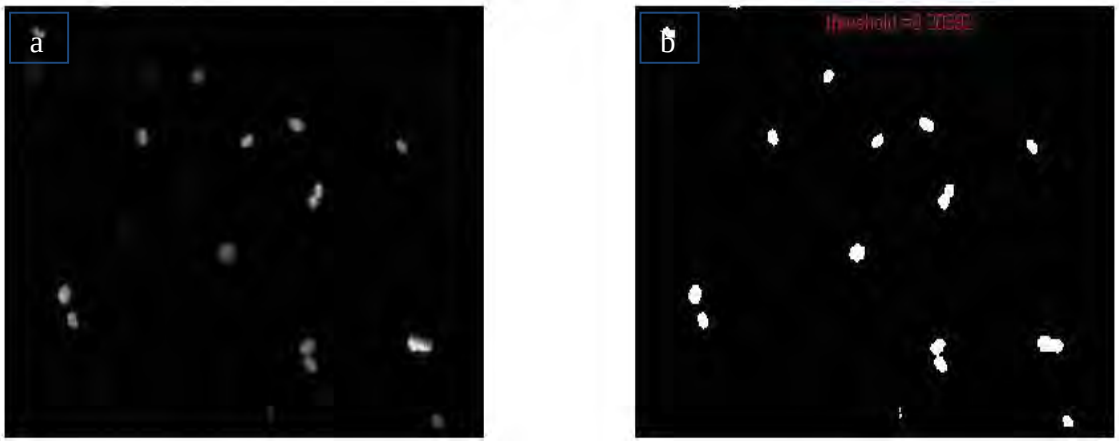


Figure 3. 7 Illustration of auto-thresholding, a: original gray scale image; b: binary image after thresholding

3.2.2.2 Maximum curvature determination

As shown in the images above, cellular nuclei that make contact are recognized as a single nucleus. Therefore an algorithm to separate nuclei was required. There are two categories of algorithms proposed in literature for splitting merged cells(Makkapati and Naik, 2009). The first method fits a region, such as an ellipse(Gonzalez, 2008), to segment the dumbbell shape into two separate nuclei. The other method detects the maximum curvature of an edge circumscribing the thresholded binary image. In this thesis, the maximum curvature method(Makkapati and Naik, 2009) was applied.

Edge detection on a binary image results in an edge that circumscribes the object or blob(Gonzalez, 2008). But for phase contrast cell images, edge detection algorithms seldom are as successful as the human visual system. The 'Canny' edge detection algorithm explained in detail in chapter 3.3.2 was used in this study.

Figure 3. 8 Edges (Red) of blobs detected on black-white binary image

Once the edges of each identified blob is segmented, and the curvature of each points along the edges are calculated using the equation shown below(Gonzalez, 2008),

$$k = \frac{|x'y'' - y'x''|}{(x'^2 + y'^2)^{3/2}} \qquad \textbf{Equation 3. 5}$$

where x and y are coordinates of the boundary points. The derivatives are computed by convoluting the boundary with Gaussian derivatives. Examples of detected points with curvature whose k values are larger than the threshold are shown in figure 3.9.

Figure 3. 9 Points detected with the maximum curvature value (blue points)

The cleft between touching circles results in an edge with maximal curvature. All the maximum curvature points are then connected in pairs. Only those chords below a specified length constraint are used to separate merged nuclei (figure 3.10).

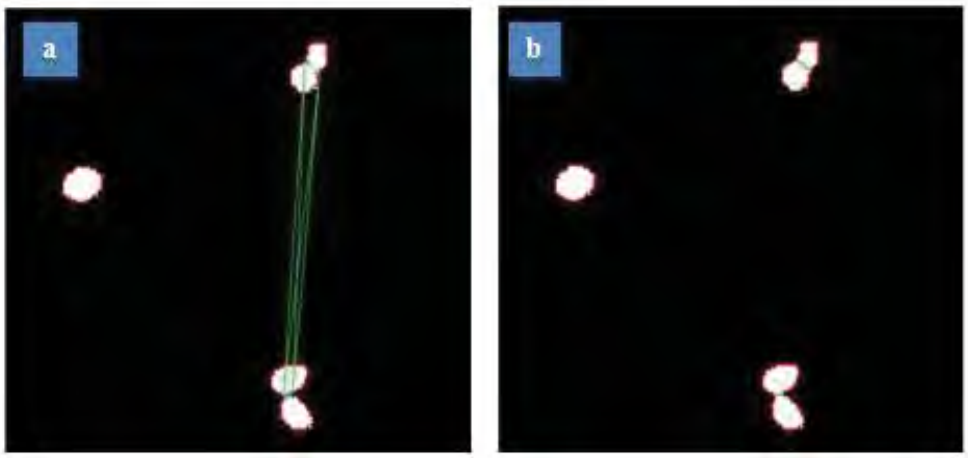


Figure 3. 10 Illustration of blob splitting algorithm a: connection of maximal curvature points b: application of length constraint

3.2.2.3 Noise removal

In this thesis image noise is defined as segmented foreground information that is not object information. In this case the object information is cellular nuclei. Most of the noise not

eliminated by thresholding of fluorescent images as above is small. For example, the cell fragment at the bottom of the image shown in figure 3.11 would have been segmented using a fluorescent threshold, however this is clearly not an intact cell and would be considered as ‘image noise’. However, for phase-contrast images, removal of image noise is a major problem. Briefly covered in Section3.2.1, image erosion(Gonzalez, 2008) is one method to remove trace amounts of noise, smoothing object edges at the same time. This however is only effective for the removal of small object noise. Increasing either the eroding structuring element or the number of erosions can remove large object noise. However, cellular objects are also eroded, thus eliminating crucial cellular details in the process. Therefore, image erosion is not an optimal method for noise removal in this study.

Removal of image noise requires a two-step process; first, blobs or components are identified and labelled using pixel connectivity algorithms(Gonzalez, 2008). Then each component is classified using its pixel characteristics i.e., area, mean intensity, centroid, radius etc. Labelled components which have ‘typical’ pixel characteristics and fall within defined constraints are accepted as objects of interest, whilst outliers are considered to be image noise. The final segmented image has image noise removed (Figure 3.11). This method is elaborated in the following sections.

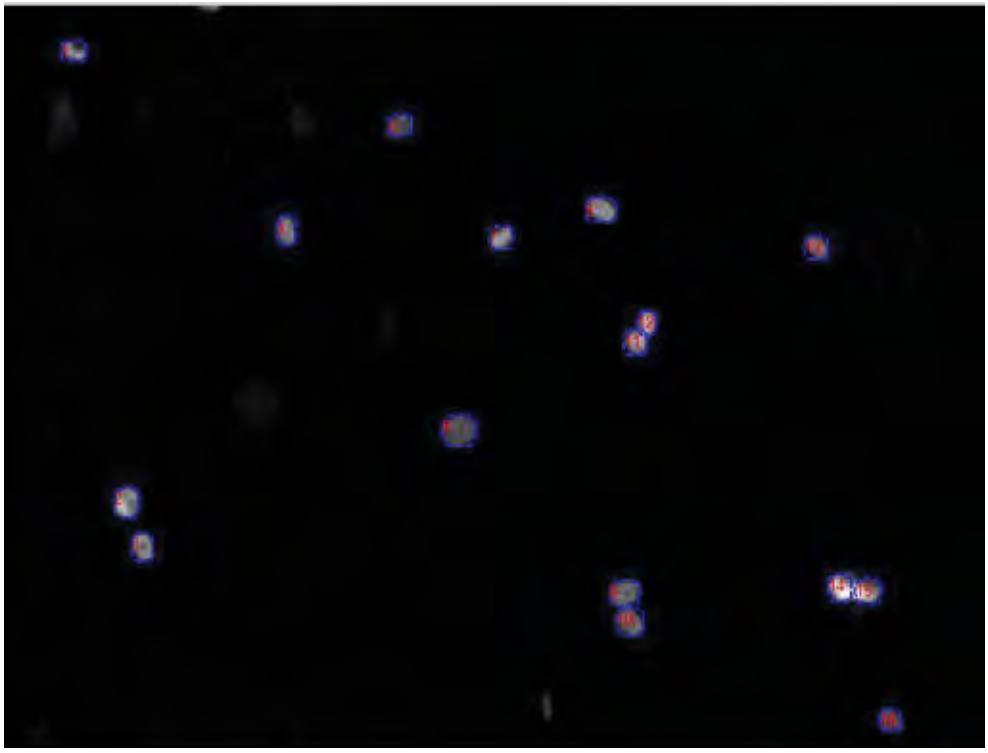


Figure 3. 11 Final segmentation result with the noise information removed

3.2.2.4. Object connectivity

Object connectivity is a fundamental process essential to binary region based image processing. It is a method used to differentiate objects based on pixel border relations(Gonzalez, 2008). A pixel in a two dimensional image is a square with four faces. In three dimensions the small image unit is a voxel, a six-faced cube. A group of pixels linked together form an image object. Object pixels are “connected” or linked either at the faces or the corners, hence the term connectivity. Standard image connectivity is defined as either “4” or “8-connectivity”(Gonzalez, 2008). Pixel connectivity is evaluated in the manner shown below:

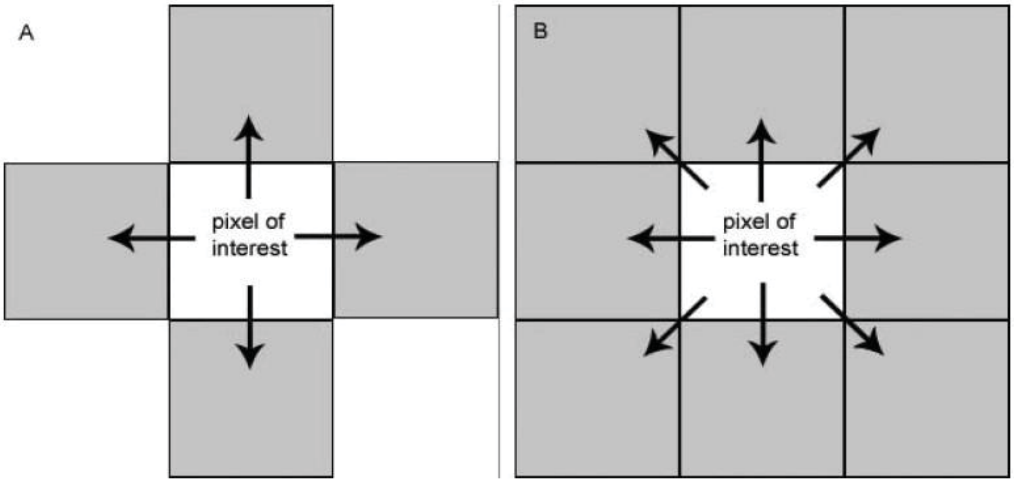


Figure 3. 12 Connectivity (a) 4-connectivity of interest pixel; (b) 8-connectivity of interest pixel

Any pixels adjacent to the edges of the pixels of interest or corners are considered “4-connected” figure 3.12(a). In the other case of figure 3.12 (b), object pixels either adjacent to the corners or edges of the pixel of interest are designated “8-connected” pixels.

3.2.2.5. Component labelling

As stated in the previous section, 8 and 4-connectedness are two pixel classification methods employed in 2D image processing. Adoption of either 8 or 4-connectedness will result in quite different regions as shown in figure 3.13(b-c).

The choice of “8-connectivity” to label the original object in 3.13 (a) extracts one single object, formed by two apparent squares. In contrast, “4-connectivity” labelling produces two separate objects or squares numerically labelled in fig 3.13 (a). The discrepancy is a very subtle problem, but has quite a pronounced effect on the final results, as illustrated in 3.13 (b-c).

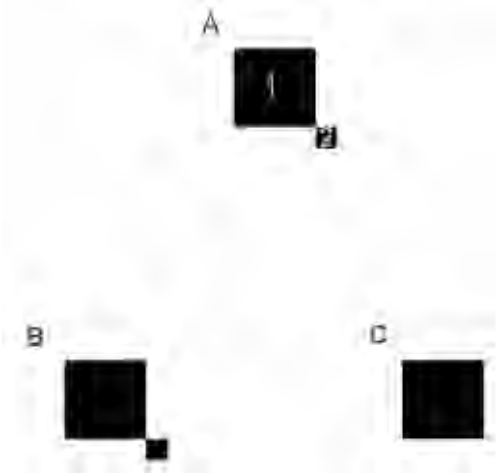


Figure 3. 13 Component labeling; (a) connected components, (b) component labelling by 8-connectivities, (c) component labelling by 4-connectivities

Many component labelling algorithms require the implementation of split/merge algorithms to group pixels defined by their "connectedness" in a specific region together. The image is searched and the pixels assigned to single objects using the connectivity method are labelled with consecutive numerical values. After the labelling of each image object, mainly arithmetic and geometric descriptive measures can be calculated to characterize these objects. The most useful characteristic in image noise removal is pixel area defined as the number of pixels forming a certain object; other useful characteristics will be described in the next section. Object size is an important parameter because the majority of the noise present in the image foreground segmented using a simple threshold are objects with surface area much smaller than the size of a cell or nucleus. In this study, 4-connectivity is chosen because it excludes image noise. The final segmented cellular image is displayed above in fig 3.13.

3.2.2.6 Component feature measurement

After segmentation, 6 features of each nuclei were determined for each labeled component ('regionprops' function in Matlab image analysis toolbox).

1. Centroid (x_{ij}, y_{ij}) is the most common feature obtained to track cellular migration. It is essentially the centre of mass of the pixel distribution within the nuclear object body, figure 3.14(a).
2. a_{ij} is the number of pixels belongin to the component and represents cell volume, figure 3.14(a) .
3. To interpret the other features, an ellipse model is fitted to the segmented cell regions. l_{ij} is the length of long axis. o_{ij} is the angle between the long axis of fitted ellipse and the horizontal line, denoted in figure 3.14b. e_{ij} is eccentricity of each cell, which is denoted as the ratio of the length of the long and short axis, figure 3.14(a).
4. M_{ij} is the average pixel intensity of each segmented blob in figure 3.15. It is calculated on the combined images, which is the multiplication of gray scale images and segmented binary images.

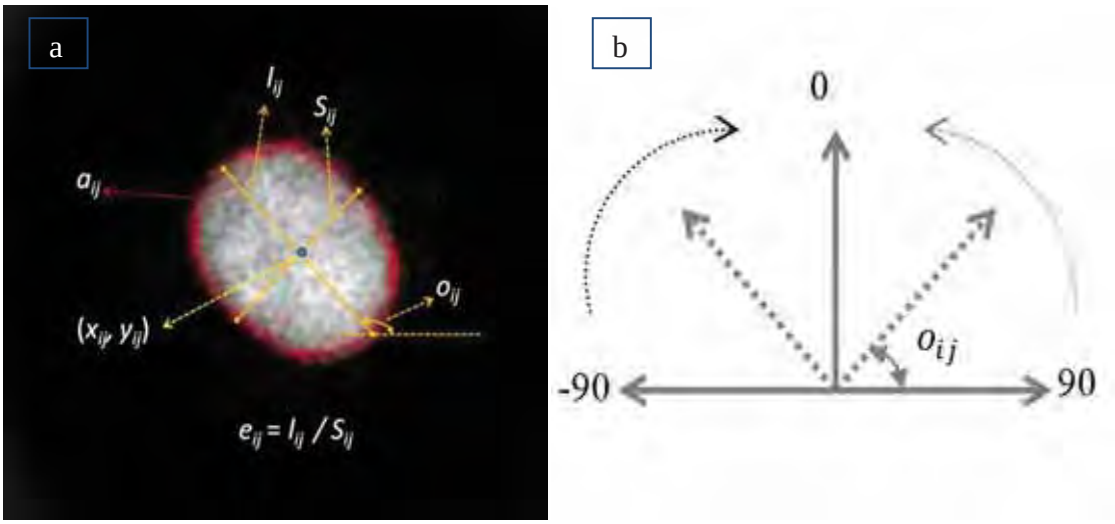


Figure 3. 14 Feature measurement (a) calculation of component parameters, (b) calculation of orientation angle

The six features are stored as a Matlab structure array (figure 3.15).

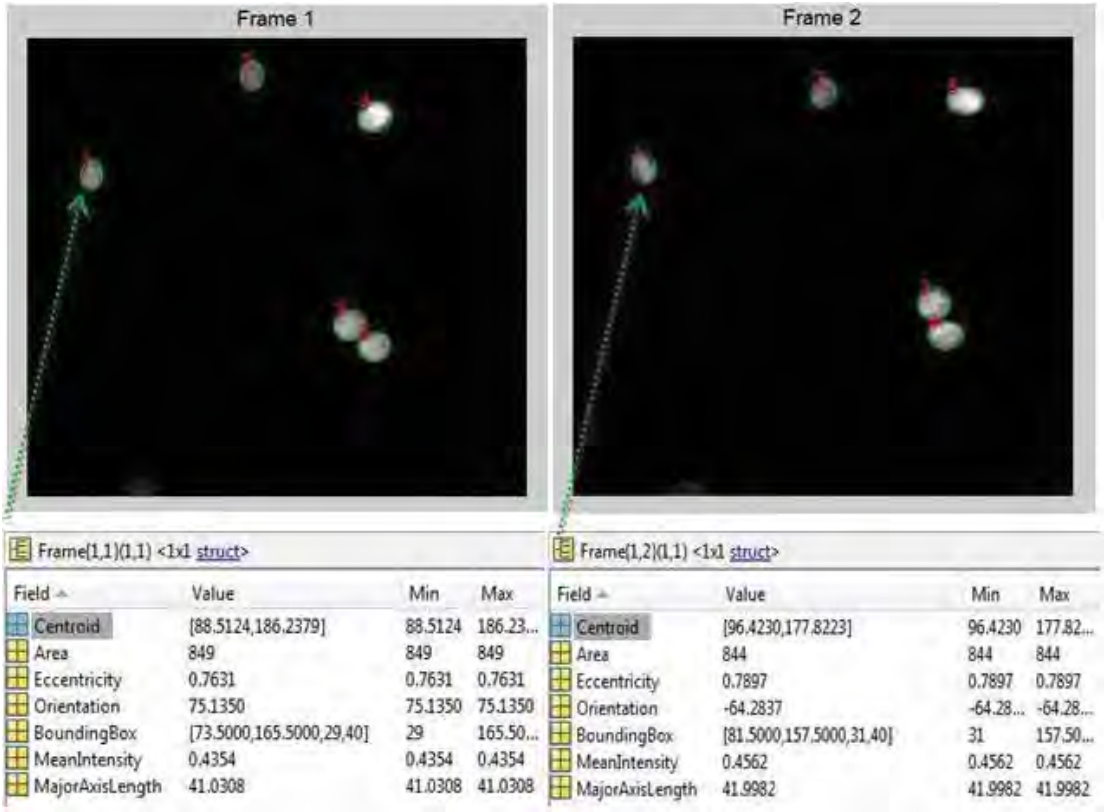


Figure 3. 15 Segmented cells in two consecutive frames and features recorded

3.3 Cell trajectory mapping by feature matching

Segmentation generates a sequence of images with segmented objects classified by their features as shown in the Matlab table (figure 3.15). The identification of specific cell states requires high quality of cell images and high segmentation accuracy. In our experiment, to track cells over time, a simple three-step feature matching method was adopted: i) Construction of likelihood-score matrices for all cells in consecutive images. ii) To calculate frame-to-frame correspondence matrices and iii) To identify cell divisions and assign mother-daughter relationships. Suppose that the image sequence has P frames, image j has m cells and image j+1 has n cells. A forward score matrix (m by n) and backward score matrix (n by m) describes the likelihood of a cell in image t corresponding to a cell in image j+1 and that of a cell in image j+1 corresponding to a cell in image j, respectively. These matrices are further combined and transformed into a binary correspondence matrix with 1 indicating a correspondence between a

cell in image t and a cell in image j+ 1. These matrices are further transformed into a tracks-information matrix and a cell-life matrix. The tracks-information matrix adds morphological descriptors for each time set in the track, and the cell-life matrix annotates cell division, apoptosis or a 'lost' cell. Cell division is identified by tracking the image sequence backwards to identify two tracks that merge to one track. Cell death is defined as a sudden stop of track within the tracking period where cells are observed to become non-adherent and disappear from the field of view. A lost cell is one that migrates outside of the field of view.

To construct forward and backward scores for each pair of cells, a likelihood score is assigned to cell pairs. The likelihood calculator for cellular tracking has been well developed by Al-kofahi in 2006 (Al-Kofahi et al., 2006). For cell i on frame j, features f_{ij} are associated using parameter μ and ϵ, which are the previously estimated mean and covariance of a difference vector *f* calculated using a correctly (manually) labeled training set. The likelihood of cell movement between cell i in frame j and cell k in frame j+1 is calculated using a multivariate normal probability distribution, where N is the number of features examined, and the difference vector $d_{ik}(j) = |f_{ij} - f_{k(j+1)}|$, is the difference in the features of cell pairs.

$$\boldsymbol{P\left(f_{ij}, f_{k(j+1)}\right) = \frac{1}{\sqrt{(2\pi)^N|\epsilon|}} \exp\left[-\frac{1}{2}(d_{ik}(j) - \mu)^T \epsilon^{-1}(d_{ik}(j) - \mu)\right]} \qquad \textbf{Equation 3. 6}$$

In our application, eight features $f_{ij} = (x_{ij}, y_{ij}, a_{ij}, e_{ij}, l_{ij}, o_{ij}, m_{ij})$ described in section 3.2.2.6, are sufficient to characterize each cell. Cells segmented in two consecutive frames are compared in pairs. The difference of features between each pair of cells in two consecutive frames is calculated below,

$$\boldsymbol{d_{ik}(j) = \begin{bmatrix} |a_{k(j+1)} - a_{ij}| \\ |x_{k(j+1)} - x_{ij}| \\ |y_{k(j+1)} - y_{ij}| \\ |e_{k(j+1)} - e_{ij}| \\ |l_{k(j+1)} - l_{ij}| \\ |o_{k(j+1)} - o_{ij}| \\ |m_{k(j+1)} - m_{ij}| \end{bmatrix}} \qquad \textbf{Equation 3. 7}$$

A similar formula is used to calculate a backward difference vector which is the difference between features in frame i and i+1. Note that the forward and backward difference vector are not always equal. To combine information from the forward and backward likelihood matrices into a single correspondence matrix, we start by noting that a cell in image t can disappear (die), grow, or divide into two cells in image j+1. So, a cell in image j can only correspond to 0, 1, or 2 cells in image j+1. We also assume that a cell in image j+1 can correspond to at most 0 or 1 cells in image j, which restates the fact that a cell is an orphan or a descendent of a cell in image j. We then combine the two score matrices using the dot product (element wise multiplication, not matrix multiplication) and calculate the row and column maxima in the combined score matrix; our algorithm then scans the combined score matrix to locate values that are both row and column maxima and assigns matches for corresponding cell pairs. During this process, if more than two cells in image j+1 can be assigned to one cell in image j, the backward score matrix is referenced so that the two cells with best scores are used to assign the matches. We repeat this process while excluding the cells in image j+1 that are already matched and enforcing the cell growth event restraints until no more correspondences can be found.

For example, cells i in frame 1 (figure 3.15(a)) is associated to cell k in frame 2 (figure 3.15(b)), and stored in forward matrix $P_f(i, k)$,and backward matrix $P_b(k,i)$ in figure 3.16. The scores marked by red circle are the scores higher than the threshold. In the backward score matrix, the row maxima is the only one be chosen. While in forward score matrix, one or two cells can be found in each row. However, in the ‘Combine’ matrix where the backward and forward matrix multiplied together, the false matches were discarded by searching the row and column maxima. In figure 3.16(d), the combine matrix is depicted as a pixel heat map with good matches having warm colours (red).

Backword <5x5 double>

	1	2	3	4	5
1	0.9545	0.6544	0.3384	0.4295	0.2886
2	0.6289	0.9105	0.4845	0.7230	0.4355
3	0.3189	0.3830	0.9196	0.7122	0.9102
4	0.2671	0.3371	0.9091	0.6464	0.9399
5	0.3365	0.6109	0.5827	0.911	0.5747

Forward <5x5 double>

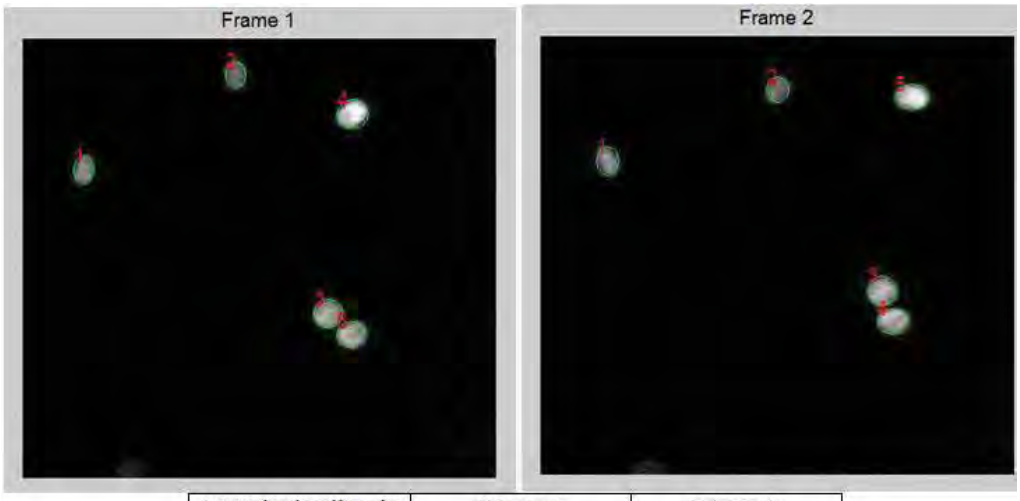
	1	2	3	4	5
1	0.9543	0.6289	0.3189	0.2671	0.3365
2	0.6544	0.9105	0.3830	0.3371	0.6109
3	0.3384	0.4845	0.9196	0.9091	0.5827
4	0.4295	0.7230	0.7122	0.6464	0.9113
5	0.2886	0.4355	0.9102	0.9399	0.5747

Combine <5x5 double>

	1	2	3	4	5
1	0.9111	0.4115	0.1079	0.1147	0.0971
2	0.4115	0.8291	0.1856	0.2437	0.2660
3	0.1079	0.1856	0.8457	0.6475	0.5304
4	0.1147	0.2437	0.6475	0.4179	0.8565
5	0.0971	0.2660	0.5304	0.8565	0.3303

Figure 3. 16 Sample scoring matrix calculated for cells in figure 3.15

Once the correspondence matrices for all consecutive images are generated, we reconstruct the tracks-information matrix by linking the sequence of correspondence matrices. Individual tracks are obtained by backward linking the cells identified in the last frame to the cells in the first frame. For example, cells in figure 3.15 are finally associated in figure 3.17:



Matched cell pair	Frame 1	Frame 2
Pair 1	Cell1	Cell1
Pair 2	Cell2	Cell2
Pair 3	Cell3	Cell3
Pair 4	Cell4	Cell5
Pair 5	Cell5	Cell4

Figure 3. 17 Feature matching results (a) frame 1 (b) Frame2 (c) associated pairs of cell

3.4 Phase contrast image segmentation

This section describes the methodology used to segment cell division pedigrees from phase contrast image stacks.

3.4.1 Contrast normalization and background bias correction

Normalization is also called contrast stretching, which changes the range of pixel intensity values. The purpose of dynamic range expansion is to achieve consistency in dynamic range for a sequence of images and to maximise the range of pixel intensity values. For cell tracking normalisation of grey scale range is particularly important when using phase contrast images(Gonzalez, 2008).

For our image data, the image to the intensity range was normalised to an 8 bit image intensity range (0 to 255) after correction for background bias. The method used for background bias correction was polynomial surface fitting described in section 3.2.1.2. Figure 3.18 compares the original image, subtracted background bias, and image after correction and normalization.

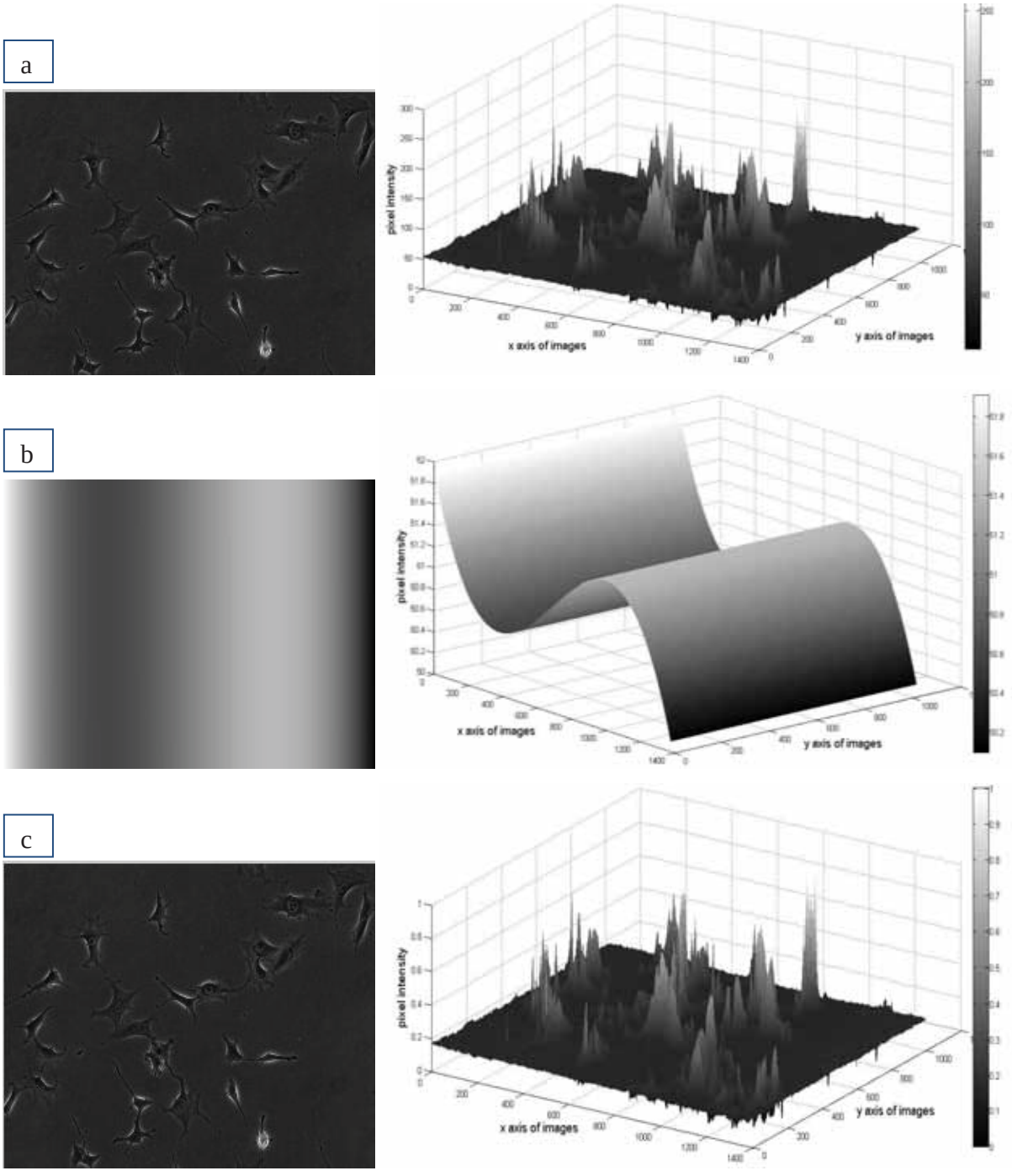


Figure 3. 18 Comparison of images before and after processing, a: original image; b: Background estimated by polynomial fitting; c: Image after background correction and normalization

3.4.2 Background subtraction

Phase contrast images generate contrast from changes in phase that occur due to refraction through the cell body, or the way light is reflected at the substratum (resulting in phase reversal and dark regions). Phase contrast images of adherent cell growth have common features; cell attachment to the substratum is darker than background, whereas cell thickness or rounding creates a signal which is much brighter than background. When a cell undergoes mitosis or division, three major phases can be observed: first, the cell will detach from the surface and begin to round up; second, the cytoplasm of the dividing cell separates into two distinct regions forming a characteristic dumbbell shape; and lastly, the two daughter cells are created from a single dividing cell. Following cell division and plating, cells will interact with their surroundings by forming attachments via integrin receptors with the underlying surface. Upon maturation and accumulation of these attachments the cells will begin to elongate or spread. Once fully spread, the cells will migrate along the underlying surface.

Distinction between spreading and motile cells is sometimes difficult, especially when fully spread cells develop mature contacts to commence migration. Migration of cells is not possible without cell spreading because spreading forms many of the immature contacts with the substratum necessary to initiate cellular migration. The major distinguishing factor between spread/motile cells from dividing cells is a characteristic bright halo enclosing a dark cell body of low pixel intensity.

Dividing cells are easily distinguished because a bright halo forms at cytokinesis. This may be due to the increased height of the cell, and refractive index differences between the cell body and the surrounding fluid. When observed in time lapse, divisions are easily visualized as transient bright spots. Since bright cells are distinct from spreading and motile cells, separate thresholds and segmentation methods are implemented for cell spreading and mitosis.

3.4.2.1 Edge detection

Many edge detection methods have been developed because of its central importance in image processing. Region boundaries and edges are closely related, since there is often a sharp adjustment in intensity at the region boundaries. The basic principle behind most cellular edge detectors is to locate regions of maximum change in pixel intensity. The regions of maximum change can be determined from “zero-crossings” of the first derivative of the intensity profile (Gonzalez, 2008).

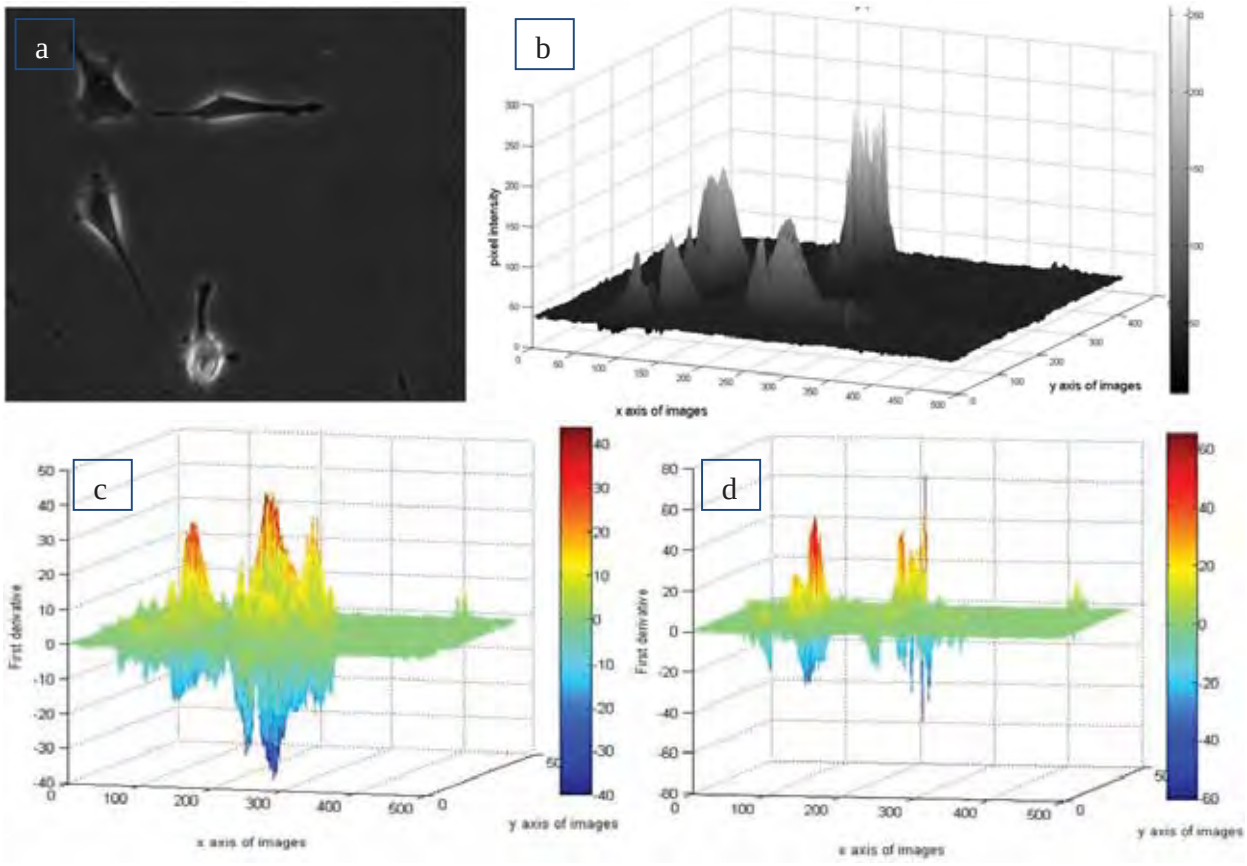


Figure 3. 19 (a): Cropped demo image (432x495); (b): The image intensity profile; (c): The first derivative of image in x direction; (d): The first derivative of image in y direction

The 2D- first order partial derivative of the image intensity profile is called the image gradient. In fig 3.19 (c) and (d), the gradient of image is saturated with zero crossings due to the fluctuation in background pixel intensity. But the zero-crossings of interest are adjacent to local

maxima. Zero-crossings associated with each maximum correspond to the increasing pixel intensity caused by the prominent bright cell halo, and to a lesser extent the darker cell body. Sensitivity parameters are incorporated into the edge detection method to only detect fluctuations within a specific range and to filter out ubiquitous zero-crossings contributed by variation in background pixel intensity, Uneven illumination produces additional fluctuations leading to errors in edge detection(Gonzalez, 2008).

A Canny edge detector was utilized in this study to identify cellular edges. Application of the Canny edge detector is pictured below in fig 3.20 (a) & (b). The Canny edge detector has three primary goals: low error rate, edge points should be well localized, and the region boundary should be localised to a single edge(Gonzalez, 2008). For a mathematical discussion of this algorithm see Canny's publication (Canny, 1986, Gonzalez, 2008). The algorithm can be summarized as the following steps:

1. Smooth the image with a Gaussian filter
2. Compute the gradient magnitude and angle images by calculating the 2D-first order partial derivatives of each pixel.
3. Apply nonmaxima suppression to the gradient magnitude image
4. Use high-and-low level thresholding (figure 3.20 (b)): a high threshold for low edge sensitivity and a low threshold for high edge sensitivity. Edge starts with the low sensitivity result and then grows it to include connected edge pixels from the high sensitivity result.
5. 8-connectivity analysis (chapter 3.3.3) to detect and link edges

As illustrated in figure 3.20(b), Canny edge detection together with well defined two level threshold can isolate regions of large contrast between the cellular membrane and the background; however, the imaging dilemma described in the literature review poses many barriers for complete and continuous edge detection of the cellular boundaries. Hence, the deficiencies of cellular edge detection can be offset by combination with other image processing methods. We combined the results of edge detection and bright halo thresholding together in figure 3.20.

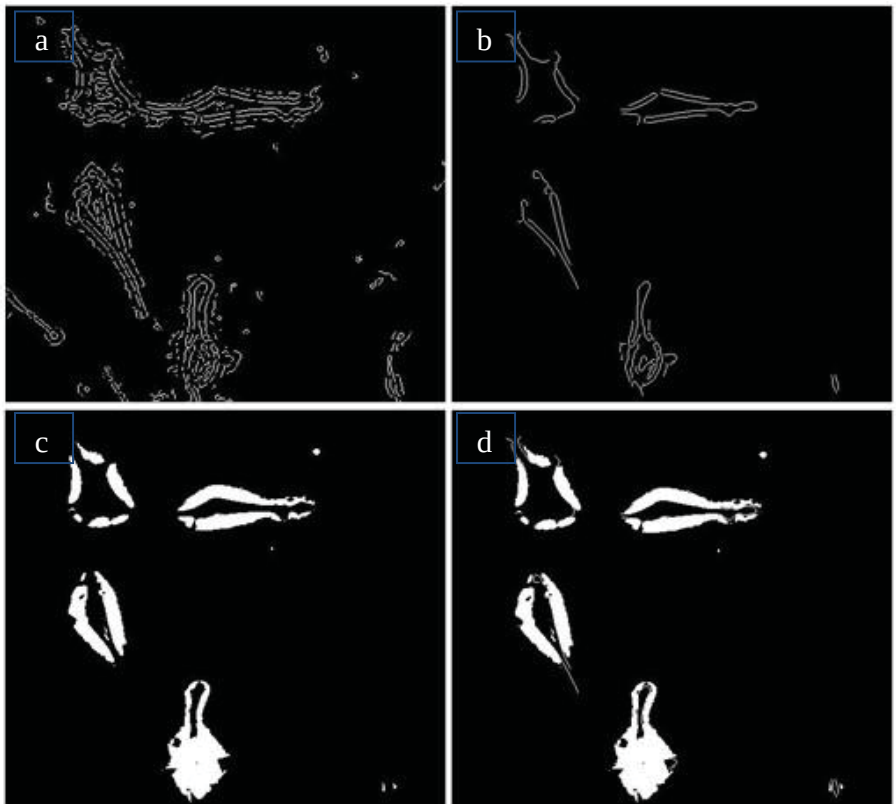


Figure 3. 20 Comparison of Canny edge detection and thresholding (a) Canny edge detection, (b) Canny edge detection with high-low threshold, (c) thresholding of bright halo, (d) combination of (b) and(c)

3.4.2.2 Object closing and filling

Although the combination of edge and thresholded information recovers much of the image information of the cell bodies, quite a bit of information is still "lost". Once again, the majority of the information is lost in the cellular boundaries of low contrast. Image object closing(Gonzalez, 2008) can be used to roughly interpolate regions of discontinuous cellular boundaries. As described in section 3.2.1.2, object closing consists of binary image dilation followed by erosion, which closes or interpolates the gaps within or in between objects governed by structuring elements (SE). Binary image dilation uses an SE or a binary mask similar to the one described before to determine local pixel intensity variation, however, the mask used in image dilation, more commonly referred to as an SE, can assume many basic geometric shapes.

Certain SE shapes are suited for specific image processing conditions. The size of SEs may also vary, depending on the type of segmentation desired and also the size of the objects(Gonzalez, 2008). Similar to the rolling ball filter (3.2.1.1), the binary SE is moved and centered along each pixel. In dilation, the mask assigns the maximum value determined within the mask region to the pixel of interest. In contrast, erosion assigns the minimum value. The size and construction of the SE structure and binary element values govern the image segmentation.

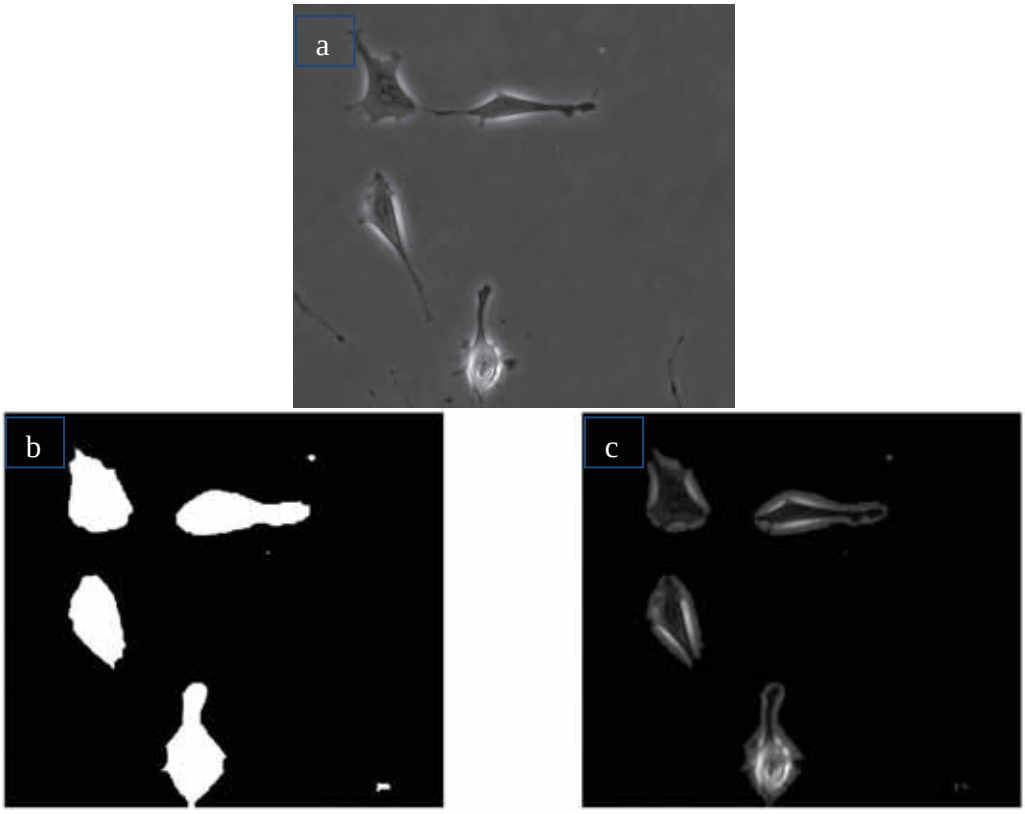


Figure 3. 21 Segmented cells, a: gray scale image; b: blobs founded by image filling; c: images have the background 'zeroed'

3.4.2.3 Further processing

Application of the background subtraction eliminates the majority of the background pixels, isolating approximate regions which contain the following information:

1. Majority of the pixels comprising the cell body

2. Discontinuous cell edges

3. Bright cell halos

The combination of edge detection and histogram thresholding excluded discontinuous cell edges due to the low contrast of these structures. As previously discussed, poor image contrast of structures creates an overlap in pixel intensity between pixels of interest and the background. However, since the majority of background pixels have been zeroed, the overlap in pixel intensity between background and discontinuous cell edge pixels becomes less pronounced. The separation of background pixels from the pixels of interest is accentuated in the intensity maps displayed below in fig 3.22 (b & c).

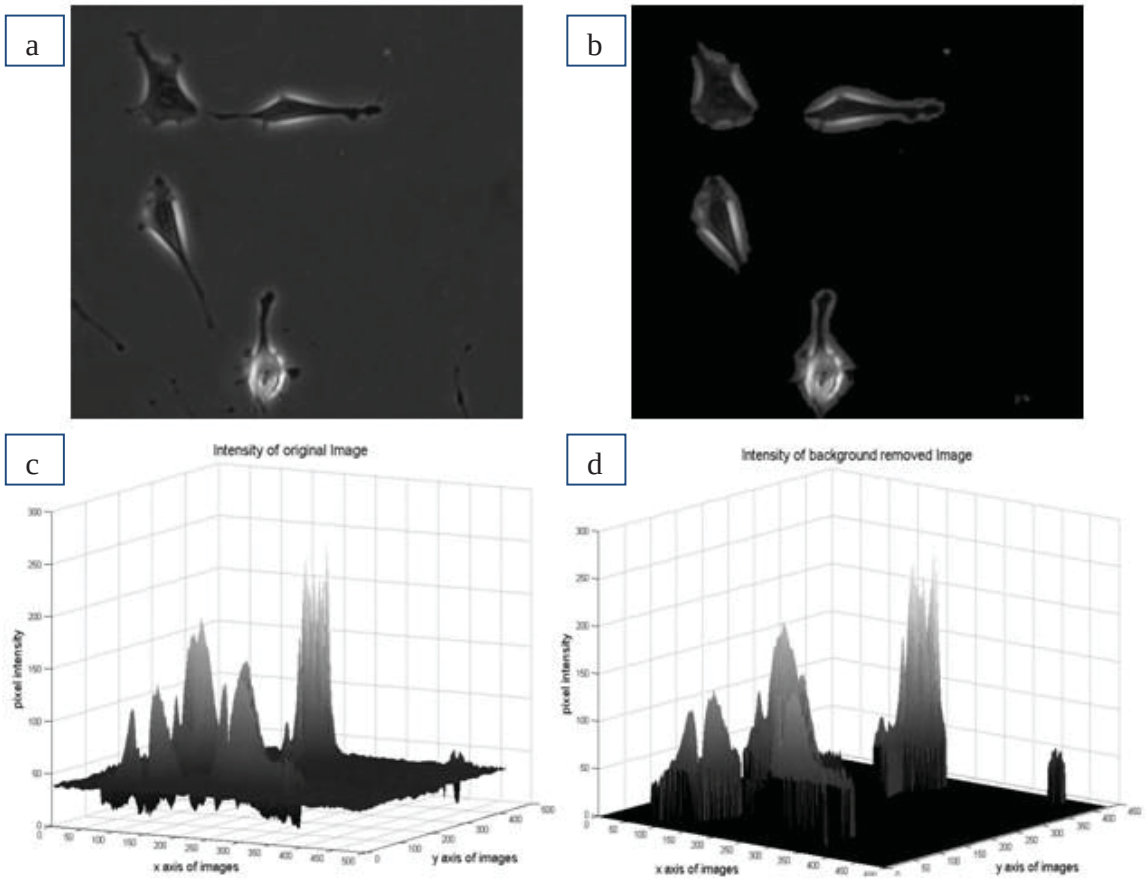


Figure 3. 22 Gray scale image and image with background mask in 2D and their 3D intensity profile, (a) gray scale image, (b) gray scale image with background mask, (c) , (d)

Now that a large separation between the background and the pixels of interest exists, a final two-level global threshold range can be utilized to isolate the cell bodies. The lower limit of the threshold range is set to a pixel intensity of 1, since the background is zero. The upper limit is set to the previously manually determined upper threshold. Everything falling in between the lower and upper threshold limits is incorporated. This new threshold range is adjusted and expanded to incorporate the discontinuous cell edges producing a more complete segmentation of the desired cell bodies, figure 3.23(b).

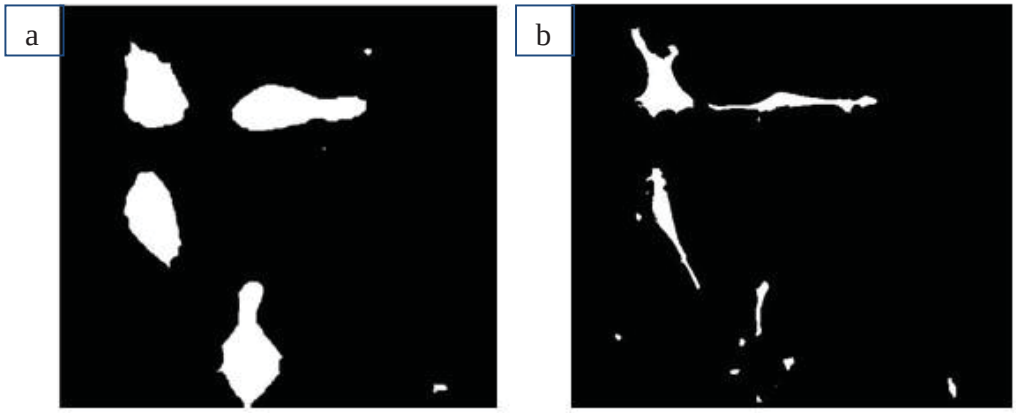


Figure 3. 23 Compare of segmented object by background mask (a) and second global thresholding (b)

As described earlier, dividing cells (fig 3.24(a)) can be easily segmented via global thresholding due to the pronounced bright pixels comprising the dividing cell, and also the extremely well defined boundary. The large separation between the background and bright pixel intensity, as illustrated in fig 3.24 (c), makes simple global thresholding tremendously effective, therefore the creation of a background mask as described in section 3.4.2.2 is needless. Parts of the bright cell are not incorporated into the thresholded cell in fig 3.24 (b), due to the dark intensity of the cell nucleus. In the case of dividing cells, binary object filling can be implemented to fill or interpolate the gaps from the thresholding because the cell boundaries are continuous and well-defined.

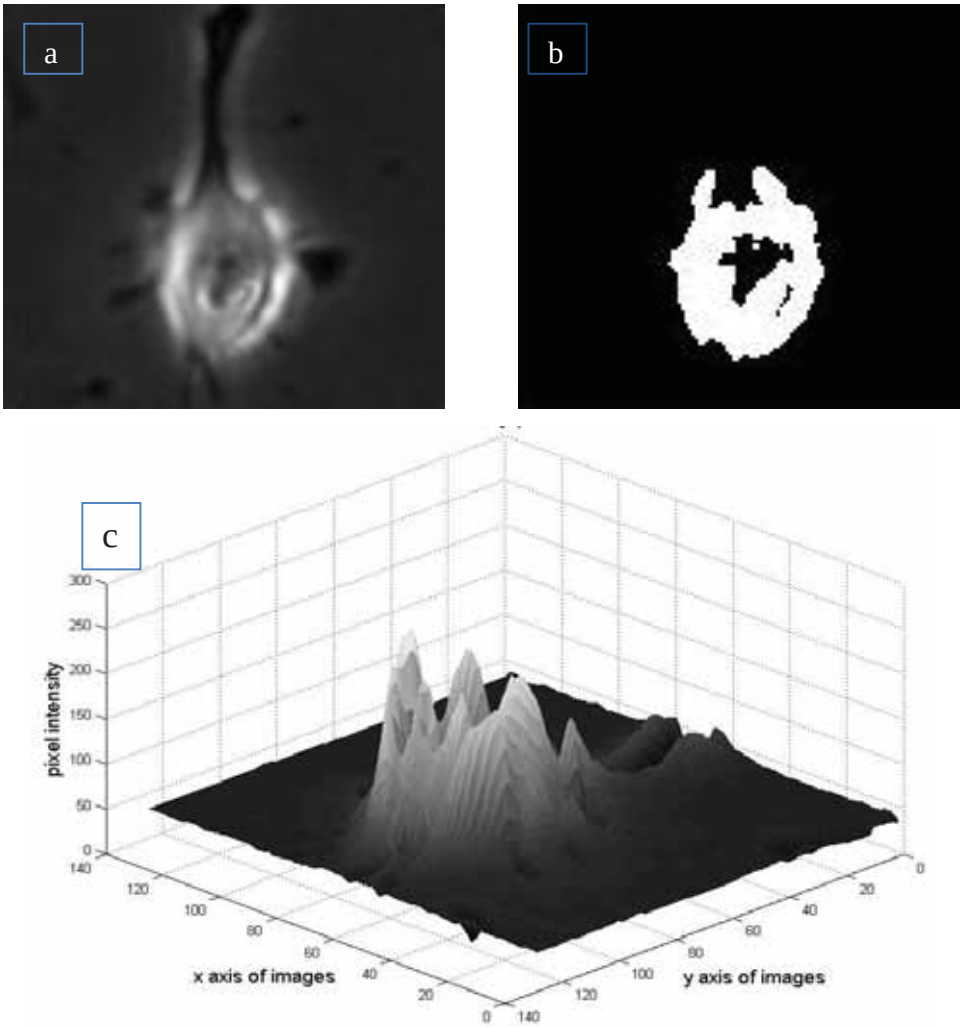


Figure 3. 24 Mitotic cell and segmentation by thresholding (a) mitotic cell, (b) binary segmentation results by thresholding, (c) intensity of gray scale image of mitotic cell

The last step of segmentation is removing the remaining noise, noted as small dots enclosed by blue lines in figure 3.25(a). Here, we use the component labelling method stated in section 3.2.2.3.2 to remove these objects to avoid the over-erosion problem, which would result if on applied erosion. The final result of cellular segmentation on phase contrast image is showed in figure 3.25(b).

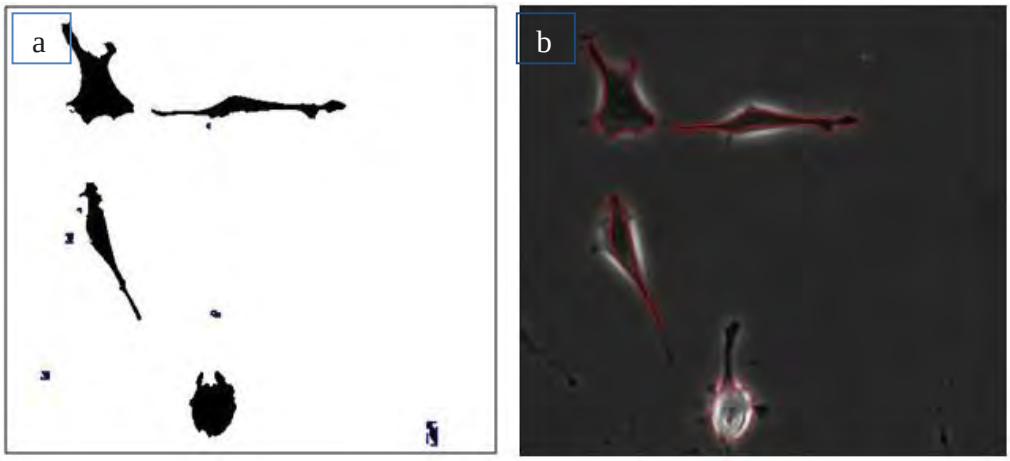


Figure 3. 25 (a) Eroded binary image with noises, (b) Final result of segmentation with red contour

3.5 Tracking by nearest neighbour determination

Although several features of segmented cells can be obtained from phase contrast images, the tracking of objects is hard to achieved using the feature matching algorithm developed for fluorescent nuclei. The main obstacles are:

1. The frequent shape changes for fibroblasts and other adherent cell types.
2. Cell contact as cells grow to confluence. There are no suitable methods for correctly splitting connected adherent cells on phase contrast images.

Because or irregular geometry of adherent cells it was not possible to define a feature mapping training set as was developed for matching fluorescent cell nuclei (section 3.3). For this study we adopted a nearest neighbor algorithm(Gonzalez, 2008) for cell centroid tracking. When cells are well separated and the frame rate of the image sequences is sufficient, cells can be distinctly separated by the distance of between cell centroids(location) for consecutive images because of small changes in cell position without large changes in cell morphology. As a result, cells have the smallest distance in two consecutive frames are associated together, which is similar to feature matching in section 3.3, but the only feature that is matched is the centroid. Also by tracking backwards, the division of cell is identified as the merging of two different tracks into

one. Even though cell shape is heterogeneous, centroid position is a robust cellular feature that is commonly used to track fibroblast.

The component labelling technique described earlier is quite useful to calculate the centroid of all the cells within an image. All the pixels of each cell are accessed individually via the labelling matrix from component labelling, and the centroid is determined individually for each cell. Alternately in this study, the distance transform (DT) can also be implemented to determine a central reference point for each cell.

3.5.1 Distance transform and quasi-center of mass

Distance transform(Gonzalez, 2008), also known as distance map or distance field, is a derived representation of a digital image. The choice of the term depends on the point of view on the object in question; whether the initial image is transformed into another representation, or it is simply endowed with an additional map or field. The map labels each pixel of the image with the distance to the nearest obstacle pixel. The most common type of obstacle pixel is a boundary pixel in a binary image.

The Euclidian distance formula(Gonzalez, 2008) is the standard method to determine distance, which calculates the straight-line distance between two pixels. But slight variations of distance determination have also been implemented. The binary image produced from the DT conveys crucial image information: the size of objects, the ratio between the major and minor object axes can be inferred, and object morphology. In addition, to the crucial image information described above, the DT also provides a means to classify pixels within an object based on distance mapping. More importantly, pixels of minimum distance value designate object boundaries because border pixels are the nearest object pixels to the background. On the other hand, pixels with maximum distance are the most distant from the background specifying a quasi-centre of mass (QCM).

The DT issued to determine the group or region of pixels farthest from the image background for each object. This group or region of pixels is usually localized in the vicinity of the cell nucleus,

and the numerical approximation of the centroid can be applied to these regions to produce the QCM.

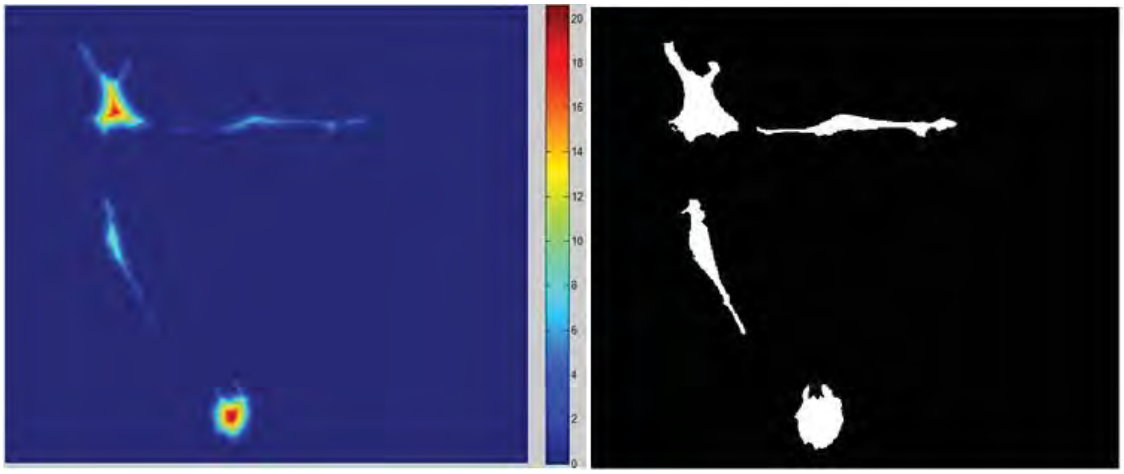


Figure 3. 26 Distance transform of the gray scale image

Referring above to fig 3.26(a), the centre of mass approximation would only be calculated for the region of maximum distance value or in this case the green to red region. On the other in hand in the case of the centroid, the centre of mass approximation would be calculated for the entire cell body in fig 3.26(b). The QCM has two advantages over the traditional centroid:

1) The position of the QCM is almost always inside the cellular boundaries, whereas the centroid sometimes falls outside of cellular boundaries due to the numerical approximation. (This problem is accounted for in our system and will be addressed later)

2) QCM can possibly identify cellular divisions, during mitosis the cellular morphology reshapes into an object which gives rise to two QCMs, as opposed to one centroid.

Although the QCM is a major improvement over the cell centroid, a single reference point sometimes is not indicative of the behaviour of the entire cell body, hence there is some inherent bias in the cellular trajectory data extracted from single cellular reference points. However, centre of mass references are indicative of gross translocations of the entire cell body.

3.6 Manual tracking and correction

Validation of high-throughput time-lapse cell segmentation and tracking algorithms is a challenging task since most cell assays have many cells and rapid acquisition times (leading to

many time points) making manual validation time-consuming and tedious. We developed a simple edit-based validation program that consists of the following elements: i) It accepts a general input independent of the segmentation/tracking algorithm ii) Annotates the image with a cell numerical identify which allocates the cell to a track iii) A region-growing tool with predefined parameters (intensity growing threshold) to manually correct segmentation results iv) A manual tracking tool for editing and adding of tracks using mouse-clicks on cell centroids, tracking one cell over a sequence of images.

3.6.1 Region growing

Region growing(Gonzalez, 2008) is a morphological tool for image segmentation and pixel classification. It is a procedure that groups pixels or sub-regions into larger regions based on the predefined criteria for growth. The growth starts from a set of 'seed' points within the growing regions, and extends to neighboring pixels which have the predefined properties similar to the seed. The first problem is the selection of similar criteria, which depends on object characteristics and the quality of image data. The most commonly used criteria is the range of pixel intensity and color, while some current applications utilized the gradient and entropy of an image(Gonzalez, 2008). Another problem is the formulation of stopping rules, which stop the growth when there are no more pixels to satisfy the criteria. The stopping limitations include thresholds of absolute or differential value of the pixel properties compared with the seed, and size, shape, or likeness of the candidate pixels related to the grown pixels which take in account of the growth 'history'(Gonzalez, 2008).

For this study intensity was main criteria for growth. The stopping rules were based on the average pixel intensity and size of the grown region. The growth procedure is stated as follow.

Let f(x,y) denote an input image array; S(x,y) denote a seed array valued 1 at the locations of seed points and 0 elsewhere. And assume that $S(x_0,y_0)$ equal to 1. Q denote the intensity threshold to be applied at each location (x,y). N denote threshold of the growing pixel number. V denote summary of growing points. Arrays V, f and S are assumed to be of the same size.

1. Let $V(x_0,y_0)=1$, Comparing the intensity threshold Q with the intensity differences between seed and its 8 connectivity shown as the template in figure. If $f(x,y)<Q$ and (x,y) is not the boundary of whole image, let V(x,y) equal to 1, else equal to 0.

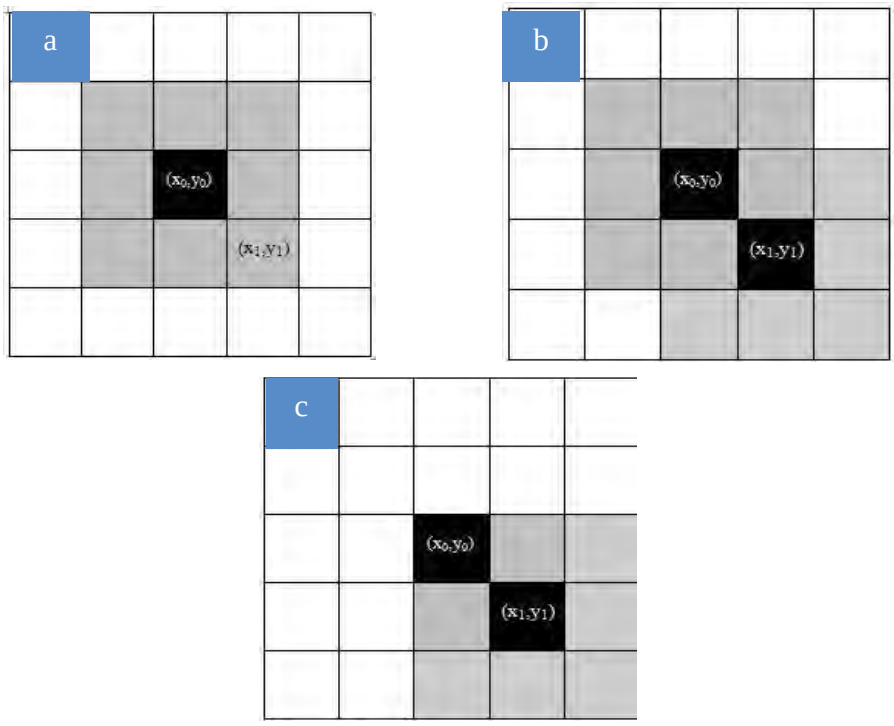


Figure 3. 27 Demo of region growing (8 connectivity)

2. Calculate the number of 1 valued pixels within V. if $Ni<N$, let each of the 1 valued neighboring points count as the new seed, eg (x_1, y_1). Repeat step 1. When all the 8 connectivity valued 0 in V, continue with step 3.
3. Label each connected blobs in V with different region label.

In addition, the region growing function program incorporates interactive seed selection and threshold determination which helps manual inspection of the growing pace. The work flow is stated as below.

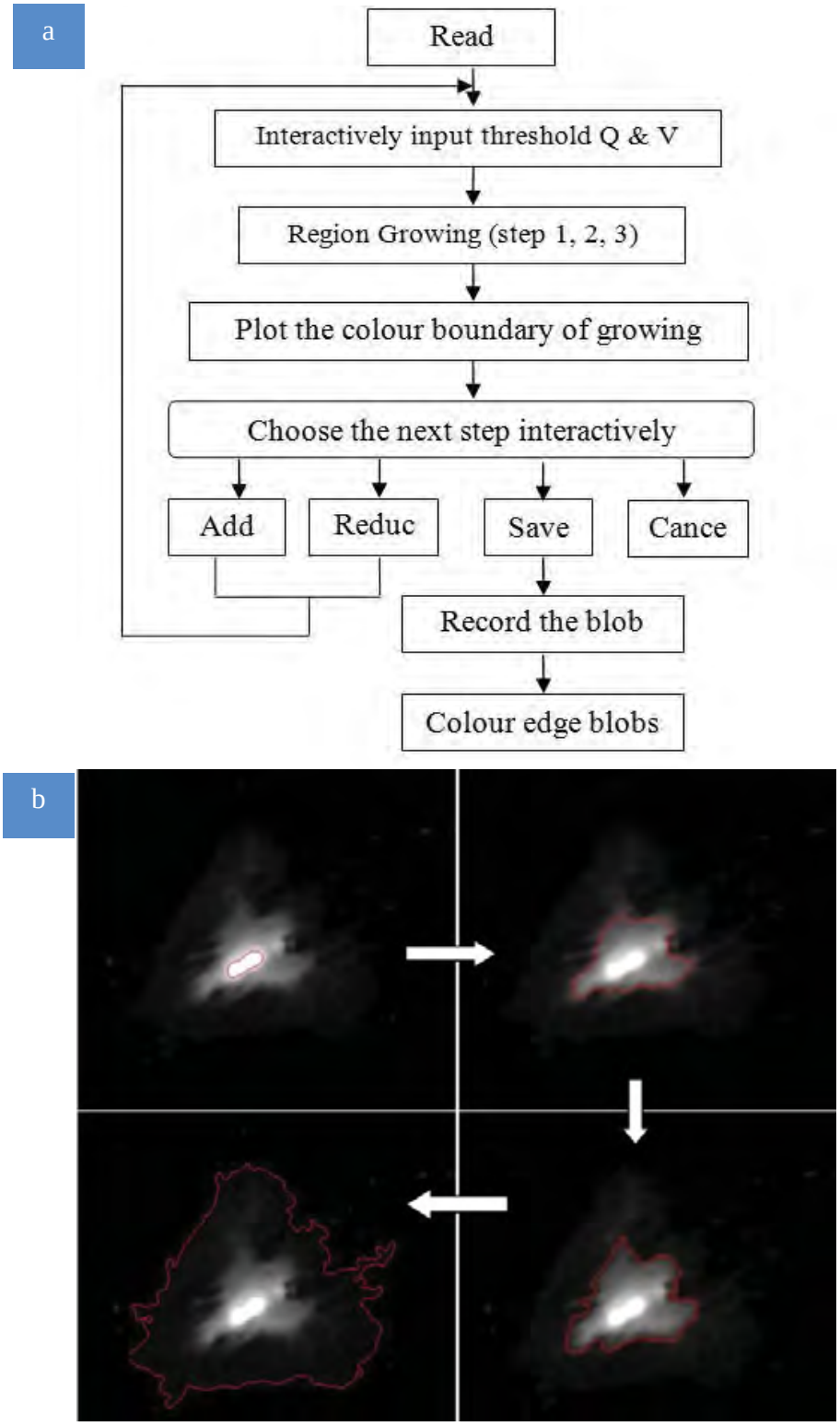


Figure 3. 28 Region growing program; (a) flow chart of program; (b) examples (β-actin cCFU-F) of region growing, cell boundary was contoured by red

3.7 Cell Cycle Analysis

It was shown over 30 years ago by Smith et al that the cell cycle time distribution follows a lag-exponential function(Smith, 1973). Using the same observation here, the cell cycle analysis is performed by estimating the age-dependent probability for mitosis and apoptosis by Kaplan-Meier (KM) analysis.

The KM statistics(Kaplan and Meier, 1958, Nordon, 2011a) is a class of survival statistics, which relate the time that passes before some event occurs to one or more covariates that may be associated with that quantity. It is an non-parametric and empirical probability estimation which includes censored data. The use of censoring overcomes the problem of incomplete observation of life history. For example, one cannot observe the mitosis of all cells because many of them dead or lost before division. As a result, it is suitable for both large sample groups and small sample groups, although the survivor curve estimated from small sample group may have a relative larger error.

The age-dependent probability of mitosis $\Pr[A_{mitosis} > t]$ is defined as the probability that a cell undergoes mitosis after time *t*. For the Smith-Martin model the age-dependent probability of mitosis is

$$\boldsymbol{Pr}[\boldsymbol{A_{mitosis}} > t] = \boldsymbol{e}^{-\boldsymbol{\lambda}(\boldsymbol{t}-\boldsymbol{L})}, \ \ \boldsymbol{t} > L \qquad \textbf{Equation 3. 8}$$

$$\boldsymbol{\log_e}(\boldsymbol{\Pr}[\boldsymbol{A_{mitosis}} > \boldsymbol{t}]) = -\ \boldsymbol{\lambda}(\boldsymbol{t}\text{-}\boldsymbol{L}) \qquad \textbf{Equation 3. 9}$$

where L is a time lag, and λ is the rate of the exponential process.

Surprising, this model often approximates the cell cycle distribution of homogeneous cell populations

As shown in figure 3.29, generation 2 cells which have been tracked for up to 100 hours are analyzed. The blue curve is the age-dependent probability distribution of cells remaining undivided (before mitosis). It is often called an empirical probability, because the probability is estimated using the data, does not make any assumption about the shape of the probability distribution and is thus a non-parametric statistic. The upper and lower dashed curves provide the 95% confidential interval. Logarithm transformation of the probability distribution enables us to

fit the probability curve to a linear function with estimation of the rate of the exponential process (the slope is $-\lambda$) as well as the x intercept which is the lag time L. The red line in figure 3.29 is the linear regression result, with a lag which indicates the smallest cycle time. The mean of an exponential variant is $\frac{1}{\lambda}$. Thus the average cell cycle time is $L+\frac{1}{\lambda}$.

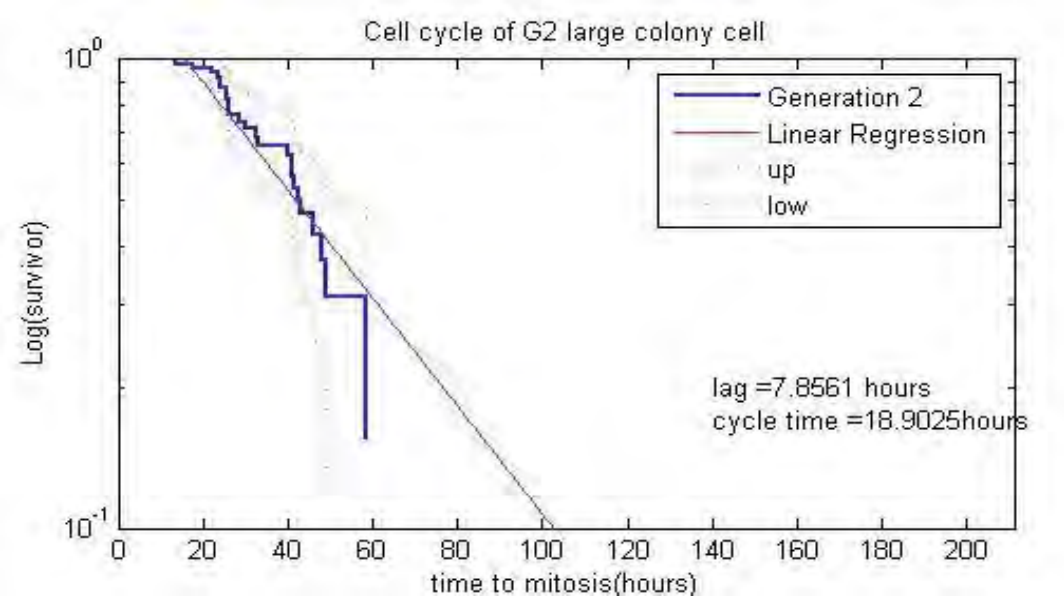


Figure 3. 29 Illustration of cell cycle analysis using cell division history and KM analysis (cardiac CFU-F isolated from a large colony, passage three, from the beta-actin GFP mouse)

3.8 Summary

Semi-automated cell tracking software was described in this chapter. The software has had great utility in this project. Firstly, it was successful applied for tracking green fluorescent protein (GFP) positive nuclei using the PDGFR-α GFP transgenic mouse. The software robustly tracked the majority of GFP+ cells in time-lapse fluorescent image sequences acquired by the live cell imaging system leading to the analysis of cardiac CFU-F development by Kaplan Meier analysis in chapter 5. Secondly, the computation time for segmenting cells was reduced to a few seconds for each image, which is comparable to other segmentation algorithms. Thirdly, the development

of an automated and robust curvature based method for separating adjacent nuclei correctly identified separate nuclei on fluorescent images. Moreover, the manual correction framework is efficient at certain times for correcting unsuccessfully cell segmentation and trajectory maps over image sequences.

Future work will address problems associated with inaccurate cell and trajectory segmentation using phase contrast images, or images with poor quality such as low fluorescence images. . These problems are discussed in more detail in chapter 6. There will also be a need to incorporate the developed algorithms into a user friendly software platform (e.g., C++) which is optimised for speed, particular during interactive editing of segmented image stacks.

Chapter 4 Division Tracking of NIH3T3

4.1 Introduction

The cell line NIH3T3 was initially studied with a custom-built live cell imaging system to evaluate tracking software and statistical methods. NIH3T3 is a mouse embryonic fibroblast cell line, and has an average doubling time of 20 hours (Jainchil.Jl et al., 1969). A photo-toxicity assay was developed to assess the effect of fluorescent imaging on fibroblast survival. Cell death was detected by imaging irreversible cell detachment. The cell cycle time distribution of NIH3T3 cells was estimated by Kaplan Meier analysis. The live cell imaging system was used to test the hypothesis that the cell cycle of mother and daughter cells are inherited by estimating the correlation coefficient for mothers/daughter and sibling cycle times.

4.2 Materials and Methods

4.2.1 Experiment setup

One liter of Iscove's Modified Dulbecco's Medium (IMDM) with 1 mM L-glutamine, 25mM HEPES adjusted to pH 7.2, (Sigma) was prepared every two months. Media was sterilized by filtration through a 0.2um Acrocap filter unit (Pall Life Sciences). Sterile filtered media was stored at 4^{o}C. Media was supplemented with 10% Fetal Bovine Serum (Invitrogen), 3.7g/L of sodium bicarbonate, and penicillin / streptomycin antibiotic solution (CSL Biosciences). NIH 3T3fibroblasts were passaged 2-3 times per week with culture media and maintained below a surface density of 100,000 cells/25cm^2. Cells were maintained in a 5% CO_2 incubator at 37^{o}C.

Cells were trypsinized (JRH Biosciences), and plated at a desired density of approximately 10,000 cells/25cm^2 8 hours before imaging. A plating density of ~10,000 cells/25cm^2 was suitable for extraction of cellular trajectories because the number of cellular collisions were minimizes, thus reducing the cellular overlapping dilemma previously described. A temperature-regulation system enclosed the entire microscope system at 37^{o}C, and a smaller gas-tight environmental chamber maintained 5% CO_2 environment around petri dishes (figure 4.1).

Uniform heating was regulated by a custom-built PID controller (M168, AVR microcontroller) for closed-loop regulation of microscope incubator temperature at 37.0 ± 0.1 C. Temperature was measured using platinum 100 temperature probes placed inside microscope incubator compartment on top of the environmental chamber. The heater with fan was housed within the microscope incubator enclosure, its duty cycle set by the PID controller to maintain temperature at the 37.0 C set point. An infrared CO_2 detector was connected to the output of the culture pod to check that the 5% CO_2 concentration was maintained within the environmental chamber. The air/ CO_2 mix was humidified by recirculation through an underwater sealed vent (figure 4.2).

During transport of tissue culture plates from the tissue culture incubator to the imaging system condensation forms under the lids of petri dishes of a drop in temperature. Condensation refracts illuminated light and diminishes resolution within a focal plane. Therefore cells were placed inside the incubator for 2 hours prior to imaging to allow temperature equilibration and evaporation of condensation. Images were taken using an inverted Olympus IX70 microscope equipped with a 10x phase contrast objective and automated x, y, z-stage. A 12-bit peltier-cooled camera (1376×1032 pixels) with firewire interface (F-View II, SIS XM10 Trigger supplied by Olympus) captured images. Cell/M software (Olympus) was used to set a protocol that scanned 4 positions every 2 minutes for 4 days.

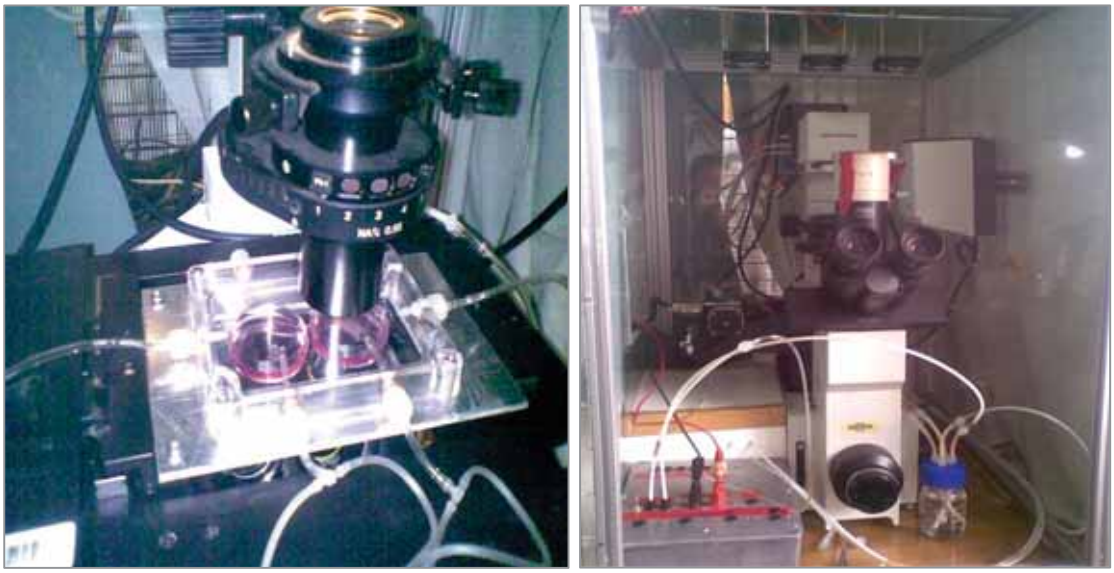


Figure 4. 1 Cell culture chamber and front view of the Cell imaging stage (a) Cell culture chamber (b) Front view of the imaging stage

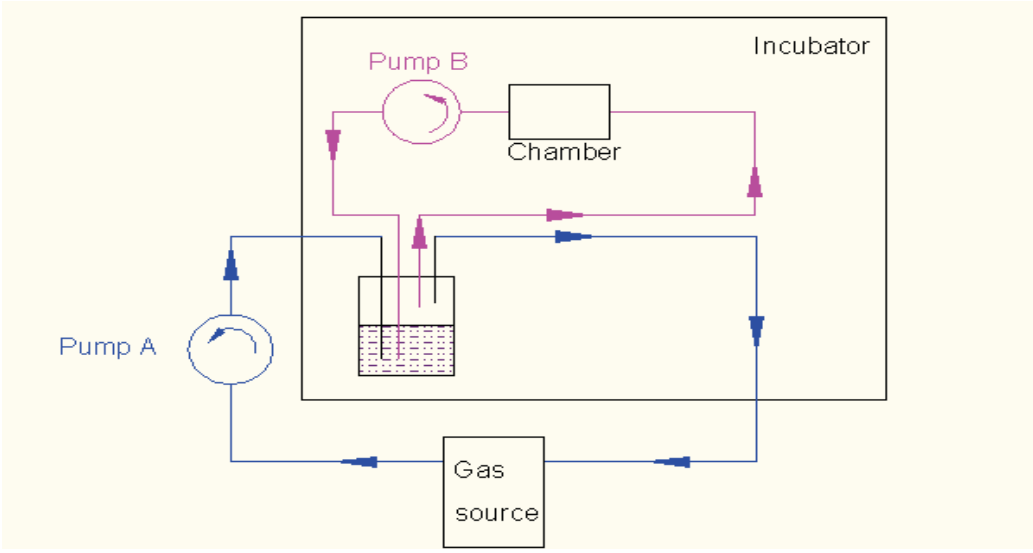


Figure 4. 2 Design of the CO_2 concentration control and hydration circuit

One of the two dishes of cells served as a control sample to test the photo-toxicity of blue light excitation (figure 4.1). Phototoxicity was measured by continuous exposure to blue light for time intervals of up to 8 minutes. A phototoxicity matrix array was created on the cultured monolayer

by programming blue light exposure times for different *x, y* positions. Cell detachment at each of these positions was imaged at just before exposure, and at 16, 24, 48 and 72 hours after exposure.

Maintaining focus was a major issue in all live cell imaging experiments, though this problem was totally resolved many months later when it was discovered that Olympus firmware was driving the z-focus coupling at a velocity that resulted in slippage. For tracking NIH3T3 cells with phase contrast imaging auto-focus software gave inconsistent results because bright or dark field edges were used by the software algorithm. By over exposure, which is in essence a thresholding technique, the autofocus algorithm in Cell/M only utilized dark field features over consecutive images (see Figure 4.3). This technique proved to be an efficient method for keeping fine focus for long term image series.

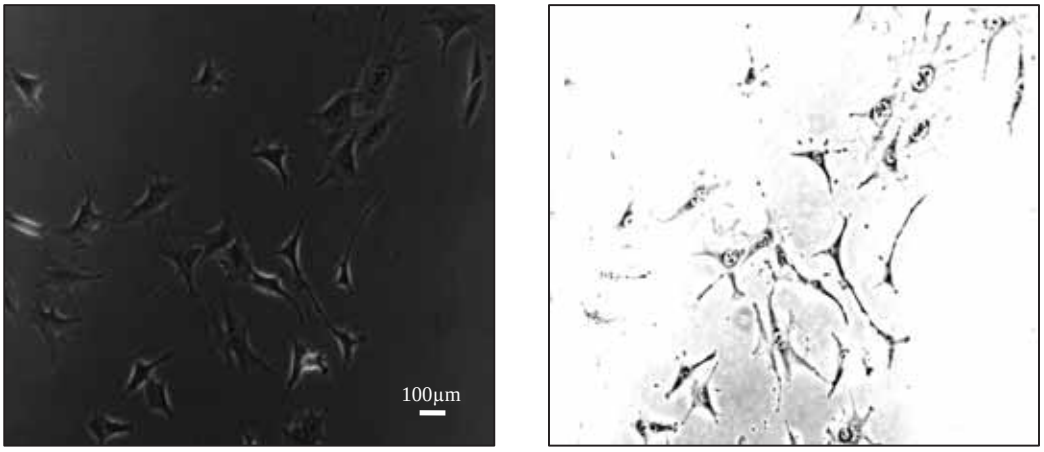


Figure 4. 3 X10 Demo of the developed auto-focus method, a: PC image of NIH3T3 cells; b: over-exposure image, ET=100ms

4.3 Results and Discussion

4.3.1 Phototoxicity test

Images were taken for 9 positions exposed to different doses of blue light. The cell detachment rate was counted for each position:

$$Death\ rate = \frac{Num(detach\ cell)}{Num(total\ cell)} \qquad \textbf{Equation 4. 1}$$

Figure 4.4 is an image montage showing micrographs of the 9 positions taken at 72 hours. Significant phototoxicity was detected after 2 minutes of blue light exposure; the number of detached cells (black inside, bright halo, round shape) accounts to more than 50% of the total cell number. Whilst this morphology is also seen during mitosis, cells will rapidly reattach to the surface after cell division.

Figure 4.5 shows the proportion of detached cells versus time for the range of blue light doses. The detachment rate was more variable after 48 hours because of cell migration into the death zone. However, it is clear that even short blue light exposures were detrimental to NIH3T3 viability. Thus there is a trade-off between the extra information obtained by fluorescent imaging and the loss of cell viability.

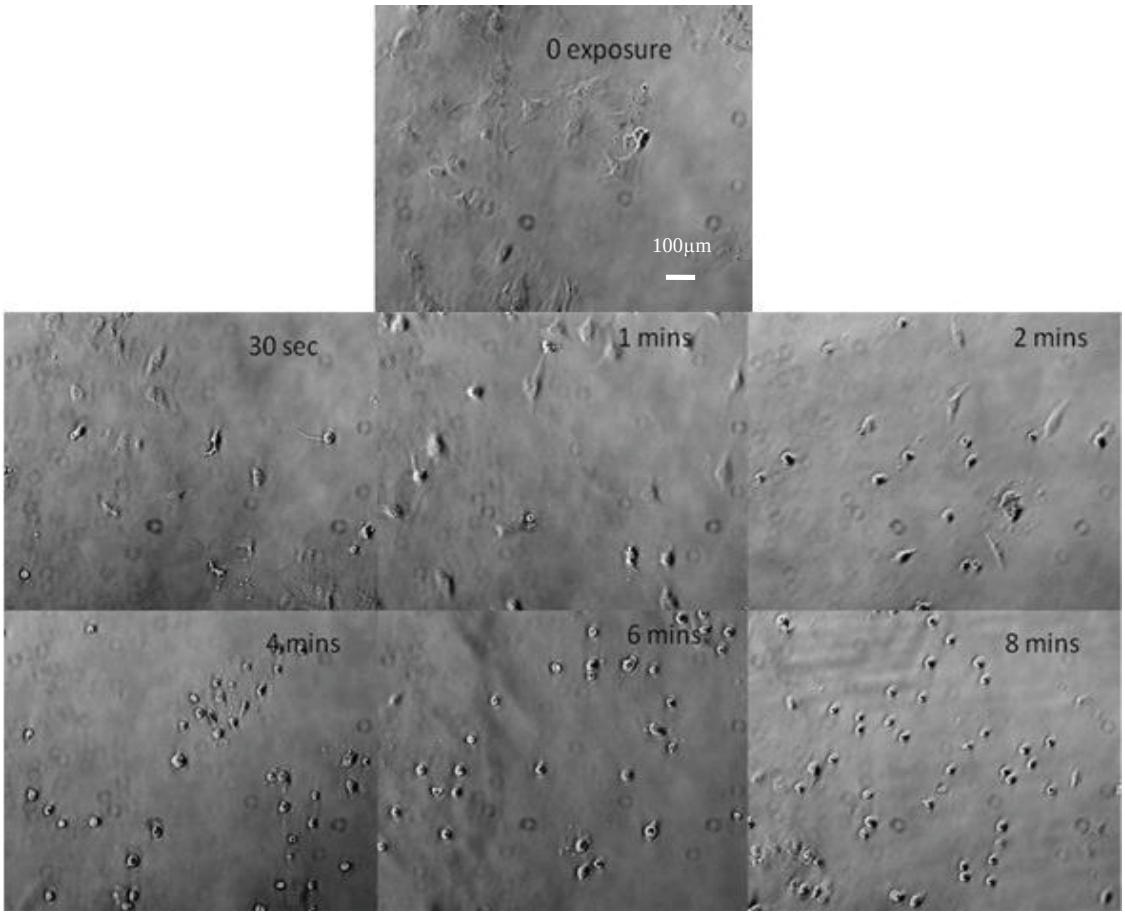


Figure 4. 4 Snapshots taken for exposure groups at the end of culture

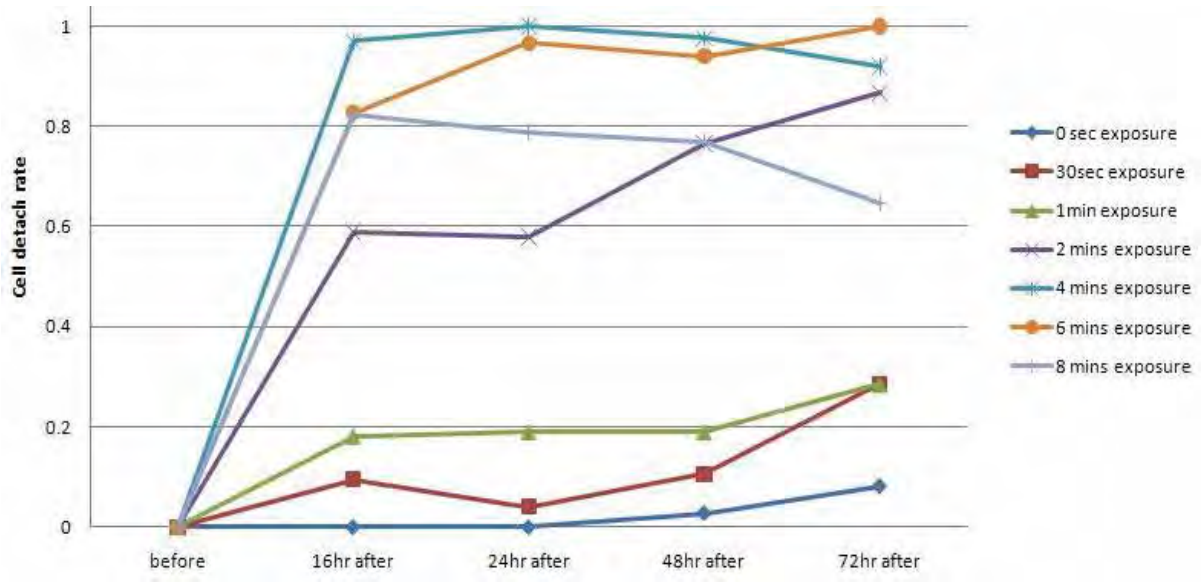


Figure 4. 5 Cell detach rate of different light exposure vs time

4.3.2 Trajectory maps

All tracking software and statistical calculations were performed using the Matlab2010a image analysis and statistics toolboxes. Cell tracking over image sequences were performed manually by Julia Yin (B.E GSBmE UNSW). She identified cell tracks using the software described in chapter 2.

The pedigrees of 44 colonies were constructed from tracking 482 cells within four fields of view over four days of culture. Figure 4.6 shows the complete trajectories of NIH3T3 at position one, where up to 4 generations are observed during the tracking experiment.

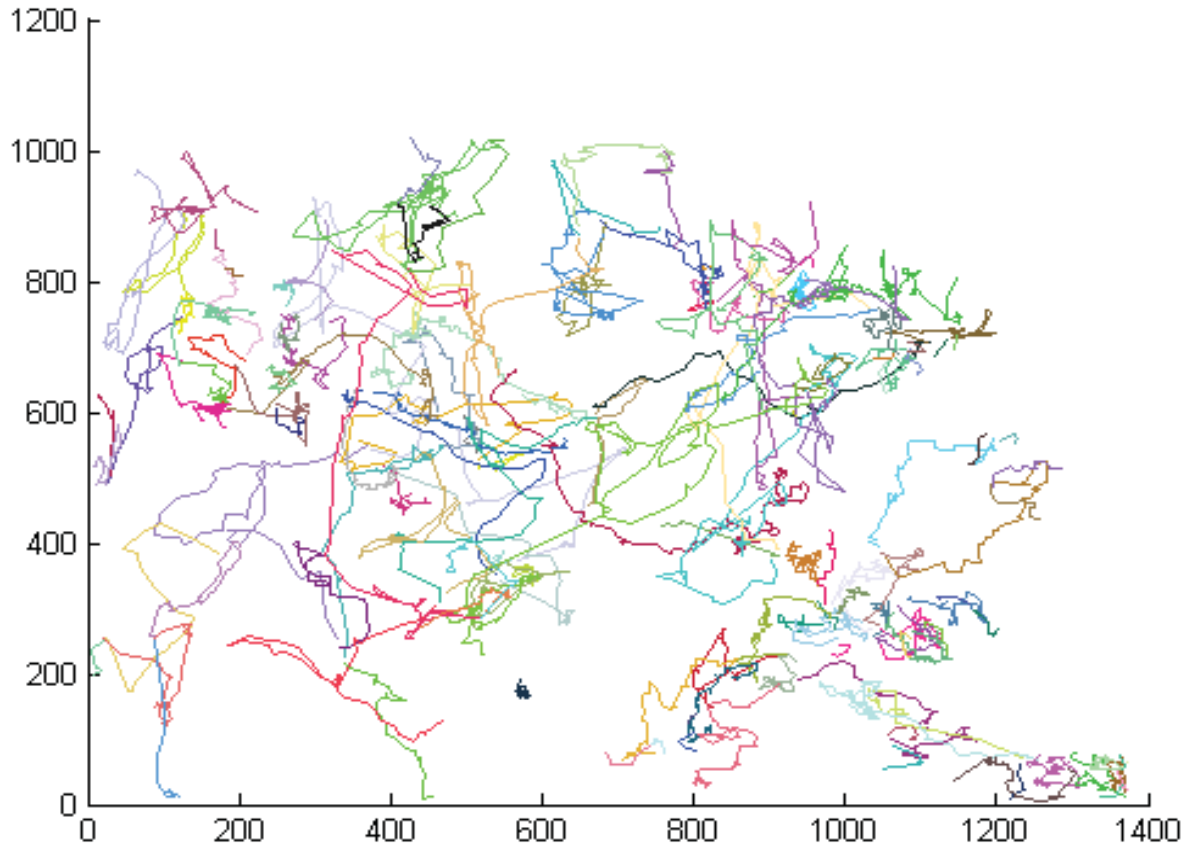


Figure 4. 6 Trajectories of 145 cell tracks in one field of view (P1) in four days culture

4.3.3 Kaplan Meier Analysis

The distribution for cell cycle lengths were estimated by the Kaplan Meier (KM) statistic for cells belonging to each generation. The KM statistic estimates the empirical probability of an event during a time interval from a time series data. The statistic is commonly used in medical research to estimate the empirical probability of survival (or response to therapy) using incomplete lifetime data.(Kaplan and Meier, 1958) KM analysis to cell lifetime data has only recently been recently applied to cell cycle analysis (Nordon, 2011a) and is a novel aspect of this thesis.

It was assumed that cell cycle time had an exponential distribution because KM plots of log of probability could be approximated by a linear function.

$$Pr[T > t] = e^{-\lambda(t-L)};$$

$$\ln\left[Pr[T > t]\right] = -\lambda(t - L); \qquad t > L \qquad \text{Equation 4. 2}$$

Here the probability that a cell divides after time *t* is *Pr*[*T*>*t*] where *T* is a random variable describing the cell cycle length. The exponential function is displaced to the right by a constant lag *L* which represents the minimum cell cycle time required for S and G1 phase of the cell cycle.

The mean of an exponentially distributed random variable with parameter l is $\frac{1}{\lambda}$. To estimate l it was assumed that cycle times were exponentially distributed (*L*=0*)*. The KM estimated with 95% confidence interval at day 4 was used to estimate l with 95% confidence intervals (CI), and the average cycle time with 95% CI was estimated by taking the reciprocal of l ± 95% CI.

Table 4.1 lists the average cell cycle time (± 95% CI) and lag time estimated for each generation, and also records the number of cells that were tracked. Figure 4.7 to Figure 4.8 shows the log transformed probability of mitosis for generation 0-3 as well as linear regression analysis of log transform data. The main findings are summarized below:

1. Generation zero cells have no lag time. This is because their last mitosis occurred before the start of the experiment, and only the residual cell cycle time was observed.
2. The lag time for generations 1-3 is between 10-15 hours.
3. Average cell cycle time increase with consecutive generations.

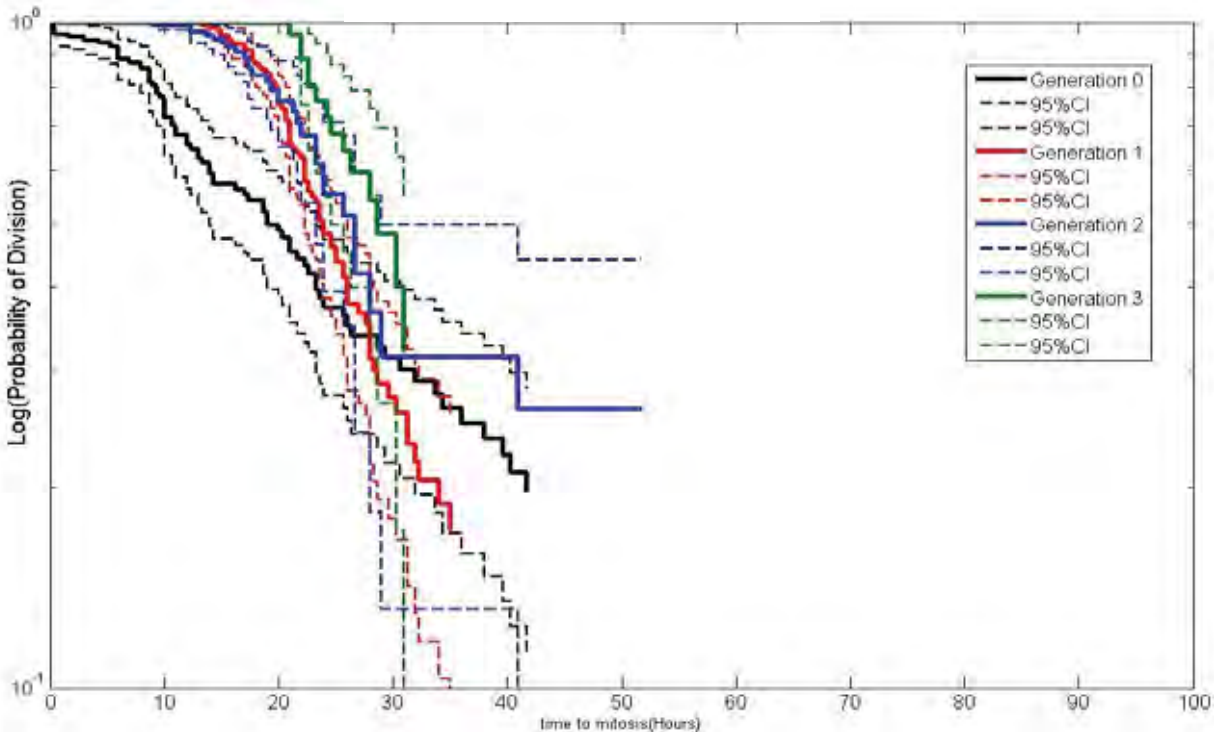


Figure 4. 7 Kaplan Meier Analysis for NIH3T3 of four generations

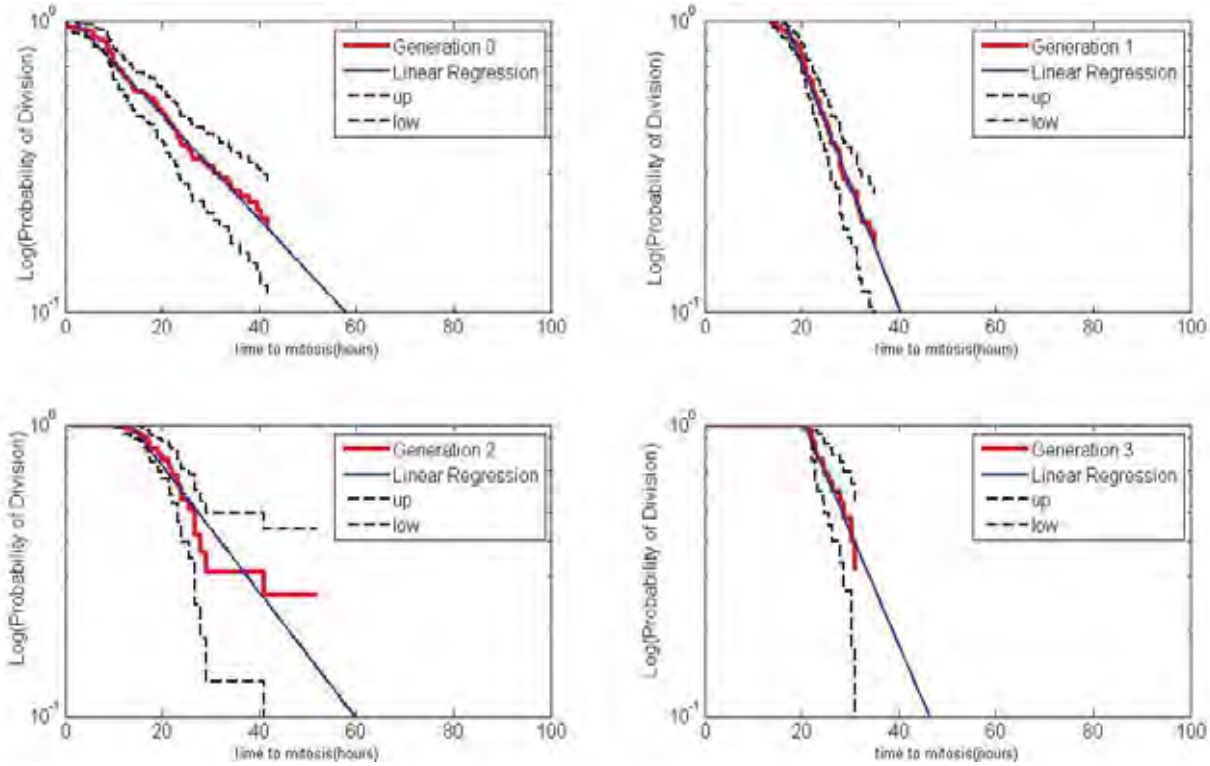


Figure 4. 8 KMA and cycle estimation for G0-G3 NIH3T3

Table 4. 1 cell cycle estimation results

Generation	% divide at day 4	Average cell cycle time (hours)	Lag time (hours)	Tracked cell number
0	80.29(7.36)	32.23(4.45)	0	101
1	80.70(9.01)	23.70(5.64)	12	140
2	73.84(17.88)	26.29(12.83)	10	154
3	68.98(14.78)	27.21(3.15)	15	58
4	0	----------	-------	29

The significance of differences between 5 generations was determined by a Cox proportional hazards model described in detail in section 5.3.7. Table 4.2 lists the estimated relative hazard, error of relative hazards, the z statistic and the probability of the null hypothesis. Here the null hypothesis was no effect of generation number on the rate of mitosis. It was rejected because $p<0.001$. Therefore the effect of generation number of cell cycle is significant.

Table 4. 2 Cox proportional regression for five generations

Predictors X (5 generations)	**Relative hazard β**	**Error of β**	**z-statistic**	**p-value**
Generations 1: G1, 2:G2, 3:G3, 4:G4, 5:G5	-0.2981	0.0803	-3.7099	2.0736e-04

4.3.4 Inheritance of cell cycle times by sisters and daughter cells

As mentioned before, it was observed that siblings divided within a short time of each other. The inheritance of cell cycle times by sisters or daughter cells was estimated by performing regression analysis on these related cells. Figure 4.9 shows that sibling generation time is

correlated and has a regression line with a slope of 1. This means that the cell cycle duration of sister cells is almost identical, and that they divide almost synchronously. The correlation coefficient was R^2=0.70, p= 1.94e-12. Red dots shown in figure 4.9 are sister cells that die by apoptosis or migrate outside of the field of view. These sister pairs were not included in the regression analysis. Moreover, there are a handful of outliers with large residuals (deviations from linear regression line) indicating that there are also some asymmetric divisions with quite different generation times.

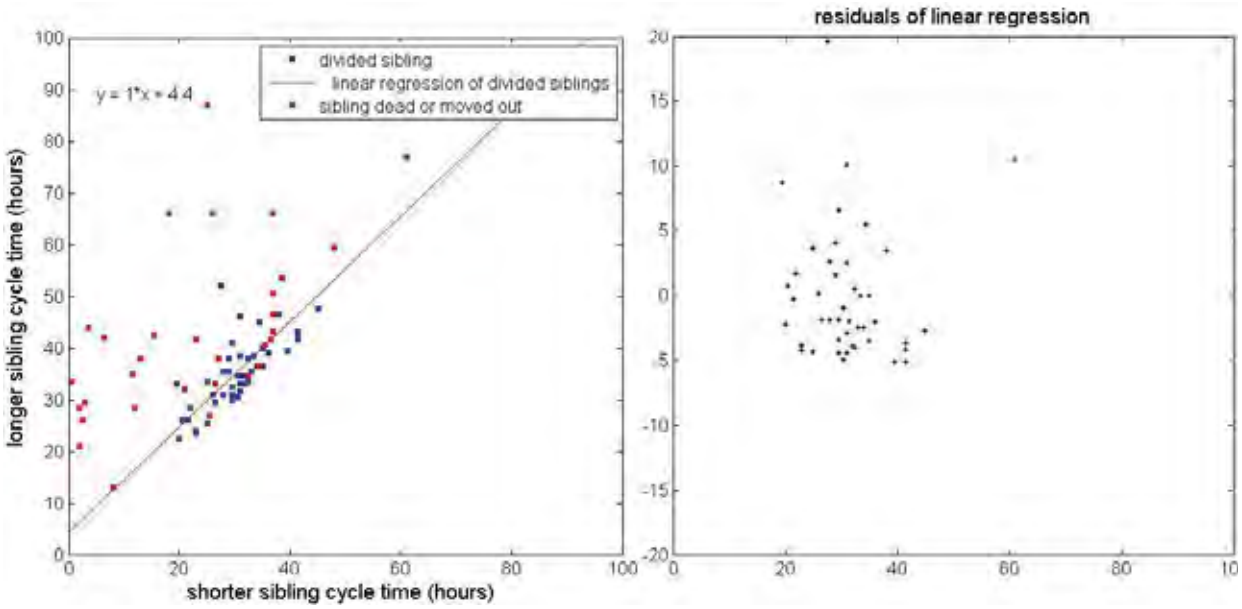


Figure 4. 9 Correlated sibling cycle (censored the missed and dead cell), with linear regression and residual plot; linear regression R^2=0.70; k=1; p= 1.94e-12

The correlation between mother and daughter cycle times was examined (figure 4.10(a)). Over 121 mother daughter pairs were analyzed. In contrast to sisters, cell cycle times were not close to equal, but daughter cells had longer cycle times compared to mothers. The slope of the regression line was 29.32 and the correlation coefficient was 0.023 (p=0.0942). Figure 4.10(b) shows that the average cell cycle time of mother cells is shorter than that of daughters.

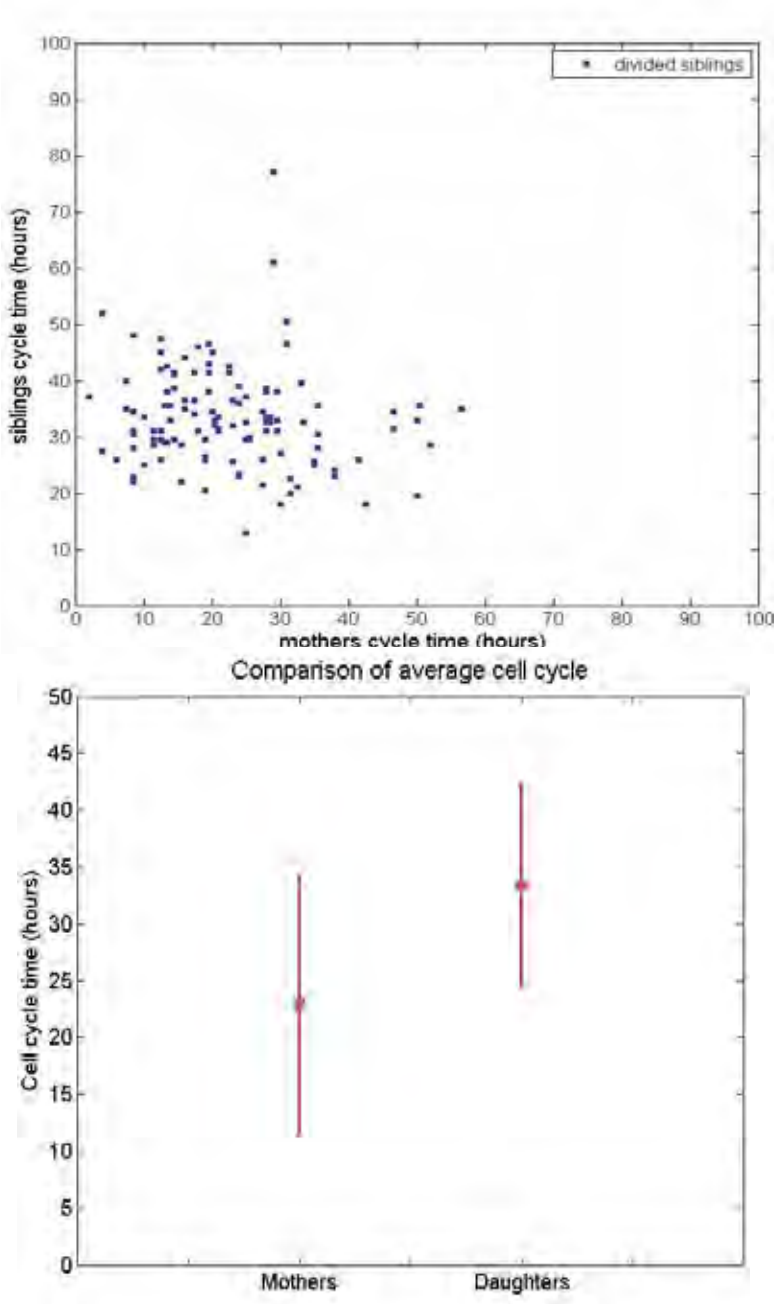


Figure 4. 10 Correlated mother-daughter cycle (censored the missed and dead daughters); a: dot plot of pairs of mother-daughter cycle; b: average cell cycle

4.4 Conclusion

The aim of this chapter was to determine the feasibility of tracking adherent cells (NIH3T3) using our live cell imaging system and to evaluate statistical methods for characterising cell cycle kinetics and the inheritance of cell generation times. A novel test for photo-toxicity was developed and showed that blue light excitation sources are quite toxic to NIH3T3 cells. Exposure to blue light for 2 minutes resulted in 50% of cells detaching. Exposure as short as 30 seconds had measurable effects on cell detachment and death.

482 NIH3T3 cells in 4 days of culture were tracked with the semi-automatic cellular tracking software developed in chapter 3. The probability of mitosis was estimated by Kaplan Meier analysis. Cell lifespan (time between successive mitoses) was approximated by an exponential distribution. The average cycle time estimated for NIH3T3 was 23.70 ±5.64 hours (± 95% CI), which was similar to the literature record of around 20 hours (Jainchil.Jl et al., 1969). Thus the growth of NIH3T3 in our culture and imaging system was found to be similar to conventional culture methods.

In addition synchronization of sibling mitoses was observed. On the other hand there does not appear to be inheritance of generation times between mother and daughters, though a program for lengthening cycle times was observed. The average cycle time of daughters was longer than that of mothers.

Overall, the cell cycle and lineage relationship results obtained in our tracking studies are close to those reported in the literature(Smith, 1973). The lag-exponential model of cell cycle was first reported by Smith-Martin, though they did not measure sibling and maternal cycle times. Nordon et al. (Nordon, 2011a) have quantified generation time inheritance in haematopoietic cell populations, though their studies showed that mother and daughter generations for granulocyte-macrophage progenitors were correlated.

Chapter 5 Tracking of Cardiac CFU-F

5.1 Introduction

The recent discovery of endogenous cardiac stem cells in the post-natal heart (Chong et al., 2009) has led support for the ambitious goal of therapeutic cardiac regeneration. Although the biological origin of these cells is currently poorly understood, the epicardium has been identified as an important source of cells for cardiac repair.(Chong et al., 2009) We employed the colony forming unit fibroblast (CFU-F) assay (Friedenstein, 1989b) to characterize a subpopulation called cardiac CFU-F(cCFU-F) isolated by fluorescent activated cell sorting from the adult murine heart, and quantified their potential for colony formation, long term self-renewal and multi-potency (properties of true stem cells), and also assessed the role of the platelet-derived growth factor receptor alpha (PDGR-a) gene in the development of cCFU-F.

Given the heterogeneous properties of cultures initiated with multipotent cells, it is hypothesized that cCFU-F is composed of cellular subsets with different potency and growth kinetics. To date cCFU-F colonies have been broadly classified according to their size. Small or 'micro' colonies generally have restricted differentiation potential and replating efficiency. On the other hand large colonies self-renew and are multipotent. In addition to having the characteristics of mesenchymal stem cells (form bone, cartilage and fat under appropriate culture conditions), cCFU also form fibromyocytes, smooth muscle, cardiomyocytes and endothelium (Harvey et al. personal communication). Therefore the first aim of this study was to identify cellular subsets from cCFU-F with specific cellular features - size, morphology, motility and PDGFR-a

expression - that correlate with growth kinetics. Platelet derived growth factor is a known mitogen, and may be responsible for regulating cardiac CFU-F cell cycle progression.

Given that large colonies when replated generate large colonies, and small or microcolonies have poor replating ability, we would like to understand the mechanism for inheritance of colony size. A mechanism for inheritance of cell cycle kinetics is by vertical transmission of cell cycle times i.e, from mother to daughter cells. Previous studies have shown that haematopoietic cell generation times are correlated in sisters or mothers and daughters (Nordon, 2011a). Therefore the second aim was to perform a detailed pedigree analysis to determine whether mother-daughter and sister cell cycle times were correlated.

The limited replating efficiency and growth potential of small colonies may relate to longer cycle times or entry into a quiescent state, or increased apoptosis. It is hypothesised that cell cycle time is an epigenetic trait that determines whether a small or large colony is formed. Therefore the final aim of this study was to determine whether colony size correlates with the cycling status of cCFU-F.

To summarise, live cell imaging with lineage tracing was used to address the following specific aims:

1. Identify cCFU-F subsets with distinctive cell cycle behavior.
2. Determine whether cell generation times are inherited from mother to daughter cells.
3. Determine whether cell cycle length is an epigenetic trait that specifies colony size.

5.2 Materials and Methods

Phase contrast images of cardiac fibroblasts have indistinct cell boundaries making cell segmentation difficult. Therefore transgenic mice that report gene expression using green fluorescent protein (GFP) were utilized to i) improve segmentation accuracy using a simple fluorescence threshold and ii) investigate the role of PDGFR- α signaling in regulation of cCFU-F cell cycle. We first used PDGFR- α -GFP cells isolated from hearts of PDGFR- α-GFP transgenic mice, which have one of the PDGFR-α genes replaced by enhanced GFP (EGFP). The expression pattern of the H2B-EGFP fusion protein mimics that of endogenous PDGFRα gene (Hamilton et al., 2003). To address aim 1, i.e., to identify cCFU-F subsets that correlate

with cell cycle behavior, the cell cycle characteristics of cells expressing endogenous PDGFRα (JAX® Mice, B6.129S4-*PDGFRA*$^{tm11(EGFP)Sor}$/J (Hamilton et al., 2003)) as detected by GFP expression was determined. Cells were imaged using a combination of phase contrast and fluorescent microscopy. In a second study both PDGFRα+ and PDGFRAα- cells were tracked using cCFU-F isolated from a transgenic mouse that ubiquitously expresses green fluorescent protein (β-actin-GFP) (JAX® Mice, CByJ.B6-Tg(CAG-EGFP)1Osb/J(Amano et al., 2004)). All cells were tracked in these studies making it possible to detect cell cycle differences between PDGFRα + and PDGFRα - cells. In addition to PDGFRα expression cell size, morphology and motility were quantified, to see whether these cellular characteristics correlate with cell cycle kinetics.

Cardiac CFU-F was continuously tracked over a 5 day period. Kaplan Meier analysis was used to determine the cell cycle time distribution of gated cells. An immediate finding of this analysis was that large immotile cells which often had two nuclei had a longer cell cycle time compared to highly motile spindle-shaped cells.

To address aim 2, the trajectories of mother and daughter cells were analyzed to determine if cycle times were correlated. To address the hypothesis that cell cycle length is the epigenetic trait that determines colony size (aim 3), individual cells were taken from large and small colonies and replated to determine if the cell cycle distribution differed for these two cellular sources. A Cox proportional hazards regression model (Matlab) was used to estimate the significance of the effect of colony size (large versus small) and cell subsets (high motility small cells versus low motility large cells) on cell cycle length.

Cardiac CFU-F cells were isolated from the heart of PDGFR- α -GFP mice (JAX® Mice, B6.129S4-*PDGFRA*$^{tm11(EGFP)Sor}$/J) or β-actin-GFP mice (JAX® Mice, CByJ.B6-Tg(CAG-EGFP)1Osb/J), with collagenase II (Worthington Biochemical Co.), and were labeled with anti-Sca-1 and anti PECAM-1 (BD Biosciences). The interstitial cell fraction of hearts were sorted by a FACS BDAria (BD Bioscences) collecting the Scal-1+ PECAM-1- population. 50,000 cells were innoculated on 75mm^2 culture flasks. Cells were cultured in MEM α (Gibco) with 2200mg/L Sodium Bicarbonate, 20% FCS, 1% Penicillin, 1% L-glutamine in a 37C incubator with 5% CO_2. Culture medium was changed every 3 to 4 days, due to the fast medium consumption of these cells. Passaging was performed every 10-12 days ensuring that cells did

not grow beyond 60% confluence. Passage one harvested cells were frozen in -40^{o}C and stored in liquid nitrogen.

For PDGFR-α GFP mice, frozen passage 3 cells were thawed and cultured as described above. Note that GFP is expressed in the nucleus, and is not localized at the cellular membrane (figure 5.1). 5000 cells were inoculated onto 35 mm^2 culture dishes and grown in a standard tissue culture incubator for 24 hours prior to tracking inside the microscope incubator chamber (figure 5.2). Media was replaced just prior to transfer into the microscope incubator.

For β-actin GFP mice 2 large colonies and 2 small colonies from passage 1 cultures (e.g., Figure 5.3) were picked up by O-ring and expanded separately under CFU-F culture condition. At day 13 of passage 2, cells were tripsinized and 50 cells were inoculated onto two glass well-plates coated with Polyerystin, which have better transparency than plastic culture dishes. These 35 mm^2 dishes were grown in a standard tissue incubator for 24 hours prior to tracking inside the microscope incubator.

The graphical user interface of the experiment manager of Cell/M software (Olympus) was used to program a large-field scanning protocol. Images were taken using an inverted Olympus IX70 microscope equipped with a 10x (for PDGFR-α-GFP cell experiments, figure 5.1) or 4x (β-actin GFP experiments, figure 5.4) phase contrast objective, a fluorescent illumination system and automated x, y, z-stage (Optiscan, Prior Scientific). A 12-bit peltier-cooled camera (1376×1032 pixels) with firewire interface (F-View II, SIS XM10 Trigger supplied by Olympus) captured images. PDGFR-α cells were imaged every 30 minutes for 5 days from a 5×6 contiguous region (figure 5.2). Likewise β-actin GFP cells were imaged every 30 minutes for 5 days, however there were only 50 cells at the start of culture, so 33 regions were selected so that one or two cells were initially at the centre of each scanning region. GFP was excited at 488 nM; PDGFR-α cells were exposed for 300 ms (figure 5.1), whilst β-actin GFP cells were exposed for 100 ms (figure 5.4).

At day 4, medium was replaced and image reference positions were re-aligned.

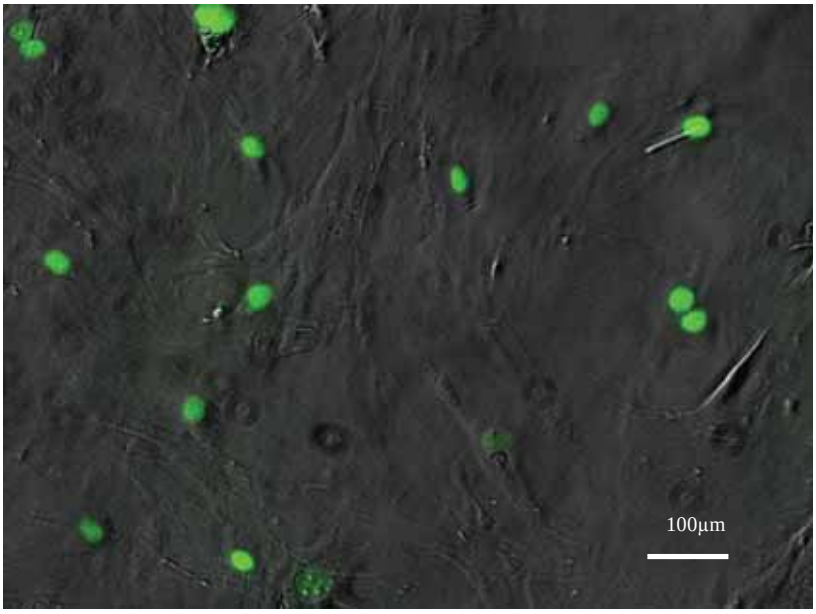


Figure 5. 1 X10 Superimposed GFP &phase contrast images of passage 3 PDGFR−α - GFP CFU-Fs Frame 127 @ position 17

The second petri dish culture served as a control for the effect of phototoxicity (figure 5.2) and was imaged by fluorescence once a day as described in chapter 4.

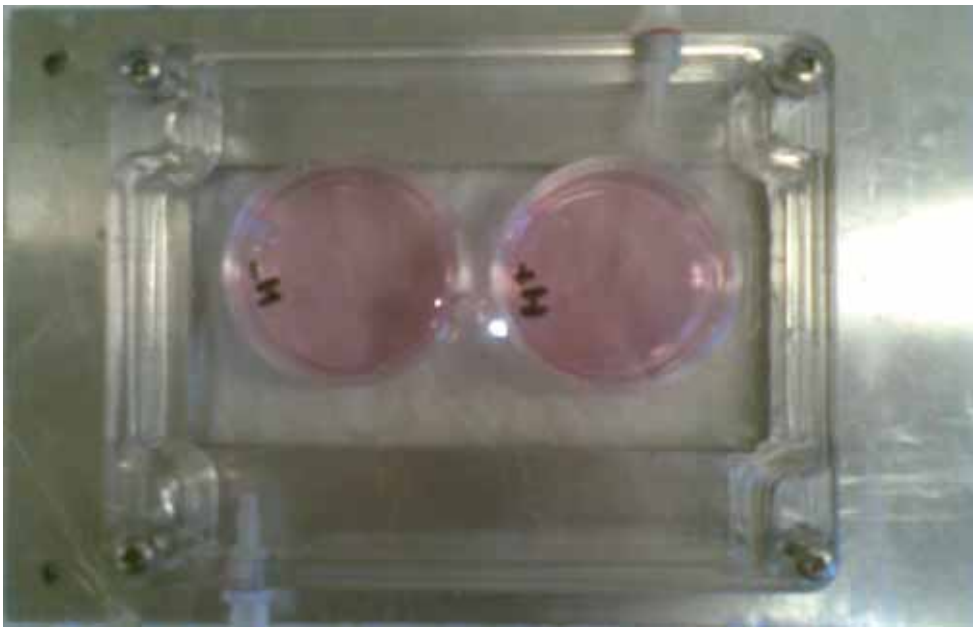


Figure 5. 2 snapshots of Environmental chamber and two petri-dish culture NIH3T3 cells at the end of experiment, H+: tracking sample; H-: control sample

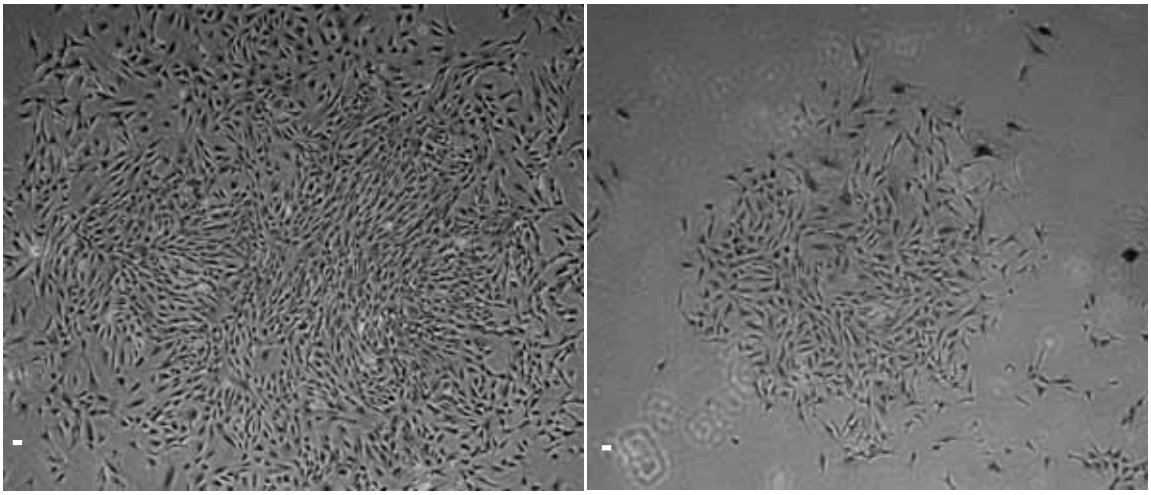


Figure 5. 3 Large colony and small colony X2 Frame

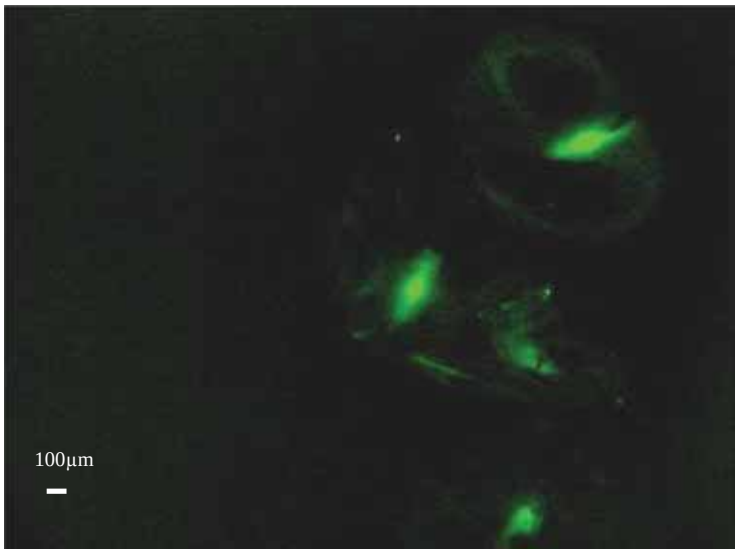


Figure 5. 4 X4 Frame 211(1675X2225 μm) @ position 15 of β-actin-GFP CFU-Fs

5.3 Results and Discussion

5.3.1 Trajectory maps for PDGFR-α GFP cCFU-F

Tracking of passage 3 PDGFR-α-GFP cCFU-F from image sequences was performed by the automatic fluorescent image tracking program, followed by manual confirmation and correction. The automatic and manual confirmation software was described in Chapter 2.

137 lineage maps were constructed from cell tracking data accumulated over five days of culture. In Figure 5.5, montage images of a large field of view (5x6) are aligned, and phase contrast cellular images and green fluorescent nuclei images have been superimposed. Different cell tracks are marked with different colours. The cell trajectories are also plotted as 3D maps, in which the x and y axis are accordingly the x, y axis of 2D image and z axis is the tracking time points. In figure 5.5(b), a division occurs at approximately the 17^{th} frame and can be clearly identified, as the splitting of one track into two tracks of different colours.

By visual inspection and manual correction of the automatic tracking results, we identified two distinctive cell morphologies in culture; large flattened-cells with low motility and spindle-shaped cells. Illustrated in figure 5.6, a large flattened-cell is in the upper left corner of the image and the nucleus is circled in red. Spindle-shaped cells are widely distributed in the image and the nuclei are circled in green. The expression of PDGFR-α-GFP is in the nucleus, and does not localise at the cell membrane as a fusion cell surface receptor. Prominent trends are that large flattened-cells have a higher expression of PDGFR-α compared to small, spindle-shaped cells. The motility of spindle shaped cells is also significantly higher than flattened cells.

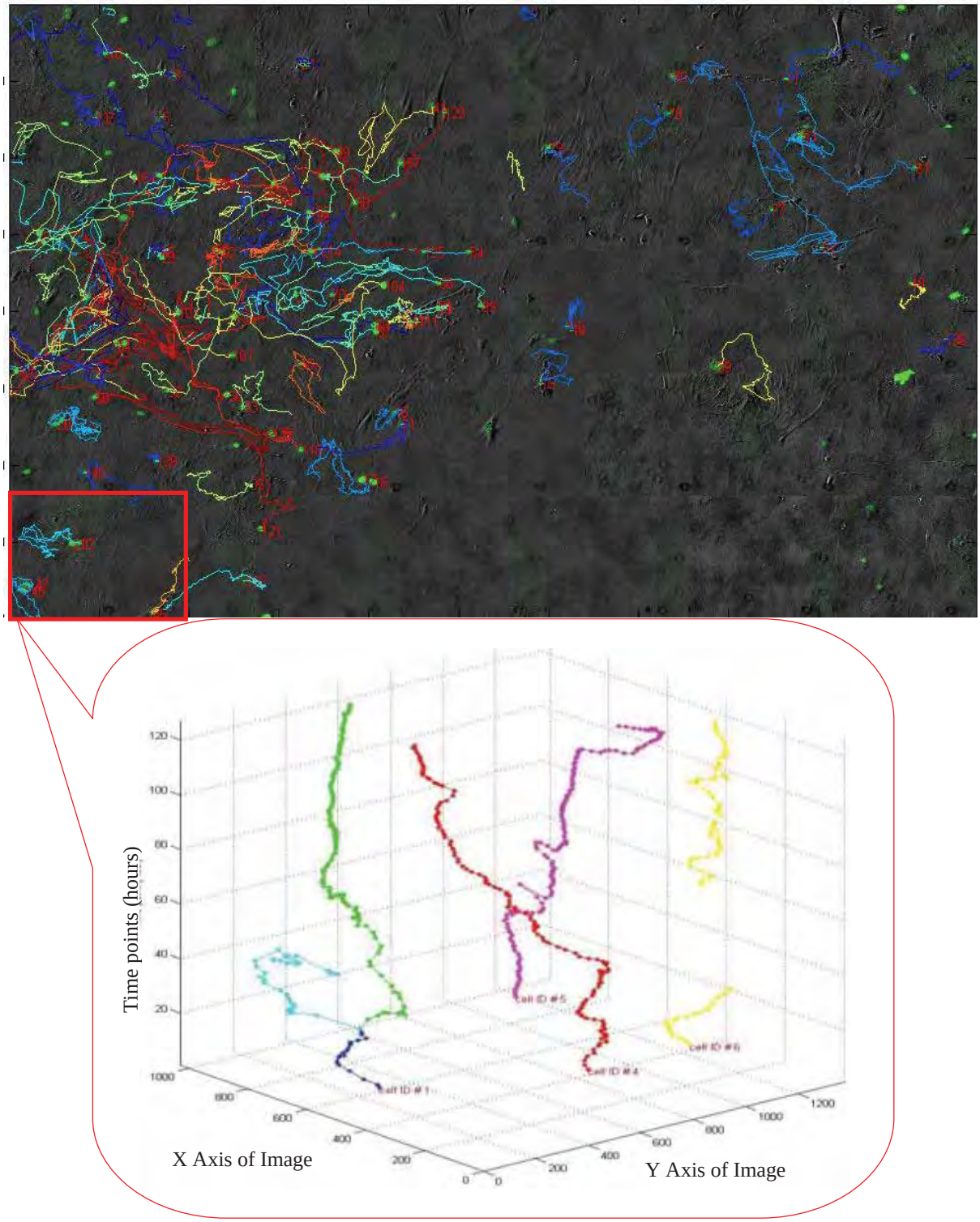


Figure 5. 5 Trajectories of PDGFR- PDGFR- α-GFP+ cells over five day culture (a) Trajectories of 137 PDGFR- α-GFP+ cells in large field of view; (b) 3D-trajectories of cells in contoured region

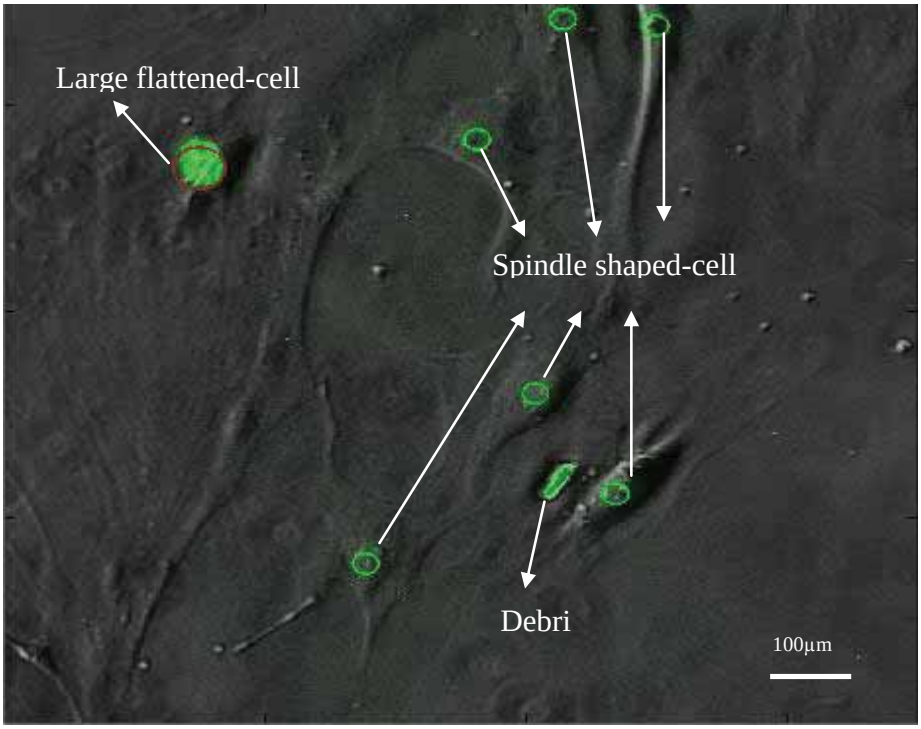


Figure 5. 6 (X10frame (670X890µm)) Identification of large flattened-cell and spindle shaped cells

5.3.2 Motility analysis (PDGFR-α GFP mouse)

As illustrated in the dot plot in figure 5.7, *x* and *y* axes are cell displacement measured in pixels per hour for *x* and *y* directions, respectively. Blue dots which indicate spindle cells, are mostly distributed within the blue polygonal region, while red dots representing large flat cells are centred within the red region. Spindle cells clearly have greater motility compared to flat cells.. This conclusion is further illustrated in figure 5.8, which compares the average velocity (pixels/hour) with standard error of the mean (SEM) for cells with different shapes (flat red, spindle blue), and division histories. During the tracking period, cells which have been observed to undergo at least one division appear to have greater motility compared to cells which undergo no divisions (quiescent cells). The effect of quiescence on motility was not significant as determined by a two way ANOVA (table 5.1). The effect of observing no divisions during the culture experiment did not significantly influence motility, (p=0.2267) but cell size was highly

significant (p=0.0005). However, flat cells with no PDGFR-a GFP expression were excluded from this analysis, because these cells boundaries are not distinct on phase contrast images. Experiments with beta-actin GFP mice resolved this problem (see below).

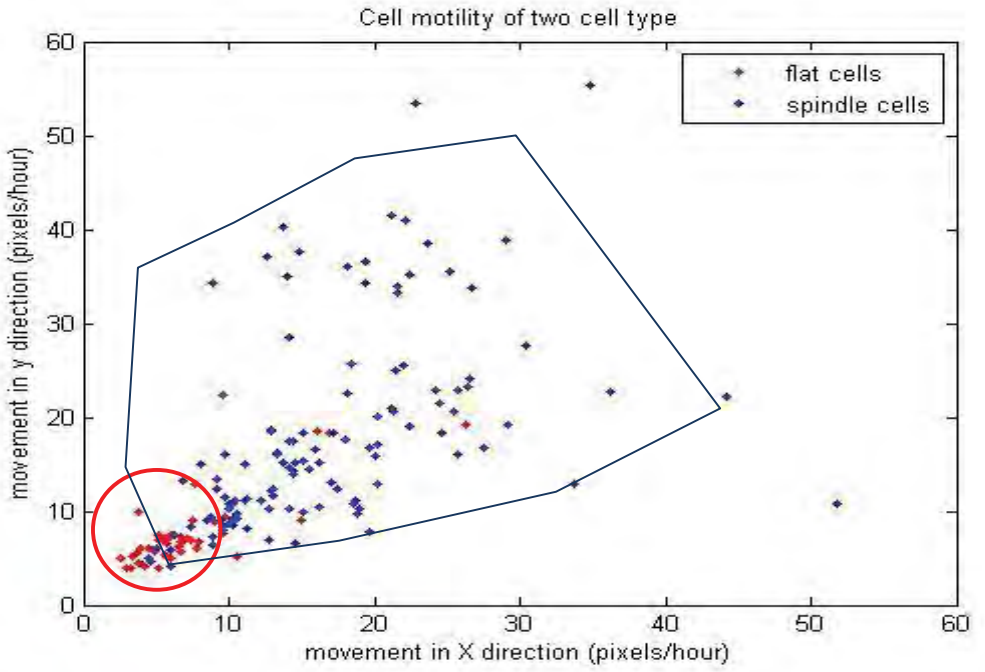


Figure 5. 7 Comparison of motility of flat and spindle cells

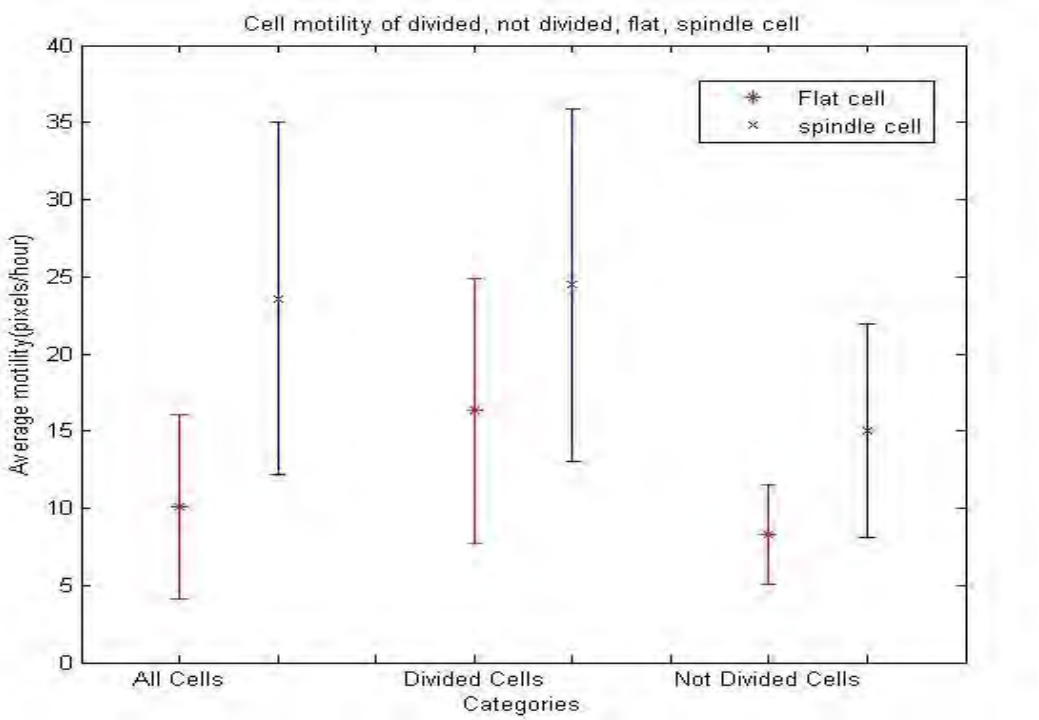


Figure 5. 8 Comparison of cell motilities between cells with different shapes, and division histories

Table 5. 1 Two-way ANOVA to determine the significance of size and division history affecting motility

ANOVA Table

Source	SS	df	MS	F	Prob>F
Columns	949.7	1	949.74	1.52	0.2261
Rows	9258.5	1	9258.54	14.79	0.0005
Interaction	574.1	1	574.13	0.92	0.3447
Error	22540	36	626.11		
Total	33322.5	39			

5.3.3 Kaplan Meier analysis (PDGFR-α GFP mouse)

Based on the observation of cell morphologies with differences in cell motility, the cell cycle time distribution of these subsets was analyzed by Kaplan Meier analysis (KMA) as described in section 3.6. Illustrated in figure 5.9, figure 5.10 and figure 5.11, the probability of division (y-axis) after a specified cell age (x-axis), was estimated using the Kaplan Meier (KM) statistic. It was assumed that cell cycle time had an exponential distribution because KM plots of log of probability could be approximated by a linear function.

$$Pr[T > t] = e^{-\lambda(t-L)};$$

$$\ln\big[Pr[T > t]\big] = -\lambda(t - L); \qquad t > L \qquad \text{Equation 5. 1}$$

Here the probability that a cell divides after time t is $Pr[T>t]$ where T is a random variable describing the cell cycle length. The exponential function is displaced to the right by a constant lag L which represents the minimum cell cycle time required for S and G1 phase of the cell cycle.

The mean of an exponentially distributed random variable with parameter l is $\frac{1}{\lambda}$. To estimate l it was assumed that cycle times were exponentially distributed (L=0). The KM estimated with 95% confidence interval at day 5 was used to estimate l with 95% confidence intervals (CI), and the average cycle time with 95% CI was estimated by taking the reciprocal of l ± 95% CI.

It was not possible to know the birth time of cells that were plated at the start of the experiment. So-called generation zero cells are therefore analysed separately. For generation zero one assumes that the experiment commenced at uniformly sampled random times after cell division. The rate l of an exponentially distributed generation time should be equal to the rate of the residual time before cell division because exponential variables have independent increments(Bremaud, 1999) . The average cell cycle times for flattened-large and spindle-shaped cells were listed in table 5.2 below. From this data it appears that the cycle time of flattened-large cells is much longer than spindle-shaped cells. Cox regression analysis was later used to determine the significance of the observed differences (see below). Due to the low turnover rate of flat cells very few divisions were observed, so the cell cycle time was not estimated.

Table 5. 2 Estimated Cell Cycle of Passage 3 CFU-F

Cell Birth	Cell Type	% divide at day 5	Cell cycle (hours)	Lag time (hours)	Tracked cell number
Cell birth prior to experiment	**Flat**	83.2 (11)	214.44(103.24)	1.49	48
	Spindle	28.6(20)	21.60(8.13)	-1.17	22
Cell birth during experiment	**Flat**	----------	----------	5.75	10
	Spindle	53.55 (17)	135.89(29.91)	3.19	84

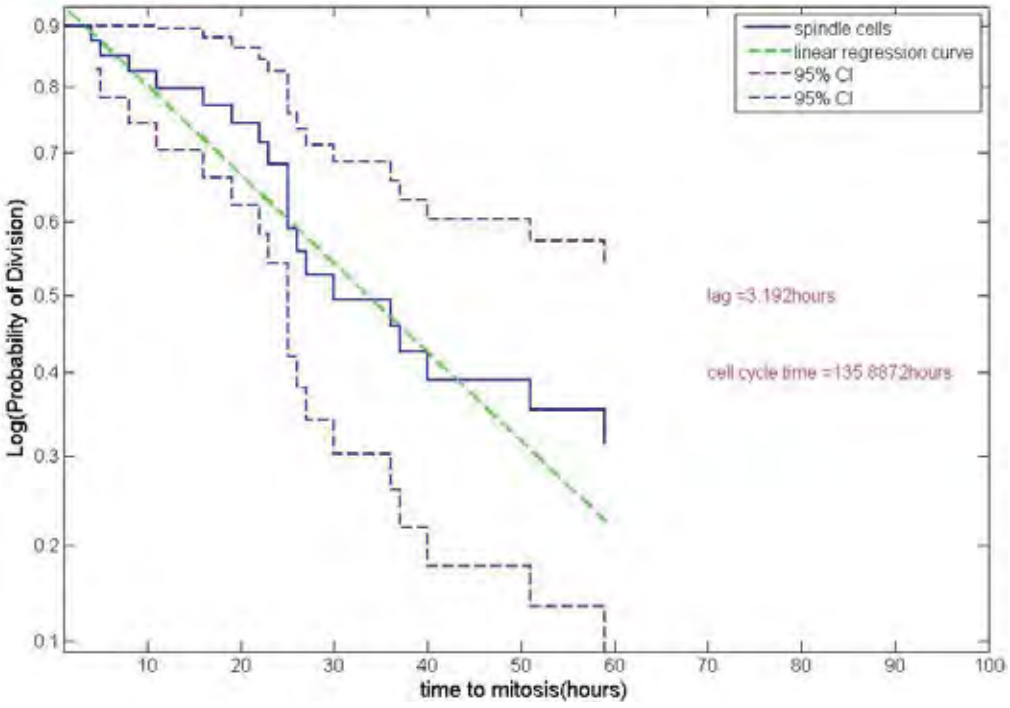


Figure 5. 9 Kaplan Meier analysis of spindle cell cycle time distribution (birth during experiment)

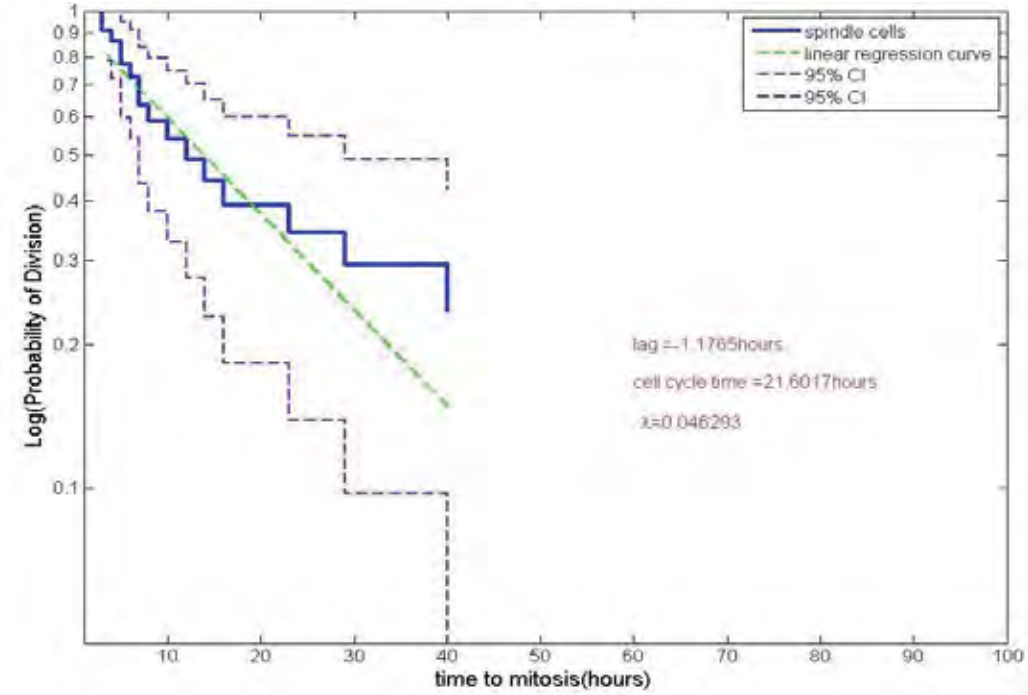


Figure 5. 10 Kaplan Meier analysis of spindle cell cycle time distribution (birth before experiment)

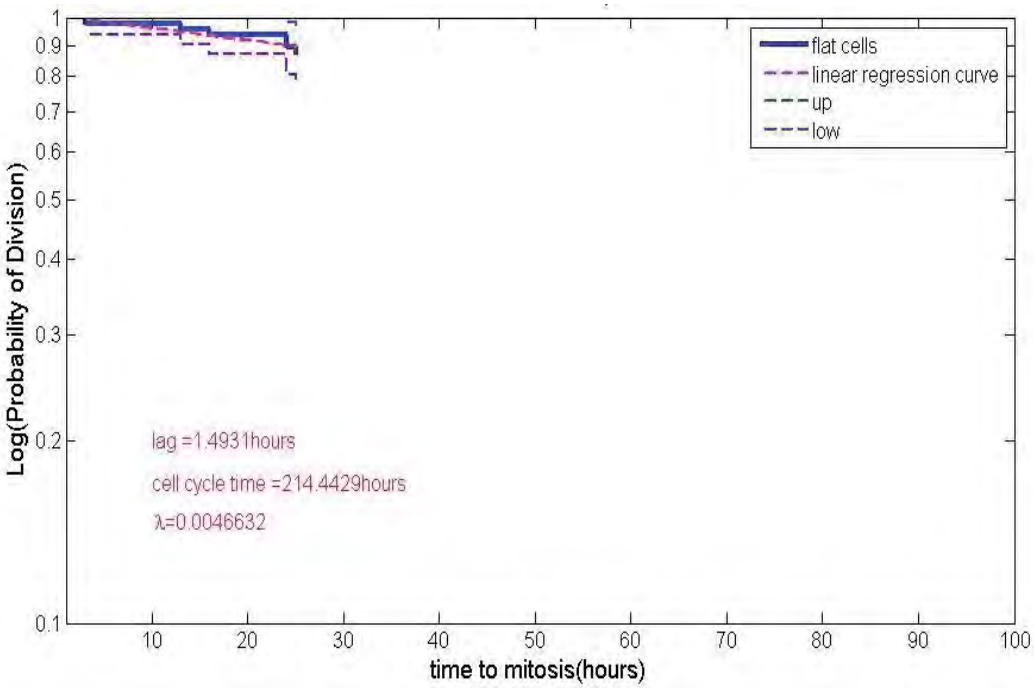


Figure 5. 11 Kaplan Meier analysis of flat cell cycle time distribution (birth before experiment)

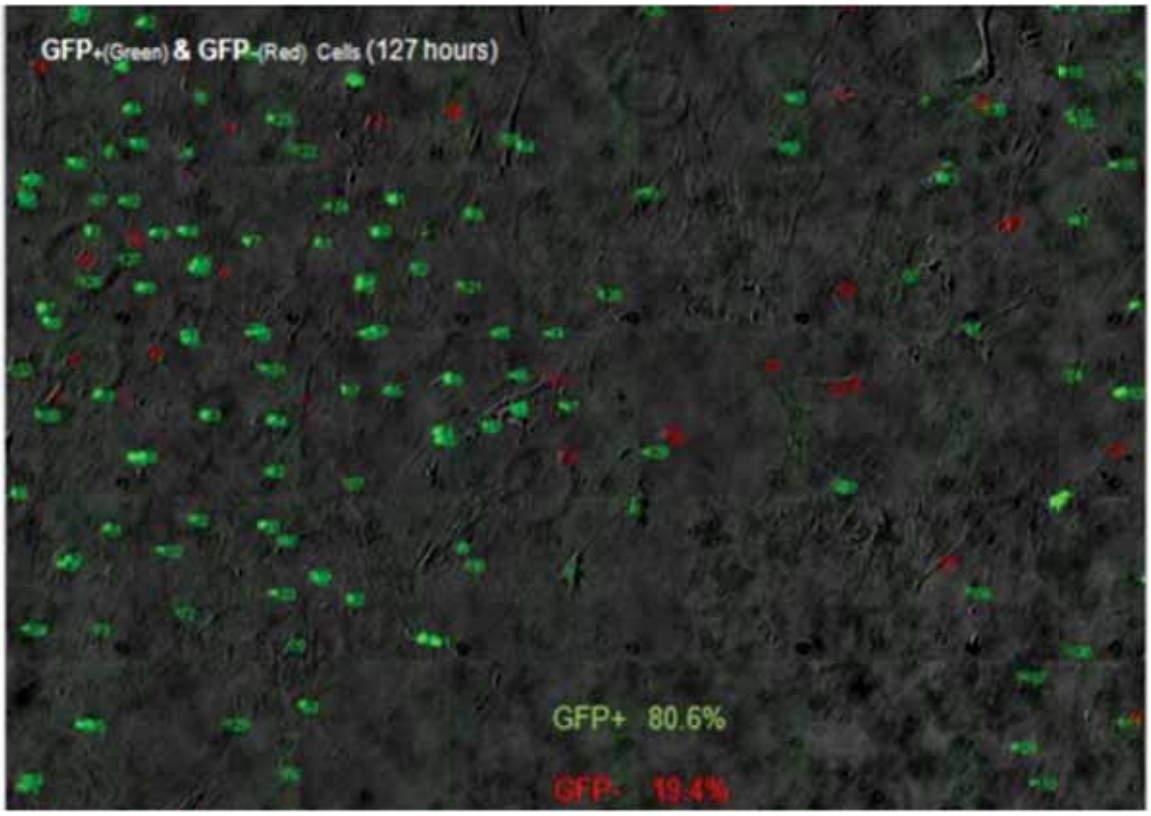


Figure 5. 12 Proportion of GFP expression @127 frame of the large field of view

It was observed that the majority of the flattened cells have a stronger expression of PDGFR- α in comparison to small-spindled cells. In figure 5.12, the proportion of GFP positive and negative cells were counted at the end of the experiment (frame 127). GFP positive cells are marked and numbered in green, and GFP negative cells are noted in red; 80.6% of all cells remain GFP positive and 19.4% of them have completely lost GFP expression. Thus PDGFR- α was expressed in the majority of cells and KM estimates of cell generation times for the flat and spindle shaped cells may be close to total population values. At the conclusion of these experiments a number of technical shortcomings were noted:

1. The 5x6 contiguous scanning field at x10 generated a large number of images. A smaller magnification (x5) would have been sufficient given the large size of cells 100 microns. In some extreme cases, a large cell occupied 1 to 2 fields-of-view at x10 magnification. As a result, the number of cells that were observed in one scanning experiment was limited, and the extension of scanning area to larger areas poses problems with data storage and processing.

2. For these experiments the culture dish was scanned every hour. It was possible to track cells manually using visual inspection, though the sampling rate was not high enough for automated tracking of small cells with high motility and division rates. The scanning interval should be reduced to less than 30 minutes with x10.
3. Large flattened cells were observed to form binucleated cells. Thus a more general definition of mitosis is required. Some cell types undergo the process of endomitosis (megakaryocytes, smooth muscle, skeletal muscle) without the division of the cell body. Thus the live cell classification should be based on division of the nucleus, rather than separation of cell bodies.
4. Phase contrast images of cell boundaries were poorly defined because of the thin and spreading property of cCFU-Fs. Techniques for labeling actins with fluorescent proteins or staining with chemicals, which can increase the image contrast of cell boundary are desired.
5. The initial seeding density of cCFU-Fs may affect cell growth rate, cycle time and differentiation outcomes. Studies reported in the literature (Masur et al., 1996) show that cultured cCFU-Fs at extremely low density differentiate into smooth muscle cells. As a result, a systematic study of the influence of seeding density on differentiation is required.

Therefore in subsequent experiments cCFU-F was tracked at 30 minute intervals at x4 using cCFU-F isolated from the beta-actin GFP mouse.

5.3.4 Trajectory maps for beta-actin GFP cCFU-F

352 cells were tracked over five days of culture at 30 minute sampling intervals (211 frames) using the developed automated cell trajectory mapping software. Pedigrees for 48 colonies were constructed from cell trajectories.. Figure 5.13(a) and (b) show a complete trajectory map for a colony of spindle-shaped cells where up to 4 divisions are observed during the tracking experiment. From the 3D trajectories in figure 5.13(b), we can clearly identify the timing of divisions on the z axis. Notably siblings (same generation) have a similar lifespan and almost divide synchronously. The synchronicity between siblings will be examined in more detail later. Moreover, these tracking experiments confirm the observations made with the PDGFR-a mouse

with identification of two distinctive cell morphologies; large flattened-cells and spindle-shaped cells.

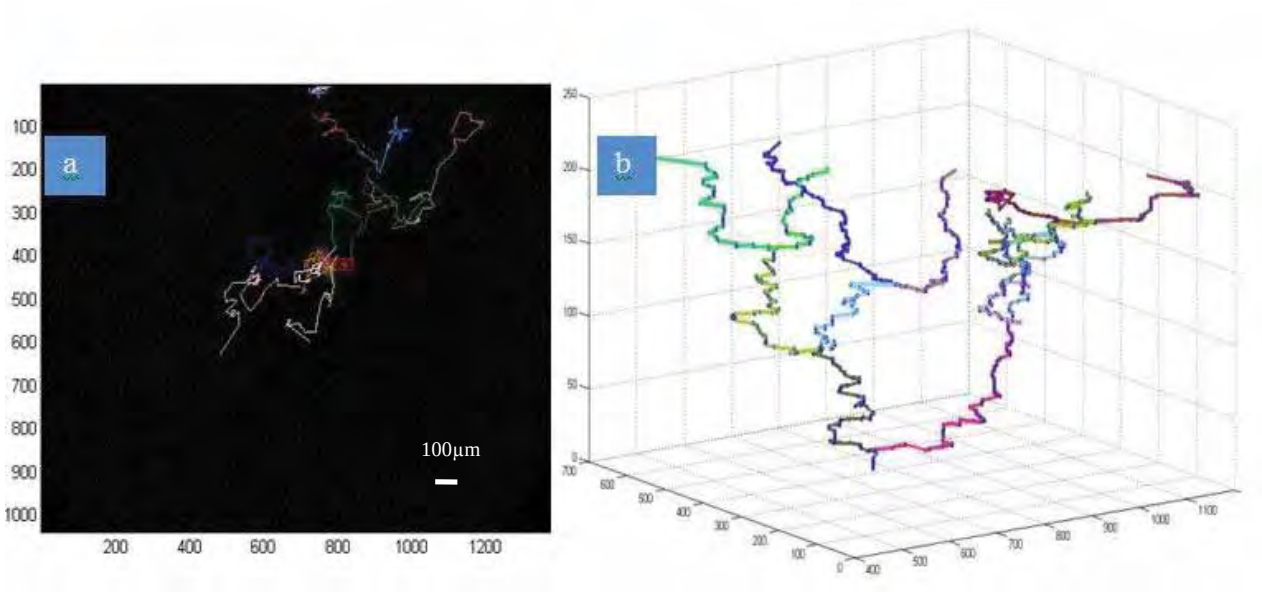


Figure 5. 13 X4 Trajectories of colony @position 15, a: 2D trajectories; b: 3D trajectories, x y corresponds to the x, y axis of image frame, z axis is the frame number from 1 to 211, which corresponding to the time point from 1 to 105 hours.

5.3.5 Motility analysis (beta-actin GFP mouse)

Figure 5.14 (a) is a dot plot showing for cell motility in x- and y- directions expressed in units of pixels/hour. Figure 5.14(b) compared the average speed of flat and spindle cells under different conditions. It is clear that the motility of large flattened cells is significantly lower than that of spindle-shaped cells, confirming the observations made with the PDGFR- α-GFP mouse.

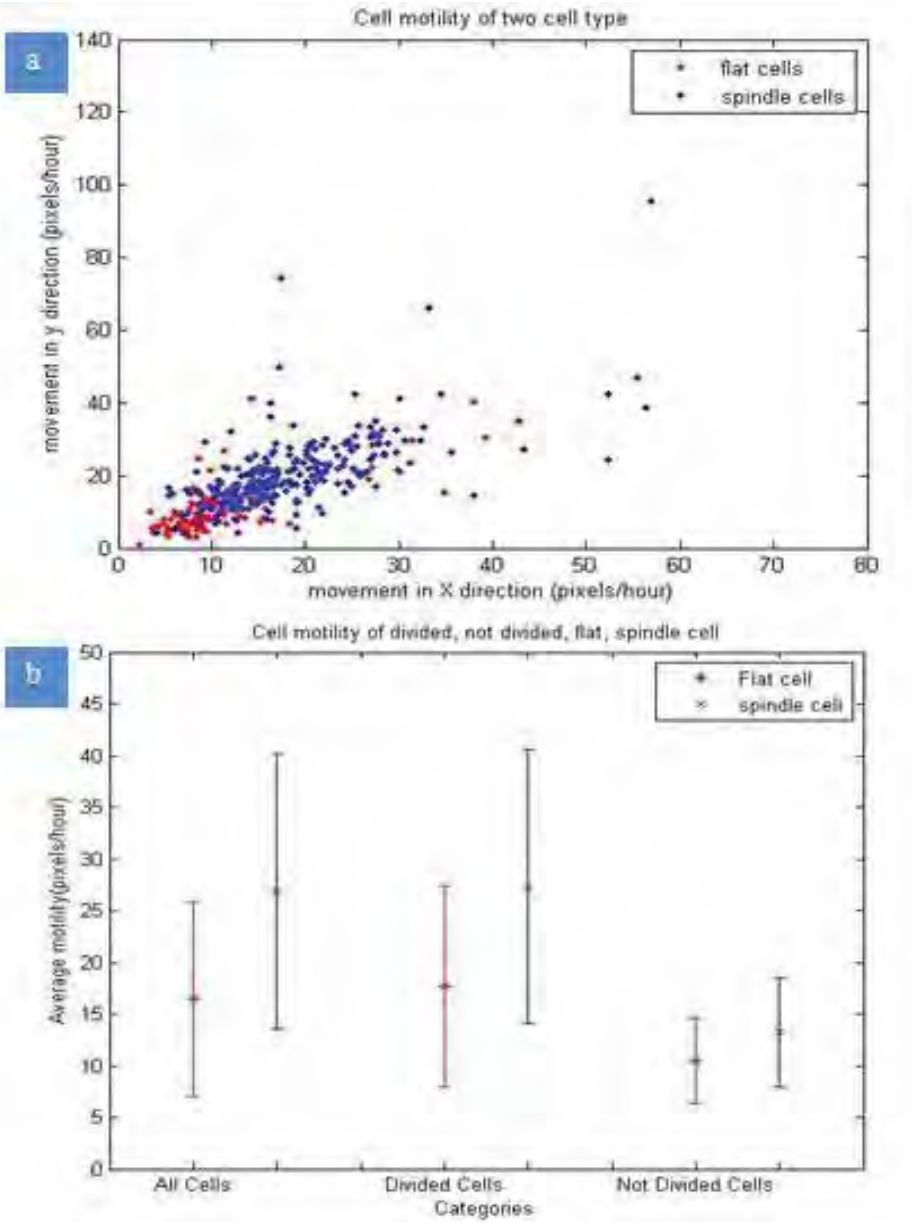


Figure 5. 14 Comparison of motility of cells with different shape and division history with error-bar

5.3.6 Kaplan Meier analysis (beta-actin GFP mouse)

The distribution of cells ages at mitosis was determined by Kaplan Meier analysis. Data for cCFU-F derived from small or large colonies was pooled. Figure 5.15 only shows the influence of cell size on cell cycle kinetics. During the five days of tracking, large cells (figure 5.15(a)) only have 3 generations, while small cells(figure 5.15(b)) have undergone 5 divisions in culture. It is shown later that large flattened cells, not matter whether they were selected from large or small colonies, have slower turn-over rate and lower growth potential than small spindle cells.

To further clarify the colony forming potential, cells from large colony and small colony are analysed separately. For this analysis data for small spindle cells and large flat cells are pooled. Figure 5.16 (a) displays the probability of mitosis of cells from large colony, and they have formed 5 generations in culture, meanwhile, cells from small colony only have 3 generations (figure 5.16(b)). This means that cells originating from large colonies have potential shorter cycle time compared to small colonies, regardless of cell size. Table 5.3 lists the average death rate of cell from different size colonies; cells from large colonies have a lower apoptosis rate.

Table 5. 3 Average cell death rate in 5 days tracking

Cell origin	Average death in 5 days (%)
Large colony	3.28
Small colony	16.67
Flat Cell	4.35
Spindle Cell	5.06

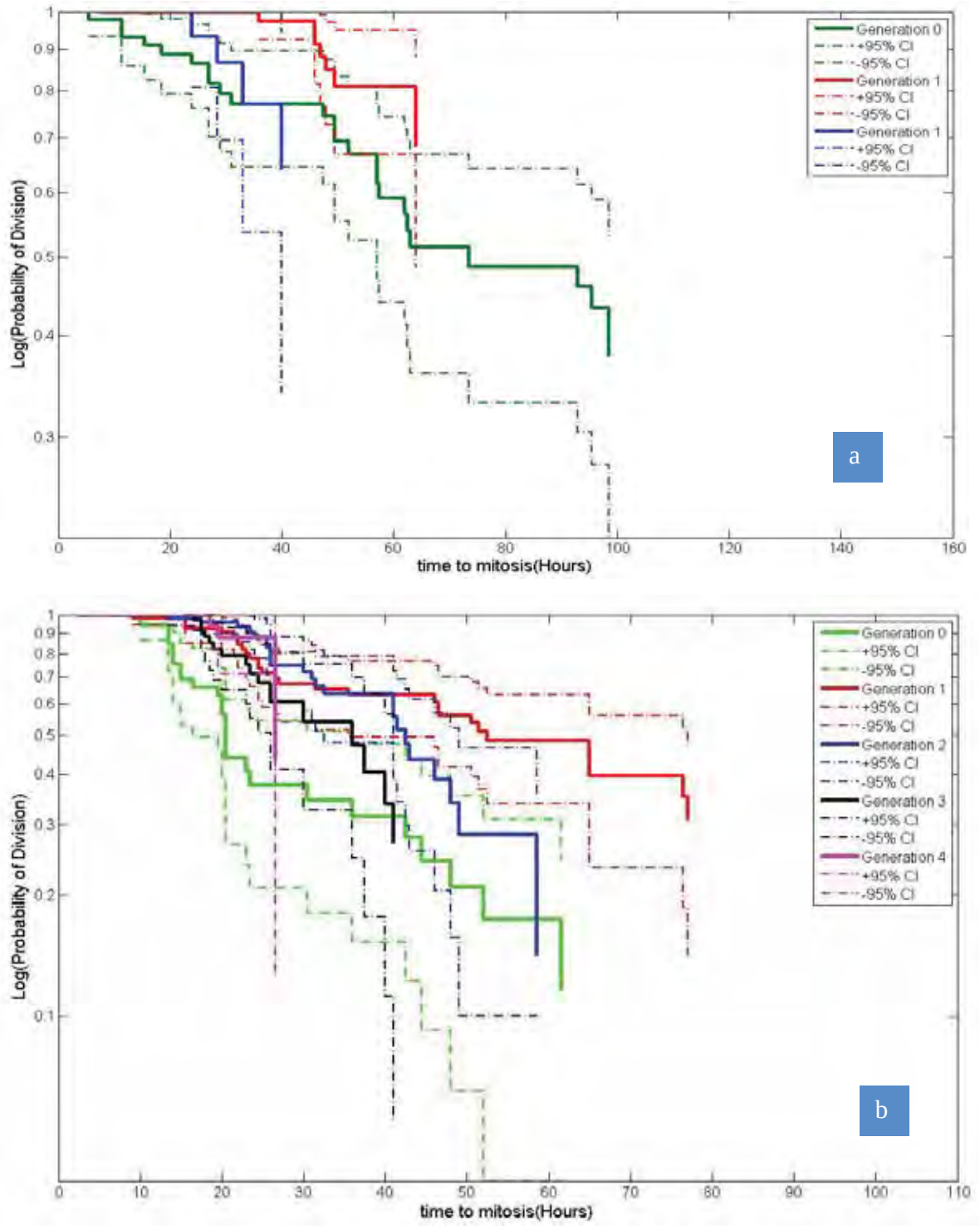


Figure 5. 15 Kaplan Meier Analysis cell division age showing the influence of cell size and shape; a: large and flat cells; b: small and spindle cells

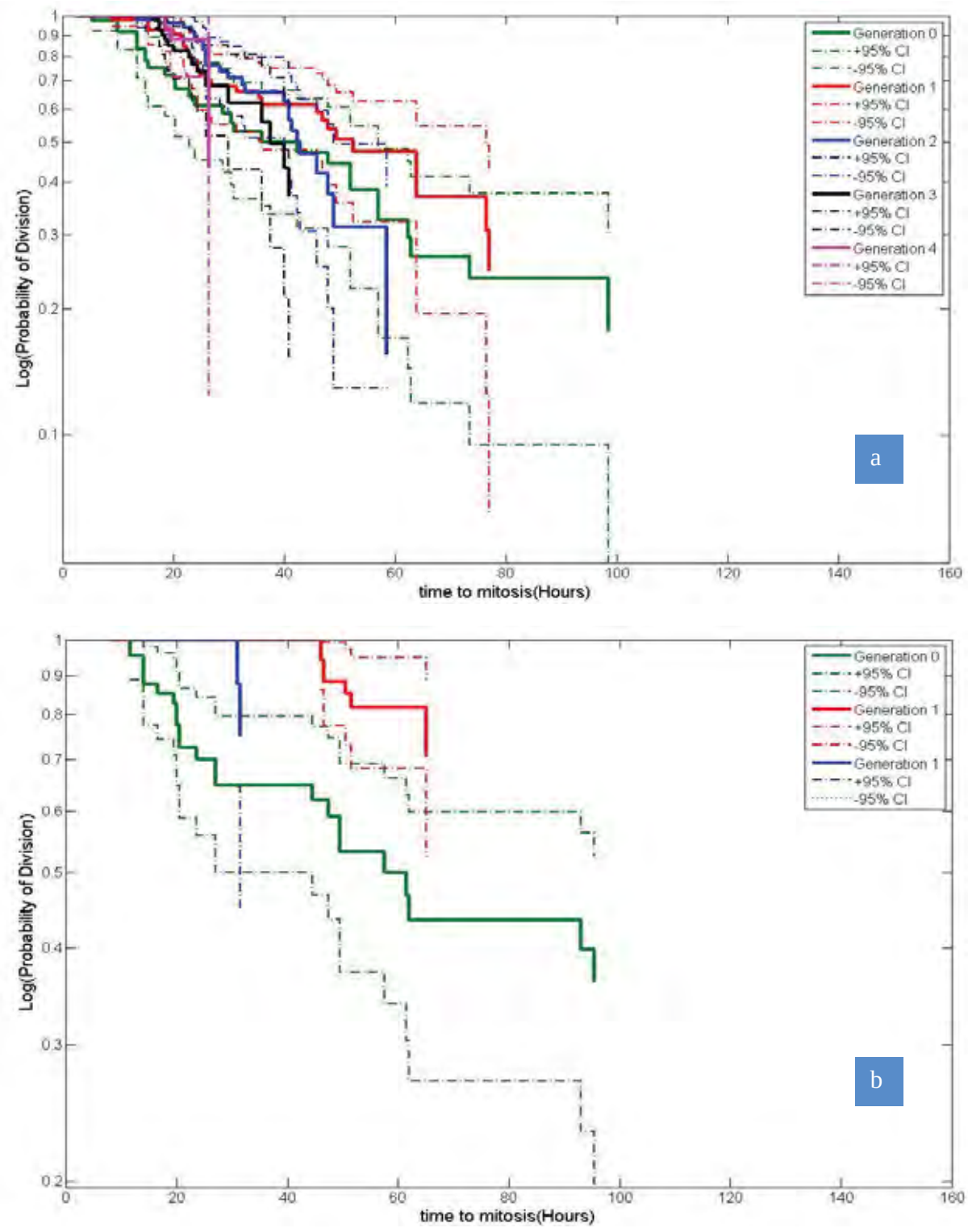


Figure 5. 16 Kaplan Meier analysis of cells with different origins, a: cells picked from large colonies; b: cells picked from small colonies

Thus cell cycle is related to both cell size and colony size. A more robust statistical comparison of cell lifetime data is made by performing Cox-regression analysis in section 5.3.7

Table 5.4 lists the cell cycle (± 95% CI), lag time estimated for large and small cells for each generation from large and small colonies, respectively, and records the number of tracked cells. It appears that cell cycle distribution varies with generation number. A number of patterns were observed: Excluding generation 0 cells where total life span of cells was not observed, generation 1 cells sometimes had a longer cycle time than subsequent generations (small colonies, small cells or large colonies, large cells). There was also a tendency for small cells from large colonies to increase their cell cycle length. The period of observation was too short to make comments about more than one generation of large cells.

Figure 5.17 shows KM plots for generation 0 cells where cell lifetime data has been pooled using categories based on cell size (small-spindle versus large-flat) and colony origin (small versus large). Cell cycle length is shortest in small cells from large colonies, followed by large cells from large colonies, small cells from small colonies and large cells from small colonies. Figure 5.18 shows the same trends using data pooled from generations 1-5.

Thus it is hypothesized cell cycle length is independently related to cell size and colony size. Cox regression analysis was used to test this hypothesis.

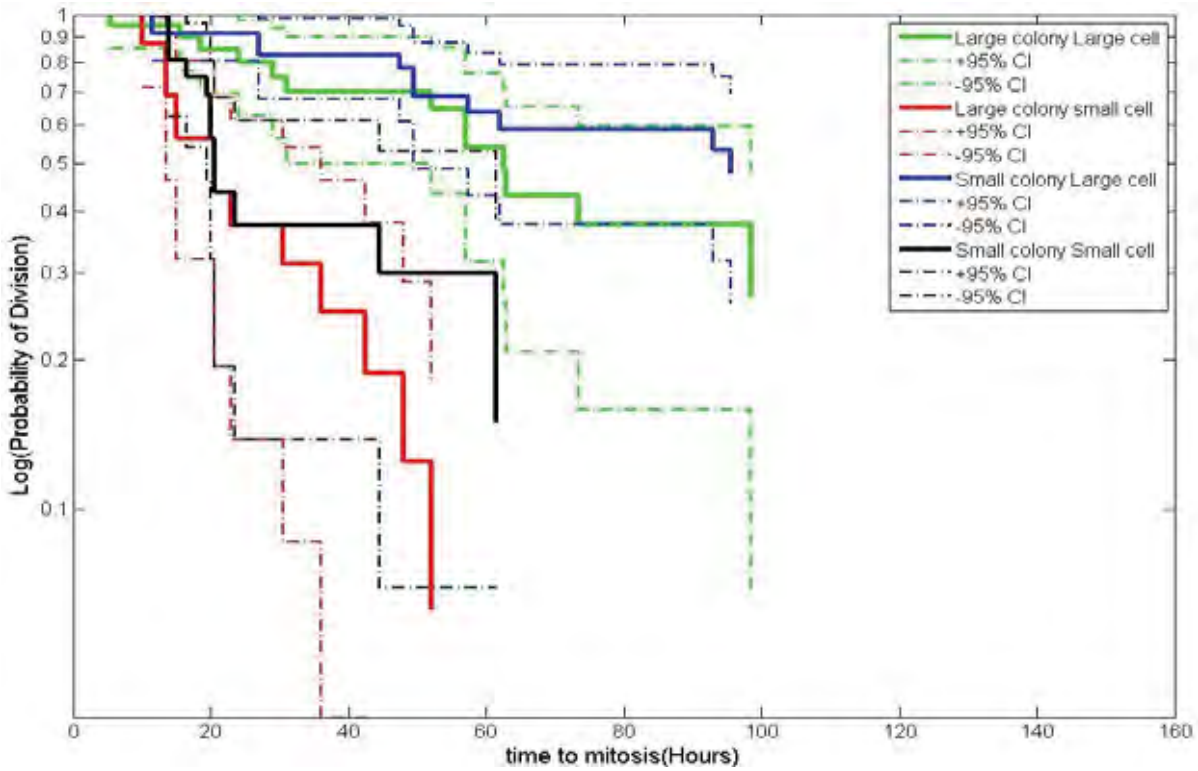


Figure 5. 17 KMA of G0 cells with different size and origin

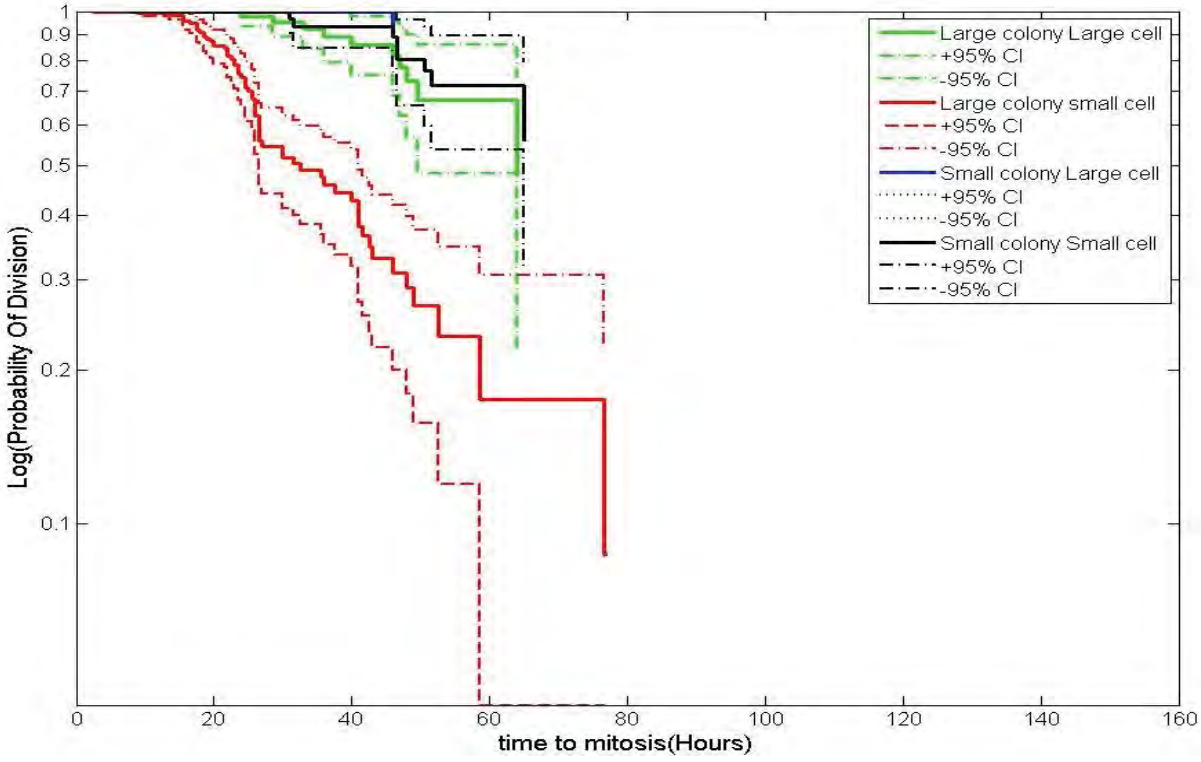


Figure 5. 18 KMA of G1-G5 cells with different size and origin

Table 5. 4 Estimated Cell Cycle Time

Cell origin	Cell Type	Cell age	% divide @ day 5	Cell cycle (hours)	Lag time (hours)	Tracked cell num
Small Colony	Small Cell	**Generation 0**	66.7 (11)	14.19(5.87)	2.75	18
		Generation 1	21.7(16)	120.33(53.67)	5.97	23
		Generation 2	20.0(30)	63.25(97.14)	5.75	10
		Generation 3	0	------	------	4
	Large Cell	**Generation 0**	45.80 (20)	64.73(24.21)	2.04	24
		Generation 1	0	------	-----	19
Large Colony	Small Cell	**Generation 0**	93.75 (9)	10.33(4.23)	2.95	16
		Generation 1	70.83(18)	15.73(15.62)	1.56	24
		Generation 2	47.05(22)	16.76(8.39)	7.10	34
		Generation 3	43.75(17)	12.02(4.51)	6.25	32
		Generation 4	21.42(32)	80.96(108.81)	1.40	28
		Generation 5	0	-------	------	12
	Large Cell	**Generation 0**	70.00 (10)	37.95(18.00)	3.07	20
		Generation 1	22.72(30)	81.19(103.43)	5.98	22
		Generation 2	40.00(25)	14.64(4.23)	11.18	10
		Generation 3	0	--------	--------	8

5.3.7 Cox proportional hazards regression applied to cell survival data

To determine the statistical significance of the observed differences between cell size and colony size, we tested our data by a proportional hazards model. Proportional hazards models and Kaplan Meier Analysis (Chapter 3.6) are commonly used to estimate survival in medical research. The application here to estimate the covariates associated with cell survival is somewhat novel. Cox regression analysis is used to assess the influence of a covariate (treatment or condition) that is relate to the time that passes before some event occurs. In medical research the covariate is commonly a treatment for cancer, and the event is disease-free survival. As applied here, the covariates are cell size and size of colony, and the event is cell mitosis. The proportional hazards model is used to determine whether the rate of mitosis is related to cell size or the size of colony origin.

Proportional hazards model is a semi-parametric method. On one hand, the determination of the baseline hazards function is empirically performed without parameters using KM analysis. On the other hand, a parametric model is used to examine hazard rates and covariates. In a proportional hazards model, the unique effect of a unit increase in a covariate is multiplicative with respect to the hazard rate which is illustrated in equation 5.2.

$$\boldsymbol{h(t|X) = h(t)\exp\left(X_1\beta_1 + \cdots + X_p\beta_p\right)} \qquad \textbf{Equation 5. 2}$$

where β is the relative hazards, and X are covariates, $h(t)$ is the non-parametric baseline hazards function.

As a result, the hazard ratio $HR(x)$is defined as the hazard function with given covariates x divided by the overall hazard function. The logarithm of the hazard ratio is then a linear regression model where β, the relative hazard, is to be estimated .

$$\boldsymbol{\log[HR(x)] = \log\left(\frac{h((t|x))}{h(t)}\right) = X_1\beta_1 + \cdots + X_p\beta_p} \qquad \textbf{Equation 5. 3}$$

In our application, the proportional hazards model can give an estimate of the relative mitosis hazard (β), their error, and p-value for different covariates. The covariates are colony size and

cell size. If p<0.01 there is a significant effect of cell size or colony, or both. For example, if large cells are defined by the categorical variable X_1=1, (small cells X_1=0), then exp(β_1) is the mitosis hazard ratio for large cells with respect to small cells. If the hazard ratio is significantly higher than 1, large cells have higher mitosis hazards than small cells.

Our data were tested by Cox-hazards regression with different combination of covariates (predictors), the estimated relative hazards, error of hazard, z-test value and p-test value are listed in table 5.5 and table 5.6.

Table 5. 5 Proportional hazards regression analysis for two covariates ($\boldsymbol{\beta_1}$ $\boldsymbol{and}$ $\boldsymbol{\beta_2}$)

Gen	**Predictors X** (2 groups)	**Relative hazard** β	**Error of** β	**z-statistic**	**p-value**
G0	**Colony** (1: Large, 0:Small)	0.9952	0.2490	3.9966	6.4259e-05
	Cell size (1: Large, 0:Small)	-0.9085	0.2707	-3.3570	7.8792e-04
G1-G5	**Colony** (1: Large, 0:Small)	0.8493	0.1898	4.4740	7.6787e-06
	Cell size (1: Large, 0:Small)	-1.6119	0.3082	-5.2294	1.7008e-07

In table 5.6, the regression model results for 4 covariates are presented. The covariates are (X_1, X_2, X_3, X_4). If the colony size is large and the cell size small, then $X_1 = 1, X_2 = 0, X_3 = 0, X_4 = 0$, if the colony size is large and the cell size large, then $X_1 = 0, X_2 = 1, X_3 = 0, X_4 = 0$, and so on. The rate of mitosis is significant different for all 4 subsets for generation 1-5, but only significant for large colonies and small cells, and small colonies and small cells in generation 0. By calculating the hazard ratio (exp(β)) it is evident that small cells and large colony cell have higher mitosis hazards than large cells and small colonies, respectively.

Moreover, the hazard ratio calculated from table 5.6 confirmed the observation of the cell cycle differences between cells with large or small size, from large or small colonies. The higher

mitosis hazards, indicates higher probability of mitosis and shorter cycle time. The Generation 0 cells with mitosis hazards from high to low are large colony-small cells, small colony-small cells, large colony-large cells, and small colony-large cells. Late generation cells have similar mitosis hazards order, but have exchange of position between large colony-large cells and small colony-small cells.

Table 5. 6 Proportional hazards regression with 4 covariates ($\boldsymbol{\beta_1 - \beta_4}$)

Gen	Predictors X (4 subsets)		Relative hazard β	Error of β	z-test	p-test
	Colony Size	Cell Size				
G0	Large	Small	1.3257	0.2815	4.7088	2.49e-06
	Large	Large	0.0152	0.2964	0.0514	0.9590
	Small	Small	0.8843	0.3285	2.8547	0.0043
	Small	Large	-0.4777	0.3098	-1.4542	0.1459
G1-G5	Large	Small	0.5296	0.2043	2.5916	0.0096
	Large	Large	-0.9373	0.3369	-2.7822	0.0054
	Small	Small	-1.0790	0.3636	-2.9677	0.0030
	Small	Large	-2.7129	1.0103	-2.6849	0.0073

Thus the observed differences in cell cycle for the cellular subsets defined by colony size and cell size have been confirmed by Cox regression analysis.

5.3.8 Cycle correlation analysis

As illustrated in figure 5.13(b), despite the heterogeneity of cell cycle lengths exhibited by different cell types, we found that siblings often divide within a short time of each other. A

comprehensive analysis is conducted by estimating the correlation coefficient for sibling cycle times is performed as the same method described in chapter 4.2.3

Figure 5.19 is a bivariate dot plot showing that sibling cycle times are in most cases within the sampling time interval of 30 minutes. This is further supported by figure 5.20, in which dead or lost cells (migrate outside of the field of view) were censored where the correlation coefficient was R^2=0.7853 (p=6.09e-47, m=361.49). These data indicate that the lineage relationship of sisterhood plays a role in determining cell generation time. Clearly, cell generation time is an important aspect of cell behavior regulating the growth of tissues, and is modulated during different states of lineage commitment. Bivariate analysis of mother-daughter generation times are described below.

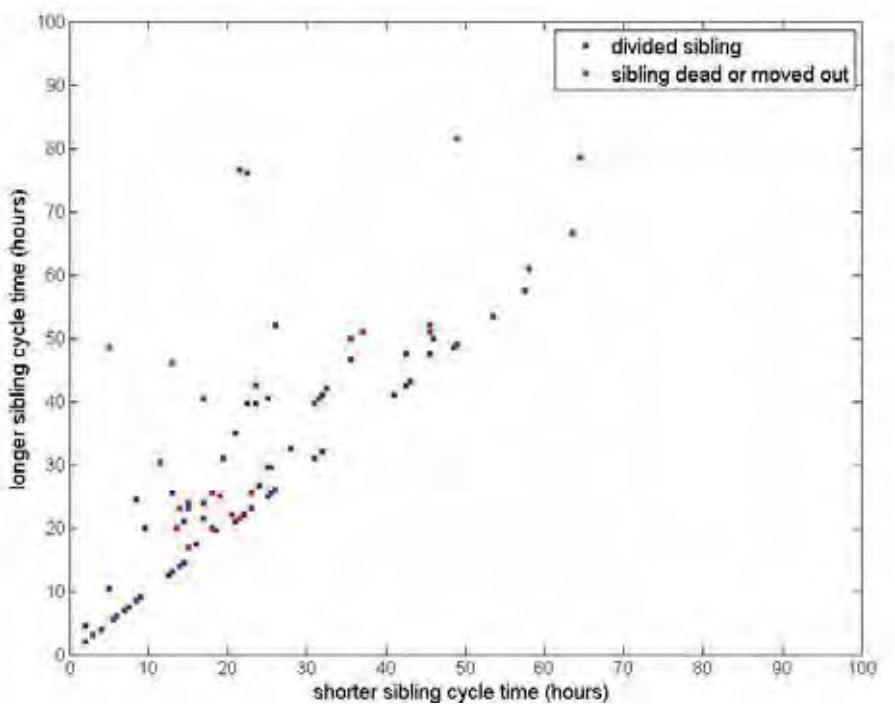


Figure 5. 19 Correlation of sibling cycle time

We have noticed that in figures 5.19 and 5.20, there are a few outliers which have large residual errors. These are two pairs of siblings with asymmetric cycle times within the 10 outliners (|Residual|>10). An example of an asymmetric division is shown in Figure 5.21, which also

provides the evidence that large flat cell can give rise to small spindle cells. In the image sequence, Cell 95 remained a flat cell during the whole tracking period. Cell 96 divided into 97 and 98 at the 27th hour, followed by division of 97 to 99 and 100 at 92nd hour. The birth cell 98 did not divide any more, but grew to be large and flat. These observations support the hypothesis that the cell cycle length is related to cell size because asymmetric division gave rise to two different sized siblings, the smaller sibling having shorter generation time.

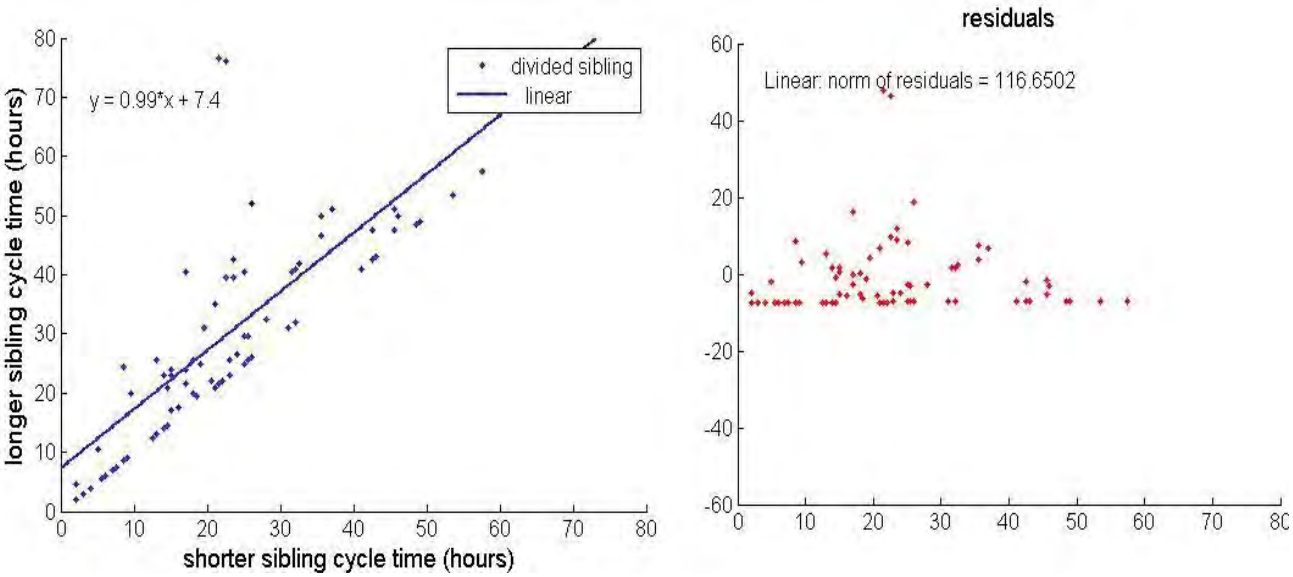


Figure 5. 20 Linear regression of sibling cycle with censoring of lost or and apoptotic cells inindicated in figure 5.19). Correlation R^2=0.7853, slope K=0.99

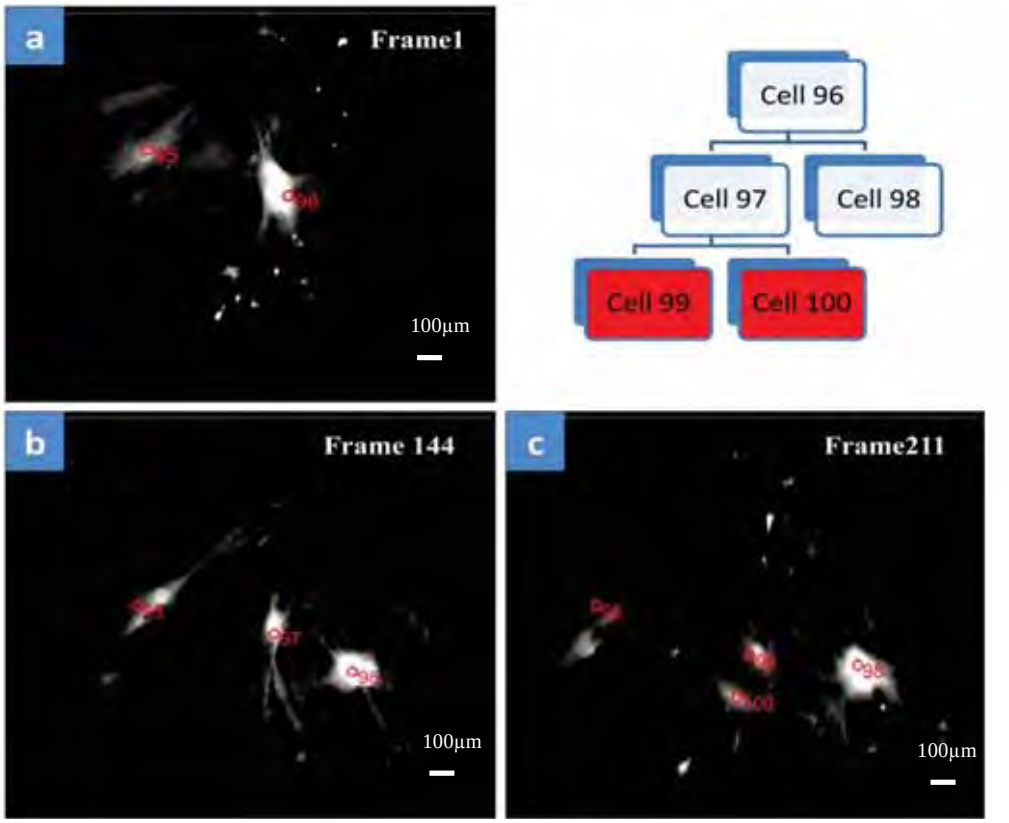


Figure 5. 21 Evidence of a flat cell giving birth to a spindle cell. Cell 95 remained a flat cell during the whole tracking period. Cell 96 divided into 97 and 98 at 27th hour, followed by division of 97 to 99 and 100 at 92th hour. Cell 98 did not divide any more, but grew to be large and flat.

The relationship between mother and daughter cycle times is shown by a dot plot (figure 5.22a); 201 pairs of mothers and daughters were identified. The dot distribution pattern is not along a line with gradient 1, but with a much higher gradient 72.89 (m=887.9, r^2=0.0113, p=0.3325). This means that cell cycle times are increasing in progeny, so overall, the cycle time distribution should be moving to the right (longer generation times). Interestingly maternal cells with an average cycle time of around 20 hours give rise to progeny with a much larger range of cycle times (10 to 90 hours, average 33 hours, figure 5.22). A similar phenomenon was observed in NIH3T3s in chapter 4, another fibroblast cell line. Therefore cCFU-Fs that are activated by addition of serum in culture media are gradually returning to quiescence after few generations, which is most likely their resting state in vivo.

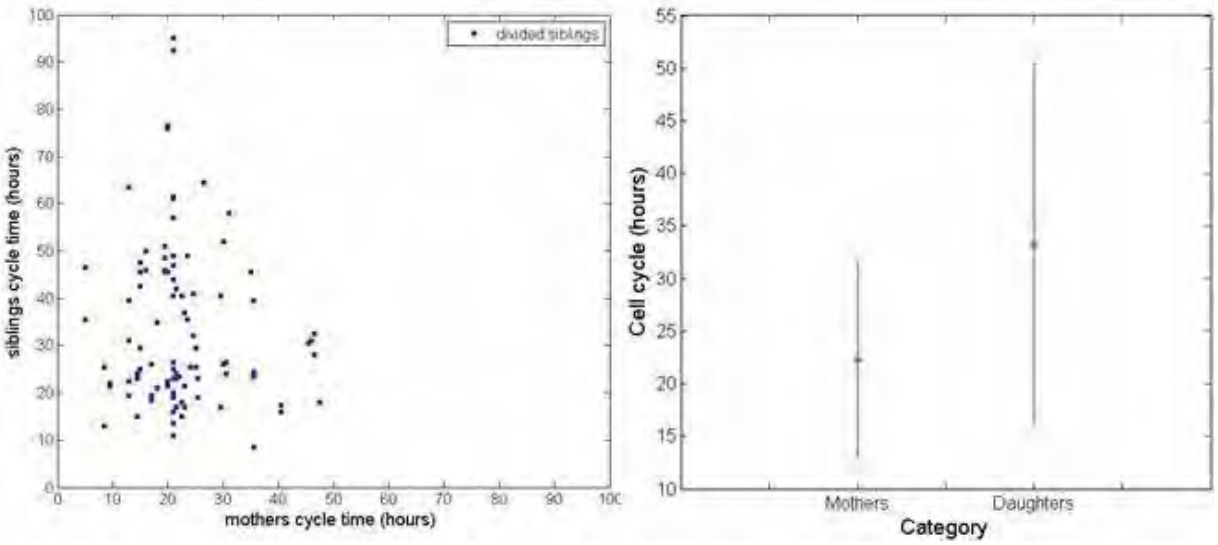


Figure 5. 22 Mother daughter cycle comparison a: Bivariate dot plot of mother daughter cell cycle length, b: Average cell cycle time.

5.4 Conclusion

Development of stem cell therapies in heart disease requires an understanding of cardiac lineage development and stem cell plasticity. Towards this end, a colony-forming assay originally developed by Friedenstein was adapted to quantify tissue resident cardiac stem cells. These cells give rise to fibroblastic colonies (cardiac colony forming units – fibroblast; cCFU-F) which have features in common with the mesenchymal stem cell phenotype, but are able to differentiate at high efficiency into cardiac and smooth muscle. The goal of this chapter was to analyse cCFU-F by live cell imaging, to detail lineage relationships, and to correlate cell phenotype with cell cycle dynamics.

A live cell imaging system and semi-automated in-house software was used to track the trajectories and divisions of cCFU-F colonies over extended culture periods. Cardiac CFU-F from the PDGFR α -GFP mouse and beta actin-GFP mouse were enriched from an interstitial cell fraction of hearts by FACS (PDGFR-a+, Sca1+/Pecam1-). Initial studies have focused on live cell imaging of GFP-positive cells (PDGFR α + and BA+), which have long-term growth potential *in vitro.*

Lineage maps of PDGFR α + cells within passage 3 colonies (164 maps) and BA+ cells within passage 3 colonies (352 maps) were constructed from tracking cells over five days of culture. Two distinctive cell morphologies were indentified; large flattened-cells with low motility and spindle-shaped cells. The probability of mitosis was estimated by Kaplan Meier analysis. Cell lifespan (time between successive mitoses) was approximated by an exponential distribution. The average cycle time of flattened- and spindle-shaped cells varies in different generations. However, small spindle shaped cells have overall shorter cell cycle than large flattened cells, and differences between two subsets are statistically significant. Thus early passage cCFU-F is composed of at least two subpopulations with markedly different morphology and cell cycle kinetics. Moreover, the growth potential and colony forming ability of large colony cells are higher than that of small colony cells. We have confirmed the hypothesis cell cycle times are inversely related to colony size and directly related to cell size.

Whilst sister cells have very similar cycle times, daughter cells have longer generation times compared to their mother so cCFU-F will return to quiescence over successive generations.

Furthermore, asymmetric divisions where large cells give rise to small spindle cells have been observed with a frequency of 2/40 providing a mechanism for generation of spindle shaped cells from large flat cells. Future work will focus on immunophenotyping these two cellular subsets to identify how cCFU-F growth and development is regulated by microenvironmental cues.

Chapter 6 Summary and Future Work

The central goal of this thesis was to establish methodology for mapping lineage pedigrees by live cell imaging. The utility of live cell imaging is the ability to incorporate cellular dynamic properties into the phenotypic classification of cell subsets, greatly enhancing understanding of stem and progenitor cell development. The system hardware and software has been validated using NIH3T3 cells and primary cardiac CFU-F isolated from PDGFR-a-GFP or β-actin-GFP mice. In addition to developing enabling methodology for lineage mapping, the thesis has led to new and exciting observations that will contribute to the understanding cardiac CFU-F growth and development. Two subpopulations (large flat and small spindle cells) were identified with quite different live cell properties. Preliminary immunophenotypic characterisation using fluorescently labelled antibodies to intracellular antigens (Nkx2.5, Myosin Heavy Chain, SMA-α, Calponin), not presented in this thesis because of their preliminary nature, indicate that slowly dividing large-flat immotile cells are smooth muscle cells whilst the spindle shaped cells have a fibromyocyte phenotype.

This thesis has addressed the following aims:

1. Customized a live cell imaging platform for long term imaging of cardiac stem cells
2. Developed a fluorescence imaging methods for long term tracking of GFP marked cells
3. Developed software to segment cardiac stem cell division trees from cell growth videos
4. Acquired enough data to characterize cardiac stem cell development
5. Applied Kaplan-Meier statistics and Cox regression analysis of cell lifetime data to identify the different cell types that are present in cardiac CFU-F based on cell cycle kinetics and other live cell imaging characteristics i.e., motility and cell size.

6. Quantified the transmission of cell cycle length from mother to daughter cells using lineage pedigrees.

Novel aspects of this study were the application of Kaplan-Meier and Cox regression analysis to cell lifetime data (Nordon, 2011b) as well as identification of two new functional phenotypes that comprise cCFU-F.

6.1 Mapping lineage pedigrees by live cell imaging

The aim of this study was to develop a microscope incubator system and software to semi-automate tracking and segmentation of cell division and motion by continuous live cell imaging (aims 1-3). The growth rate of the cell line NIH3T3 in the microscope environmental chamber was similar to that of cells grown by standard tissue culture. The combination of region based and edge based segmentation methods improved the efficiency of cellular image segmentation; the feature measurement and likelihood matching methods enabled accurate association of cells between consecutive frames. The manual editing software platform helped in validation and visualization of the results.

Despite the stated advantages of the software, there are still many problems to address. Shorter computation time will allow processing of large image stacks, as well as increasing the response time to manual edits. The method proposed in this study has similar computational speed to other systems implemented via Matlab (Cohen et al., 2010) . However, converting the algorithm to a compiled programming language (e.g., C++) would drastically improve the efficiency, especially manual editing.

Secondly, development of more automated and robust cell tracking algorithms would improve the performance of this proposed method. The current segmentation method fails to separate overlapping cells in phase contrast images. Although manual interaction is inevitable at certain times, techniques which can separate merged cells automatically are important to reduce the level of editing required to correct overlapping tracks.

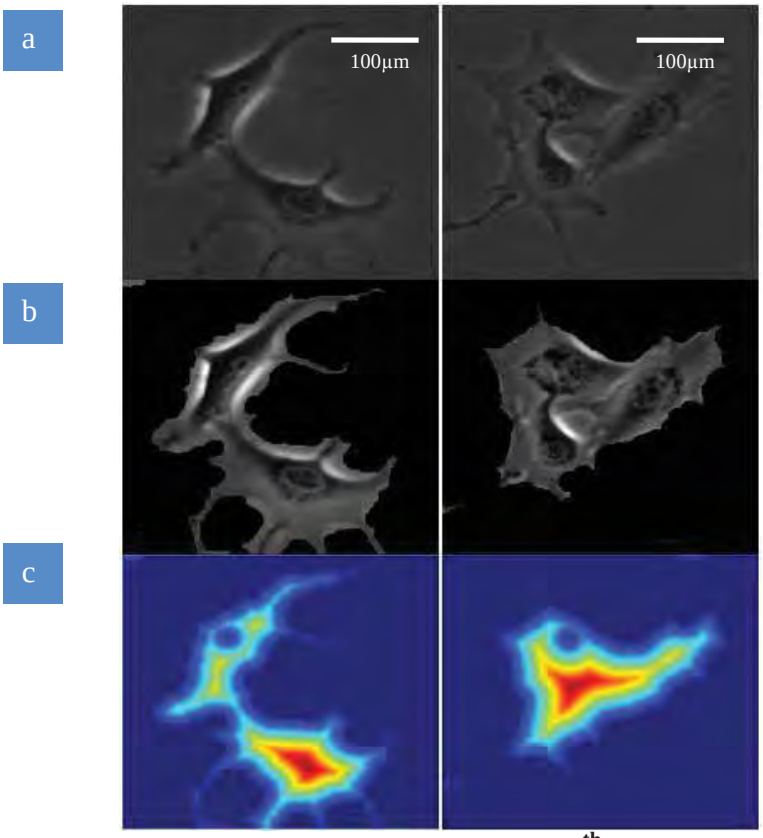


Figure 6. 1 X10 Phase contrast images taken at the 30th minute; (a) 10x phase contrast images; (b) Segmentated binary image of the foreground; (c) Distance transform using the binary masks in b.

When difficult situations are encountered such as overlapping cells (figure 6.1) alternate segmentation methods will be required. For example in figure 6.1 two or more cells were overlapping. One may apply the distance transform to separate cells in the left image, because there are two local maxima (the yellow and red regions). However, the right hand image which contains 3 cells has a single maximum, so the distance transform would not work in this situation. An interactive platform with multiple segmentation algorithms will inevitably be required, since no single method will provide a robust solution in all situations.

Surprisingly, we identified a few cases where NIH3T3 cells did not undergo cell separation after nuclear division. This process is called endomitosis and was observed quite frequently in cCFU-F (large flat cells). In figure 6.2, it appears that cells detach from the surface and underwent mitosis at the 675th minute; however, instead of fully dividing into two separate cells at the 729th minute, the nuclei reattached to the culture surface, and were incorporated as a binuclear cell at

the 780th minute, and remained undivided until the end of tracking period (6032th minutes). The reason of this abnormal division pattern is still unknown, although multi-nucleated NIH3T3 have been identified long ago (Zelikoff et al., 1986). This pattern of division was not anticipated and resulted in a tracking error. Future versions of the software will need to distinguish between mitosis and endomitosis.

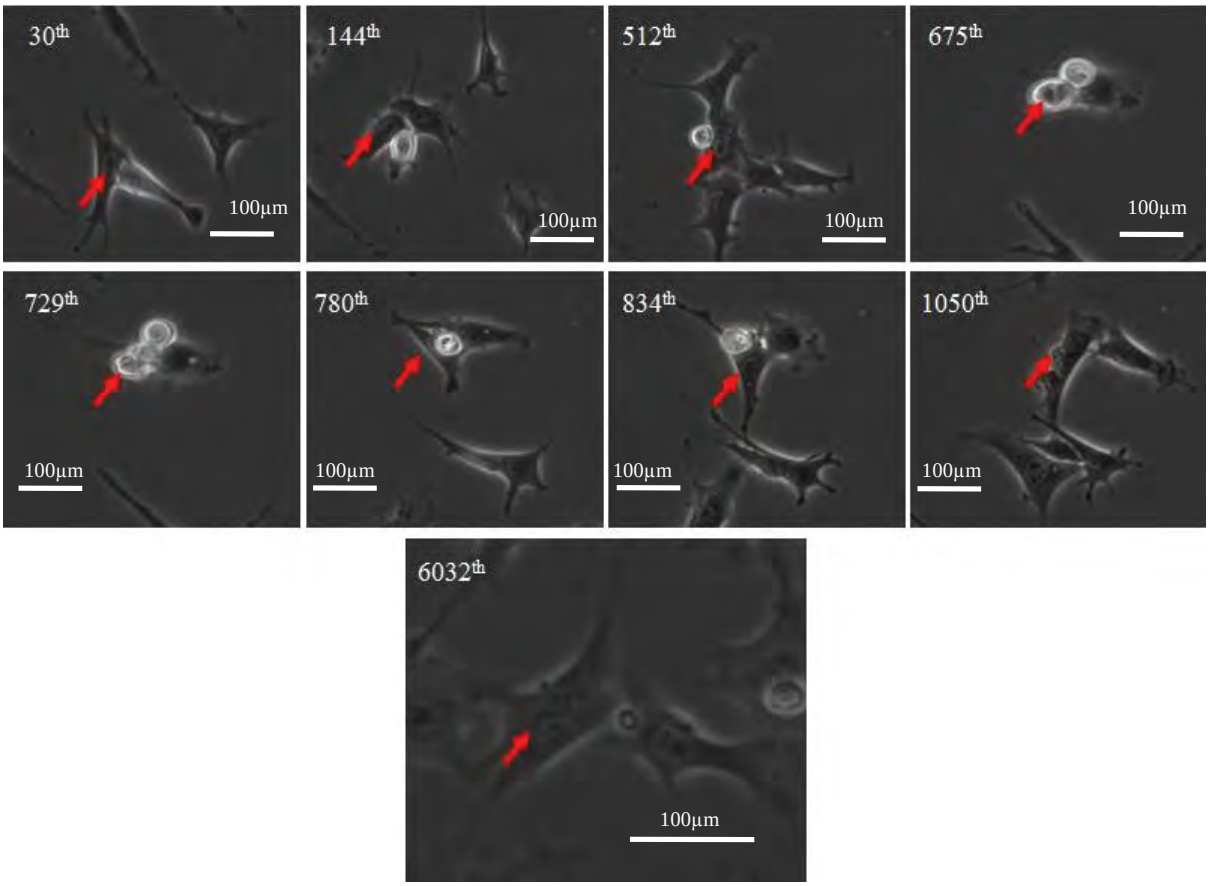


Figure 6. 2 The failed division identified on X10 phase contrast image sequences, the cell was marked by the red arrow

Lastly, a comprehensive statistical analysis of the automatic segmentation accuracy was not performed. Given that a variety of segmentation methods may be applied to an image stack, it will be necessary to benchmark each algorithm with respect to false positive and negative identification errors. Thus a tool that provides the user with feedback on how many errors were made at the end of manual correction would be invaluable for optimising segmentation algorithms and their parameters.

6.2 Tracking cCFU-F

The second goal of this study was to apply live cell imaging and lineage mapping to characterize cardiac CFU-F development (aims 4-6). Lineage maps of PDGFR-α+ cells (164 maps) and beta actin+ cells (352 maps) within passage 3 colonies were constructed by tracking cells over five days of culture. Two distinctive cell morphologies were identified; large flattened cells with low motility and spindle-shaped cells. The average cycle time of flattened- and spindle-shaped cells were estimated for each generation. Statistical analysis of cell cycle and motility confirmed significant differences between the two cell types.

Cell cycle correlation analysis of mothers and daughters confirmed the observation that lineage relationships play a role in determining division rate (Al-Kofahi et al., 2006, Nordon, 2011a). Lastly asymmetric divisions where large cells gave rise to small and large cells with different cycle times suggest a possible mechanism by which spindle shaped cells are generated. Even though the generation of large cells from small cells was not observed, more extensive studies are required to reject this possible transition. Future work will also examine the evolution of cCFU-F dynamic behavior with respect to passage number; by examining passage 0, it will be possible to determine whether the two populations already exist in vivo, or are generated by the in vitro culture system. Also the relationship between PDGFR-a gene expression and generation of cCFU-F and its subpopulations deserves a deeper analysis to establish a role in cardiac cell generation. This will be possible by crossing two transgenic mice strains i.e., PGDFR-a GFP mice with beta-actin DsRed mice, to track both PDGFR+ and PDGFR- cells using two colour fluorescence.

Reference

AHMED, W. M., GHAFOOR, A. & ROBINSON, J. P. 2007. Knowledge extraction for high-throughput biological imaging. *IEEE Multimedia,* 14**,** 52-62.

AL-KOFAHI, O., RADKE, R. J., GODERIE, S. K., SHEN, Q., TEMPLE, S. & ROYSAM, B. 2006. Automated cell lineage construction - A rapid method to analyze clonal development established with murine neural progenitor cells. *Cell Cycle,* 5**,** 327-335.

ALISON, M. R. & ISLAM, S. 2009. Attributes of adult stem cells. *Journal of Pathology,* 217**,** 144-160.

AMANO, K., OKIGAKI, M., ADACHI, Y., FUJIYAMA, S., MORI, Y., KOSAKI, A., IWASAKA, T. & MATSUBARA, H. 2004. Mechanism for IL-1 beta-mediated neovascularization unmasked by IL-1 beta knock-out mice. *Journal of Molecular and Cellular Cardiology,* 36**,** 469-480.

BARANIAK, P. R. & MCDEVITT, T. C. 2010. Stem cell paracrine actions and tissue regeneration. *Regenerative Medicine,* 5**,** 121-143.

BREMAUD, P. 1999. Markov Chains. *Gibbs fields, Monte Carlo simulation, and queues,Springer, Text in Applied Mathematics,* 31.

CAMELLITI, P., BORG, T. K. & KOHL, P. 2005. Structural and functional characterisation of cardiac fibroblasts. *Cardiovascular Research,* 65**,** 40-51.

CANNY, J. 1986. A Computational Approach To Edge Detection. *IEEE Trans. Pattern Analysis and Machine Intelligence,* 8**,** 679-698.

CHONG, J. J., XAYMARDAN, M., PRALL, O., CHANDRAKANTHAN, V., AHMED, I., ASLI, N., BOUVERET, R., SCARLETT, C., COLVIN, E., PINESE, M., BIANKIN, A. & HARVEY, R. P. 2009. Cardiac Stem Cells Defined by Colony Forming Unit Assays Are From an Epicardial/Mesodermal but Not Neural-Crest, Cardiomyocyte or Bone-Marrow Origin. *Circulation,* 120**,** S808-S808.

COHEN, A. R., GOMES, F., ROYSAM, B. & CAYOUETTE, M. 2010. Computational prediction of neural progenitor cell fates. *Nature Methods,* 7**,** 213-U75.

COOPER, G. 2000. Chapter 14: The Eukaryotic Cell Cycle". *The cell: a molecular approach (2nd ed.) Washington, D.C: ASM Press. ISBN 0-87893-106-6.*

DE BOER, R. J. & PERELSON, A. S. 2005. Estimating division and death rates from CFSE data. *Journal of Computational and Applied Mathematics,* 184**,** 140-164.

DEBEIR, O., CAMBY, I., KISS, R., VAN HAM, P. & DECAESTECKER, C. 2004. A model-based approach for automated in vitro cell tracking and chemotaxis analyses. *Cytometry Part A,* 60A**,** 29-40.

DEBEIR, O., VAN HAM, P., KISS, R. & DECAESTECKER, C. 2005. Tracking of migrating cells under phase-contrast video microscopy with combined mean-shift processes. *Ieee Transactions on Medical Imaging,* 24**,** 697-711.

DRUMMOND, S. P. & ALLEN, T. D. 2008. Chapter 6 From Live-Cell Imaging to Scanning Electron Microscopy (SEM): The Use of Green Fluorescent Protein (GFP) as a Common Label. *In:* TERENCE, D. A. (ed.) *Methods in Cell Biology.* Academic Press.

DUFOUR, A. & OLIVO-MARIN, J. C. Year. Tracking live cells in 4D microscopy: From active surfaces to active meshes. *In:* Signals, Systems and Computers, 2008 42nd Asilomar Conference on, 26-29 Oct. 2008 2008. 51-55.

DUFOUR, A., SHININ, V., TAJBAKHSH, S., GUILLEN-AGHION, N., OLIVO-MARIN, J. C. & ZIMMER, C. 2005. Segmenting and tracking fluorescent cells in dynamic 3-D microscopy with coupled active surfaces. *Ieee Transactions on Image Processing,* 14**,** 1396-1410.

DUFOUR, A., THIBEAUX, R., LABRUYERE, E., GUILLEN, N. & OLIVO, J. C. 2010. 3D Active Meshes: Fast Discrete Deformable Models for Cell Tracking in 3D Time-Lapse Microscopy. *Image Processing, IEEE Transactions on,* PP**,** 1-1.

DZYUBACHYK, O., ESSERS, J., VAN CAPPELLEN, W. A., BALDEYRON, C., INAGAKI, A., NIESSEN, W. J. & MEIJERING, E. 2010a. Automated analysis of time-lapse fluorescence microscopy images: from live cell images to intracellular foci. *Bioinformatics,* 26**,** 2424-2430.

DZYUBACHYK, O., NIESSEN, W. & MEIJERING, E. Year. Advanced level-set based multiple-cell segmentation and tracking in time-lapse fluorescence microscopy images. *In:* Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on, 2008. 185-188.

DZYUBACHYK, O., VAN CAPPELLEN, W. A., ESSERS, J., NIESSEN, W. J. & MEIJERING, E. 2010b. Advanced Level-Set-Based Cell Tracking in Time-Lapse Fluorescence Microscopy. *Ieee Transactions on Medical Imaging,* 29**,** 852-867.

EILKEN, H. M., NISHIKAWA, S. & SCHROEDER, T. Year. Blood generation from hemogenic endothelium proven by continuous single cell imaging. *In,* Jul 09-12 2008 Boston, MA. S1-S1.

FLAIBANI, M., LUNI, C., SBALCHIERO, E. & ELVASSORE, N. 2009. Flow Cytometric Cell Cycle Analysis of Muscle Precursor Cells Cultured Within 3D Scaffolds in a Perfusion Bioreactor. *Biotechnology Progress,* 25**,** 286-295.

FORD, J. W., WELLING, T. H., STANLEY, J. C. & MESSINA, L. M. 1996. PKH26 and I-125-PKH95: Characterization and efficacy as labels for in vitro and in vivo endothelial cell localization and tracking. *Journal of Surgical Research,* 62**,** 23-28.

FRIEDENS.AJ, C. R., LALYKINA, K. S 1970. Development of Fibroblast Colonies in Monolayer Cultures of Guinea-Pig Bone Marrow and Spleen Cells. *Cell and Tissue Kinetics,,* 3**,** 393.

FRIEDENSTEIN, A. 1989a. Stromal-Hematopoietic Interrelationships: Maximov's Ideas and Modern Models *Haematology and Blood Transfusion,* 32**,** 159-167.

FRIEDENSTEIN, A. 1989b. Stromal-Hematopoietic Interrelationships:

Maximov's Ideas and Modern Models *Haematology and Blood Transfusion,* 32**,** 159-167.

FRIGAULT, M. M., LACOSTE, J., SWIFT, J. L. & BROWN, C. M. 2009. Live-cell microscopy - tips and tools. *Journal of Cell Science,* 122**,** 753-767.

GODARA, P., NORDON, R. E. & MCFARLAND, C. D. 2008. Mesenchymal stem cells in tissue engineering. *Journal of Chemical Technology and Biotechnology,* 83**,** 397-407.

GONZALES, C. & PEDRAZZINI, T. 2009. Progenitor cell therapy for heart disease. *Experimental Cell Research,* 315**,** 3077-3085.

GONZALEZ, R. C. (3ed.) 2008. *Digital Image Processing,* Upper Saddle River, N.J.: Prentice Hall.

GREEN, K. J. 2002. Improving understanding of exercise effects on in vitro T-lymphocyte function - The role of fluorescent cell division tracking. *Exercise Immunology Review,* 8**,** 101-115.

GUSTAVSSON, T., ALTHOFF, K., DEGERMAN, J., OLSSON, T., THORESON, A. C., THORLIN, T. & ERIKSSON, P. Year. Time-lapse microscopy and image processing for stem cell research modeling cell migration. *In:* SONKA, M. & FITZPATRICK, J. M., eds., 2003 San Diego, CA. 1-15.

HALTER, M., ELLIOTT, J. T., HUBBARD, J. B., TONA, A. & PLANT, A. L. 2009. Cell volume distributions reveal cell growth rates and division times. *Journal of Theoretical Biology,* 257**,** 124-130.

HAMILTON, T. G., KLINGHOFFER, R. A., CORRIN, P. D. & SORIANO, P. 2003. Evolutionary divergence of platelet-derived growth factor alpha receptor signaling mechanisms. *Molecular and Cellular Biology,* 23**,** 4013-4025.

HANSSON, E. M., LINDSAY, M. E. & CHIEN, K. R. 2009. Regeneration Next: Toward Heart Stem Cell Therapeutics. *Cell Stem Cell,* 5**,** 364-377.

HARDER, N., MORA-BERMUDEZ, F., GODINEZ, W. J., WUNSCHE, A., EILS, R., ELLENBERG, J. & ROHR, K. 2009. Automatic analysis of dividing cells in live cell movies to detect mitotic delays and correlate phenotypes in time. *Genome Research,* 19**,** 2113-2124.

HARVEY, R. P., MEILHAC, S. M. & BUCKINGHAM, M. E. 2009. Landmarks and Lineages in the Developing Heart. *Circulation Research,* 104**,** 1235-1237.

HASBOLD, J. & HODGKIN, P. D. 2000. Flow cytometric cell division tracking using nuclei. *Cytometry,* 40**,** 230-237.

HATTORI, F. & FUKUDA, K. 2010. Strategies for ensuring that regenerative cardiomyocytes function properly and in cooperation with the host myocardium. *Experimental and Molecular Medicine,* 42**,** 155-165.

HOEBE, R. A., VAN OVEN, C. H., GADELLA, T. W. J., DHONUKSHE, P. B., VAN NOORDEN, C. J. F. & MANDERS, E. M. M. 2007. Controlled light-exposure microscopy reduces photobleaching and phototoxicity in fluorescence live-cell imaging. *Nature Biotechnology,* 25**,** 249-253.

JAINCHIL.JL, AARONSON, S. A. & TODARO, G. J. 1969. MURINE SARCOMA AND LEUKEMIA VIRUSES - ASSAY USING CLONAL LINES OF CONTACT-INHIBITED MOUSE CELLS. *Journal of Virology,* 4**,** 549-&.

JOGGERST, S. J. & HATZOPOULOS, A. K. 2009. Stem cell therapy for cardiac repair: benefits and barriers. *Expert Reviews in Molecular Medicine,* 11.

JUN, Y., XIAOBO, Z., QIONG, Y., NING, L., QIANSHENG, C. & WONG, S. T. C. Year. An Effective System for Optical Microscopy Cell Image Segmentation, Tracking and Cell Phase Identification. *In:* Image Processing, 2006 IEEE International Conference on, 8-11 Oct. 2006 2006. 1917-1920.

KABLE, E. 2006. Keeping Cells Alive Under the Microscope. *Optical Imaging Techniques in Cell Biology.* CRC Press.

KAPLAN, E. L. & MEIER, P. 1958. NONPARAMETRIC-ESTIMATION FROM INCOMPLETE OBSERVATIONS. *Journal of the American Statistical Association,* 53**,** 457-481.

KELLEHER, C. M. & VACANTI, J. P. 2010. Engineering extracellular matrix through nanotechnology. *Journal of the Royal Society Interface,* 7**,** S717-S729.

KHOSHMANESH, K., KOUZANI, A. Z., NAHAVANDI, S., BARATCHI, S. & KANWAR, J. R. 2008. At a glance: Cellular biology for engineers. *Computational Biology and Chemistry,* 32**,** 315-331.

KO, K. H., ODELL, R. & NORDON, R. E. 2007. Analysis of cell differentiation by division tracking cytornetry. *Cytometry Part A,* 71A**,** 773-782.

KOHARA, H. & TABATA, Y. 2010. Review: Tissue Engineering Technology to Enhance Cell Recruitment for Regeneration Therapy. *Journal of Medical and Biological Engineering,* 30**,** 267-276.

KRENNING, G., ZEISBERG, E. M. & KALLURI, R. 2010. The Origin of Fibroblasts and Mechanism of Cardiac Fibrosis. *Journal of Cellular Physiology,* 225**,** 631-637.

LEE, G. M., FONG, S. S., OH, D. J., FRANCIS, K. & PALSSON, B. O. 2002. Characterization and efficacy of PKH26 as a probe to study the replication history of the human hematopoietic KG1a progenitor cell line. *In Vitro Cellular & Developmental Biology-Animal,* 38**,** 90-96.

LEVI, V. & GRATTON, E. 2007. Exploring dynamics in living cells by tracking single particles. *Cell Biochemistry and Biophysics,* 48**,** 1-15.

LI, F. H., ZHOU, X. B., MA, J. W. & WONG, S. T. C. 2010. Multiple Nuclei Tracking Using Integer Programming for Quantitative Cancer Cell Cycle Analysis. *Ieee Transactions on Medical Imaging,* 29**,** 96-105.

LI, K., MILLER, E. D., CHEN, M., KANADE, T., WEISS, L. E. & CAMPBELL, P. G. Year. Cell population tracking and lineage construction with spatiotemporal context. *In,* Oct 26-Nov 02 2007 Brisbane, AUSTRALIA. 546-566.

LI, K., MILLER, E. D., CHEN, M., KANADE, T., WEISS, L. E. & CAMPBELL, P. G. 2008. Cell population tracking and lineage construction with spatiotemporal context. *Medical Image Analysis,* 12**,** 546-566.

LIAO, S., CHAN, C. K. & RAMAKRISHNA, S. 2008. Stem cells and biomimetic materials strategies for tissue engineering. *Materials Science & Engineering C-Biomimetic and Supramolecular Systems,* 28**,** 1189-1202.

LOBUTOVA, E., RESAGK, C., RANK, R. & MULLER, D. Year. Extended Three Dimensional Particle Tracking Velocimetry for Large Enclosures. *In:* NITSCHE, W. & DOBRILOFF, C., eds., May 14-15 2009 Berlin, GERMANY. 113-124.

MAKKAPATI, V. V. & NAIK, S. K. Year. Clump splitting based on detection of dominant points from contours. *In:* Automation Science and Engineering, 2009. CASE 2009. IEEE International Conference on, 22-25 Aug. 2009 2009. 197-201.

MASUR, S. K., DEWAL, H. S., DINH, T. T., ERENBURG, I. & PETRIDOU, S. 1996. Myofibroblasts differentiate from fibroblasts when plated at low density. *Proceedings of the National Academy of Sciences of the United States of America,* 93**,** 4219-4223.

MEIJERING, E., DZYUBACHYK, O., SMAL, I. & VAN CAPPELLEN, W. A. 2009. Tracking in cell and developmental biology. *Seminars in Cell & Developmental Biology,* 20**,** 894-902.

MULLASSERY, D., HORTON, C. A., WOOD, C. D. & WHITE, M. R. H. 2008. Single live-cell imaging for systems biology. *Essays in Biochemistry: Systems Biology, Vol 45.*

MURRY, C. E., FIELD, L. J. & MENASCHE, P. 2005. Cell-based cardiac repair - Reflections at the 10-year point. *Circulation,* 112**,** 3174-3183.

MUZZEY, D. & VAN OUDENAARDEN, A. 2009. Quantitative Time-Lapse Fluorescence Microscopy in Single Cells. *Annual Review of Cell and Developmental Biology,* 25**,** 301-327.

NISHIGAKI, T., WOOD, C. D., SHIBA, K., BABA, S. A. & DARSZON, A. 2006. Stroboscopic illumination using light-emitting diodes reduces phototoxicity in fluorescence cell imaging. *Biotechniques,* 41**,** 191-197.

NORDON, R. E., GINSBERG, S. S. & EAVES, C. J. 1997. High-resolution cell division tracking demonstrates the Flt3-ligand-dependence of human marrow CD34+CD38- cell production *in vitro*. *British Journal of Haematology,* 98**,** 528-539.

NORDON, R. E., K.-H. KO, R. ODELL, AND T. SCHROEDER 2011a. Multi-type branching models to describe cell differentiation programs. *Journal of Theoretical Biology***,** p. In press.

NORDON, R. E., KO, K., NAKAMURA, M. & ODELL, R. Year. Mathematical modeling of differentiation and proliferation from flow cytometric division tracking of blood stem cells. *In*, Nov 06-09 2005 Queensland, AUSTRIA. 406-406.

NORDON, R. E., RIEGER, M.A., KO, K-H., LORD, M.S., ODELL, R.A., SCHROEDER, T 2011b. Systematic analysis of progenitor pedigrees to quantify differentiation programs. *Submitted*.

OSHER, S. & SETHIAN, J. A. 1988. Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. *J. Comput. Phys.,* 79**,** 12-49.

OTSU, N. 1979. A threshold selection method from gray level histograms. *IEEE Trans Syst.Man Cybern,* 9**,** 62-66.

PADFIELD, D., RITTSCHER, J. & ROYSAM, B. Year. Spatio-temporal cell segmentation and tracking for automated screening. *In:* Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on, 14-17 May 2008 2008. 376-379.

PADFIELD, D., RITTSCHER, J., THOMAS, N. & ROYSAM, B. 2009. Spatio-temporal cell cycle phase analysis using level sets and fast marching methods. *Medical Image Analysis,* 13**,** 143-155.

PADFIELD, D. R., RITTSCHER, J., SEBASTIAN, T., THOMAS, N. & ROYSAM, B. Year. Spatio-temporal cell cycle analysis using 3D level set segmentation of unstained nuclei in line scan confocal fluorescence images. *In:* Biomedical Imaging: Nano to Macro, 2006. 3rd IEEE International Symposium on, 6-9 April 2006 2006. 1036-1039.

PLONSEY, R. 1999. The Biomedical Engineering Handbook. *The Biomedical Engineering Handbook, Second Edition. 2 Volume Set.* CRC Press.

PORTER, K. E. & TURNER, N. A. 2009. Cardiac fibroblasts: At the heart of myocardial remodeling. *Pharmacology & Therapeutics,* 123**,** 255-278.

POTTER, A. J. & WENER, M. H. 2005. Flow cytometric analysis of fluorescence in situ hybridization with dye dilution and DNA staining (flow-FISH-DDD) to determine telomere length dynamics in proliferating cells. *Cytometry Part A,* 68A**,** 53-58.

SAGE, D., NEUMANN, F. R., HEDIGER, F., GASSER, S. M. & UNSER, M. 2005. Automatic tracking of individual fluorescence particles: Application to the study of chromosome dynamics. *Ieee Transactions on Image Processing,* 14**,** 1372-1383.

SBALZARINI, I. F. & KOUMOUTSAKOS, P. 2005. Feature point tracking and trajectory analysis for video imaging in cell biology. *Journal of Structural Biology,* 151**,** 182-195.

SCHROEDER, T. 2008. Imaging stem-cell-driven regeneration in mammals. *Nature,* 453**,** 345-351.

SHOKEIR, A. A., HARRAZ, A. M. & EL-DIN, A. B. S. 2010. Tissue engineering and stem cells: Basic principles and applications in urology. *International Journal of Urology,* 17**,** 964-973.

SMAL, I., NIESSEN, W. & MEIJERING, E. Year. Particle Filtering for Multiple Object Tracking in Molecular Cell Biology. *In:* Nonlinear Statistical Signal Processing Workshop, 2006 IEEE, 13-15 Sept. 2006 2006. 129-132.

SMITH, J. A. M., L. 1973. Do cells cycle? *Proceedings of the National Academy of Sciences of the United States of America,* 70**,** 7.

SOUDERS, C. A., BOWERS, S. L. K. & BAUDINO, T. A. 2009. Cardiac Fibroblast The Renaissance Cell. *Circulation Research,* 105**,** 1164-1176.

STERNBERG.S 1983. Biomedical image processing. *IEEE Computer,* 16**,** 22-34.

SWEDLOW, J. R. & ELICEIRI, K. W. 2009. Open source bioimage informatics for cell biology. *Trends in Cell Biology,* 19**,** 656-660.

SWEDLOW, J. R., GOLDBERG, I. G., ELICEIRI, K. W. & CONSORTIUM, O. M. E. 2009. Bioimage Informatics for Experimental Biology. *Annual Review of Biophysics,* 38**,** 327-346.

TAKAHASHI, K. & YAMANAKA, S. 2006. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell,* 126**,** 663-676.

TATEISHI, K., TAKEHARA, N., MATSUBARA, H. & OH, H. 2008. Stemming heart failure with cardiac- or reprogrammed-stem cells. *Journal of Cellular and Molecular Medicine,* 12**,** 2217-2232.

THEISE, N. D. 2010. Stem cell plasticity: Recapping the decade, mapping the future. *Experimental Hematology,* 38**,** 529-539.

THOMAS, N. 2010. High-Content Screening: A Decade of Evolution. *Journal of Biomolecular Screening,* 15**,** 1-9.

WALLACE, P. K., TARIO, J. D., FISHER, J. L., WALLACE, S. S., ERNSTOFF, M. S. & MUIRHEAD, K. A. 2008. Tracking Antigen-Driven Responses by Flow Cytometry: Monitoring Proliferation by Dye Dilution. *Cytometry Part A,* 73A**,** 1019-1034.

WANG, X. D., POSSOZ, C. & SHERRATT, D. J. 2005. Dancing around the divisome: asymmetric chromosome segregation in Escherichia coli. *Genes & Development,* 19**,** 2367-2377.

WANG, Y. L., HAHN, K. M., MURPHY, R. F. & HORWITZ, A. F. 2006. From imaging to understanding: Frontiers in Live Cell Imaging, Bethesda, MD, April 19-21, 2006. *Journal of Cell Biology,* 174**,** 481-484.

WHITE, M. R. H., NELSON, D. E., JOHNSON, J. R., SEE, V., ELLIOTT, M., HORTON, C. A., GIBNEY, C. A., FOREMAN, B. E., NELSON, G. & SPILLER, D. G. Year. Real-time imaging of cell division and apoptosis. *In,* Mar 19-23 2005 Warwick, ENGLAND. Elsevier Ireland Ltd, 205-206.

WIRTZ, D. 2009. Particle-Tracking Microrheology of Living Cells: Principles and Applications. *Annual Review of Biophysics,* 38**,** 301-326.

WOJAKOWSKI, W. & TENDERA, M. 2010. New Concepts in Cardiac Stem Cell Therapy. *Hellenic Journal of Cardiology,* 51**,** 10-14.

WU, P.-H., AGARWAL, A., HESS, H., KHARGONEKAR, P. P. & TSENG, Y. 2010. Analysis of Video-Based Microscopic Particle Trajectories Using Kalman Filtering. *Biophysical Journal,* 98**,** 2822-2830.

XINYU, L., YIFEI, W. & YU, S. 2009. Cell Contour Tracking and Data Synchronization for Real-Time, High-Accuracy Micropipette Aspiration. *Automation Science and Engineering, IEEE Transactions on,* 6**,** 536-543.

YANG, X. D., LI, H. Q. & ZHOU, X. B. 2006. Nuclei segmentation using marker-controlled watershed, tracking using mean-shift, and Kalman filter in time-lapse microscopy. *Ieee Transactions on Circuits and Systems I-Regular Papers,* 53**,** 2405-2414.

ZEISBERG, E. M. & KALLURI, R. 2010. Origins of Cardiac Fibroblasts. *Circulation Research,* 107**,** 1304-1312.

ZELIKOFF, J. T., ATKINS, N. M. & BELMAN, S. 1986. Stimulation of cell growth and proliferation in NIH-3T3 cells by onion and garlic oils. *Cell Biology and Toxicology,* 2**,** 369-378.

ZHANG, Y. V., WHITE, B. S., SHALLOWAY, D. I. & TUMBAR, T. 2010. Stem cell dynamics in mouse hair follicles A story from cell division counting and single cell lineage tracing. *Cell Cycle,* 9**,** 1504-1510.

ZHOU, X. & WONG, S. T. C. 2006. Informatics challenges of high-throughput microscopy. *IEEE Signal Processing Magazine,* 23**,** 63-72.

ZIMMER, C., LABRUYERE, E., MEAS-YEDID, V., GUILLEN, N. & OLIVO-MARIN, J. C. 2002a. Segmentation and tracking of migrating cells in videomicroscopy with parametric active contours: A tool for cell-based drug testing. *Ieee Transactions on Medical Imaging,* 21**,** 1212-1221.

ZIMMER, C., LABRUYERE, E., MEAS-YEDID, V., GUILLEN, V. M. & OLIVO-MARIN, J. C. Year. Improving active contours for segmentation and tracking of motile cells in videomicroscopy. *In:* KASTURI, R., LAURENDEAU, D. & SUEN, C., eds., Aug 11-15 2002b Quebec City, Canada. 286-289.

ZIMMER, C. & OLIVO-MARIN, J. C. 2005. Coupled parametric active contours. *Ieee Transactions on Pattern Analysis and Machine Intelligence,* 27**,** 1838-1842.