KNOWLEDGE DISCOVERY OF CELL-CELL AND CELL-SURFACE INTERACTIONS

A Thesis
Presented to
The Academic Faculty

by

Jing Su

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
Department of Biomedical Engineering

Georgia Institute of Technology
April 2008

KNOWLEDGE DISCOVERY OF CELL-CELL AND CELL-SURFACE INTERACTIONS

Approved by:

J. Carson Meredith, Committee Chair
Department of Biomedical Engineering
Georgia Institute of Technology

J. Carson Meredith, Advisor
Department of Biomedical Engineering
Georgia Institute of Technology

Larry McIntire
Department of Biomedical Engineering
Georgia Institute of Technology

Zorina Galis
Department of Biomedical Engineering
Georgia Institute of Technology

Andrés García
Department of Biomedical Engineering
Georgia Institute of Technology

Mark Prausnitz
Department of Biomedical Engineering
Georgia Institute of Technology

Date Approved: April 2008

Dedicate to my wife, my parents, and my whole family, especially to my twenty-month-old daughter Zhanqi, who explores routine life as an adventurer, shares with me her novel points of view of the world, inspires me by describing ordinary events in novel metrics, and renews me with new knowledge she discovered everyday.

ACKNOWLEDGEMENTS

First of all, I would like to thank Dr. Meredith for his patience, inspiring guides and continuous support. Also I would like to thank Dr. Galis and Dr. Lessner, who led me into the world of cell biology, and Dr. McIntire and Dr. Eskin, for their intelligent instruction, and Dr. García for his generous help with fluorescent microscopy. I would like to thank Dr. Mark Prausnitz as well, for dedicating his valuable time to serve on my committee.

I would like to give my special thanks to Dr. Sung for his help and support during our collaboration for more than two years. I would like to gratefully acknowledge our research group, particularly Dr. Chattopadhyay and Mr. Zapata, for the inspiring discussion and generous help. I wish to thank to everyone who inspires me, helps me and supports me. I am grateful for that. Thank you all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS OR ABBREVIATIONS

A	Cell class: all cells.
$Adj\,R^2$	adjusted coefficient of determination.
A_{image}	the area of the observing window used in $MVNCDF_{inhibit}\left(A_{image}\right)$.
$ANOVA$	Analysis of variance.
$APTES$	3-aminopropyl-triethoxysilane.
$\bar{L}$	Arithmetic mean of PCL size. $\bar{L} = \frac{SumL}{N_{PCL}}$.
$\bar{S}$	Arithmetic mean of PCL area.
$\bar{S}_S$	Area-weighted mean area of PCL islands. $\bar{S}_S = \frac{\sum_i S_i}{\sum_i S_i}$.
bin_{dist}	a set of distance bins defined on $Scale_{dist}$ $bin_{dist} = \{[d_0, d_1), [d_1, d_2), \ldots, [d_{N-1}, d_N]\}$.
BrdU	5-bromo-2-ϕ-deoxyuridine.
C	PCL islands.
ΔTH	The difference of threshold between the last two iterations.
$\Delta_{PC\|AC}$	The difference between $f_{PC}(\hat{d}, \widehat{dia})$ and $f_{PC}(\hat{d}, \widehat{dia})$.
dia_i	Size of the i^{th} PCL island $dia_i = \sqrt{S_i \frac{4}{\pi}}$.
dis	Cell-to-cell distance between two cells, measured between the centroids.
dis_{AA}	Cell-to-cell distance between two cells.
dis_{AC}	Cell-to-PCL-island distance between a cell and a PCL island.
dis'_{AC}	Cell-to-PCL-island distance between the centroid of a cell and the edge of a PCL island.
dis_{AP}	Cell-to-cell distance between a cell and a proliferated cell.
dis_{AR}	Cell-to-cell distance between a cell and a non-proliferated cell.

$dis_{div,i,j}$ the distance from the the daughter cell of the i^{th} cell to the i^{th} cell along the j^{th} candidate direction in simulation.

dis_{excl} the excluded distance, the shortest allowable distance between two cells in simulation.

dis_{PA} Cell-to-cell distance between a proliferated cell and another cell.

dis_{PC} Cell-to-PCL-island distance between a proliferated cell and a PCL island.

dis'_{PC} Cell-to-PCL-island distance between the centroid of a proliferated cell and a the edge of PCL island.

dis_{PP} Cell-to-cell distance between two proliferated cells.

dis_{PR} Cell-to-cell distance between a proliferated cell and a non-proliferated cell.

$DispS$ Dispersion of PCL phase area. $DispS = \frac{\bar{S}}{\bar{S}_S}$.

dis_{RA} Cell-to-cell distance between a non-proliferated cell and another cell.

dis_{RC} Cell-to-PCL-island distance between a non-proliferated cell and a PCL island.

dis_{RP} Cell-to-cell distance between a non-proliferated cell and a proliferated cell.

dis_{RR} Cell-to-cell distance between two non-proliferated cells.

DMEM Dulbecco's Modification of Eagle's Medium.

DPBS Dulbecco's Phosphate-Buffered Saline.

f_{AA} the frequency functions, or cell environments, for cell-to-cell distances AA.

$f_{AC}(\hat{d}, \widehat{dia})$ the PCL island background of the A class cells observed from the grid window $(\hat{d}, \widehat{dia})$.

FDA Food and Drug Administration.

f_{PA} the frequency functions, or cell environments, for cell-to-cell distances PA.

$f_{PC}(\hat{d}, \widehat{dia})$ the PCL island background of the P class cells observed from the grid window $(\hat{d}, \widehat{dia})$.

$f_{PR}(\hat{d})$	the R type cell background of the P class cells observed from the window $\hat{d}$.
f_{std}	the standard frequency distribution.
$grid$	a set of distance bins defined on $Scale_{dist}$ $bin_{dist} = \{[d_0, d_1), [d_1, d_2), \ldots, [d_{N-1}, d_N]\}$.
$\hat{d}$	centroid set for bin_{dist} $\hat{d} = \{\hat{d}_1, \hat{d}_1, \ldots, \hat{d}_N\}$.
$\hat{d}_i$	centroid of each interval in bin_{dist} for bin_{dist}.
I_{avg}	Average of signal intensities I_i from region of a given cell nucleus. $I_{avg} = \frac{\sum_i I_i}{\sum_i 1}$ for $i \vert i \in$ region of given cell nucleus.
I_{sum}	Summation of signal intensities I_i from region of a given cell nucleus. $I_{sum} = \sum_i I_i$ for $i \vert i \in$ region of given cell nucleus.
$LCFA$	local cell-feature analysis.
LCST	low critical solution temperature.
L_i	Size of the i^{th} PCL island.
MC	the *Monte Carlo* method.
$MCWS$	Marker-controlled watershed segmentation.
μ_{avoid}	the mean as a parameter of $N_{avoid}\,(dis \vert \mu_{avoid}, \sigma_{avoid})$.
$MVNCDF_{inhibit}\left(A_{image}\right)$	the cumulative distribution function of $MVN_{inhibit}\,(pos \vert pos_o, \Sigma_{inhibit})$ with the observing window of area A_{image}.
$MVN_{inhibit}\,(pos \vert pos_o, \Sigma_{inhibit})$	an assumed, multivariate normal distribution based *Monte Carlo* possibility function used in simulation for addressing the contact inhibition effects.
N	resolution of $Scale_{dist}$.
$N_{avoid}\,(dis \vert \mu_{avoid}, \sigma_{avoid})$	an assumed, normal distribution based *Monte Carlo* possibility function used in simulation for addressiing the self avoiding effects.
N_{Cell}	Number of cell per image.

$N_{div}\,(dis\|\mu_{div}, \sigma_{div})$	an assumed, normal distribution based *Monte Carlo* possibility function used in simulation for addressing the effects of spatial preference of cell division .
N_{PCL}	Number of PCL islands.
n_{Pk}	The number of P-class in the the k^{th} image.
$N_{PR}\|_{\hat{d}}$	the number of elements of set PR that are in the interval $[d_{i-1}, di)$ (which is centered at $\hat{d}_i$).
$N_{PR}(\hat{d})$	an N-bin histogram generated from PR which is sorted by $\hat{d}$ $N_{PR}(\hat{d}) = h_N(PR\|_{bin_{dist}}) = h_N(PR\|_{\hat{d}})$.
N_{Prolif}	Number of proliferating cells per image.
n_{Rk}	The number of R-class in the the k^{th} image.
$\overrightarrow{pos_i}$	the position of the centroid of the i^{th} cell.
P	Cell class: proliferating cells.
p_{avoid}	possibility that a cell will be removed in simulation according to self avoiding effects.
PBS	Phosphate-Buffered Saline.
PC	the set of all distance PC_{ijk} in all images $PC = \bigcup_{i,j,k} PC_{ijk}$.
PCA	Principal component analysis.
PC_{ijk}	the distance between the centroids of the nucleus of the i^{th} P-cell and the j^{th} PCL island in the k^{th} image.
PCL	poly(ε-caprolactone).
$PerS$	Percentage of PCL phase area. $PerS = \dfrac{SumS}{S_{Image}}$.
ϕ	weight composition of PCL in a polymer blends.
ϕ_o	initial and the lowest weight composition of PCL in a polymer blends in a composition gradient.
ϕ_t	final and the highest weight composition of PCL in a polymer blends in a composition gradient.
$p_{inhibit}$	possibility that a cell will not proliferate in simulation according to contact inhibition effects.

$p_{inhibit,div,i,j}$	the strength of the contact inhibiton of cell division if the the daughter cell of the i^{th} cell is set at the candidate location $pos_{div,i,j}$ along the j^{th} candidate direction in simulation.
PLGA	poly(DL-lactide-co-glycolide).
PO	posterior odds.
$pos_{div,i,j}$	the candidate position for the daughter cell of the i^{th} cell along the j^{th} candidate direction in simulation.
PR	the set of all distance PR_{ijk} in all images $PR = \bigcup_k PR_k$.
PR_{ijk}	the distance between the centroids of the nuclei of the i^{th} P-cell and the j^{th} R-cell in the k^{th} image.
PR_k	the set of all distance PR_{ijk} in the k^{th} image $PR_k = \{PR_{ijk} \mid i = 1, 2, \ldots, n_{Pk};\ j = 1, 2, \ldots, n_{Pk}\}$.
$Prol$	Proliferation ratio. $Prol = \frac{N_{Prolif}}{N_{Cell}}$.
$Prol_o$	the probability of proliferation without effects of contact inhibitionin in the simulation.
R	Cell class: non-proliferating cells.
r	the dice value, a random number generated from a uniform distribution and used to compare with possibility function used in the *Metropolis Monte Carlo algorithm*.
R^2	coefficient of determination.
$R_{\hat{d}_i}$	the feature variable of R type cell background.
r_{div}	the division ratio used in the simulation.
$RMSE$	root mean squared error.
S_i	Area of the i^{th} PCL island.
σ_{avoid}	the variance as a parameter of $N_{avoid}\ (dis \mid \mu_{avoid}, \sigma_{avoid})$.
$\sigma_{inhibit}$	the variance as a parameter in $\Sigma_{inhibit}$ in $MVN_{inhibit}$.
$SMSP$	Surface Micro-structure Pattern.
SSE	sum of squared error.
$SumL$	Sum of the PCL island Sizes in an images. $SumL = \sum_i L_i$.

$SumS$	Total PCL area in an image. $SumS = \sum_i S_i$.
T	temperature.
TCPS	tissue culture polystyrene.
TH	Threshold used for segmentation.
$theta_j$	the j^{the} candidate direction based on $theta_o$ and used in simulation for determining the cell dividing direction.
$theta_o$	the base direction used in simulation for determining the cell dividing direction.
TH_{final}	Final threshold for classifying cell proliferation status by BrdU staining.
$TH_{final,avg}$	Final threshold for classifying cell proliferation status by BrdU staining using I_{avg}.
$TH_{final,sum}$	Final threshold for classifying cell proliferation status by BrdU staining using I_{sum}.
TH_o	Threshold determined by original iterative selection method.
$VAIS$	Variation-adjusted iterative selection.

GLOSSARY

ANOVA Analysis of variance, as known as Fisher's analysis of variance. The null hypothesizes of ANOVA is that the observed variance is due to a group of explanatory factors. ANOVA is a collection of statistical methods that partitions the overall variance into components which corresponding to the hypothesized explanatory factors and evaluates the importance of the explanatory factors by hypothesis testing., p. 58.

local cell-feature analysis a novel data analysis method that uses the viewpoint of individual cells, describes local cell environment with local metrics, classifies cells according to cell behaviors, and analyzes the influence of surface features on cell behaviors., p. 6.

morphological closing the image processing that a dilation followed by an erosion with the same structuring element is called a morphological closing process., p. 24.

morphological opening the image processing that an erosion followed by a dilation with the same structuring element is called a morphological opening process., p. 23.

surface micro-structure pattern specific and observable three-dimensional geometric characteristics which can be represented by motifs and distributions of such motifs., p. 1.

LIST OF LISTINGS

SUMMARY

High-throughput cell culture is an emerging technology that shows promise as a tool for research in tissue engineering, drug discovery, and medical diagnostics. An important, but overlooked, challenge is the integration of experimental methods with information processing suitable for handling large databases of cell-cell and cell-substrate interactions. In this work the traditional global descriptions of cell behaviors and surface characteristics was shown insufficient for investigating short-distance cell-to-cell and cell-to-surface interactions. Traditional summary metrics cannot distinguish information of cell near neighborhood from the average, global features, thus often is not suitable for studying distance-sensitive cell behaviors.

The problem of traditional summary metrics was addressed by introducing *individual-cell based local metrics* that emphasize cell local environment. An *individual-cell based local data analysis* method was established. Contact inhibition of cell proliferation was used as a benchmark for the effectiveness of the local metrics and the method. Where global, summary metrics were unsuccessful, the local metrics successfully and quantitatively distinguished the contact inhibition effects of MC3T3-E1 cells on *PLGA*, *PCL*, and *TCPS* surfaces.

In order to test the new metrics and analysis method in detail, a model of cell contact inhibition was proposed. *Monte Carlo* simulation was performed for validating the individual-cell based local data analysis method as well as the cell model itself. The simulation results well matched with the experimental observations. The parameters used in the cell model provided new descriptions of both cell behaviors and surface characteristics.

Based on the viewpoint of individual cells, the local metrics and local data analysis

method were extended to the investigation of cell-surface interactions, and a new high-throughput screening and knowledge discovery method on combinatorial libraries, *local cell-feature analysis*, was developed. *PLGA*/*PCL* combinatorial libraries were used as a prototype and a "shaper and holder" phenomenon involving MC3T3-E1 cells interacting with PCL islands was discovered.

In summary, the viewpoint of individual cells casts new light on the study of cell-cell and cell-surface interactions and represents a *novel methodology* for developing new data analysis and knowledge discovery methods.

The results of contact inhibition study and the "shaper and holder" model provide new knowledge, while the local data analysis method as well as the cell model of contact inhibition suggested novel approaches to study cell-cell and cell-surface interactions.

CHAPTER I

INTRODUCTION

Cell-cell and cell-surface interactions are critical in life science, tissue engineering and medicine. Among them, cell contact inhibition of proliferation and *surface micro-structure pattern*s ($SMSP$) are of special importance. Both have been investigated for decades, but the underlying mechanisms are still not well known.

In the following sections, the current state of art is summarized, main obstacles are analyzed, and a different point of view, the *'individual-cell viewpoint'*, is proposed for phenomena description, data analysis, modeling and simulation. In this work a novel methodology based on the individual-cell viewpoint is developed and examined for the sensitivity of both cell contact inhibition of proliferation and cell-$SMSP$ interactions.

1.1 Cell-cell Interactions: Cell Contact

Cell-cell recognition is critical to a wide range of problems in biology and medicine. The development of biochemical assay for protein binding and transcription events associated with cell-cell recognition has seen many advances. However, little has been achieved regarding algorithms for the detection of the effects of cell-cell interactions from microscopy images. Such methods might offer complimentary benefits to biochemical assay, due to the relative ease of collecting microscopy data from cell cultures, as well as from tissues. Two important cell-cell recognition effects of adhesion dependent cells, contact inhibition of cell migration[3, 1] and cell growth (Ambrose and Sheppley's personal communication in 1964[84]) have been known for more than four decades. These phenomena are related to the mutation, development, invasion, and metastasis of cancer cells[2], and play an important role in cell differentiation and tissue development.

Furthermore, understanding cell-cell interactions, especially cell contact effects, is critical for studying the effects of other factors on cell behaviors. For example, the effects of cell-surface interactions on cell proliferation and differentiation are often overlapped and entangled so much with the effects of cell contact that it is difficult to study other effects quantitatively without a valid mathematical model of cell contact phenomenon.

However, the underlying mechanisms relating cell contact, cell proliferation, migration, and differentiation are not completely understood. Cell adhesion molecules have been shown to play a major role[16]. Particularly, cadherins, e.g. VE-cadherin [20] and N-cadherin [61, 62, 63, 90], have been found to initiate contact inhibition in some cell lines. For example, the N-cadherin-triggered signaling pathway initiates upon cell-to-cell contact a cascade of reactions that lead to the nucleus and ultimately modulates cell cycle patterns[69, 26, 25] Some cancer cells show simultaneously reduced cadherin expression and loss of contact inhibition of growth, which in turn promote cancer metastasis[13, 24]. Conversely, it has also been shown that cell-cell contact under specific conditions can promote cell proliferation, known as "density-dependent" contact stimulation of cell proliferation[18, 17, 48, 77].

The ambiguity of contact inhibition mechanisms may be due in part to the traditional methods of data presentation and analysis, which is discussed in details in the next section.

1.2 Data Analysis Methods for Cell Contact

The effects of cell density on proliferation have been studied mainly as relationships between global descriptions, such as cell density, proliferation rate, and protein expression level[84, 19, 77, 76]. These general, summary-statistic descriptions may not illuminate all of the information available from the data for addressing cell contact phenomena. Cell-to-cell contacts are local interactions and are very sensitive to short-range cell-to-cell distance. When summary indicators of average values (cell density) are used, all cell-to-cell

distances are treated equally. Critical information pertaining to nearest-neighbor cell-to-cell distances is "diluted" by the many other cell-to-cell distances in the data set. These other longer-range distances are not expected to have an effect and consequently introduce noise into the data. Therefore, the global effect of cell contact on cell behaviors can only be observed between cell cultures with dramatically different seeding density. Furthermore, the stochastic characteristics of cell behaviors and the limited sample space from traditional experiment design and data acquisition methods results in significant noise so that only the major trends of cell response to neighbors is distinguishable (e.g., [90]).

In order to address the "diluting" effect mentioned above, artificially designed cell attachable patterns were used to specifically and directly study the contact stimulation effects of growth of a pair of cells by decoupling the effects of cell-cell contact from others[77] and consistent conclusions were achieved. However, this method was limited by:

1. the difficulty in mimicking high cell density necessary for studying contact inhibition,
2. the limited sample space, and
3. the cell attachable area being artificially designed (design itself may affect cell behaviors via cell-surface contact interactions).

In this work we introduce a novel localized description system that emphasizes individual-cell properties (like contact inhibition), but allows sampling from a large number of cells. The method is based upon defining local cell feature metrics, which are histograms of local cell properties. The use of these local variables expands the sample space considerably and allows separation of short- and long-range effects. We show how the local cell feature histograms are then incorporated into a Bayesian model. The new method and model are examined quantitatively and compared with traditional summary approaches in a study of contact inhibition of cell osteoblast proliferation. The noise due to the random features of cell behavior was significantly suppressed as a result of the mass sample space.

1.3 Modeling and Simulation of Contact Inhibition

Quantitative models of contact inhibition fall into two main groups: differential equation models and *Monte Carlo* simulation models. For example, cell cycle dynamics was studied as age-structured population models with classic partial differential equations[72, 35, 33, 34, 60] and was simplified into differential delay equations[74, 75, 67, 9, 8]. Bonhoeffer and co-workers studied cell proliferation kinetics with partial differential equation models[11]. Partial differential equations models are challenging when studying complicated systems, and, practically, compromising assumptions are often necessary to simplify the original partial differential equations. Individual-cell based *Monte Carlo* simulation approaches, initiated by Loeffler, Drasdo and co-workers[64, 38], have been shown promising for mimicking cell-cell and cell-surface interactions[41]. Based on different theoretical models, such approaches have been used for studying the growth dynamics of epidermal stem cells[64], growth and folding of one-layered tissues[39, 40, 41], development of tumors[37, 41], and epithelial cell populations[46].

The data description methods used in the previous simulations were mainly general and global, e.g., average density, total number of cells, and cluster size. In contrast, *Monte Carlo* simulations are based on local, close-range interactions of individual cells. The major reason for using only global (summary statistic) descriptors was the lack of individual-cell based (local statistics) experimental data for validating simulation results. In a previous study we developed a set of robust individual cell-based local variables for cell-cell interactions. Using an experimental system of osteoblasts on various polymer surfaces, this study showed that local metrics have significantly higher sensitivity to distance-dependent cell behaviors, relative to global variables. It follows naturally that local metrics may be a useful choice for comparing between experiments and models of cell interaction phenomena.

Monte Carlo simulations of cell contact inhibition are important for validating the new data representations, like local metrics, as well as for testing candidate hypotheses and mechanisms, and designing experiments. The explanations of data analysis results can be

tested by artificial assignment of rules representing supposed physical meaning for *Monte Carlo* simulations. Candidate mechanisms can be screened by comparing simulation results with experimental data. The individual-cell based *Monte Carlo* simulation and local metrics for data analysis form a consistent system for modeling, testing, and validating cell behaviors from the point of view of individual cells.

1.4 Screening Surface Micro-structure Patterns According to Cell Response

Both chemical and physical micro- and nanometer surface features affect cell attachment survival, differentiation and functions on polymeric biomaterials[79, 36, 49, 52, 29, 94, 89, 80, 66, 12, 32, 65, 47, 59]. In general cell-surface interactions, integrin binding and downstream events play a central role in regulating cell responses to surface chemical properties[47, 53, 54, 55].

Compared to homogeneous surfaces, patterned distributions of chemically distinct domains provide new dimensions for flexible and reliable control of cell responses. The promise of this technology is "fine-tuned" cell behavior by using engineered $SMSP$. This ability would provide useful advances in tissue engineering constructs, diagnostic devices, and cell culture surfaces.

One significant limitation in the development of patterned biomaterial surfaces is the challenge inherent to experiments aimed at understanding cell-material interactions. In particular, the mechanisms by which material physical surface features interact with cells remains enigmatic, although important aspects have been delineated. These include the effect of local mechanical stresses on cell shape[44, 21, 22, 23]. nuclear shape[36], formation of focal adhesion[27, 28, 45], internal organization of the cytoskeleton[29, 51], signaling[50], and selection of gene programs[27, 51].

Lithographically-patterned surfaces are very useful for controlled experiments varying a single parameter. However, natural tissue surfaces do not typically exhibit these regular, ordered features. Self-organization phenomena that lead to tissue structures usually

produce randomly-distributed or semi-regular patterns. In addition, natural self-organizing phenomena (crystallization, phase separation) are not inherently limited to two-dimensions. However, such pattern-generating phenomena are inherently complex and depend on multiple parameters. To address this problem, combinatorial libraries of phase-separated biodegradable polymers have been used for high-throughput assay of both osteoblast[71, 85, 91] and aortic smooth muscle cells[86]. Others have developed experimental strategies for combinatorial synthesis of polymer libraries that investigate chemical differences based on monomer content[15, 14]. A "microarray"-type technique for investigating cell-polymer interactions has been reported also[4, 6]. One major conclusion of this previous experimental work is the ongoing and critical need for an effective informatics platform for analyzing the large multi-variate data sets.

Traditional analysis of data in materials science and cell biology relies upon comparisons of three summary statistics. When summary statistics are taken from combinatorial data sets, modeling and data mining techniques have been applied to discover correlations in cell-material interactions[56, 57, 83, 82, 88, 92]. Two important considerations not easily managed in global (summary) analysis are the inherently local, distance-dependent nature of cell-material interactions and the multivariate, combinatorial relationships between patterned surface chemistry, physical properties and cell responses.

In this work, from the point of view of individual cells, we applied the *localized description system* on cell-$SMSP$ interactions, extended the one-dimensional data analysis method mentioned in Section 1.1 to a novel two-dimensional screening method, and established a simple *local cell-feature analysis* ($LCFA$) method, based on histograms of micro-structural and cell images. We demonstrate that $LCFA$ is ideally suited for screening combinatorial cell-material data sets, where large amounts of data from a broad sample space often "mask" the detection of effects. $LCFA$ overcomes this limitation by enhancing sensitivity to each cell's "neighborhood", composed of other cells and microstructures within a range ($< 100\mu m$) most likely to affect that cell. Local feature histograms are also

directly applicable to Bayesian analysis, a powerful data-mining method used widely in multi-variate inference problems. With few exceptions, the local histogram approach has not been routinely applied in the analysis of cell-material interactions. Frequency-based metrics for surface roughness have been shown to be significantly better at explaining surface topography effects on cell functions[7]. However, the local feature analysis proposed here incorporates not only material surface properties, but also describes quantitatively the local relationships between surface features and cell position.

CHAPTER II

SETUP OF EXPERIMENTAL SYSTEM

A comprehensive experimental system has been established for studying contact inhibition of cell proliferation and for screening the effects of $SMSP$ on cell spreading and proliferation.

2.1 Materials

Three criteria are used for screening and choosing material models for surface micro-structure pattern library construction and cell responses testing:

1. Controllable phase separation. Phase separation driven surface micro-structure pattern construction, as most surface phenomena, is very sensitive to background noise such as the slight change of temperature, air flow, humidity, and so on. Therefore, system robustness becomes an important consideration for selecting candidate polymer blends.

2. Convenient pattern characterization. Since multi-channel microscopy is required for simultaneously examining both surface micro-structure patterns and cell behaviors, as model systems, surface micro-structure pattern characterization should be convenient to measure *in situ*, and be independent to bio-assay.

3. Potential application. The chosen polymer system should be potentially useful in tissue engineering, bioscience, or medical fields.

Since the compatibility of poly (DL-lactic-glycolic acid) (*PLGA*) and poly (ε- caprolactone) (*PCL*) is moderate for phase separation process, and the temperature sensitivity of this blend is appropriate for controlling the size of surface micro-structure patterns, the

PLGA/*PCL* blend system is robust during preparation. The crystallinity of *PCL* phase, in contrast to the amorphous *PLGA* phase, allows using crossed polarized microscopy for $SMSP$ characterization. Since most *in situ* bio-assay methods employ fluorescent microscopy, the contrast of crystallinity of *PLGA* phase and *PCL* phase minimizes the potential interference between pattern characterization and bio-assays. More important, both *PLGA* and *PCL* are approved by the Food and Drug Administration (FDA) for biomedical application, and are widely used biomaterials in tissue engineering. Therefore, the *PLGA*/*PCL* system was selected as the material model for the proposed work.

The tissue culture polystyrene (*TCPS*, Corning®) surface is used as a control surface for *PLGA* and *PCL* in the cell contact inhibition study since $TCPS$ is a widely used cell culture surface for adhesion dependent cell lines, and all cells used in this work were passaged on and harvested from $TCPS$ surfaces. The $TCPS$ surface is modified from polystyrene surface by corona discharge treatment, which grafted highly energetic oxygen ions into the polystyrene. Therefore, comparing with original polystyrene surface, $TCPS$ is negatively charged and becomes more hydrophilic, which improves cell attachment and growth[5, 78].

2.2 PLGA and PCL Polymer Surfaces Preparation

Poly (DL-lactic-glycolic acid) (*PLGA*, block copolymer, 50:50 ratio of poly-lactic acid and poly-glycolic acid, 40,000 75,000 Da) and poly (ε-caprolactone) (*PCL*, 114,000 Da, Mw/Mn = 1.43) were obtained from Sigma Aldrich, St Louis, MO. *PLGA* and *PCL*, respectively, were dissolved in chloroform to 8% and 5% mass and spin coated on silicon chips ($22mm \times 22mm$ to $25mm \times 25mm$). To provide adhesion of these polymers to the silicon during cell culture, silicon chips were pretreated in Pirañha etch[1] at 60 °C for half an hour, followed by one minute etching in a hydrofluoric acid bath and a final rinse in DI water (filtered at $0.2\mu m$). Hydrofluoric acid treatment striped oxidized silicon and hence the resultant silicon surface was hydrophobic, which stabilized *PLGA* and *PCL* attachment.

[1]The Piraña etch used in this work is a 3:1 mixture of 96% sulfuric acid and 10% hydrogen peroxide.

Without such treatment, polymer films would peel off from the substrate during cell culture. Since the hydrofluoric acid treated surface is sensitive to oxygen, chips were prepared freshly, stored in water and coated in half an hour after hydrofluoric acid etching.

2.3 Combinatorial Library Preparation

Silicon chips were cleaned with a Pirañha etch at 60 °C for half an hour, followed by $APTES$ (3-aminopropyl-triethoxysilane, Sigma Aldrich, St Louis, MO) modification (4% *APTES* in anhydrous hexane) to obtain a hydrophobic surface. $APTES$ modification instead of hydrofluoric acid treatment was used because library preparation is time consuming, and surfaces prepared by the hydrofluoric acid etching are time sensitive.

The polymer system, $PLGA/PCL$ polymer blend, was selected as the model system for biomaterials. Based on this polymer blend, composition-annealing temperature (ϕ/T) two-dimensional libraries were prepared[70].

2.3.1 Solvent Casting: Establishment of Composition Gradient

Solvent casting procedures were used to form controllable composition gradients on libraries. As shown in Figure 1, initially an given amount of polymer *PLGA* was infused into the vial. Two computer controlled syringes *A* (for infusing polymer *PCL* into the vial) and *B* (for withdrawing the mixed solution from the vial) operated synchronously to create a composition gradient of *PCL* (from $\phi_o = 0$ to a given final composition ϕ_t) as a function of time. This gradient solution was sampled continuously into syringe S so that the composition gradient with respect to time in the vial was transformed to a similar composition gradient along the syringe needle length(shown in Figure 1 (a)). Afterward the composition gradient in the needle was applied on a suitable substrate (a silicon chip or a glass microscope slide). Now a stripe with the same gradient was formed on the substrate (Figure 1 (b)). At last, the stripe was spread by a blade on the silicon substrate and a film with the original composition gradient of the mixture in the vial with respect to sampling time was re-produced on the silicon chip along one direction (Figure 1 (c)).

Maintaining the original composition gradient formed in the vial during the procedure was crucial for achieving polymer films with repeatable and controllable *PCL* composition gradient. Due to the low diffusivities of polymer solutions (about $10 \sim 8cm^2/S$), the diffusion driving by composition gradient during processing was ignorable. In order to minimize the human-caused error, the coating process was automatically performed and controlled by LabView™programs. Quality of library composition gradient was validated by *FTIR* examination.

2.3.2 Annealing and Phase Separation: Establish Temperature Gradient

An annealing bed with temperature gradient, as Figure 2 shows, was utilized to achieve an annealing temperature gradient on libraries. Libraries with composition gradient was mounted on this annealing bed, with the direction of their composition gradient perpendicular to the direction of the annealing bed T gradient. Driven by crystallization and *LCST* (low critical solution temperature) phase separation mechanisms, *PLGA* and *PCL* separated to form surface micro-phases as a function of distinct combinations of T and ϕ. Again, during annealing process, diffusion of melting polymer resulted from composition gradient could be ignored due to polymer's high viscosity and low diffusivity as well as the thin film thickness. Annealed libraries were quenched to room temperature. Annealing temperature gradient on annealing bed and library surface was controlled and monitored by thermocouples.

After solvent casting and annealing, all combination of composition and annealing temperature in the set range were represented on *APTES* treated silicon chips. Two-dimensional combinatorial libraries of composition and annealing temperature were therefore built up.

Surface micro-structure patterns were established during library annealing. Blend composition, annealing temperature, and annealing time are main determinant factors for controlling the size, shape and distribution of *PLGA* and *PCL* phases. Since all combinations

of annealing temperature and blend composition within the screening ranges were examined and all resultant surface *PCL* phase patterns were established, these chips are called *combinatorial surface micro-structure pattern libraries*.

Hence, accompanying with the construction of ϕ/T libraries, surface micro-structure pattern libraries were also established. Such libraries can be described with micro-structure phase patterns (crossed polarized microscopy, Figure 3), surface topology (atomic force microscopy, Figure 4), and average pattern size(Figure 5).

Annealing temperature (screening range from 93 to 130 °C) and *PCL* composition (screening range from $\phi_o = 0$ to $\phi_t = 0.7$ mass fraction) gradients were generated in this work.

2.4 Cell Culture and Assay

Established from newborn mouse calvaria, the MC3T3-E1 cell line has been shown capable of differentiating into osteoblast and osteocytes *in vitro*. MC3T3-E1 cells have been shown to exhibit specific bone related protein expression patterns, under different developmental stages, similar to primary mouse calvaria cells. This cell line is thus a suitable *in vitro* model for investigating cell behaviors, regulations of such behaviors, and underlying mechanisms in different osteoblast maturation stages. Since the original MC3T3-E1 cell line has been found phenotypically heterogeneous with regard to cell differentiation, more homogeneous subclones have been established. The purity of cell line is crucial in this study, since the model system itself is already very complicated and the methods developed in this study are sensitive enough to clearly recognize the phenotypical heterogeneity introduced by impure cell lines. Therefore, in this work, MC3T3-E1 subclone 4 (from ATCC, VA), which shows homogeneous capabilities of osteogenesis both *in vitro* and *in vivo*, was chosen in order to minimize variations due to phenotypical heterogeneity.

Cell proliferation was assayed by BrdU (5-bromo-2-ϕ-deoxyuridine) immunohistochemistry. Briefly, *PLGA*- and *PCL*-coated wafers were mounted into Costar© 6-Well

TC-Treated Microplates (Corning, NY). The tissue culture treated polystyrene ($TCPS$) surfaces of the microplate wells were used as controls. After sterilization (70% ethanol solution, 30min), MC3T3-E1 cells (passage 6) were seeded onto the coated wafers at $4 \times 10^4 cells/cm^2$. This relatively high seeding density was selected to highlight the effects of contact inhibition of cell growth and other space-sensitive cell-to-cell interactions. We were using such seeding density for five reasons:

1. In order to emphasize cell contact inhibition phenomenon, researchers working on cell contact effects often use high cell density. For example, to highlight the cell-cell contact effects on the calcium response of MC3T3-E1 cells, an even higher cell seeding density, $2.34 \times 10^5 cells/cm^2$, was used. (For example, the seeding density of $1.17 \times 10^4 cells/cm^2$ was used as sparse condition and $2.34 \times 10^5 cells/cm^2$ was called a dense condition[58].

2. Cell behaviors at 50% to 90% of confluence are of specific interest. During this transient stage cells have finished a few cycles, have stabilized, and begin to communicate significantly via cell-cell contacts and soluble factors to achieve a more homogenous phenotype. The cell density observed after BrdU incorporation was about $2 \times 10^4 \sim 3 \times 10^4 cells/cm^2$.

3. We mounted polymer films on silicon chips, which were about 1mm higher than the inner bottom surface of the culture plate. During seeding, cell density on the exposed culture plate bottom surface was much higher than that on the top of the silicon chip. Therefore, the actual cell density on the silicon after seeding is much lower than the seeding density; In order to achieve a more uniform cell seeding, we used a shaker to mildly shake the culture plate while seeding cells, which significantly improved the seeding homogeneity. The fluid dynamic effect probably drove more cells to distribute on the bottom of the exposed culture dish surface instead of the top of the silicon chip.

After seeding, microplates were shaken mildly for 10min on a shaker to obtain uniform seeding. Cells were cultured in DMEM (Cellgro© Dulbecco's Modification of Eagle's Medium, Mediatech, Inc., VA) with 10% fetal bovine serum (ATCC© SCRC-1002, ATCC, VA), L-glutamine and streptomycin at 37°C in a humidified 5% CO_2 atmosphere. Five hours after seeding, surfaces were washed with Dulbecco's Phosphate-Buffered Saline (DPBS, with Ca and Mg) to remove non-attached cells, and fresh culture medium was then added. Eighteen hours after seeding, 2mM BrdU (5-bromo-2-ϕ-deoxyuridine, Sigma, MO) in Phosphate-Buffered Saline (PBS) was added to the culture medium to reach a final concentration of $20\mu m$. After 6 hour BrdU incorporation, cells were fixed with 3.6% paraformaldehyde and BrdU incorporation was assayed by immunohistochemistry (primary antibody: mouse anti-BrdU, BD Biosciences, CA; secondary antibody: goat anti-mouse, Rhodamine conjugated, Rockland Immunochemicals, Inc., PA; counter staining: Hoechst 33342, Molecular Probes, Invitrogen™Corporation, CA).

F-actin was stained by fluorescent-labeled phalloidin (Alexa Fluor® 488 phalloidin or rhodamine phalloidin, Molecular Probes, Invitrogen™Corporation, CA) for cell shape and cytoskeleton structure. Vinculin was selected as the marker of cell focal adhesion and was stained by immunohistochemistry (secondary antibody: goat anti-mouse, Rhodamine conjugated, Rockland Immunochemicals, Inc., PA)

2.5 Image Acquisition

A preliminary version of a multi-channel microscopy image acquisition (IMAQ) system was established based on an customized Olympus BX41 microscope. This IMAQ system is composed of three parts: a robotic translation stage, a multi-channel microscope, and an image acquisition system (Figure 6). The three hardware subsystems are controlled and integrated by a software platform based on *Microsoft™Visual Studio™* [2].

Most commercially available microscopes do not have both crossed polarized microscopy

[2]The first prototype was developed based on *Visual Basic™6.0*, the current prototype was based on *Visual C#.NET 2003*

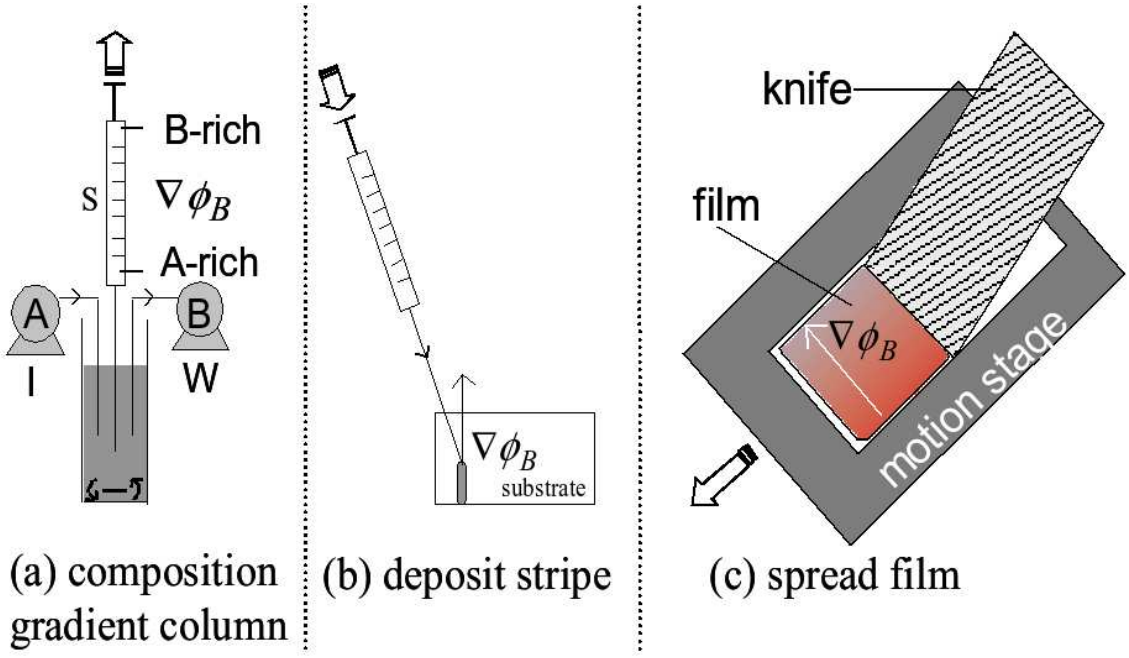
and fluorescent microscopy, because the strong UV light intensity required by fluorescent labels will also bleach common polarizers. To address this problem, a UV cut-off filter and a series of neutral density filters[3] were used to protect polarizers from strong UV light. Then the channel of crossed polarized light was combined with conventional fluorescent channels[4], which allowed features of surface micro-structure patterns and fluorescent-labeled cell response markers being examined *in situ* simultaneously.

An Optronics® MicroFire™ monochrome digital microscope camera is used for IMAQ. To assort with the monochrome camera, all fluorescent channels were designed as "band-pass" to avoid interference. Gray-scaled images were stored for further image analysis. For comprehensively visualizing information from different channels, pseudo colors were assigned to involved channels.

Cell locations and proliferation were quantified using fluorescent microscopy (Olympus BX51 Clinical Microscope). A robotic translation stage was used to image predetermined locations on each culture surface using a MicroFire© monochromatic digital camera (SKU S99826, Optronics, CA). The image locations were fixed on a 16×20 grid with horizontal and vertical spacing of $1280\mu m$ and vertical spacing of $960\mu m$. For each location a $1189 \times 892\mu m^2$ BrdU staining image and Hoechst counter staining image were acquired at a resolution of 1600×1200 $pixels^2$. All images and contextual information were organized and stored in an Oracle™ 10g (Oracle, CA) database for further image processing and data analysis.

[3] a combination of ND6, ND25 and ND50

[4] Currently three channels are used: $DAPI$, $FITC$ and $TRITC$



(a) Composition gradient column (b) Deposit stripe (c) Spread film

Figure 1: Solvent casting procedures

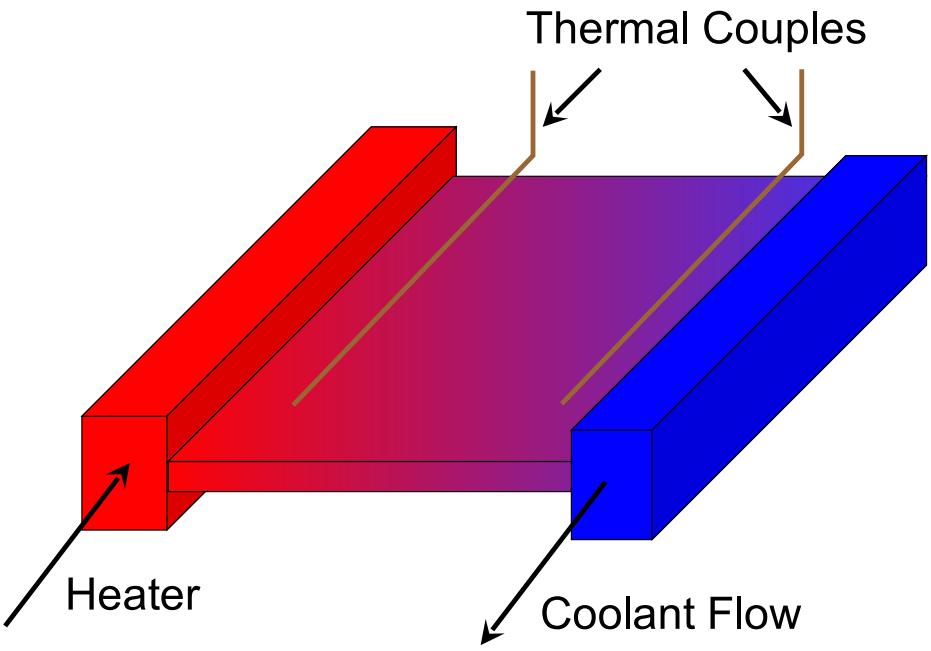


Figure 2: Annealing bed

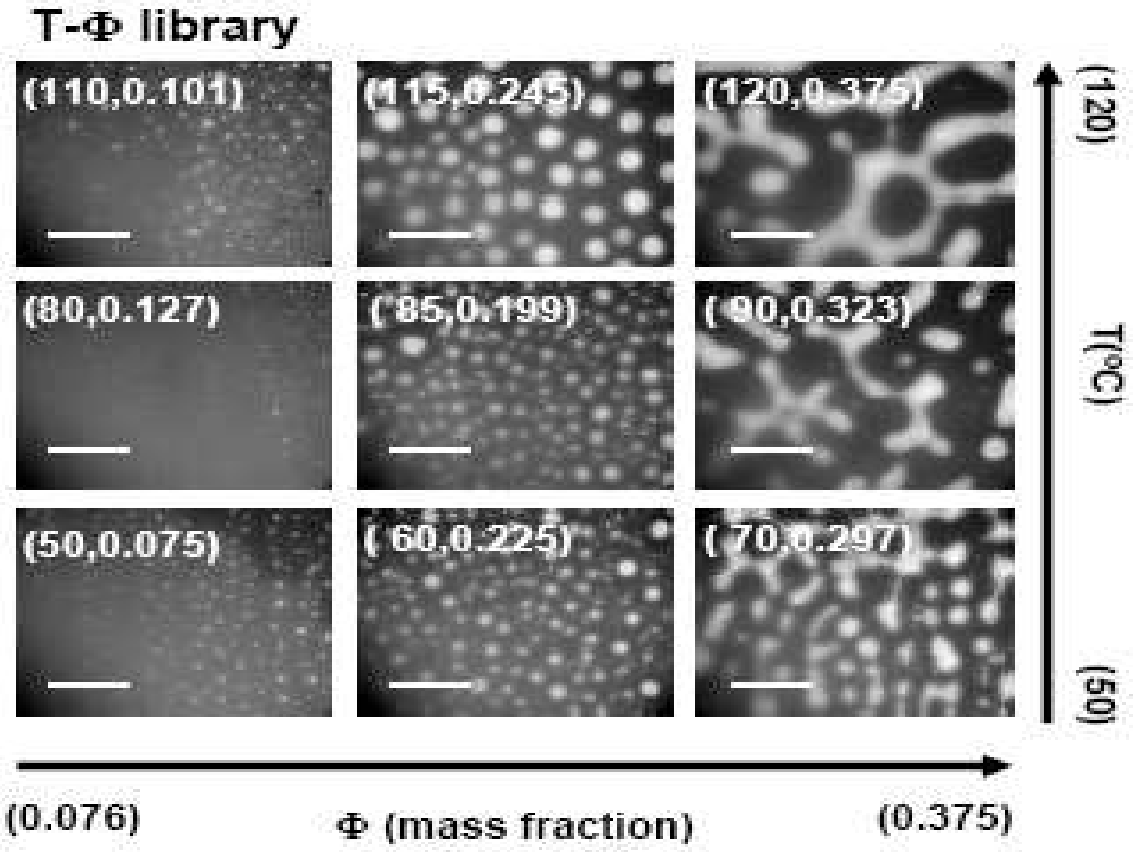


Figure 3: Polarized microscopy for *PLGA*/*PCL* surface micro-structure pattern library (scale bar = 400μm)

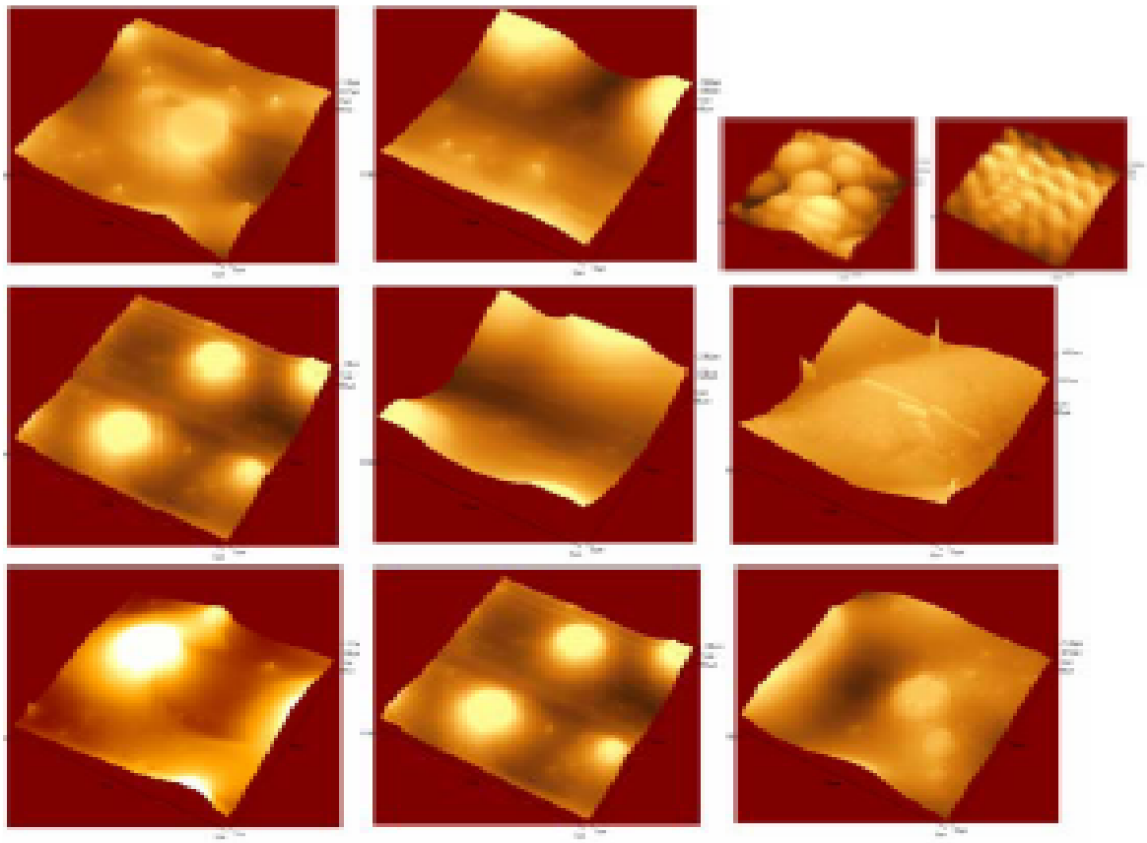


Figure 4: AFM for $PLGA/PCL$ surface micro-structure pattern library
(size of image: $100\mu m \times 100\mu m$)

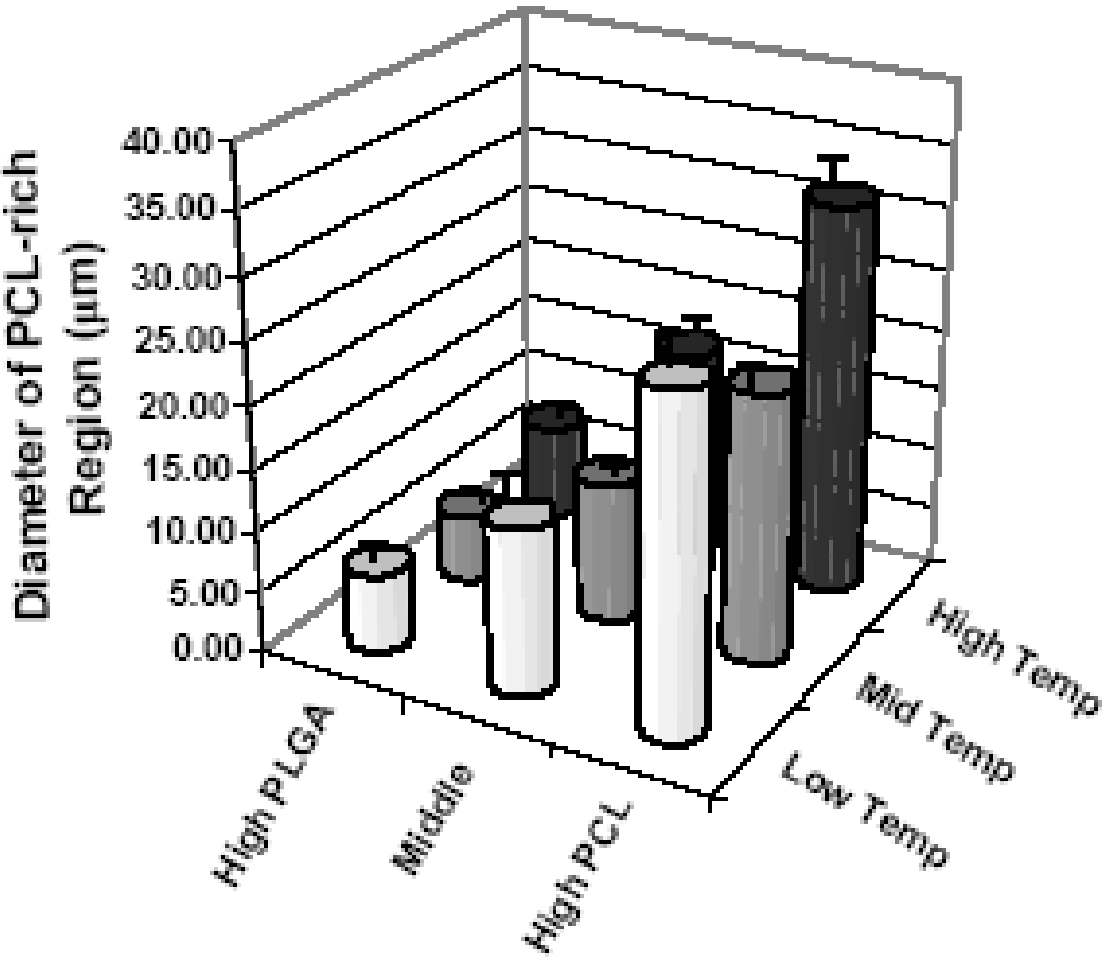


Figure 5: PCL phase diameter on library

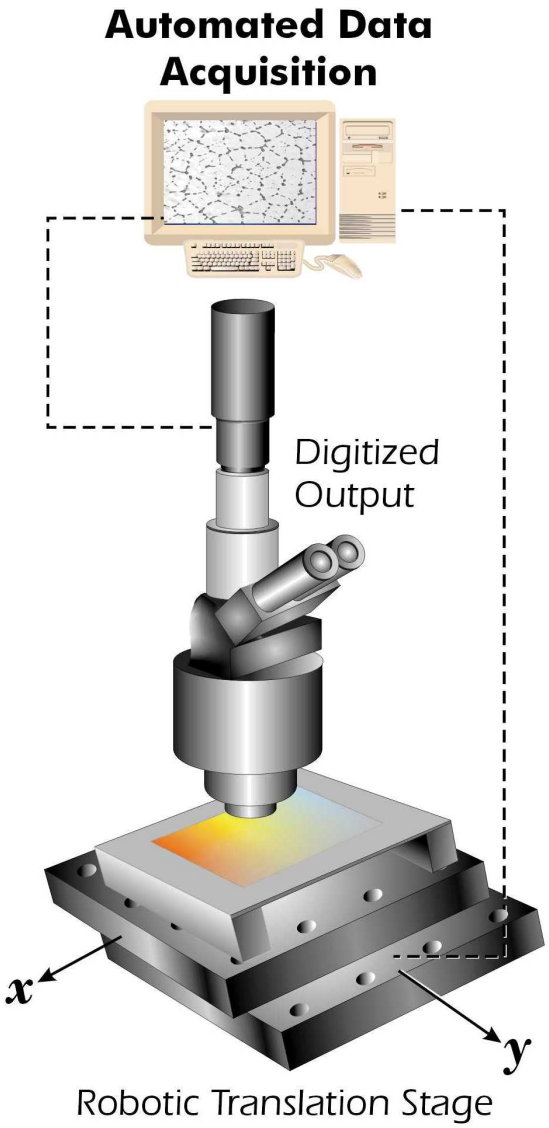


Figure 6: Multi-channel microscopy image acquisition system

CHAPTER III

IMAGE PROCESSING AND ANALYSIS

The *Image Processing Toolbox* of Matlab™ R14 (MathWorks, MA) was employed for image processing and analysis.

The enormous volume of image data space acquired in this work requires automated image enhancing, processing and analyzing methods and dynamic, self-adapting algorithms. As mentioned in Section 2.5, several hundreds to more than one thousand images was obtained from each chip[1]. It is practically impossible to manually count cell numbers and assay cell proliferation ratio. Common image processing software such as *ImageJ*[2] is sufficient for general image enhancement and complicated tasks such as cell counting and segmentation. However, such image processing and analysis software is designed for general purposes and has not been optimized for specific applications in this work. Furthermore, key parameters used in such common software are often required to be manually adjusted and the performance manually evaluated for each image, which is impractical in this work. Therefore, it is necessary to develop self-adaptive image processing algorithms and optimize them to meet the specific requirements of image processing and analysis in this study.

In following sections the overall workflow of image processing and analyzing procedures as well as details of critical image processing steps and strategies were illustrated.

[1]There were 300 ~ 400 different locations scanned for each chip, at each location two to four channels were used, and at least one image was taken from each channel at each location.

[2]ImageJ is a public domain, Java-based image processing program developed at the National Institutes of Health.

3.1 Overall Workflow

Original gray-scale images are classified as Hoechst counter staining, BrdU staining and *PCL* crossed polarized images. Images in each class went through the following steps:

- quality check;
- image enhancement;
- image processing and analysis; and
- supervision of final results.

Procedures and algorithms were comprehensively designed and coordinated with staining protocols, microscope setup and image acquisition strategies used in this work.

3.1.1 Hoechst Counter Staining for Cell Nuclei

The shape and location of cell nuclei were the major information of interest during processing Hoechst counter staining images. The following steps were used for this purpose:

Leveling. The *background* images was obtained from originals by "*morphological opening*"[3] of raw images with a disk-shape structural element[4]. The disk-like shape was selected to match the shape of cell nuclei. The diameter of the structural element is significantly larger than the cell nuclei[5]. The leveled images were obtained by subtracting the background images from originals. Most global and low frequency noises were remove during leveling while the signal of cell nuclei were reserved and relatively enhanced.

Thresholding segmentation. After leveling, black-and-white images were attained by histogram-based *thresholding segmentation*. The optimal threshold TH of each image

[3]The image processing that an erosion followed by a dilation with the same structuring element is called a morphological opening process.

[4]Structural elements are lines, surfaces or volumes used in *structural analysis*.

[5]The diameter of the disk-like structural element was 31 pixel since the sizes of most cell nuclei were smaller than 20 pixel.

was determined by the *variation-adjusted iterative selection* method (*VAIS*) which is discussed in details in Section 3.2.

Filling. The *morphological closing*[6] and holes filling[7] were performed to achieve solid cell nuclei with smooth outlines. Smooth outlines are important in watershed segmentation step for controlling over-segmentation.

Despeckling. Then noise spots were *despeckled*[8] according to the area of particles.

Separating connecting nuclei. Connecting cell nuclei were segmented by the *marker-controlled watershed* method, which is explained in details in 3.3. Images were despeckled again based on area of particles[9] after water-shed segmentation.

The resultant black-and-white images are called *cell nuclei masks*, which would be used to determine cell proliferation.

The centroids of cell nuclei as well as the total cell number on each image were determined by *regional property analysis*, and results indexed and stored.

The performance of image processing algorithms and strategies was monitored by logging image names with warning tags for

- images of abnormal cell numbers;
- images of significant decreasing of cell numbers after each despeckling;
- images of significant increasing of cell numbers after water-shed segmentation;

, and logged images were manually examined. Resultant images were also randomly sampled for evaluating the error introduced by image processing. Algorithms and strategies were further modified when necessary to ensure that cell numbers determined by manually counting were within the 95% confidence interval of that determined by programs.

[6]The image processing that a dilation following by an erosion of the same structuring element is called a morphological closing process.

[7]A disk-shaped structure element of 5 pixel diameter was used for hole-filling.

[8]Despeckle means to remove noise by filtering off small particles

[9]Particles smaller than 20 pixels were discarded as noise.

3.1.2 BrdU Staining for Cell Proliferation

The proliferating status of each cell stained by BrdU immunohistochemistry was the major information to be determined in this process. BrdU stained images were "masked" by the *cell nuclei masks* generated in 3.1.1 so that only BrdU staining signals inside cell nuclei were considered. The total amount of BrdU signal intensity (I_{sum}) and the average signal intensity (I_{avg}) inside a cell nucleus were calculated for each cell in every image. The *variation-adjusted iterative selection method* was employed to determine the *threshold* (TH_{final}) for both I_{sum} and I_{avg}. The corresponding TH_{final}'s, denoted as $TH_{final,sum}$ and $TH_{final,avg}$, were both used for determining whether a cell was in proliferative status or not. Cell proliferation ratios for each image were calculated. If the proliferative statuses of a cell determined by the two thresholds ($TH_{final,sum}$ and $TH_{final,avg}$) were not consistent, this cell is labeled as "inconsistent".

The performance of programs in this process was monitored by logging image names with warning tags for:

- images of which the proliferation ratios were outliers;
- images of which the number of cells labeled as "inconsistent" exceeded 5% of the number of cells in proliferation status determined by $TH_{final,sum}$;

and the same performance evaluation and controlled mentioned in Section 3.1.1 was employed.

Cell proliferating data attained from I_{sum} were used for further analysis.

3.2 Variation-adjusted Iterative Selection Method (VAIS)

The self-adaptive threshold-determining method, *variation-adjusted iterative selection(VAIS)* method, was developed from the original iterative selection method[68, 87, 81].

Shown in Figure 7, the histogram of BrdU staining intensity per nucleus (from all images on a chip) is composed of two major peaks: the low intensity peak (background)

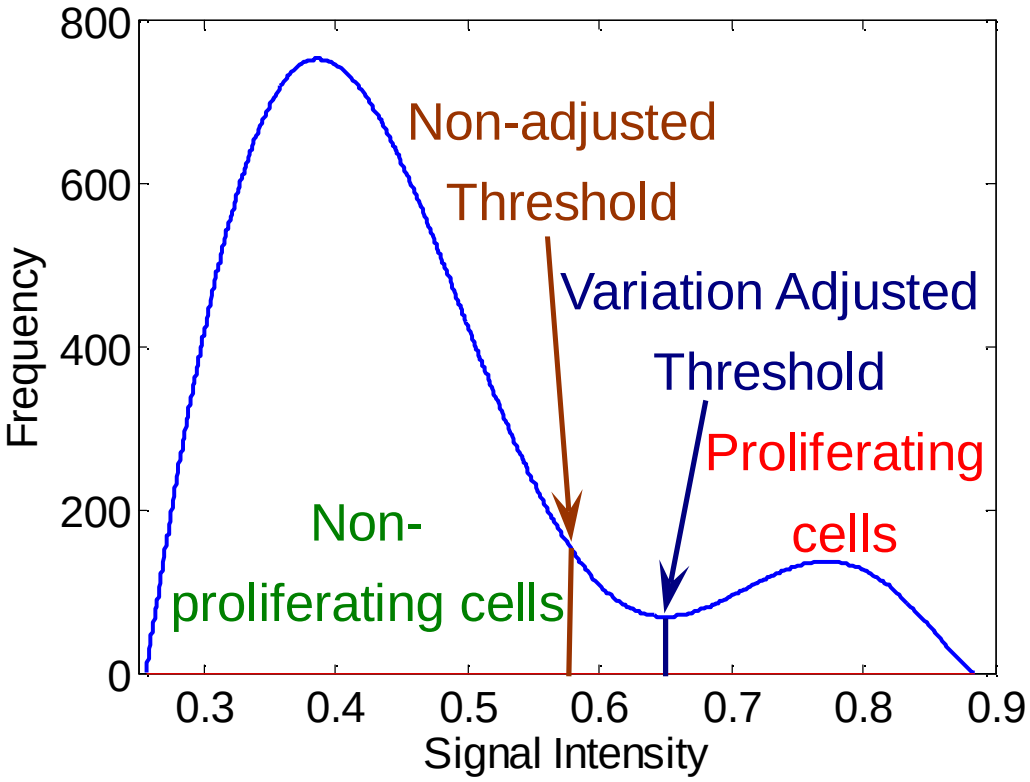


Figure 7: Variation-adjusted iterative segmentation: BrdU staining case

represents cells at rest, while the high intensity peak (foreground) indicates proliferating cells. The optimal threshold between these two peaks was determined automatically by the VAIS method.

Briefly, starting at an initial threshold TH_o determined by

$$TH_o = \frac{1}{2} \tag{1}$$

(that is, $i = 0$), the histogram was divided into resting (background) and proliferating (foreground) parts. Means and standard deviations of the foreground and background, respectively denoted as M_{bo}, M_{fp}, σ_{bo}, and σ_{fo}, were determined by fitting peaks in each section to a Gaussian curve. A new threshold TH_1 was calculated from TH_o as

$$TH_{i+1} = \frac{\sigma_{bi} M_{bi} + \sigma_{fi} M_{fi}}{\sigma_{bi} + \sigma_{fi}} \tag{2}$$

where $i = 0$. This process was repeated until convergence on a stable threshold. Most often, Equation (2) converged to a single stable threshold, indicated by the difference of the last two iteration ΔTH

$$\Delta TH = abs\,(TH_{i+1} - TH_i) < 1\text{bit}, \tag{3}$$

and the final threshold TH_{final} was determined by

$$TH_{final} = \frac{TH_i + TH_{i+1}}{2}. \tag{4}$$

In rare cases, system reached a bi-stable state[10]. In this case, when

$$\Delta TH_{f1} = abs\,(TH_{i+2} - TH_i) < 1\text{bit} \tag{5}$$

and

$$\Delta TH_{f2} = abs\,(TH_{i+3} - TH_{i+1}) < 1\text{bit}, \tag{6}$$

system was defined to be in stable status. Two stable thresholds can be calculated by

$$TH_{final,1} = \frac{TH_i + TH_{i+2}}{2} \tag{7}$$

[10]Bi-stable state: TH_i is jumping back and forth from one stable TH to another one

and

$$TH_{final,2} = \frac{TH_{i+1} + TH_{i+3}}{2}, \tag{8}$$

and the final threshold was defined as

$$TH_{final} = \frac{TH_{final,1} + TH_{final,2}}{2}. \tag{9}$$

Compared with the original iterative selection methods, which use a simple mean intensity

$$TH_{i+1} = \frac{M_{bi} + M_{fi}}{2} \tag{10}$$

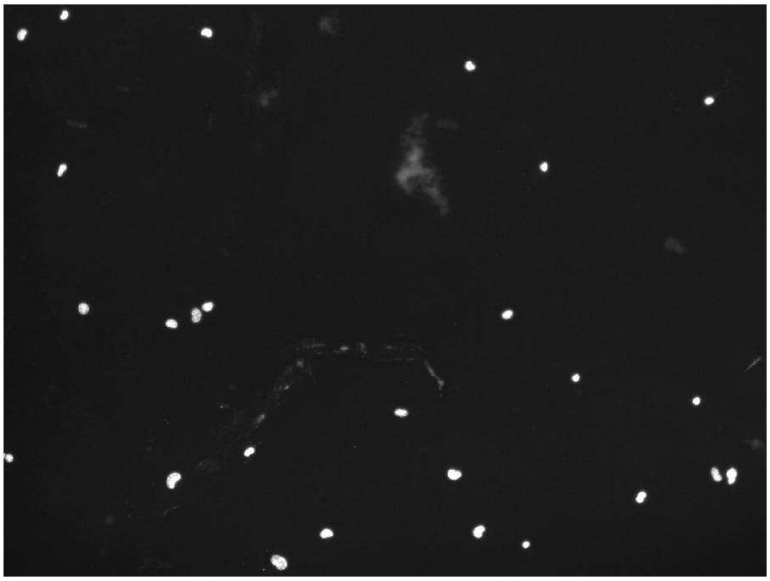
the modified $VAIS$ procedure is more robust when background and foreground intensities have different variances.

Figure 7 shows a typical example of overall histogram of BrdU staining intensities on a chip. In this case, the standard variation of the BrdU signals from non-proliferating cells was significantly greater than that of the proliferating cells[11]. The threshold determined by original iterative selection method noticeably biased toward the wider peak (the background peak in this case) while the one determnined by the variation-adjusted method is much closer to the optimal one.

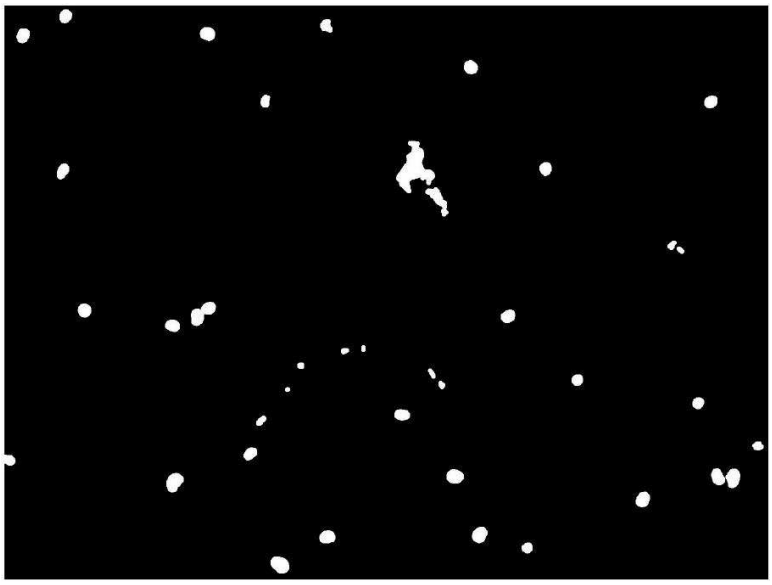
Figure 8 illustrates an example of Hoechst staining case. The original image (Figure 8 (a)) shows considerable background noise, which means a wide variation in the background section. Some background noises were maintained after segmentation by the iterative selection method (Figure 8 (b)) since the background part peak is wider than the forground one. The original iterative selection tends to generate a threshold toward the wider background peak and mistakenly classifies some noises as signals. Contrastingly, most noises were filtered off by the variation-adjusted selection method (Figure 8 (c)).

As illustrated above, the $VAIS$ method is a self-adapt algorithm which recognizes the characteristics of the image histogram, generates optimal thresholds, and addresses the specific image processing needs of this work. This modification is critical to the later data

[11] Background peaks were generally wider than foreground peaks in this work, therefore the modification of the original method is necessary.



(a) Original image



(b) Iterative selection segmentation

Figure 8: Variation-adjusted iterative segmentation: Hoechst counter staining case

(c) Variation-adjusted iterative selection segmentation

Figure 8 (continued) Variation-adjusted iterative selection segmentation: Hoechst counter staining case

analysis, while no common or commercial image processing software provides similar function.

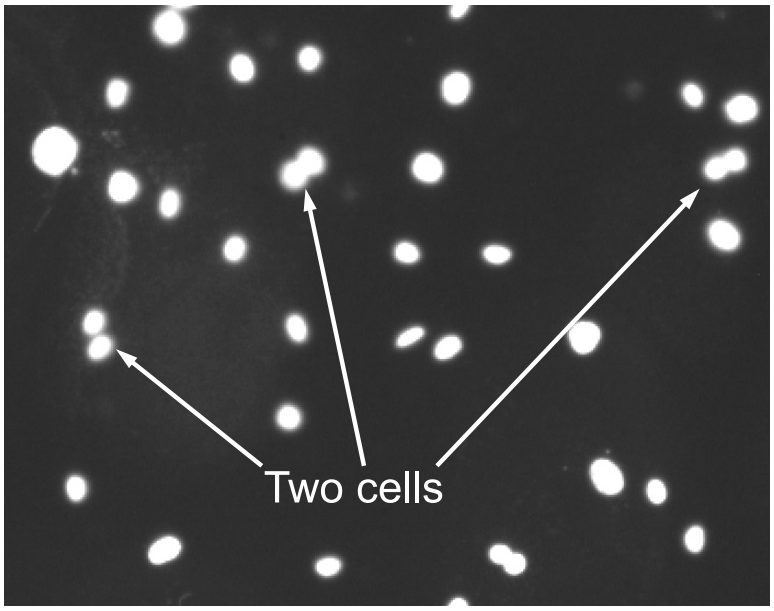
3.3 Marker-controlled Watershed Segmentation (MCWS) Method

Segmentation of contacting cell nuclei is critical in this work. The nuclei of a pair of recently-divided cells were often too close to be distinguished with thresholding segmentation alone. Comparing with other common segmentation algorithms such as thresholding methods, edge-detection based methods and clustering methods, watershed segmentation method[10], as a region-based method, is specifically efficient for separating contacting objects, and therefore widely used in bio-fields for cell nuclei segmentation and cell boundary detection. However, when addressing complex cases, *over-segmentation* effects, known as the major problem of original watershed methods, becomes disastrous. Over-segmentation causes a single cell being segmented into several to several hundred fragments, which are recognized as false "cells". False "cells" in such "cell clusters" are of very close cell-to-cell distances and entangles with the effects of cell contact inhibition. Marker-controlled watershed segmentation significantly suppresses over-segmentation by guiding the watershed processing with *a priori* "markers" as "seeds", and makes the detection of contact inhibition possible.

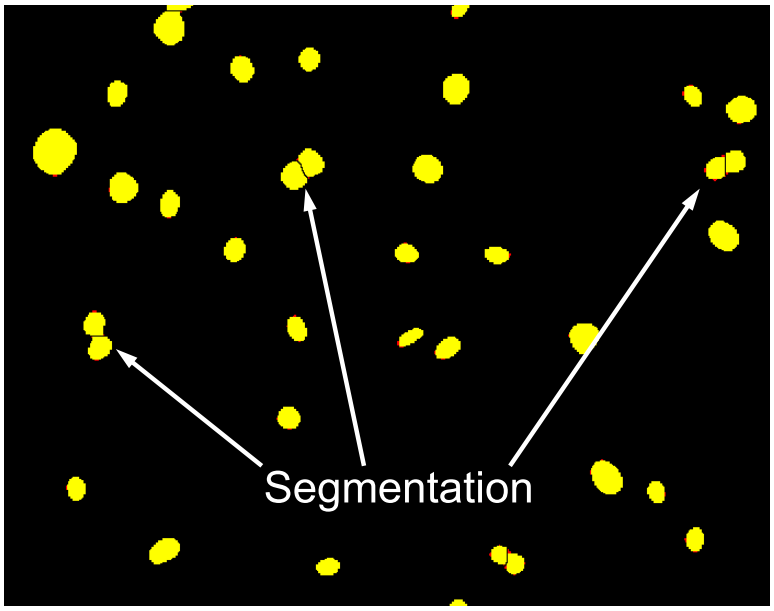
As shown in Figure 9, original gray-scale images of contacting cell nuclei (labeled in Figure 9(a)) were thresholded into binary images, then contacting cell nuclei were further segmented by the *marker-controlled watershed segmentation* (*MCWS*) method [73] to separate images of closely-spaced cell nuclei(labeled in Figure 9(b)).

Images were further defragmented after MCWS process. Comparing with the original gray-scale image(Figure 10(a)), Figure 10(b) shows the removal of fragments (red particles) generated during the MCWS process.

As the critical image processing step, the parameters used in MCWS method were

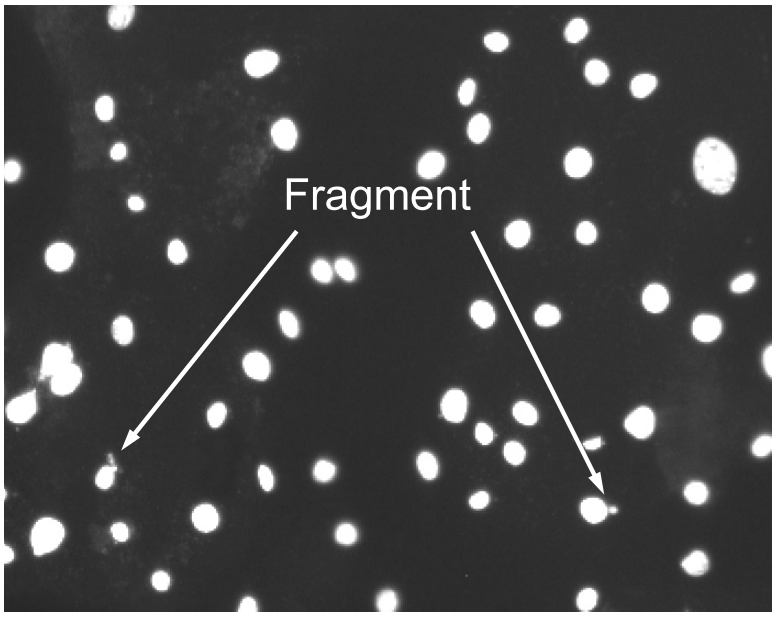


(a) Segmentation: original image

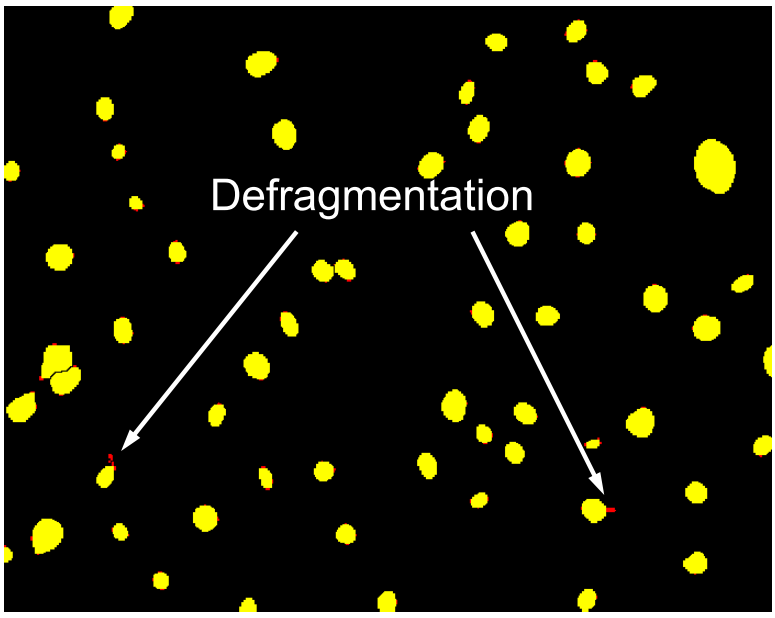


(b)Segmentation: MCWS

Figure 9: Marker-controlled watershed segmentation: segmentation. White: original signals; yellow: final results; red: fragments removed.



(a) De-fragmentation: original image



(b)Defragmenation: MCWS

Figure 10: Marker-controlled water-shed: defragmentation. White: original signals; yellow: final results; red: fragments removed.

fine tuned and the efficacy were strictly validated during image processing, and Over-segmentation and under-segmentation were closely monitored. Images with significant over- or -under-segmentation (more than 5% of segmentation cases) were discarded. If more than 5% images of good quality were discarded, the algorithms were further modified for better performance.

CHAPTER IV

METHOD: INDIVIDUAL-CELL BASED LOCALIZED ANALYSIS

As discussed in Section 1.2, in order to address the characteristics of cell-to-cell and cell-to-surface interactions, especially the sensitiveness of cells to the near neighborhood, a novel and localized data analysis method based on the point of view of individual cells are developed.

In this chapter, using contact inhibition of cell proliferation as an example, the limitation of traditional global descriptions of cell environments and cell behaviors are shown, the concepts of individual-cell based localized data analysis method proposed, mathematical forms established, and basic analysis strategies illustrated.

4.1 Global Metrics

Figure 11(a) cartooned an image of *PCL* islands (blue), nuclei of proliferating cells(red) and cells at rest (green). Cell density and proliferation were described with summary statistics such as number of cells at rest and proliferated computed for each image. This approach provides a set of global metrics for features in each image. As indicated in Figure 11(a) and Table 1, global metrics are most naturally understood in terms of conventional summary-statistics, exploratory data analysis, and well-known methods for estimating confidence and significance levels based on an assumed probability distribution.

Traditional regression and statistical inference would test to relationships between global features and global cell behaviors, from one image to another.

4.1.1 Contact Inhibition Study

In the pure *PLGA*, *PCL*, and *TCPS* cases, the ability to detect contact inhibition of cell proliferation, a known phenomenon, was used as an indicator of the effectiveness of the

Table 1: Global metrics of cell environment and behaviors

Symbol	Name	Definition
N_{Cell}	Number of cell per image	
N_{Prolif}	Number of proliferating cells per image	
$Prol$	Proliferation ratio	$Prolif = \frac{N_{Prolif}}{N_{Cell}}$
N_{PCL}	Number of PCL islands	
S_i	Area of the i^{th} PCL island	
$SumS$	Total PCL area in an image	$SumS = \sum_i S_i$
$\bar{S}$	Arithmetic mean of PCL area	
$\bar{S}_S$	Area-weighted mean area of PCL islands	$\bar{S}_S = \frac{\sum_i S_i}{\sum_i S_i}$
L_i	Size of the i^{th} PCL island	
$SumL$	Sum of the PCL island Sizes in an images	$SumL = \sum_i L_i$
$\bar{L}$	Arithmetic mean of PCL size	$\bar{L} = \frac{SumL}{N_{PCL}}$
$PerS$	Percentage of PCL phase area	$PerS = \frac{SumS}{S_{Image}}$
$DispS$	Dispersion of PCL phase area	$DispS = \frac{\bar{S}}{\bar{S}_S}$

global metrics cell density and proliferation averages. The results and discussion are in Chapter 4 Section 4.1.

4.1.2 Library Screening Study

The application of global metrics in material informatics, especially knowledge discovery approaches on $PLGA/PCL$ combinatorial libraries for screening surface microstructure patterns that affect cell proliferation, were studied as comparison and validation of the individual-cell-based methods. *Factor importance analysis* was performed for candidate factors listed in Table 1. PCL island density (N_{PCL}), arithmetic means of PCL island size and area ($\bar{S}$), area-weighted mean area of PCL islands ($\bar{S}_S$), percentage PCL phase area in an image ($PerS$), dispersion of PCL island area distribution ($DispS$), total PCL area in an image ($SumS$), and the virtual sum of PCL island sizes in an image ($SumL$) were selected as candidate metrics. Candidate descriptors shown in Table 1 were calculated for each image from the library. Then, each candidate descriptor was standardized with respect to the mean and standard deviation. The F-test was applied as the major factor screening approach to screen out the standardized descriptors that best captured data trends in cell proliferation. The significant surface feature descriptor(s) was determined based on a significant level of $\alpha = 0.05$, and their effects on cell proliferation were visualized by linear regression.

A brief *principal component analysis* (*PCA*) was also performed and the major principal components were used for *clustering* against cell proliferation.

The results are shown and discussed in Chapter 7 Section 7.1.

4.2 Local Individual-cell-based Metrics

The concepts of local individual-cell-based data analysis was proposed and established for cell-to-cell interaction study. Contact inhibition of cell proliferation was chosen as the model system for its simplicity. The method developed in contact inhibition study, after validation and refining, was then extended as a high-throughput screening method to the

much more complicated cell-to-surface interaction study.

4.2.1 Cell-to-cell Interactions and Contact Inhibition

Local feature histograms were used to monitor the individual cell environment and relate it to proliferating status. Cell-to-cell distance was defined as the distance between the centroids of two cell nuclei. The collection of all cell-cell distances were sorted into frequency histograms. the results of the proliferation assay, cell-to-cell distance histograms were further classified as representing distances between: proliferated cells (P), cells at rest (R), and any cell (A), i.e., without distinguishing proliferating vs. non-proliferating. Therefore, 9 classes of cell-to-cell distance (dis_{PP}, dis_{PR}, dis_{PA}, dis_{RP}, dis_{RR}, dis_{RA}, dis_{AP}, dis_{AR}, and dis_{AA}) were defined. Three types of distances, dis_{PP}, dis_{PR}, and dis_{PA}, from the point of view of an individual P class cell are highlighted in Figure 11(b). The biological rationale for the use of inter-nuclear distances is the relative of ease of nuclear counter staining and the involvement of near-nuclear structures in cell-contact signaling. For example, N-cadherin cell contact signaling in fibroblasts utilizes microtubules emanating from a microtubule organization center near the cell nuclei.

Local cell metrics are naturally connected to Bayesian analysis, which is a powerful statistical method used for classification. Local cell metrics are naturally connected to a statistical method known as Bayesian analysis. The cell-to-cell distance, PR, is used here to illustrate the use of local metrics in a *naïve Bayes model* to quantify the local cell background (or environment). Specifically, the Bayesian approach allows one to quantify the *local environment of a proliferating cell*, P, as the conditional probability of finding a non-proliferating (R) cell a certain distance PR from the P cell. Assume that in the k^{th} image the number of P-class and R-class cells is n_{Pk} and n_{Rk}, the distance PR_{ijk} between the centroids of the nuclei of the i^{th} P-cell and the j^{th} R-cell can be calculated readily from the results of image analysis (the location of each cell nucleus in each image were returned as a list of coordinates of cell nucleus centroids by image analysis). On the k^{th} image, the

set of all such distance, PR_k was defined as

$$PR_k = \left\{ PR_{ijk} |\, i = 1, 2, \ldots, n_{Pk};\ j = 1, 2, \ldots, n_{Pk} \right\} \tag{11}$$

And for all images an overall set PR can be defined as

$$PR = \bigcup_k PR_k \tag{12}$$

Based on a $(N + 1)$-level distance scale,

$$Scale_{dist} = \{d_0, d_1, \ldots, d_N\}, \tag{13}$$

a set of distance bins bin_{dist} was defined as

$$bin_{dist} = \{[d_0, d_1), [d_1, d_2), \ldots, [d_{N-1}, d_N]\}. \tag{14}$$

For $i = 1, 2, \ldots, N$, centroid of each interval in bin_{dist}, $\hat{d}_i$ was defined as

$$\hat{d}_i = \begin{cases} \dfrac{d_{i-1} + d_i}{2} & \text{if } Scale_{dist} \text{ is linearly spaced} \\ \sqrt{d_{i-1} \cdot d_i} & \text{if } Scale_{dist} \text{ is logarithmically spaced} \end{cases} \tag{15}$$

and the resultant centroid set $\hat{d}$ for bin_dist is

$$\hat{d} = \left\{ \hat{d}_1, \hat{d}_1, \ldots, \hat{d}_N \right\} \tag{16}$$

was used to sort set PR into an N-bin histogram $N_{PR}(\hat{d})$:

$$N_{PR}\left(\hat{d}\right) = h_N\left(PR|_{bin_{dist}}\right) = h_N\left(PR|_{\hat{d}}\right) \tag{17}$$

where $N_{PR}|_{\hat{d}}$ is the number of elements of set PR that are in the interval $[d_{i-1}, d_i)$ (which is centered at $\hat{d}_i$). According to the needs of analysis, $Scale_{dist}$'s with specific range (represented by d_o and d_N), Resolutions(N), and spacing style (linearly or logarithmically spaced) were designed and applied. Therefore, each $Scale_{dist}$ set was used as a specific observing window for cell environments. For convenience, $\hat{d}$, an equivalent form of $Scale_{dist}$, was instead defined as an operationally *observation "window"* for a given *cell environment*.

Since the total number of elements in set PR_k is $n_{PR_k} = n_{PK} \cdot n_{RK}$, the total number of elements in set PR is

$$n_{PR} = \sum_k n_{PK} \cdot n_{RK}. \tag{18}$$

Normalized by n_{PR}, a frequency function was achieved from $N_{PR}\left(\hat{d}\right)$ as

$$f_{PR}\left(\hat{d}\right) = \frac{N_{PR}\left(\hat{d}\right)}{n_{PR}} \equiv h_f\left(PR|_{\hat{d}}\right) \tag{19}$$

$f_{PR}(\hat{d})$ was defined as the R type cell background of the P class cells observed from the window $\hat{d}$. Physical meanings of this definition will be discussed in detail in the discussion section.

Based on the definition of f_{PR}, a naïve Bayes model can be established. Given a cell, the effects of the distance from other non-proliferating cells on the possibility of proliferation of this cell to be in proliferating status is given by the following conditional probability function

$$p\left(prolif|R_{\hat{d}_1}, R_{\hat{d}_2}, \cdot, R_{\hat{d}_N}\right) \tag{20}$$

where $R_{\hat{d}_i}$, the feature variable of R type cell background, represents the possibility of occurrence of non-proliferating cells around a distance of $\hat{d}_i$. Using Bayer's theorem,

$$p\left(prolif|R_{\hat{d}_1}, R_{\hat{d}_2}, \ldots, R_{\hat{d}_N}\right) = \frac{p\left(prolif\right) \cdot p\left(R_{\hat{d}_1}, R_{\hat{d}_2}, \ldots, R_{\hat{d}_N}|prolif\right)}{p\left(R_{\hat{d}_1}, R_{\hat{d}_2}, \ldots, R_{\hat{d}_N}\right)} \tag{21}$$

In the above function, the components $p\left(R_{\hat{d}_1}, R_{\hat{d}_2}, \ldots, R_{\hat{d}_N}\right)$ and $p\left(prolif\right)$ are constants, that can be determined from their frequency in the data. The only non-constant component

the class-conditional probability,

$$
\begin{aligned}
& p\left(R_{\hat{d}_1},\ R_{\hat{d}_2}, \ldots, R_{\hat{d}_N} | prolif\right) \\
& \quad = p\left(R_{\hat{d}_1} | prolif\right) \cdot p\left(R_{\hat{d}_2}, \ldots, R_{\hat{d}_N} | prolif\right) \\
& \quad = p\left(R_{\hat{d}_1} | prolif\right) \cdot p\left(R_{\hat{d}_2} | prolif\right) \cdot p\left(R_{\hat{d}_3}, \ldots,\ R_{\hat{d}_N} | prolif, R_{\hat{d}_1}, R_{\hat{d}_2}\right) \\
& \qquad\qquad \vdots \\
& \quad = p\left(R_{\hat{d}_1} | prolif\right) \cdot p\left(R_{\hat{d}_2} | prolif\right) \cdot p\left(R_{\hat{d}_3} | prolif\right) \\
& \qquad \cdots p\left(R_{\hat{d}_N} | prolif,\ R_{\hat{d}_1}, R_{\hat{d}_2}, \ldots, R_{\hat{d}_N}\right)
\end{aligned}
\tag{22}
$$

Assuming the occurrence probabilities around the non-proliferating cell distances, denoted as $R_{\hat{d}_1}, R_{\hat{d}_2}, \ldots, R_{\hat{d}_N}$, are conditionally independent (uncorrelated) then $p\left(R_{\hat{d}_i} | prolif\right) = p\left(R_{\hat{d}_i} | prolif, R_{\hat{d}_j}\right)$ when $i \neq j$. A naïve Bayes model can be achieved as

$$
p\left(R_{\hat{d}_1}, R_{\hat{d}_2}, \ldots, R_{\hat{d}_N} | prolif\right) = \prod_{i=1}^{N} p\left(R_{\hat{d}_i} | prolif\right) \tag{23}
$$

Noticing that $p\left(R_{\hat{d}_i} | prolif\right) = f_{PR}|_{\hat{d}_i}$, and denoting $p\left(R_{\hat{d}_1}, R_{\hat{d}_2}, \ldots, R_{\hat{d}_N}\right)$ by scaling factor Z_{PR}, the original conditional probability function becomes

$$
\begin{aligned}
& p(prolif | R_{\hat{d}_1}, R_{\hat{d}_2}, \ldots, R_{\hat{d}_N}) \\
& \quad = \frac{p\,(prolif)}{Z_{PR}} \cdot \prod_{i=1}^{N} f_{PR}|_{\hat{d}_i} \\
& \quad \propto \prod_{i=1}^{N} f_{PR}|_{\hat{d}_i}
\end{aligned}
\tag{24}
$$

Therefore, the naïve Bayes model for the effects of R class cells on cell proliferation can be sufficiently determined by products of elements in $f_{PR}\left(\hat{d}\right)$, which is easily computed from a training data set.

Frequency functions, or cell environments, denoted as $f_{PP}\left(\hat{d}\right)$, $f_{AA}\left(\hat{d}\right)$, $f_{RR}\left(\hat{d}\right)$, and $f_{PA}\left(\hat{d}\right)$, were also calculated for cell-to-cell distances PP, AA, RR, and PA observed from the window $\hat{d}$ in similar manner.

To obtain a meaningful posterior conditional probability, further normalization is necessary to relate observed occurrences to random occurrences. Given the finite image size

and non-square, the distribution of random cell occurrences is not. The random distribution curve of cell-cell distance,

$$f_{std}\left(\hat{d}\right) = h_{f_{std}}\left(AA|_{\hat{d}}\right) \tag{25}$$

was calculated by a *Monte Carlo* method using 1×10^{10} randomly-distributed spots (non-interacting cells) on one image of 1600 pixels by 1200 pixels. The normalized cell histogram $\tilde{f_{PR}}\left(\hat{d}\right)$ is

$$\tilde{f}_{PR}\left(\hat{d}\right) = \frac{f_{PR}\left(\hat{d}\right)}{f_{std}\left(\hat{d}\right)} \qquad \text{for} \qquad i = 1, 2, \ldots N \tag{26}$$

and the normalized conditional probability model becomes

$$\begin{aligned} &\tilde{p}(prolif|R_{\hat{d}_1}, R_{\hat{d}_2}, \ldots, R_{\hat{d}_N}) \\ &\quad \propto \prod_{i=1}^{N} \tilde{f}_{PR}|_{\hat{d}_i} \\ &\quad = \prod_{i=1}^{N} \frac{f_{PR}\left(\hat{d}\right)}{f_{std}\left(\hat{d}\right)} \\ &\quad = \frac{\prod_{i=1}^{N} f_{PR}\left(\hat{d}\right)}{\prod_{i=1}^{N} f_{std}\left(\hat{d}\right)} \end{aligned} \tag{27}$$

Other cell histograms were also normalized similarly. Normalization of cell histograms not only allows direct comparisons of different types of cell distances on different surfaces, but also the normalized cell histograms themselves provide important information on patterns of cell attachment and migration. Practically, ratios between cell histogram functions

were used also in our analysis, in which case the standardized (random) histogram cancels,

$$\begin{aligned} r_{PR|PA}\left(\hat{d}\right) &= \frac{\tilde{f}_{PR}\left(\hat{d}\right)}{\tilde{f}_{PA}\left(\hat{d}\right)} \\ &= \frac{\prod_{i=1}^{N} f_{PR}\left(\hat{d}\right) \Big/ \prod_{i=1}^{N} f_{std}\left(\hat{d}\right)}{\prod_{i=1}^{N} f_{PA}\left(\hat{d}\right) \Big/ \prod_{i=1}^{N} f_{std}\left(\hat{d}\right)} \\ &= \frac{\prod_{i=1}^{N} f_{PR}\left(\hat{d}\right)}{\prod_{i=1}^{N} f_{PA}\left(\hat{d}\right)} \\ &= \frac{f_{PR}\left(\hat{d}\right)}{f_{PA}\left(\hat{d}\right)} \end{aligned} \tag{28}$$

More importantly, $r_{PR|PA}$ highlights the specific effects of non-proliferated cell on the central proliferating cell (P) relative to any given cell. Thus, the probability of cell responses under different cell environments can be compared meaningfully.Furthermore, each set of cell-to-cell distances can be decomposed into subsets, which allows investigations of contributions of each subset to the overall effects. For example, observed from the same $\hat{d}$, since

$$PA = PP \cup PR \tag{29}$$

and

$$PP \cap RR = \emptyset \tag{30}$$

instead of using $r_{PR|PA}$, the effects of non-proliferating cells on cell proliferation can be better presented by removing the PR component from PA, i.e. replacing PA with PP. Obviously, after removing the shared component from the denominator, $r_{PR|RR}$ has higher contrast than $r_{PR|PA}$. The method of optimizing contrast of comparison of the two sets by analyzing and manipulating compositions of sub-components is called *decomposition analysis*.

Therefore, ratios of cell backgrounds were constructed and used as classifiers for screening and identifying significant cell environment patterns. In this study, $r_{PA|AA}$, $r_{PP|RR}$, and $r_{PR|RR}$ were used for decoupling cell contact inhibition effects on cell proliferation from others, and a preliminary model was achieved.

Furthermore, $r_{PA|AA}$ and $r_{PR|RR}$ represent specific and important physical and mathematical meanings. This two descriptors defined either posterior odds or Bayes factors of proliferation behaviors. Considering the case that there is only one image acquired ($k = 1$),

$$\begin{aligned} r_{PA|AA} &= \frac{f_{PA}|_{\hat{d}_i}}{f_{AA}|_{\hat{d}_i}} \\ &= \frac{N_{PA}|_{\hat{d}_i} \big/ n_{PA}}{N_{AA}|_{\hat{d}_i} \big/ n_{AA}} \\ &= \frac{N_{PA}|_{\hat{d}_i} \big/ N_{AA}|_{\hat{d}_i}}{n_{PA} / n_{AA}} \end{aligned} \tag{31}$$

where

$$\frac{n_{PA}}{n_{AA}} = \frac{\sum\limits_k (n_{Pk} \cdot n_{Ak})}{\sum\limits_k (n_{Ak} \cdot n_{Ak})} = \frac{n_P}{n_A} = Prolif|_{average} \tag{32}$$

Also define a set of vectors pointed from one cell toward another

$$\mathbb{AA}|_{\hat{d}_i} \equiv \left\{ \overrightarrow{\left(A_i, A_j\right)} \,|\, dist\left(A_i, A_j\right) \in [d_{l-1}, d_l),\ i \neq j \right\} \tag{33}$$

and its subset

$$\mathbb{PA}|_{\hat{d}_i} \equiv \left\{ \overrightarrow{\left(P_i, A_j\right)} \,|\, dist\left(P_i, A_j\right) \in [d_{l-1}, d_l),\ i \neq j \right\} \tag{34}$$

and two sets of cells involved in these two vector sets

$$A|_{\hat{d}_i} \equiv \left\{ A_i \,|\, A_i \text{ involved in } \mathbb{AA}|_{\hat{d}_i} \right\} \tag{35}$$

and

$$P|_{\hat{d}_i} \equiv \left\{ P_i \,|\, P_i \text{ involved in } \mathbb{PA}|_{\hat{d}_i} \right\} \tag{36}$$

with numbers of elements of $AA|_{\hat{d}_i}$ and $PA|_{\hat{d}_i}$, respectively. Then, and described the numbers

of elements in sets and , respectively:

$$N_{AA}|_{\hat{d}_i} = N_A|_{\hat{d}_i} \cdot N_A|_{\hat{d}_i}$$
$$N_{PA}|_{\hat{d}_i} = N_P|_{\hat{d}_i} \cdot N_A|_{\hat{d}_i} \tag{37}$$

and the proliferation ratio is defined as

$$\frac{N_{PA}|_{\hat{d}_i}}{N_{AA}|_{\hat{d}_i}} = \frac{N_P|_{\hat{d}_i} \cdot N_A|_{\hat{d}_i}}{N_A|_{\hat{d}_i} \cdot N_A|_{\hat{d}_i}} = \frac{N_P|_{\hat{d}_i}}{N_A|_{\hat{d}_i}} \equiv Prol|_{\hat{d}_i} \tag{38}$$

Applying Equation (38) and Equation (32) to Equation (31),

$$r_{PA|AA} = \frac{N_{PA}|_{\hat{d}_i} \Big/ N_{AA}|_{\hat{d}_i}}{n_{PA} / n_{AA}} = \frac{Prol|_{\hat{d}_i}}{Prol|_{average}} \tag{39}$$

which represents how many folds the cell proliferation ratio has been promoted when two cell have a distance around $\hat{d}_i$. Extending this concept into the $k > 1$ cases will not change the physical meaning, i.e.,

$$r_{PA|AA} = \frac{Prol|_{\hat{d}_i}}{Prol|_{average}} \tag{40}$$

Let d_{max} denote the maximum distance possible in one image, and establish a virtual image composed of all images acquired from a surface, separated with each other on any direction by a distance larger than d_{max}, and only focus on the cell-to-cell distance within the range of $[0, d_{max}]$, Equation (40) still holds.

Similarly, define another descriptor for proliferation behaviors as the posterior odds

$$Prol_{Odd} = \frac{n_P}{n_R} \tag{41}$$

$r_{PR|RR}$ can be physically explained as

$$r_{PA|AA} = \frac{Prol_{Odd}|_{\hat{d}_i}}{Prol_{Odd}|_{average}} \tag{42}$$

Significant mathematical meanings can be obtained when these two conclusions (Equation (40) and Equation (42)) are re-expressed with terminology of Bayes modeling. When $Prol$ and $Prol_{Odd}$ are used as descriptors for cell proliferation, $Prol|_{average}$ and $Prol_{Odd}|_{average}$ are prior odds, $Prol|_{\hat{d}_i}$ and $Prol_{Odd}|_{\hat{d}_i}$ are posterior odds at $d = \hat{d}$, $r_{PA|AA}$ and $r_{PR|RR}$ *Bayes*

factors at $d = \hat{d}$, respectively. The models can be further simplified by normalizing *Prol* and $Prol_{Odd}$ against $Prol|_{average}$ and $Prol_{Odd}|_{average}$. In this context,

$$Odd_{prior} = 1, \tag{43}$$

$$Odd_{post}\,(Prol) = \frac{Prol|_{\hat{d}}}{Prol|_{average}} = r_{PA|AA} \tag{44}$$

and

$$Odd_{post}\,(Prol_{Odd}) = \frac{Prol_{Odd}|_{\hat{d}}}{Prol_{Odd}|_{average}} = r_{PR|RR}. \tag{45}$$

The normalized forms are used in this paper for discussion due to its simplicity.

4.2.2 Cell-to-surface Interactions: Screening and Knowledge Discovery

Based on the same conceptual and mathematical foundation, the method developed for cell-to-cell interactions and especially for study of contact inhibition of cell proliferation (Section 4.2.1) was extended to cell-surface interaction cases. The *local cell-feature analysis*, a novel high-throughput screening method for exploring surface microstructure features that affect cell proliferation on $PLGA/PCL$ combinatorial libraries was developed.

Besides the three cell classes used in Section 4.2.1, PCL islands on the $PLGA/PCL$ combinatorial libraries were treated as additional class of particles, represented as class C. Three cell-to-PCL-island distances, dis_{PC}, dis_{RC}, and dis_{AC}, were defined as the centroid-to-centroid distance between proliferated (P), non-proliferated (R), or average (A) cells and PCL islands(C). Another important modification in the $LCFA$ method is that PCL size dia_i, defined as

$$dia_i \equiv \sqrt{S_i \frac{4}{\pi}} \tag{46}$$

(slightly different with the traditional descriptor $\bar{L}$ defined in Table 1), was also used in modeling. The introduction of PCL island size extends the original one-dimensional factor system (in which the particle-to-particle distance is the only factor and frequency functions are curve lines with respect to dis) to a two-dimensional factor system (in which frequency functions are curve surfaces defined on the $dis \times dia$ space). The definitions of the naïve

Bayes model for cell-surface-micro-structure interactions are briefly summarized below with the distance set PC as an example.

As the definition of PR, PC_{ijk} is defined as the distance between the centroids of the nucleus of the i^{th} P-cell and the j^{th} PCL island in the k^{th} image, and the set PC is defined as the set of

$$PC = \bigcup_{i,j,k} PC_{ijk} \tag{47}$$

Besides the $Scale_{dist}$ defined in Section 4.2.1 (used as cell-to-PCL-island distance), another scale, $Scale_{dia}$ is defined for PCL island size. The two scales were used for defining the two-dimensional space $grid$

$$grid_{dist,dia} = \begin{Bmatrix} [d_0, d_1), [dia_0, dia_1) & [d_0, d_1), [dia_1, dia_2) & \dots & [d_0, d_1), [dia_{M-1}, dia_M) \\ ([d_1, d_2), [dia_0, dia_1)) & ([d_1, d_2), [dia_1, dia_2)) & \dots & ([d_1, d_2), [dia_{M-1}, dia_M) \\ \multicolumn{4}{c}{\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots} \\ ([d_{N-1}, d_N), [dia_0, dia_1)) & ([d_{N-1}, d_N), [dia_1, dia_2)) & \dots & ([d_{N-1}, d_N), [dia_{M-1}, dia_M) \end{Bmatrix} \tag{48}$$

After similar histogram processing based on $grid$ and normalization based on amount of elements in the set PC, the frequency function $f_{PC}(\hat{d}, \widehat{dia})$,

$$f_{PC}\left(\hat{d}, \widehat{dia}\right) == \frac{N_{PC}|grid_{\hat{d},\widehat{dia}}}{n_{PC}|grid} \equiv h_f\left(grid_{\hat{d},\widehat{dia}}\right) \tag{49}$$

is achieved and is defined as *the surface pattern background* of the P class cells. The definition of $\hat{d}$ is the same as in Equation (15), and $\widehat{dia}$ is defined in a similar way as

$$\widehat{dia}_i = \begin{cases} \dfrac{dia_{i-1} + dia_i}{2} & \text{if } Scale_{dia} \text{ is linearly spaced} \\ \sqrt{dia_{i-1} \cdot dia_i} & \text{if } Scale_{dia} \text{ is logarithmically spaced} \end{cases} \tag{50}$$

And the conditional probability model under the naïve Bayes assumption is

$$\begin{aligned} p\left(prolif|grid_{dist,dia}\right) &= \frac{pprolif}{Z_{PC}} \cdot \sum_{i=1,j=1}^{M,N} f_{PC}|_{d_i,dia_i} \\ &\propto \sum_{i=1,j=1}^{M,N} f_{PC}|_{d_i,dia_i} \end{aligned} \tag{51}$$

The $f_{AC}(\hat{d}, \widehat{dia})$ was also calculated by the same method

$$f_{AC}\left(\hat{d}, \widehat{dia}\right) = \frac{N_{AC}|grid_{\hat{d},\widehat{dia}}}{n_{AC}|grid} \equiv h_f\left(grid_{\hat{d},\widehat{dia}}\right) \tag{52}$$

Instead of the ratio $r_{PC|AC}$ between $f_{PC}(\hat{d}, \widehat{dia})$ and $f_{AC}(\hat{d}, \widehat{dia})$, which is denoted as

$$r_{PC|AC}\left(grid_{\hat{d},\widehat{dia}}\right) = \frac{f_{PC}\left(\hat{d}, \widehat{dia}\right)}{f_{AC}\left(\hat{d}, \widehat{dia}\right)} \tag{53}$$

the difference $\Delta_{PC|AC}$

$$\Delta_{PC|AC}\left(grid_{\hat{d},\widehat{dia}}\right) = \left(\hat{d}, \widehat{dia}\right) - f_{AC}\left(\hat{d}, \widehat{dia}\right) \tag{54}$$

was practically used for comparing for localized analysis of cell nearby regions. The major reason of using $\Delta_{PC|AC}$ instead of $r_{PC|AC}$ is due to the data scarcity resulted from the higher dimensions of the naïve Bayes model comparing with that for cell-cell interaction cases. The introduction of feature variable set $Scale_{dia}$ exponentially reduced the amount of data points in each grid in $grid_{dist,dia}$. Under this condition the differences showed more robust performance than ratios, given some regions of the results were slightly distorted. For the application of fast screening and material informatics, however, robustness of the screening method is crucial, while the slight distortion introduced by using the difference instead of ratio does not significantly affect the screening results.

Vinculin and *F*-actin staining for cell attachment and spreading on *PLGA*/*PCL*-libraries was used for validating the screening results.

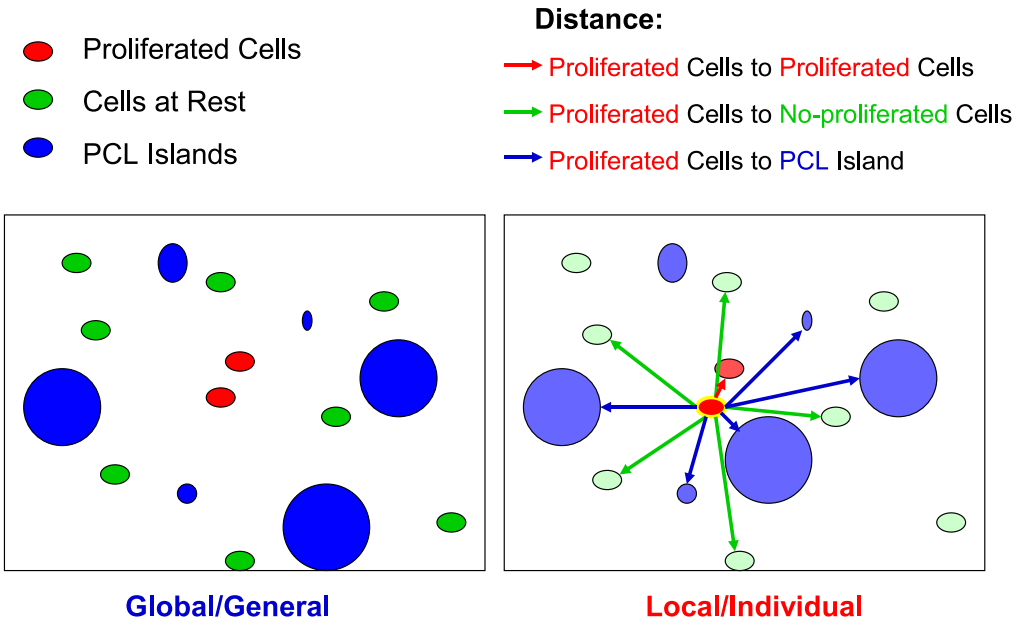


Figure 11: Global and local metrics (Blue: crossed polarized image of *PCL* islands; red: nuclei of proliferating cells; green: nuclei of cells at rest.)

CHAPTER V

LOCALIZED ANALYSIS OF CELL-TO-CELL INTERACTIONS

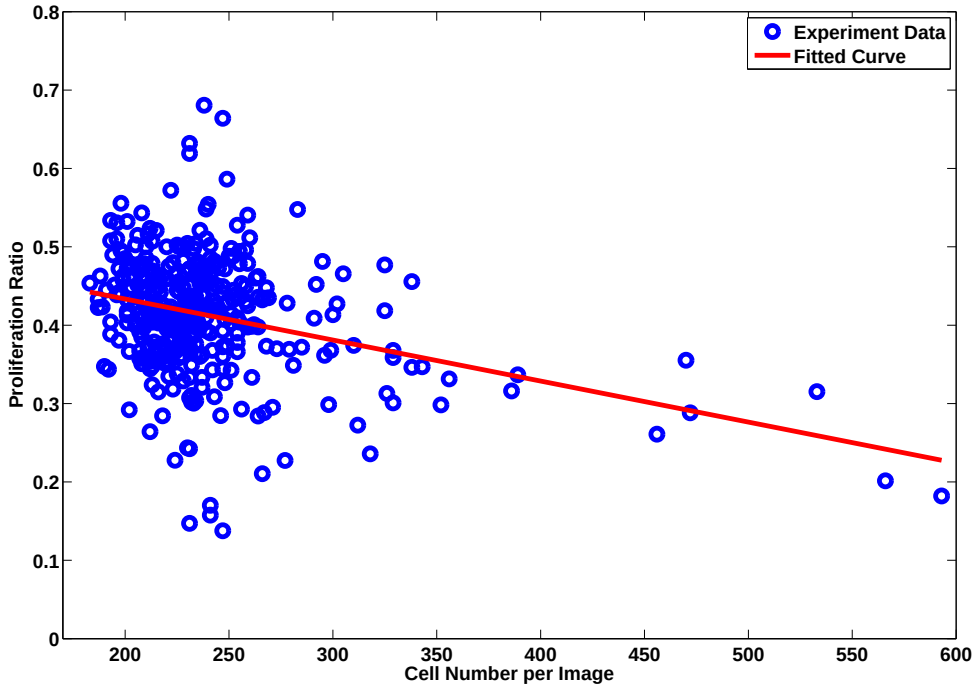
5.1 Global Metrics

To provide a benchmark for establishing the effectiveness of local metrics, contact inhibition of cell proliferation was studied using global analysis first. For each image in the database, the overall cell proliferation is plotted versus cell density, shown in Figure 12 and Tables 2. Global analysis, based on summary statistic descriptions, did not detect the contact suppression effects on cell proliferation. Linear regression (Tables 2) yields in an adjusted R^2 of 12.8% (on *PLGA*) or 10.86% (on *PCL*), indicating that the contact inhibition effect masked by "noise" in the data. This may also be related to the narrow range of the cell seeding densities. The use of a larger range may allow the global analysis to distinguish contact effects from natural variance in cell properties. However, there are drawbacks to the use of larger ranges, such as the introduction of seeding-density effects that mask or alter the cell-cell interactions.

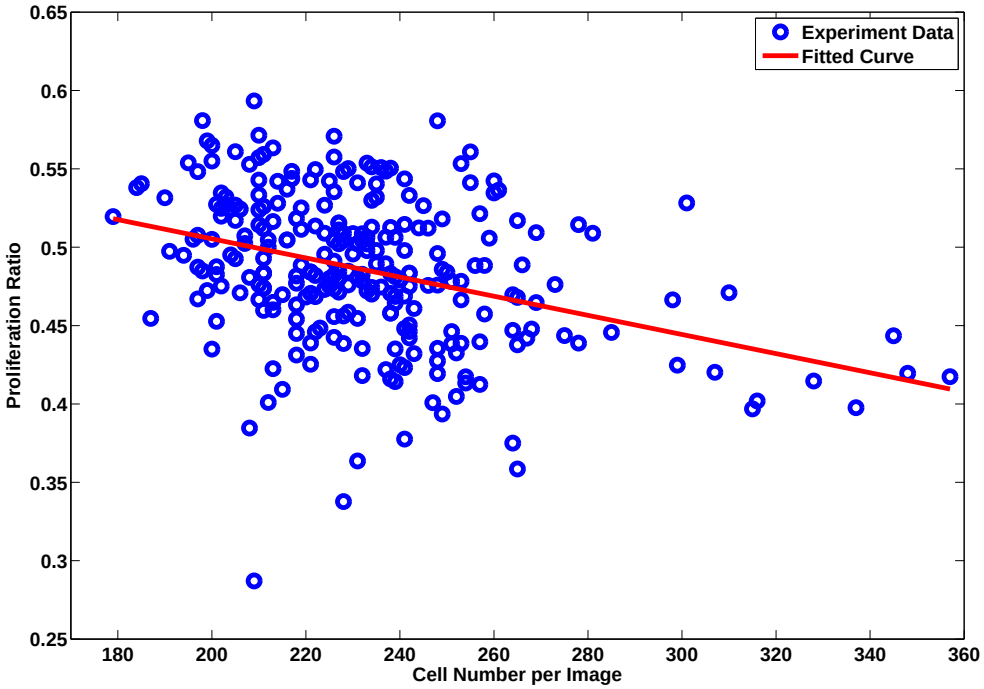
The factor importance analysis results based on global metrics are shown and discussed in Section 7.1.

5.2 Local Cell Based Metrics

For reasons discussed in Section 4.2, the noise level inherent to proliferation measurements, combined with the small seeding density range, make contact inhibition a robust test-case for comparing local `versus` global metrics. A contact phenomenon is detected when a relevant metric changes significantly relative to the data sampling noise. For global statistics, the assumed distribution (usually normal distribution) provides the random noise reference. For local metrics, the random cell-cell distance frequency distribution was calculated using



(a) *PLGA* surface



(b) *PCL* surface

Figure 12: Effects of cell density on cell proliferation for MC3T3-E1 cultured on (a) *PLGA* and (b) *PCL* surfaces

a *Monte-Carlo* approach, termed the standard frequency distribution, f_{std}. The reference f_{std} is shown in Figure 13 together with the experimental f_{AA} for MC3T3-E1 osteoblasts on *PLGA*. The profile of f_{std} is similar to a beta- or chi-distribution with asymmetry towards larger distances due to the non-overlapping nature of the nuclei centers at close distances. The computed f_{std} distribution is nearly identical to the experimental f_{AA} distribution. This is expected since f_{AA} indicates the likelihood of finding any two cells (whether proliferating or not) separated by a given distance, which should in principle be random. Figure 13 also shows the distance distribution f_{PA}, which is the likelihood of finding a proliferated cell a certain distance from any cell. If cell-cell distance has any relation to proliferating status then f_{PA} and f_{AA} should differ from one another and from f_{std}. However, the random f_{std} profile dominates all types of cell-cell distance distributions, except at very close distances ($< 80\mu m$). Hence, no noticeable effect of contact inhibition is observed unless the plot is redrawn at close distances. This has been done in Figures 14-16, which show the normalized distributions f_{PA}/f_{std}, f_{AA}/f_{std} and f_{PP}/f_{std} for the three polymer surfaces.

Figures 14- 16 indicate the non-random effects of contact inhibition when values become less than one. Contact inhibition occurs when the distance between cell nuclei becomes less than $50\mu m$. In addition, local fine structure in the contact inhibition region is seen. A local maximum peak between 10 and $20\mu m$ is observed, indicating a local enhanced proliferation at very close distance, even when overall proliferation is being inhibited. Interestingly the local peak magnitudes at $10 \sim 20\mu m$ always follow the order $f_{PP} > f_{AA} > f_{PA}$ on each of the three surfaces examined, *TCPS*, *PLGA* and *PCL*. We hypothesize that the local enhancement peak is due to two daughter cells (from the same parent cell) that are very close, not having enough time to migrate away during the BrdU staining time period. If so, then this cell division peak should appear on the f_{PP} curve but not the f_{PR} curve, which we did observe, the evidence for which is presented with the simulation data in Chapter 6 (Figure 23 (c) and (d)).

Table 2: Linear regression from global analysis results

Surface	Coeff.	SSE	R^2	$RMSE$	$Adj\,R^2$
PCL	-6.106×10^{-4}	0.5314	0.1312	0.04396	0.1280
PLGA	-5.237×10^{-4}	1.846	0.1111	0.07251	0.1086

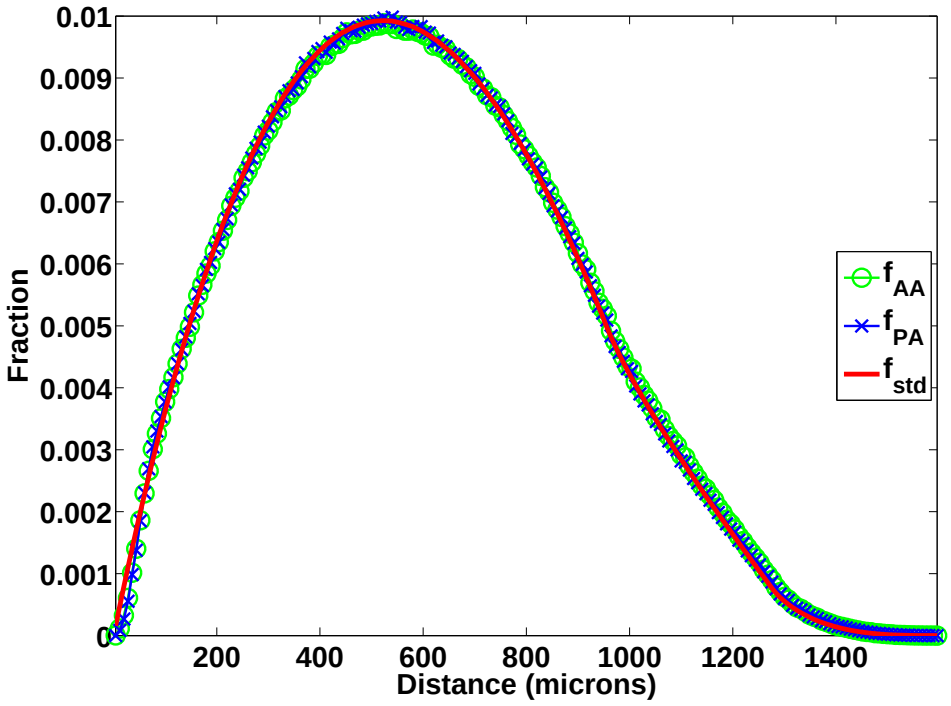


Figure 13: Comparison of the experimentally determined f_{AA} and f_{PA} for MC3T3-E1 on a *PLGA* surface and the computed standard curve f_{std}.

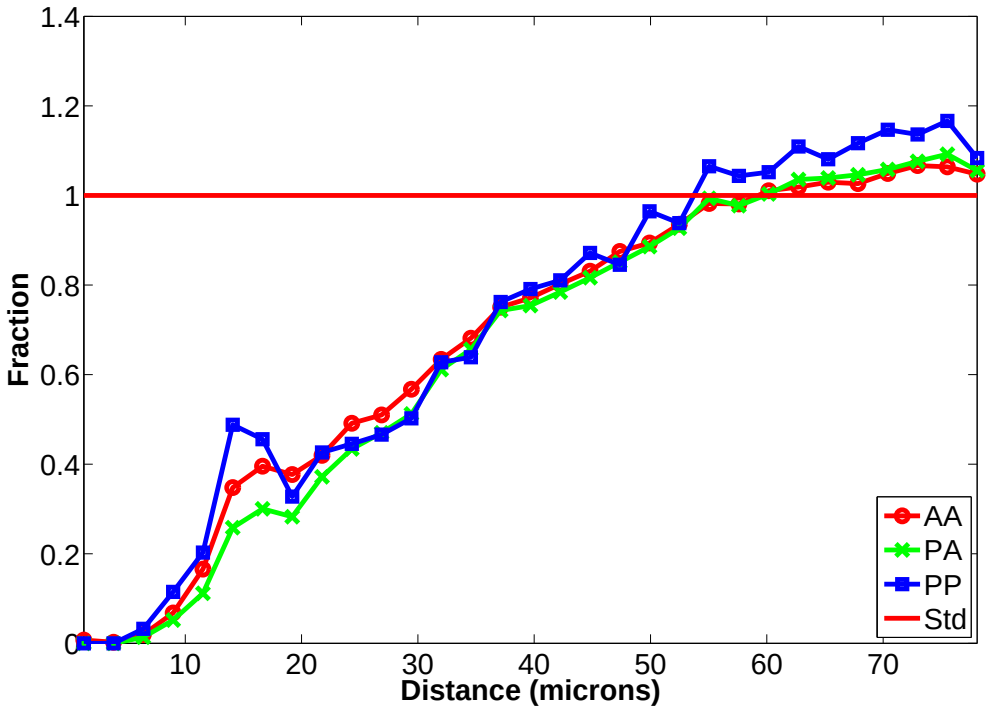


Figure 14: *PLGA* surface: $\tilde{f}_{AA}(\hat{d})$, $\tilde{f}_{PA}(\hat{d})$, and $\tilde{f}_{PP}(\hat{d})$

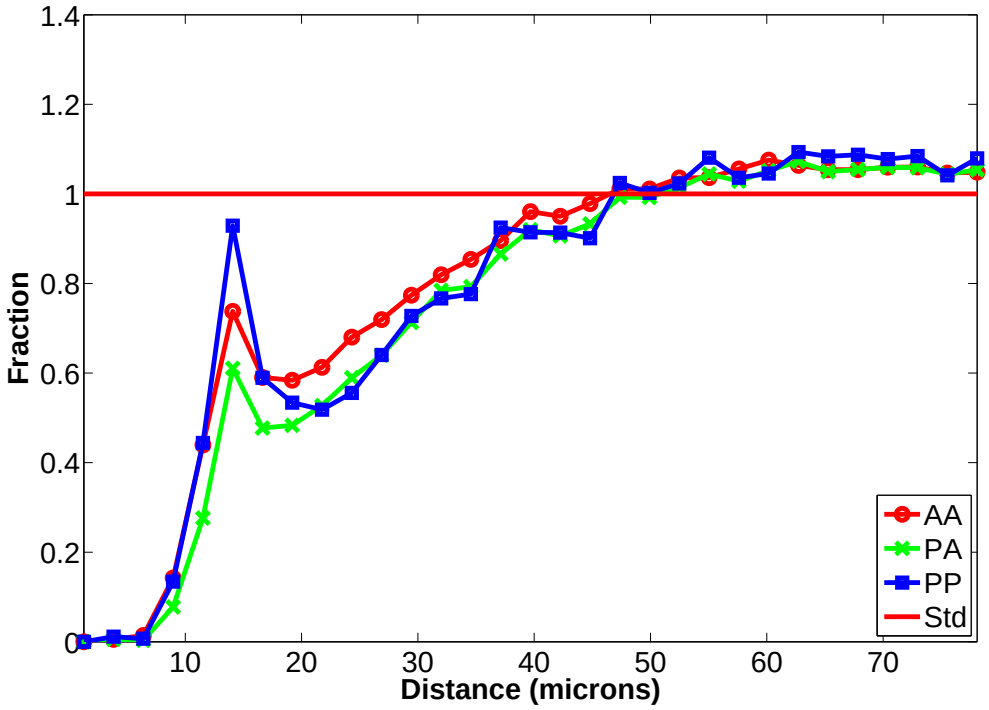


Figure 15: *PCL* surface: $\tilde{f}_{AA}(\hat{d})$, $\tilde{f}_{PA}(\hat{d})$, and $\tilde{f}_{PP}(\hat{d})$

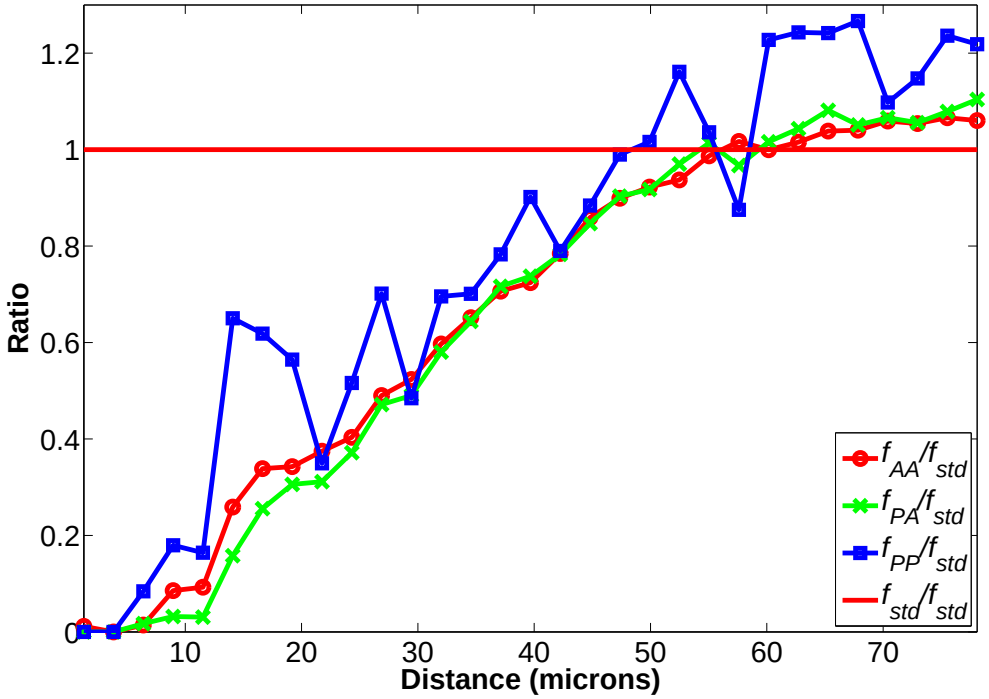


Figure 16: *TCPS* surface: $\tilde{f}_{AA}\left(\hat{d}\right)$, $\tilde{f}_{PA}\left(\hat{d}\right)$, and $\tilde{f}_{PP}\left(\hat{d}\right)$

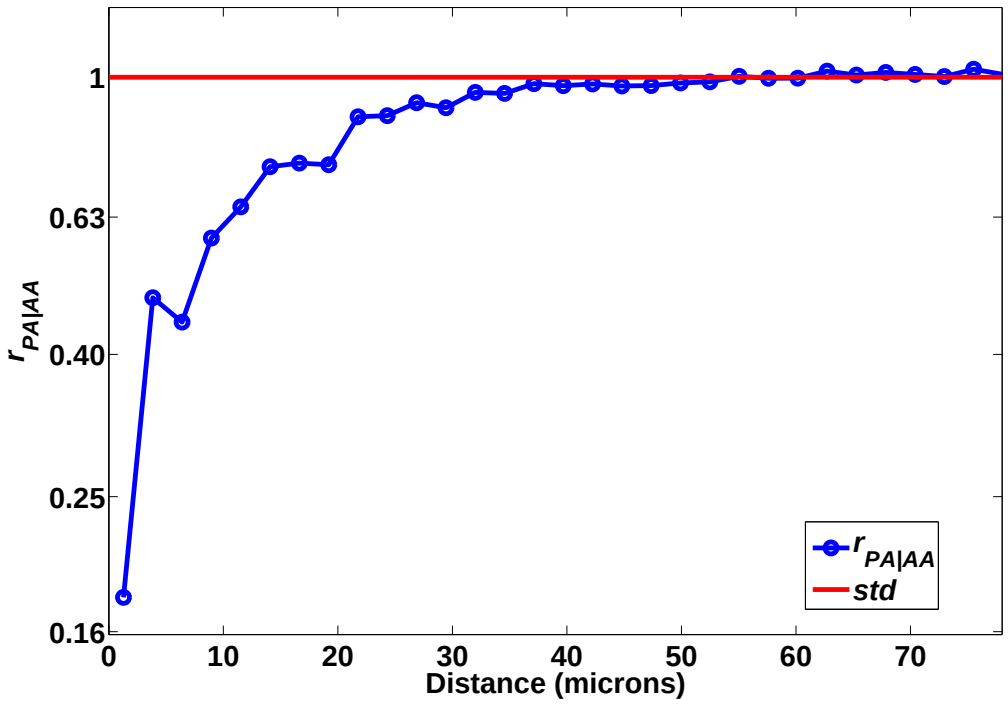


Figure 17: *PLGA* surface: localized data analysis after component decomposition: $r_{PA|AA}$.

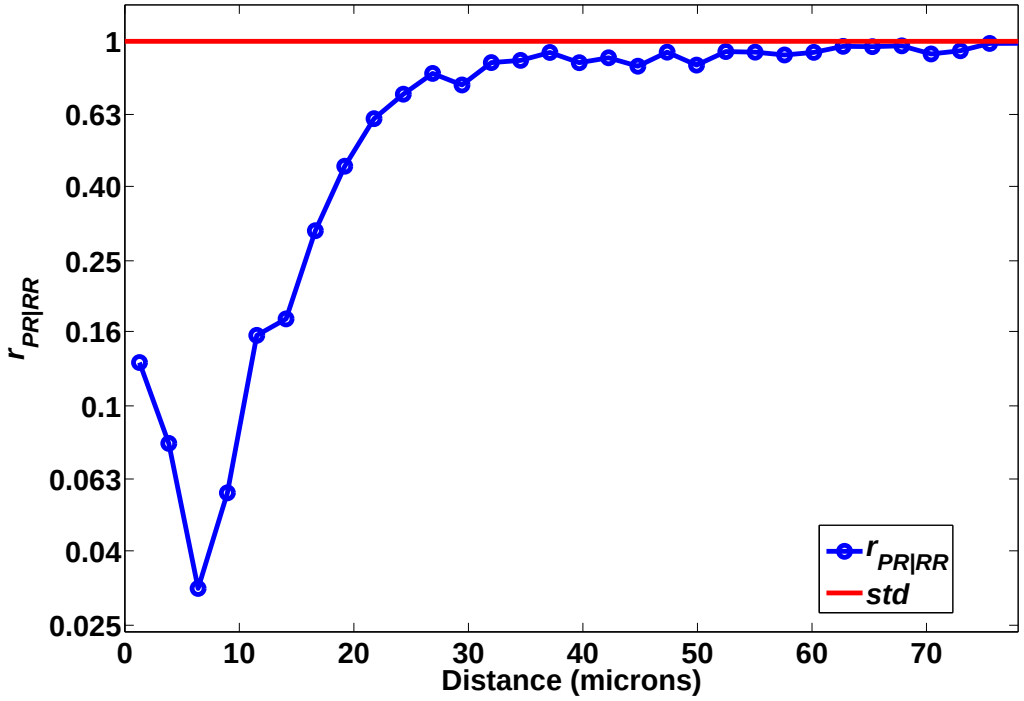


Figure 18: *PLGA* surface: model of contact inhibition of cell proliferation: $r_{PR|RR}$.

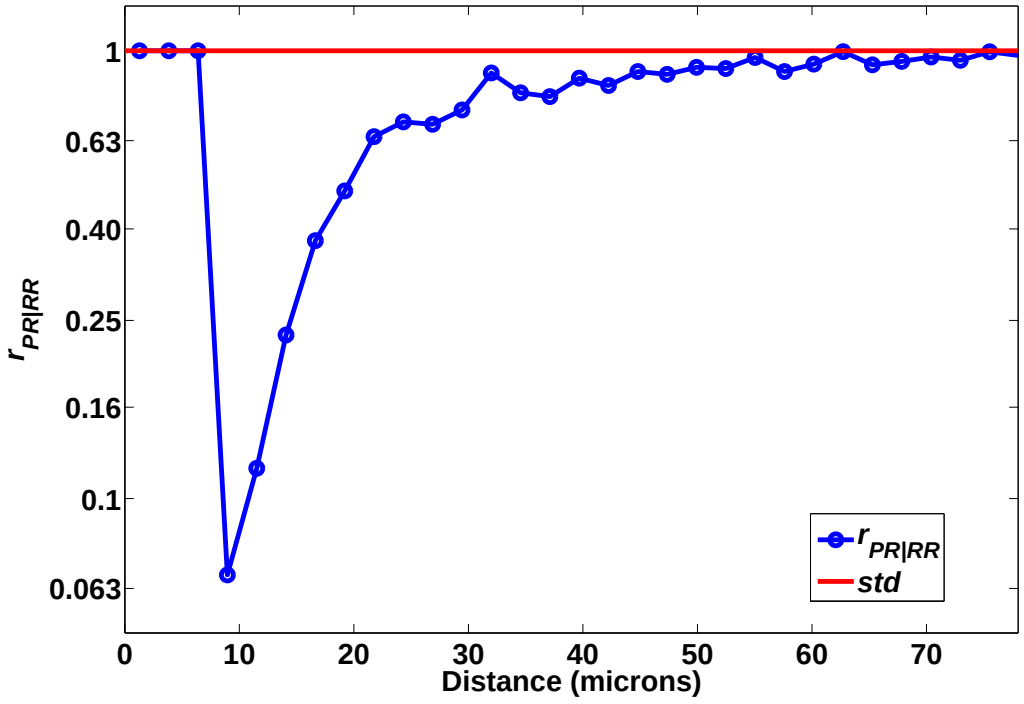


Figure 19: *PCL* surface: model of contact inhibition of cell proliferation: $r_{PR|RR}$.

Returning to the issue of detecting and quantifying contact inhibition of proliferation, direct ratios between experimentally-determined distributions are examined in Figures 17∼ 20. These are the most useful format for describing contact phenomena, since the choice of numerator and denominator can be made to isolate the phenomena of interest. Figure 17 shows the $r_{PA|AA}$ profile, which is classified into two regions: the proliferation suppressed region ($0 \sim 40\mu m$) where cell proliferation was suppressed up to 6 fold ($r_{PA|AA}$ falls to $\sim 1/6$) and the null region (beyond $40\mu m$) where cell proliferation was not noticeably affected by the contact of other cells. Based on analysis in Section 4.2 and the discussion for Figure 17, the ratio $r_{PA|AA}$ does not fully decouple the division and daughter-cell migration phenomena (indicated by PP) from the proliferation phenomena (indicated by PR), since $PP \subseteq PA$. We illustrate in details how separation of the PP and PR components enhances the detection of contact inhibition of proliferation. By definition the various distances are logically related as follows

$$PA = PP \cup PR \tag{55}$$

$$RA = RR \cup PR \tag{56}$$

$$AA = (PP \cup RR) \cup PR \tag{57}$$

with

$$PR \cap RR = \emptyset \tag{58}$$

$$PP \cap PR = \emptyset \tag{59}$$

the two shared components of AA and PA are PP and PR. Removal of the PP component from $r_{PA|AA}$ leads to $r_{PR|RR}$, shown in Figure 18 for the $PLGA$ surface. The ratio is classified into two regions: the *NaN region* (below $5\mu m$) where few pairs occur, and the *contact inhibition region* ($5 \sim 40\mu m$). In the contact inhibition region, a clear trend of decreasing probability of finding a neighboring cell is seen as the distance between cells decreases. A minimum is observed at $d_{min} = 8\mu m$, where contact inhibition effects are maximized. To

our knowledge, this is the first time both the pattern and the range of *contact inhibition of cell proliferation* have been determined quantitatively in a single function. By using the ratio of the two distance distributions, $r_{PR|RR}$, one can determine the posterior odds (PO) of proliferation at different separation distances. For example, consider two cells that are well-separated at $40\mu m$, and another two cells that are at a close distance of $8\mu m$, where the extreme in contact inhibition behavior is found. The posterior odds that one of the two cells has proliferated is

$$PO_{\frac{PP}{PR}} = \frac{r_{PR|RR}\left(\hat{d} = 40\mu m\right)}{r_{PR|RR}\left(\hat{d} = 8\mu m\right)} = 8,$$

, i.e., 8 folds lower at $8\mu m$ than at a distance of $40\mu m$.

The profiles of $r_{PP/RR}$ from the other polymer surfaces are shown in Figure 19 (PCL surface) and Figure 20 ($TCPS$ surface). The $r_{PP/RR}$ curves shows very similar shape, but with different magnitudes for the minimum point. Table 2 summarizes the variation of $r_{PR|RR,min}$ and d_{min} on different surfaces. The different location and strength of contact inhibition might be due to surface features such as roughness, crystallinity, hydrophobicity, surface charge, protein adsorption, and so on. For example, the surface roughness is increasing in the order of $TCPS$, $PLGA$, and PCL, and at the same time the $PO_{PP/PR}$ is decreasing, and the d_{min} is increasing. While purely correlative, this observation suggests that surface roughness may act to *"screen"* cell contacts, preventing as close contact as is possible on smoother surfaces. At the least, these types of comparisons suggest potentially fruitful avenues of investigation that may be of value to biomaterials research.

The sensitivity of the localized individual-cell based data analysis method for detecting cell responses to cell-to-cell distance was also evaluated by analysis of variance (ANOVA) with F-test. As shown in Table 3, the null hypothesis that cell-to-cell distance strongly affects cell behaviors is valid. The significantly large F-values suggests that the correlation of cell-to-cell distance and cell behavior is very strong, and supports that the localized individual-cell based data analysis method is very sensitive to cell-to-cell distance.

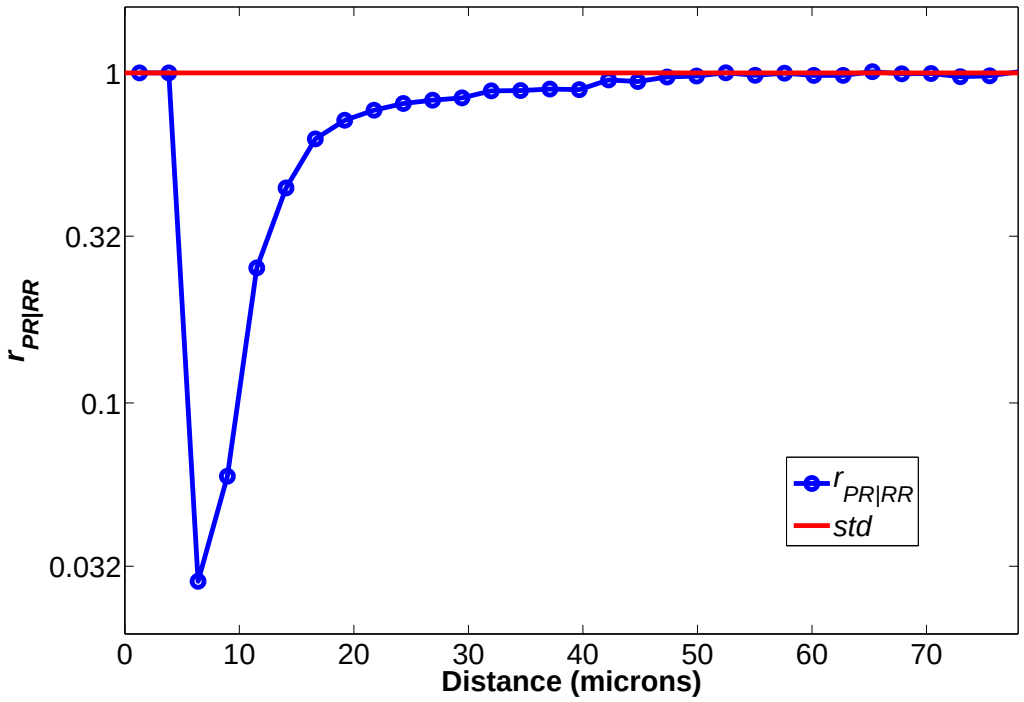


Figure 20: *TCPS* surface: model of contact inhibition of cell proliferation: $r_{PR|RR}$.

Table 3: *ANOVA* of cell-to-cell distance on *PLGA*, *PCL*, and *TCPS* surfaces

		$\tilde{f}_{AA}(\hat{d})$	$\tilde{f}_{PA}(\hat{d})$	$\tilde{f}_{PP}(\hat{d})$	$\tilde{f}_{PR}(\hat{d})$	$\tilde{f}_{RR}(\hat{d})$
PLGA	F-value	417.6	628.0	205.7	832.4	87.2
	p-value	0	0	0	0	0
PCL	F-value	241.7	444.3	257.7	364.8	101.4
	p-value	0	0	0	0	0
TCPS	F-value	759.3	728.7	99.1	726.2	546.8
	p-value	0	0	0	0	0

CHAPTER VI

MODELING AND SIMULATION OF CELL-TO-CELL INTERACTIONS

6.1 Model of Contact Inhibition

A comprehensive model of cell attachment, spatial distribution, proliferation, and division was developed. The simulations serve as a well-controlled source of data for which the interaction parameters of the (virtual) cells are set *a priori*, in order to test the effectiveness of local-feature analysis.

6.1.1 Hypotheses

Major hypotheses of this model are:

- *Self avoiding.* Cells tend to spread into void areas and avoid significant overlap with one other.
- *Contact inhibition.* A cell's proliferation is increasingly suppressed as it becomes surrounded by neighboring cells.
- *Spatial preference of cell division.* Once a cell has proliferated, the daughter cell is placed in the neighboring position that allows the most unoccupied space.

These concepts form a minimum set of constraints required to achieve cell-to-cell distance distributions in agreement with experimental data. In order to keep the model simple, other phenomena such as migration, spreading, morphology, and viability were not considered. Each of the three primary hypotheses are associated with one or more adjustable parameters and an assumed probability distribution from which cell positions are sampled in an unbiased manner using *Monte Carlo* methods.

6.1.2 Assumptions

Assumptions used in the simulations are:

- The stochastic variation in behaviors of individual cells can be simulated by a *MC* method, involving sampling of cell configurations from an assumed distribution function.
- The distribution function, describing the distance dependence of the three cell-cell interaction hypotheses above, is given by a multivariate normal distribution.
- The centroid positions of cell nuclei are a valid metric for describing cell-cell interactions.
- The contact inhibition effects can be well illustrated by a *one-round* simulation

According to the experimental settings, the last assumption obviously over-simplified the simulation by using *static* and *time independent* processes to mimic the *dynamic* and *history dependent* cell responses observed in experiments. That is, experimental data represents a collection of cells of different cell cycle phases with a history of multiple cell cycles. Cell-to-cell distance distribution $\tilde{f_{AA}}\left(\hat{d}\right)$ and $\tilde{f_{RR}}\left(\hat{d}\right)$ shows not only the effects of self avoidance, but also the effects of contact inhibition and spatial preference of cell division happened after the seeding, so do $\tilde{f_{PA}}\left(\hat{d}\right)$, $\tilde{f_{PR}}\left(\hat{d}\right)$, and $\tilde{f_{PP}}\left(\hat{d}\right)$. *Multi-round* approaches such as *Markov chain MC* simulation should represent the characteristics of the cumulative effects of cell behaviors better than the *one-round* strategy used in this research. However, since the multi-simulation is very time-consuming, and only the effects of contact inhibition of cell proliferation are of interest in this study, the one-round simulation strategy is preferable as long as the contact inhibition effects can be well mimicked.

The evaluation of the *one-round simulation* assumption is very important for assuring the simulation is valid.

6.2 Monte Carlo Simulation

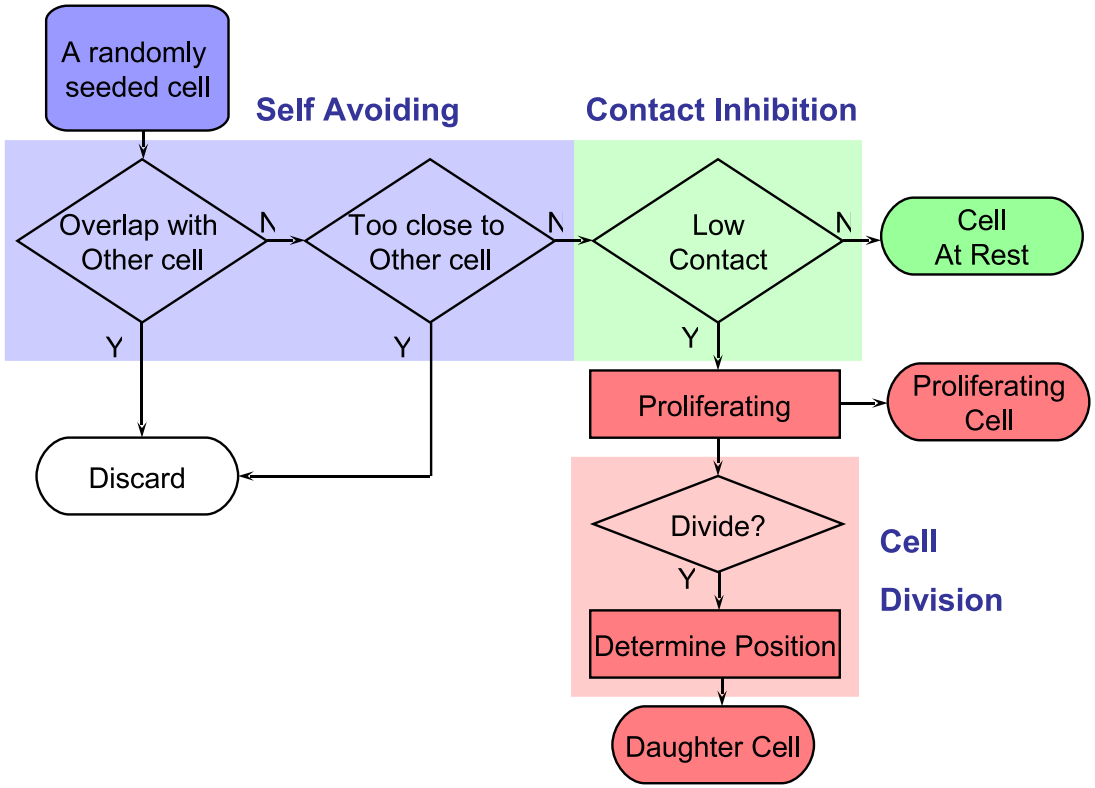


Figure 21: Flowchart of Simulation of Cell Contact Inhibition.

The strategy of simulation of cell contact inhibition based on these concepts and assumptions is summarized in the flowchart in Figure 21. The simulation was performed by three steps.

Step I: Self avoiding Cell position was simulated using the self-avoiding hypothesis. Totally N_{Cell} cells were seeded at positions chosen randomly from a uniform distribution into a rectangle with area $w \times l$. Actual values employed for this paper were $N_{cell} = 200$, $l = 1600 pixels$, and $w = 1200 pixels$, representing the actual area of a microscope image from our experimental data, $1.06 mm^2$, or $1189 \mu m$-by-$892 \mu m$. The insertion was rejected if a cell overlaps with another cell at distance dis according to a probability of avoidance, p_{avoid}, given below.

$$p_{avoid} = \begin{cases} N_{avoid}\left(dis_{min} | \mu_{avoid}, \sigma_{avoid}\right) & \text{if } dis_{min} > dis_{excl} \\ 1 & \text{if } dis_{min} \leqslant dis_{excl} \end{cases} \tag{60}$$

$N_{avoid}\left(dis | \mu_{avoid}, \sigma_{avoid}\right)$is an assumed normal distribution of cell-to-cell distances, dis, with mean (μ_{avoid}) and variance (σ_{avoid}), and dis_{excl}, the excluded distance, is the shortest allowable distance between two cells. Hence, the self avoiding effect is determined by three adjustable parameters μ_{avoid}, σ_{avoid}, and dis_{excl}. According to the *Metropolis Monte Carlo algorithm*, a random number r, known as the value of a dice, was generated from a uniform distribution and compared with provide. If $r < p_{avoid}$, this cell was removed, otherwise it remained at its randomly-chosen position. This process was repeated for each virtual cell. In each simulation, 360 virtual images were generated, and all the properties discussed below were averaged from these.

Step II: contact inhibition Cell proliferation was simulated based on the contact inhibition hypothesis. A multivariate normal distribution, $MVN_{inhibit}\left(pos | pos_o, \Sigma_{inhibit}\right)$, was defined for addressing the effects on cell proliferation of both cell-to-cell distance and local cell density around a given cell. For the given i^{th} cell , the effects of its neighbor cells was taken to be an inhibition probability, $p_{inhibit}$, which was

determined based on the position of the cell ($\overrightarrow{pos_i}$) and the possibility distribution $MVN_{inhibit}\left(pos|pos_o, \Sigma_{inhibit}\right)$:

$$p_{inhibit} = \frac{\sum_{j}^{j\neq i} MVN_{inhibit}\left(\overrightarrow{pos_j}|\overrightarrow{pos_i}, \Sigma_{inhibit}\right)}{MVNCDF_{inhibit}\left(A_{image}|\overrightarrow{pos_i}, \Sigma_{inhibit}\right)} \cdot \frac{A_{image}}{\sum_{j}^{j\neq i} 1} \cdot C \tag{61}$$

where

$$\Sigma_{inhibit} = \begin{pmatrix} \sigma_{inhibit} & 0 \\ 0 & \sigma_{inhibit} \end{pmatrix} \tag{62}$$

describes the sensitivity of cell proliferation to the cell-to-cell distance, without spatial preference, $MVNCDF_{inhibit}\left(A_{image}\right)$ is the cumulative distribution function of $MVN_{inhibit}\left(pos|pos_o, \Sigma_{inhibit}\right)$ and A_{image} is the area of the observing window ($1.92 \times 10^6 pixels$), "$\sum_j^{j\neq i} 1$" is the number of all neighbor cells of the given i^{th} cell in the given virtual image, and C is the normalization constant. The actual probability of proliferation was determined by

$$Prol_{actual} = e^{-p_{inhibit}} \cdot Prol_o \tag{63}$$

where $Prol_o$ is the probability of proliferation without effect of contact inhibition.

Step III: spatial preference of cell division The positions of daughter cells were determined by the hypothesis of *spatial preference*. A portion of cells in proliferation status would have divided during observation time window. Whether a cell would have divided or not was determined by comparing $1/r_{div}$[1] and a random number r with the *MC* method mentioned previously. For the i^{th} cell chosen in Step II to divide, the positions of the two daughter cells are determined as follows. One of the two daughter cells was assumed to occupy the position of the original cell. From this position, a base direction ($theta_o$) was randomly selected, and six candidate dividing

[1]The symbol r_{div} represents the division ratio

directions ($theta_j$ where $j = 1, 2, \ldots, 6$) derived from the base direction $theta_o$,

$$theta_j = theta_o + (i-1)\frac{1}{3}\pi \qquad (j = 1, 2, \ldots, 6) \tag{64}$$

were determined. Along each candidate direction $theta_j$, a candidate position, $pos_{div,i,j}$, for the other daughter cell was determined by both a random distance $dis_{div,i,j}$ which followed the normal distribution $N_{div}(dis|\mu_{div}, \sigma_{div})$ and the candidate direction $theta_j$ itself. The *"strength of contact inhibition of cell division"*, $p_{inhibit,div,i,j}$, were determined for of all six candidate positions $pos_{div,i,j}$ ($j = 1, 2, \ldots, 6$), and the candidate position of the least strength of contact inhibition was chosen. This procedure is similar in nature to the continuum configurational bias technique often used in *MC* simulation of polymers and large molecules [30][31].

6.3 Data Analysis

As previously discussed in Section 5.2, local cell feature metrics based on distance histograms are used to describe cell-cell contact phenomena. The normalized cell distance histograms, namely $\tilde{f}_{AA}(\hat{d})$, $\tilde{f}_{PA}(\hat{d})$, $\tilde{f}_{PP}(\hat{d})$, $\tilde{f}_{PR}(\hat{d})$, and $\tilde{f}_{RR}(\hat{d})$, were obtained from the simulated virtual images as well as experimental data. Briefly, these distributions are the probability of locating two cells separated by a distance dis when the two cells are in proliferative status P (proliferated), R (resting), and A (any randomly selected cell). Parameters used in the simulation were adjusted to best match experimental data. The coefficient of determination, $\tilde{f}_{AA}(\hat{d})$, was used for evaluating parameter fitting. The distribution was used for determining parameters for Step I, and $\tilde{f}_{PP}(\hat{d})$ for parameters for Step II and III. The parameters are summarized in Section 6.4.

6.4 Results and Discussion

6.4.1 Standard Curve for Normalization

The *Monte Carlo* simulation results of $\tilde{f}_{std}(\hat{d})$, achieved from 1×10^{10} pairs of points on a 1600-pixel-by-1200-pixel virtual image, is shown in Figure 22. As shown in Figure 22,

the random effects due to the *Monte Carlo* approach are negligible as the curve $\tilde{f}_{std}\left(\hat{d}\right)$ is smooth, that is, the normalization process will not introduce noticeable artificial effects or noise. The extremely large data space (1×10^{10} pairs of points) used in this process is necessary for a reliable and robust normalization of experimental or computational data at narrow range.

6.4.2 Determining Parameters

As mentioned in Section 6.3, after a series of try-and-error, the final values of parameters were optimized and listed in Table 4. These parameters were determined by fitting experimental curves ($\tilde{f}_{AA}\left(\hat{d}\right)$, $\tilde{f}_{PA}\left(\hat{d}\right)$, $\tilde{f}_{PP}\left(\hat{d}\right)$, $\tilde{f}_{PR}\left(\hat{d}\right)$, and $\tilde{f}_{RR}\left(\hat{d}\right)$) taken from proliferation assays of MC3T3-E1 cells on *PLGA*, *PCL*, and $TCPS$ surfaces. For example, the process for determining parameters for simulating cell behaviors on *PLGA* surfaces were shown in Figure 23. First of all, dis_{excl} was determined according to (Figure 23 (a)). The standard deviation σ_{avoid}, which address the effective range of cell self-avoiding interaction, was determined so that the $\tilde{f}_{AA}\left(\hat{d}\right)$ curve from simulation matched best with that from experiment data. A significant difference between the simulated data and the original observation is the peak between $10 \sim 20\mu m$.

6.4.3 Simulation of Contact Inhibition on *PLGA* Surfaces

The range of contact inhibition (indicated by $\sigma_{inhibit}$) was determined by matching the right-side part of the curves (beyond $25\mu m$ in this case) shown in Figure 23 (c), and the normalization constant C was adjusted to control the steepness of the contact inhibition curve. Afterward, according to the location and width of the sharp peak of cell division (between 8 and $25\mu m$), μ_{div} and σ_{div} were determined. Finally, r_{div} was determined so that the height of the peaks of cell division matched.

The simulated $\tilde{f}_{PA}\left(\hat{d}\right)$, $\tilde{f}_{PP}\left(\hat{d}\right)$ and $\tilde{f}_{PR}\left(\hat{d}\right)$ curves (Figure 23 (b), (c), and (d)) fitted the experimental ones with high R^2 values (97%, 92%, and 99%, respectively), and the residue analysis showed no obvious patterns. This suggests that the cell contact inhibition model

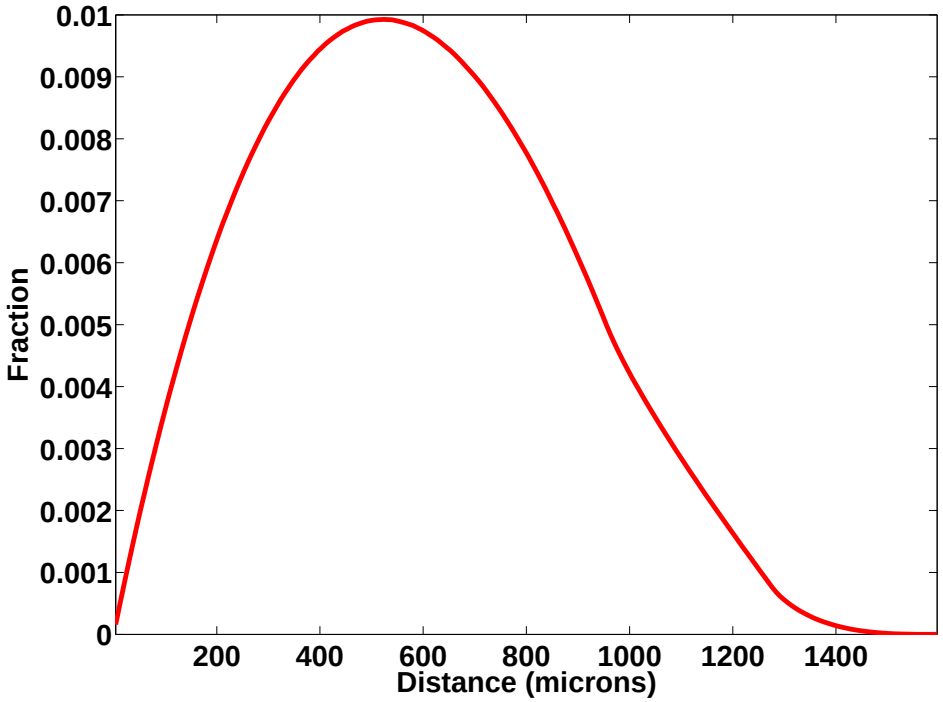


Figure 22: The normalized random distribution curve of cell-cell distance $\tilde{f}_{std}\left(\hat{d}\right)$

Table 4: Parameters used in simulation

Surface	$dis_{excl}(\mu m)$	$\sigma_{avoid}(\mu m)$	r_{div}	$\sigma_{inhibit}(\mu m)$	$\mu_{div}(\mu m)$	$\sigma_{div}(\mu m)$
PLGA	7.2	112	1/200	200	10	3
PCL	7.2	56	1/140	120	10	3
TCPS	7.2	112	1/200	200	11	3

and the simulation strategy were successful.

6.4.4 Validation of Assumptions on *PLGA* Surfaces

Although the R^2 value of the $\tilde{f}_{AA}\left(\hat{d}\right)$ case (97%) was high for simulation, a significant peak of residues was found between 10 and 20μm (Figure 23 (a)). This effect was much more significant in the $\tilde{f}_{RR}\left(\hat{d}\right)$ case (Figure 23 (e)), which showed a poor R^2 value (71%). As emphasized in Section 6.1.2, this phenomenon is expected and is due to that the simulation method is *static* and *time independent*, but the real cell responses are *dynamic* and *history dependent*. That is, the $\tilde{f}_{AA}\left(\hat{d}\right)$ and $\tilde{f}_{RR}\left(\hat{d}\right)$ distributions determined from experimental data were cumulative. Cell division during cell culture history, which would contribute to the peak between 10 and 20μm, was not recorded on the $\tilde{f}_{RR}\left(\hat{d}\right)$ curve, and only partially recorded on $\tilde{f}_{AA}\left(\hat{d}\right)$.

The simulation data show that the temporal effect was negligible in the simulation for contact inhibition. The high R^2 values of simulation of $\tilde{f}_{PA}\left(\hat{d}\right)$, $\tilde{f}_{PP}\left(\hat{d}\right)$ and especially $\tilde{f}_{PR}\left(\hat{d}\right)$, which indicated the contact inhibition effects, suggested that such simplification is valid.

6.4.5 Modeling and Simulation on *PCL* and *TCPS* Surfaces

The simulation of cell behaviors on *PCL* and *TCPS* surfaces, as shown in Figure 24 and Figure 25, also shows high R^2 values for $\tilde{f}_{PA}\left(\hat{d}\right)$, $\tilde{f}_{PP}\left(\hat{d}\right)$ and $\tilde{f}_{PR}\left(\hat{d}\right)$ curves (> 95%).

The comparable simulation quality on the *PLGA*, *PCL*, and *TCPS* surfaces suggests the hypotheses of the contact inhibition model and the assumptions used in the simulation are independent to the three surfaces investigated. Therefore, the cell model and simulation strategy developed from this study are potentially applicable for general contact inhibition phenomena on some other surfaces.

6.4.6 Explanation of Parameters

In Table 4, the parameter dis_{excl} suggests that most cells are at least $7.2\mu m$ apart after seeding, and σ_{avoid} implies more than 84% of self avoidance phenomenon happens when two cells are closer than $112\mu m$ on *PLGA* and *TCPS* surfaces, and $56\mu m$ on *PCL* surface. The value of r_div, varied from 1/140 to 1/200, implies that around 1/70 or 1/100 cells with positive BrdU staining were divided cells, which is reasonable when comparing the short BrdU staining time (6hr) with the cell cycle time. The physical meaning of $\mu_{div} = 10 \sim 11\mu m$ and $\sigma_{div} = 3\mu m$ is that the average nucleus-to-nucleus distance of two daughter cells just after division on surface is $10 \sim 11\mu m$ with a standard variation of $3\mu m$. The parameter $\sigma_{inhibit} = 120 \sim 200\mu m$ suggests that at a significant level $\alpha = 0.05$, contact inhibition of growth can be neglected when two cells are $normcdf^{-1} = 329\mu m$ apart for *PLGA* and *TCPS* cases and $197\mu m$ apart for *PCL* cases.

The dis_{excl} values larger than the defragmentation threshold (less than $3\mu m$) during image processing suggest a *hard self-avoiding* effect defined as the "cut-off" distance that almost no cell-to-cell distance shorter than that distance was observed. The similar dis_{excl} value ($7.2\mu m$) on the three different surfaces implies that the "soft self-avoiding effects" is not surface dependent for adhesion-dependent cells. Thus, when two cell nuclei are very close to each other, the effects of surface properties are overwhelmed by the self-avoiding effect to avoid overlapping of nuclei.

The distance between two daughter cells, represented by $N_{div}(dis|\mu_{div}, \sigma_{div})$, also shows surface-independence on the three polymer surfaces. This observation suggests the distance between two centrosome in the metaphase during mitosis is highly reserved.

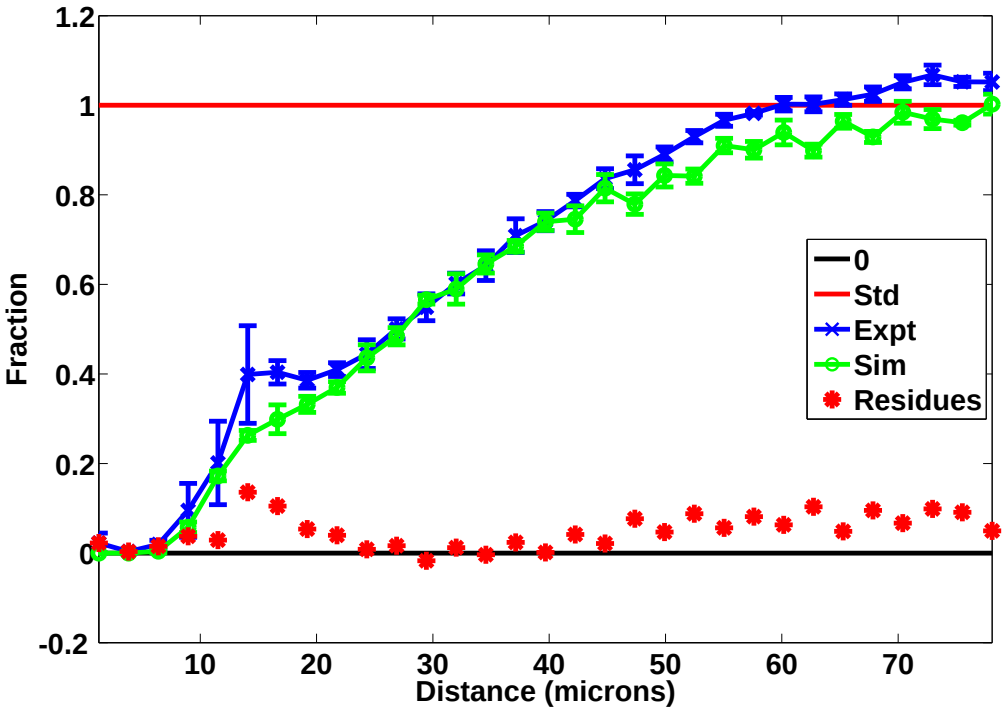
The results also imply that cells on smooth, amorphous *PLGA* and *TCPS* surfaces sense farther than those on rough, crystalline *PCL* surfaces. The *effective distance* of the *"soft self-avoiding"* effect, represented by σ_{avoid}s, are two folds longer in *PLGA* and *TCPS* cases than that in *PCL* cases, so do the *effective distance* of the *contact inhibition* effects (represented by $\sigma_{inhibit}$). This discovery is consistent to the discovery of other researchers

that cells are better spread (cover larger area)[93] and less polar [43, 42, 93] on smooth surfaces than on rough surfaces.

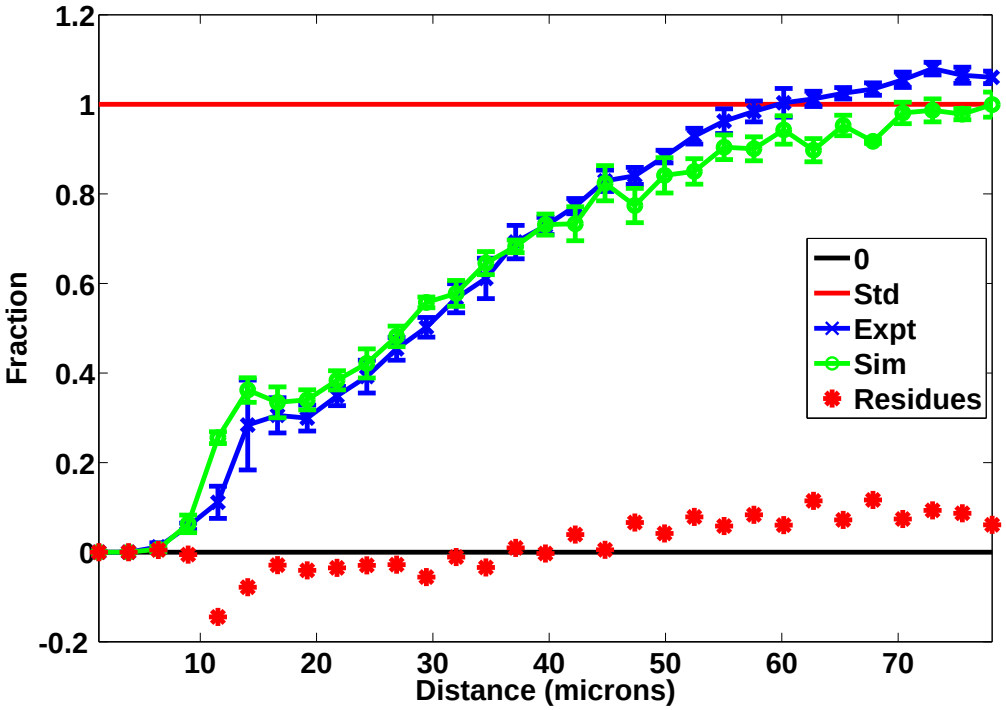
6.5 Summary

For $\tilde{f}_{AA}\left(\hat{d}\right)$, $\tilde{f}_{RR}\left(\hat{d}\right)$, and $\tilde{f}_{RR}\left(\hat{d}\right)$ cases, the simulation results well matched the experiment data, which proved the effectiveness of the hypotheses and assumptions used in the modeling. However, noticeable difference in details suggested this static and temporal independent model can be further enhanced by introducing multi-round approaches such as Markov chain *MC* simulation if other cell behaviors are also of interest.

The model and simulation worked well on three different polymer surfaces (*PLGA*, *PCL*, and *TCPS*), which suggested the cell model and the simulation method is potentially applicable to other adhesion-dependent cell lines and other surfaces. The simulation parameters determined by curve fitting are potential descriptors of characteristics of cell behaviors and effects of surface properties. Further study is required for evaluating the effectiveness of the model and simulation strategies on other cell lines and surfaces as well as the possible applications of simulation parameters in surface screening, mechanism study, tissue engineering and medical diagnosis.

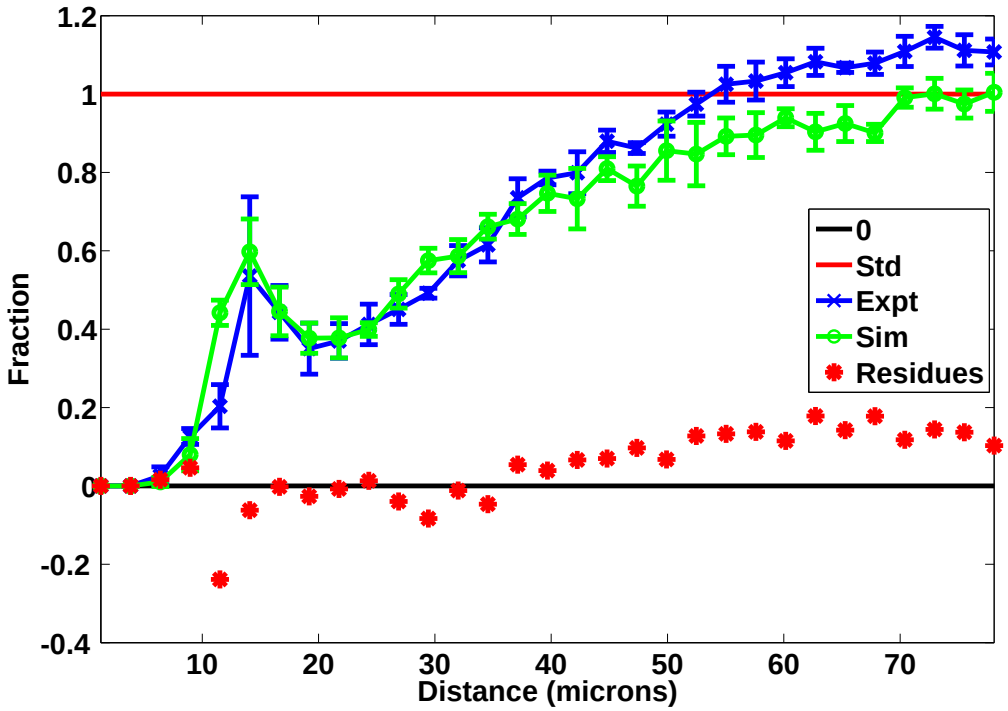


(a) $\tilde{f}_{AA}\left(\hat{d}\right)$, $R^2 = 0.9681$

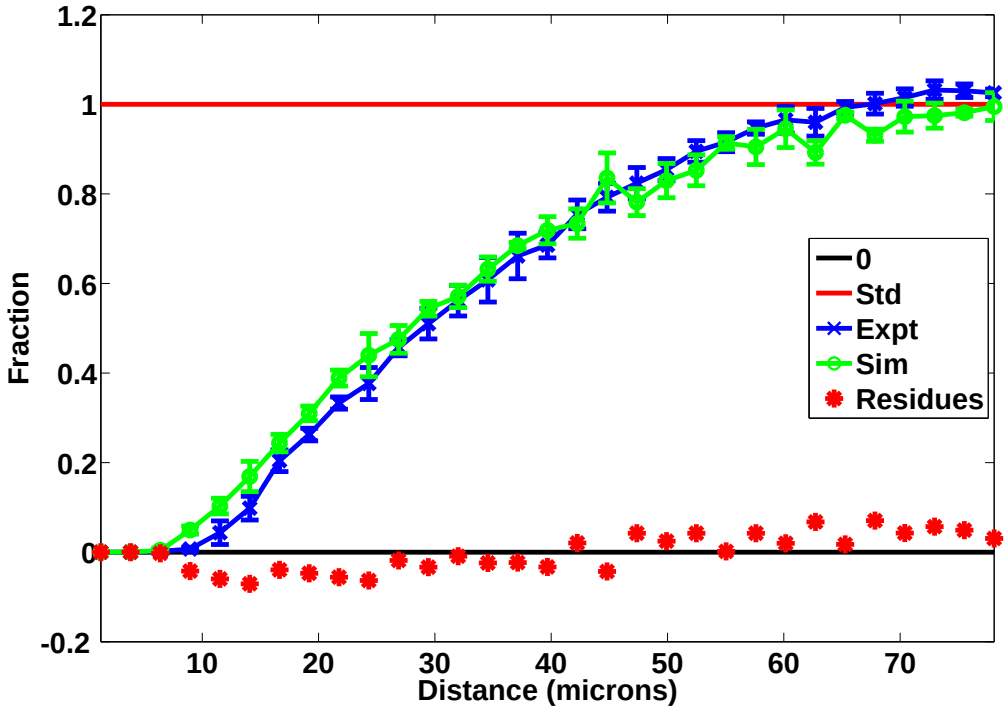


(b) $\tilde{f}_{PA}\left(\hat{d}\right)$, $R^2 = 0.9701$

Figure 23: Simulation of contact inhibition on PLGA surface (blue: experiment data; green: simulation result; red dots: residues. $Prol_o = 0.45$)

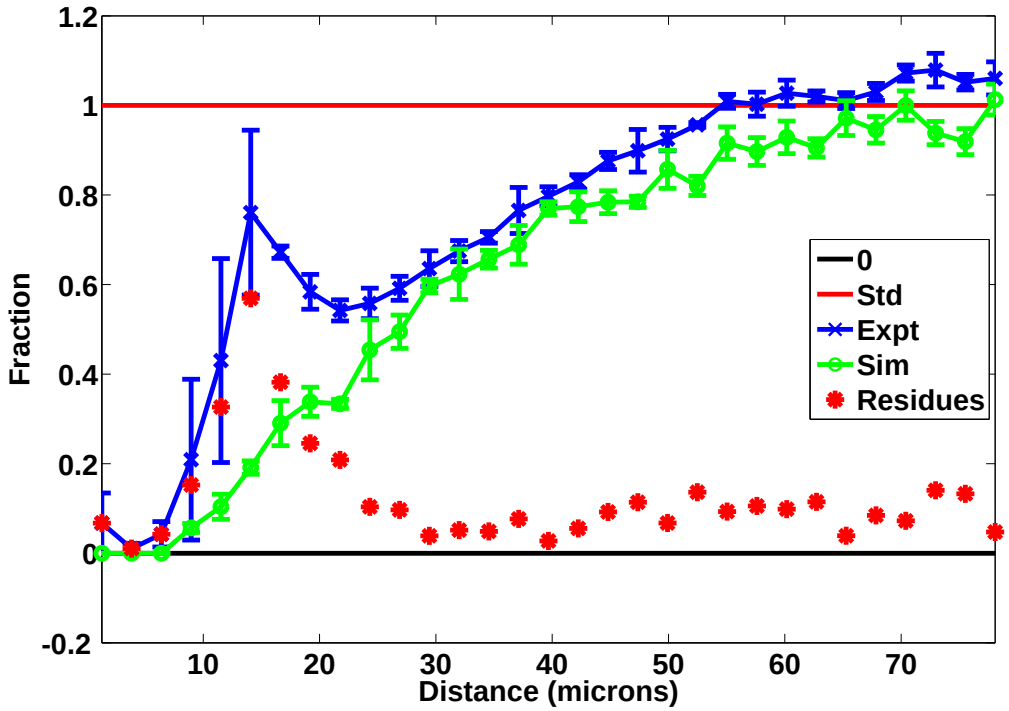


(c) $\tilde{f}_{PP}\left(\hat{d}\right)$, $R^2 = 0.9248$



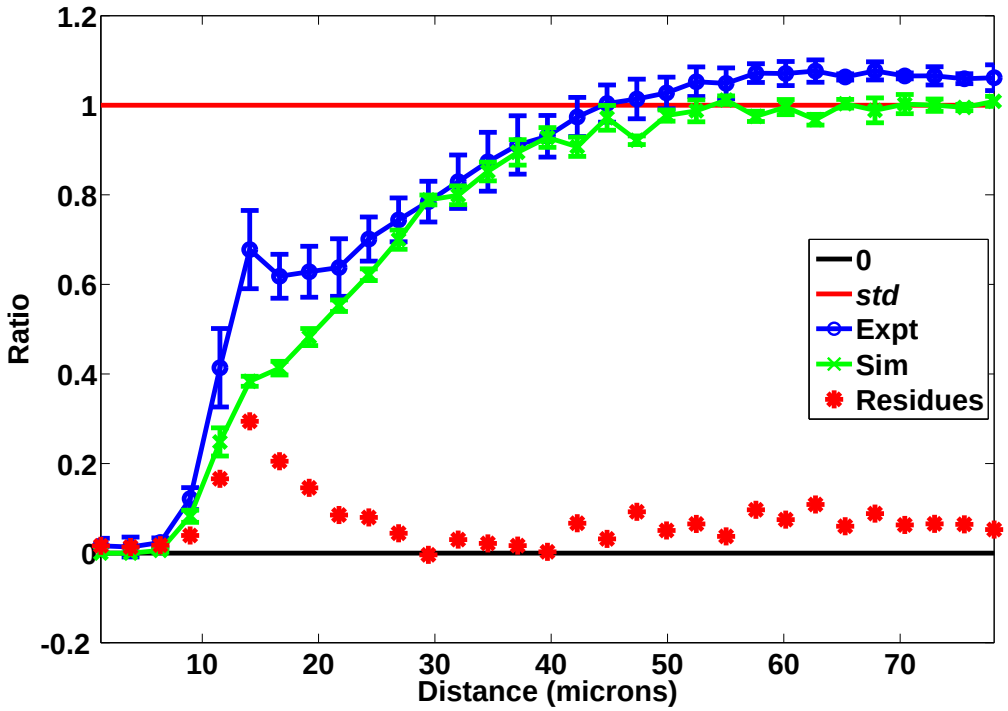
(d) $\tilde{f}_{PR}\left(\hat{d}\right)$, $R^2 = 0.9877$

Figure 23 (continued): Simulation of contact inhibition on PLGA surface

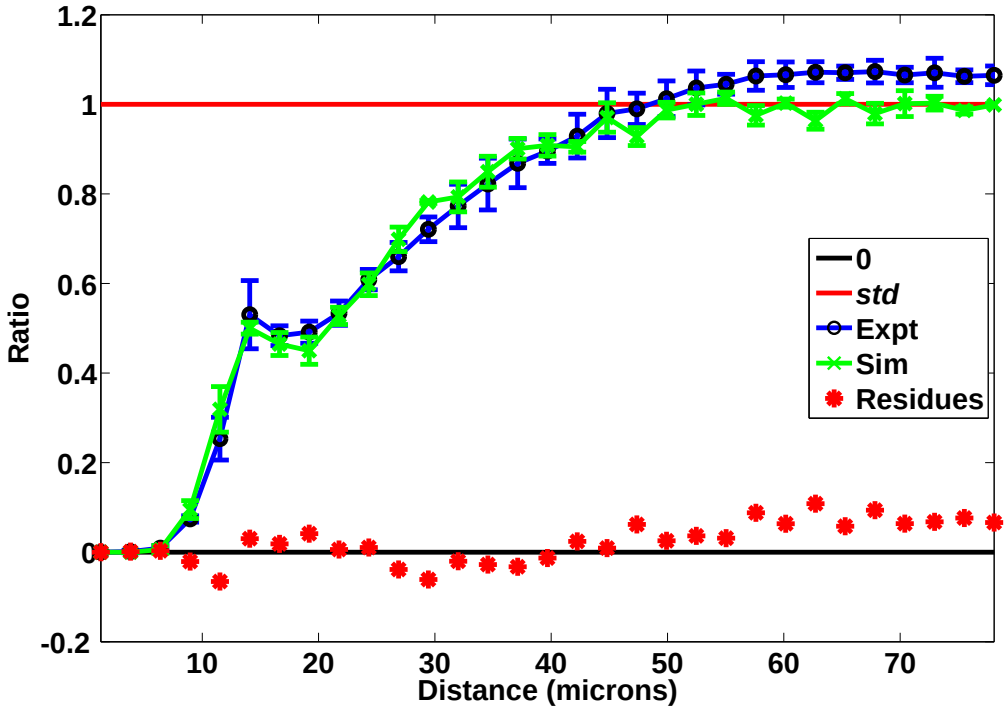


(e) $\tilde{f}_{RR}\left(\hat{d}\right)$, $R^2 = 0.7054$

Figure 23 (continued): Simulation of contact inhibition on PLGA surface

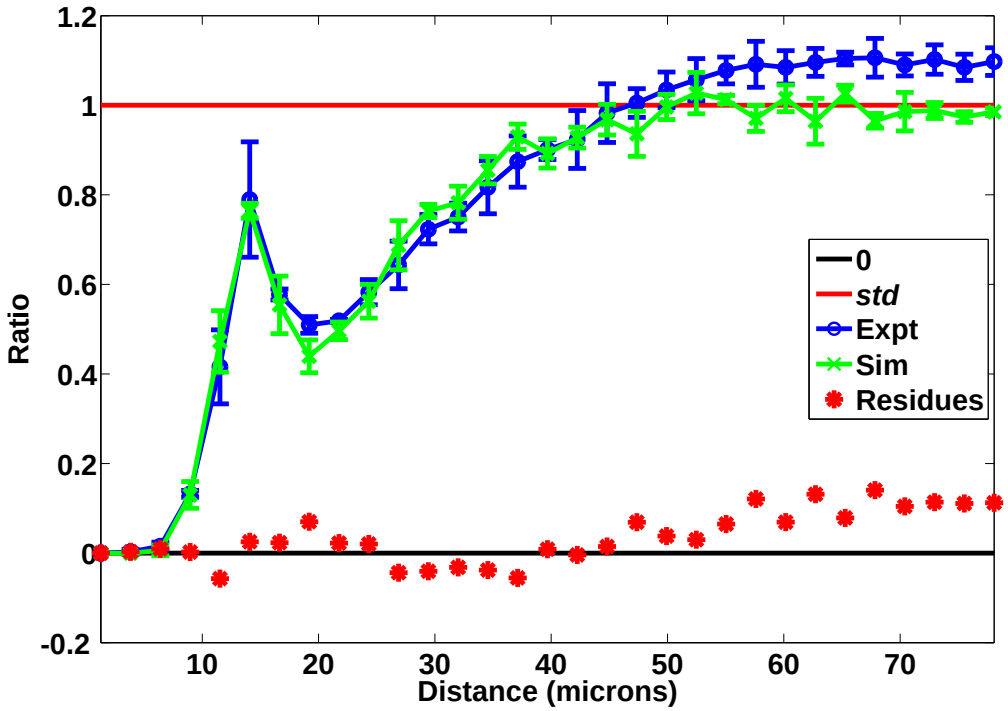


(a) $\tilde{f}_{AA}\left(\hat{d}\right)$, $R^2 = 0.9231$

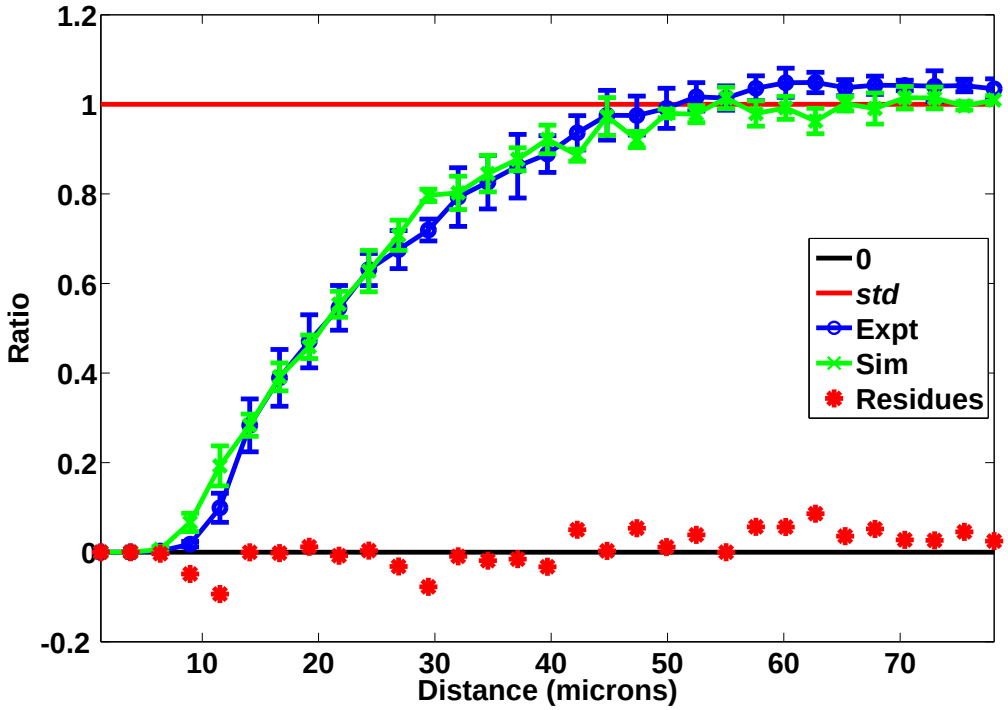


(b) $\tilde{f}_{PA}\left(\hat{d}\right)$, $R^2 = 0.9602$

Figure 24: Simulation of contact inhibition on *PCL* surface (blue: experiment data; green: simulation result; red dots: residues. $Prol_o = 0.45$)

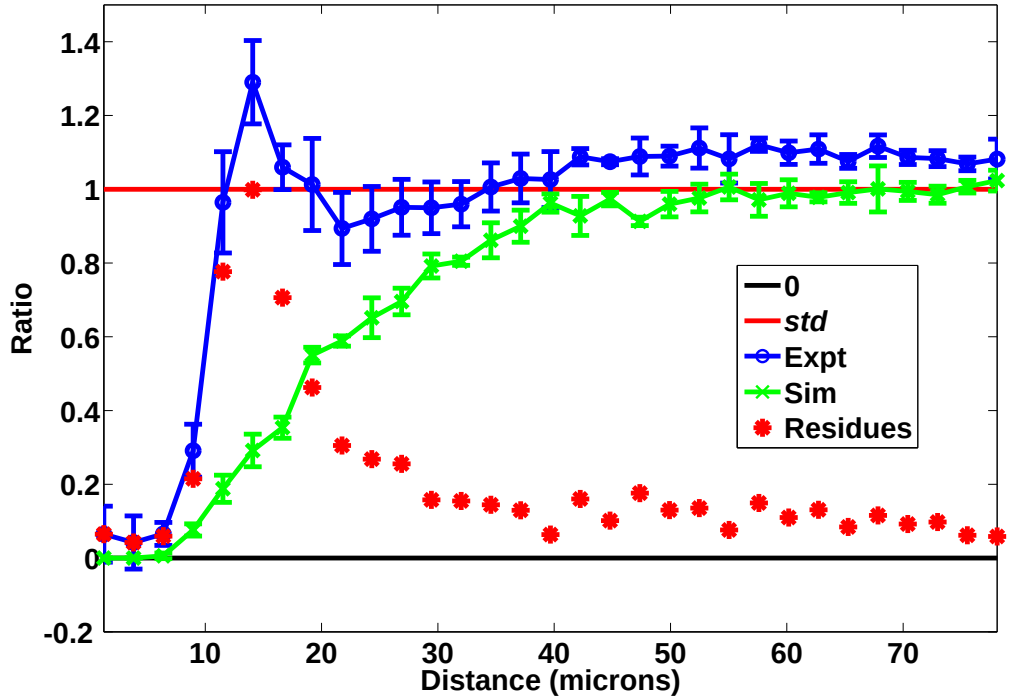


(c) $\tilde{f}_{PP}\left(\hat{d}\right)$, $R^2 = 0.9631$



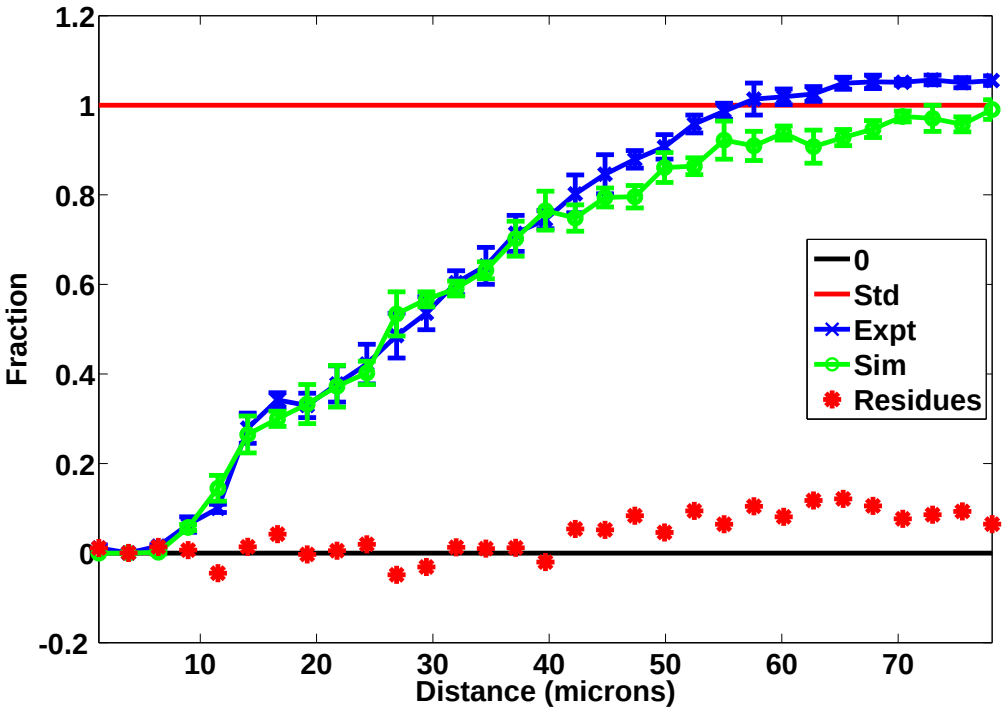
(d) $\tilde{f}_{PR}\left(\hat{d}\right)$, $R^2 = 0.9885$

Figure 24 (continued): Simulation of contact inhibition on *PCL* surface

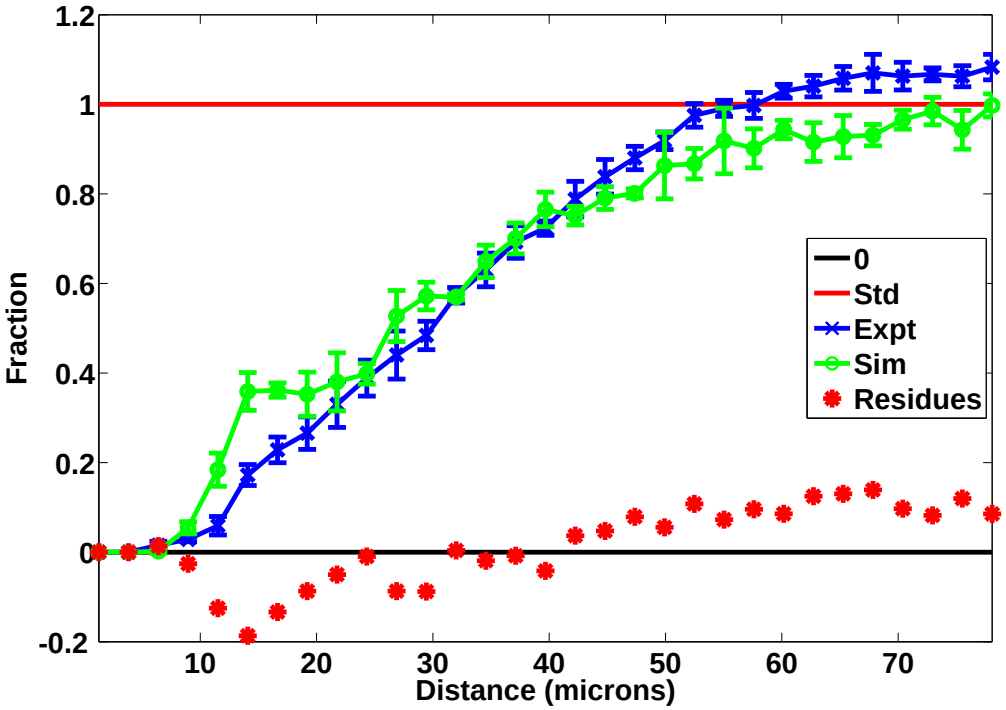


(e) $\tilde{f}_{RR}\left(\hat{d}\right)$, $R^2 = 0.1151$

Figure 24 (continued): Simulation of contact inhibition on *PCL* surface

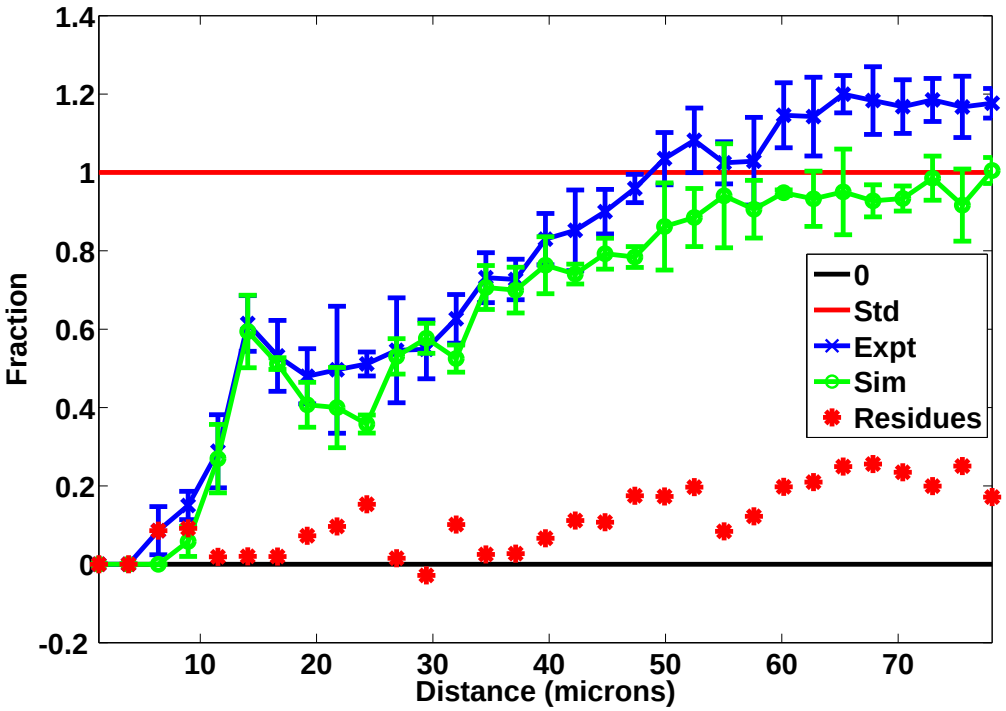


(a) $\tilde{f}_{AA}\left(\hat{d}\right)$, $R^2 = 0.9712$

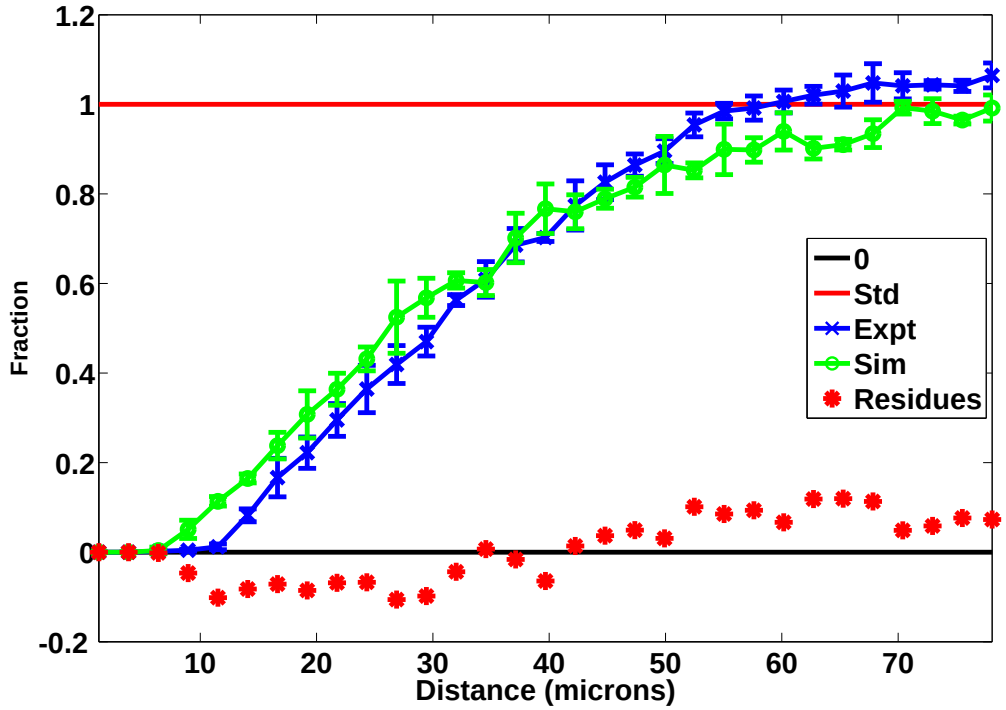


(b) $\tilde{f}_{PA}\left(\hat{d}\right)$, $R^2 = 0.9500$

Figure 25: Simulation of contact inhibition on TCPS surface (blue: experiment data; green: simulation result; red dots: residues. $Prol_o = 0.45$)

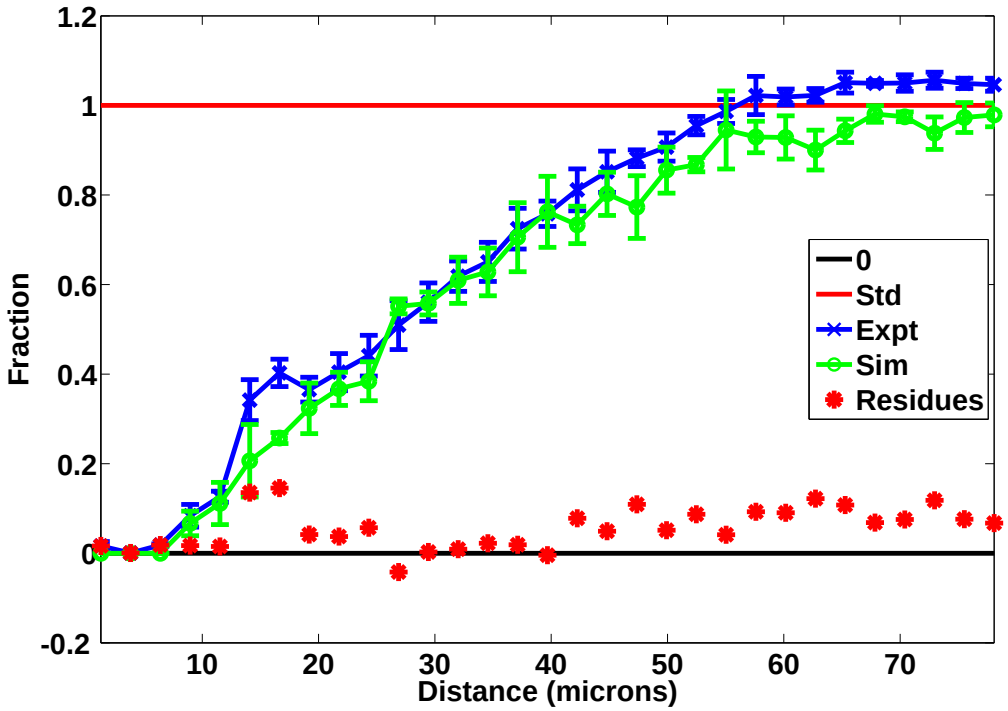


(c) $\tilde{f}_{PP}\left(\hat{d}\right)$, $R^2 = 0.8584$



(d) $\tilde{f}_{PR}\left(\hat{d}\right)$, $R^2 = 0.9661$

Figure 25 (continued): Simulation of contact inhibition on TCPS surface



(e) $\tilde{f}_{RR}\left(\hat{d}\right)$, $R^2 = 0.9596$

Figure 25 (continued): Simulation of contact inhibition on TCPS surface

CHAPTER VII

SCREENING OF CELL-SURFACE INTERACTIONS

7.1 Traditional Global Feature Analysis

Traditional global metrics were used for knowledge discovery in this study as comparison of screening results of individual-cell-based data analysis.

7.1.1 Factor Important Analysis

After standardization, distributions of candidate traditional descriptors of surface features, cell response (represented by cell proliferation ratio), and cell density (represented by cell number per image) of 126 locations on the $PLGA/PCL$ library are visualized in the boxplot in Figure 26. Cell densities on these locations are 282.15 ± 42 per image. The p-values of F-tests of these normalized surface feature descriptors, screened against $Prol$, are listed in Table 5. (Since $SumS$ and $PerS$ are linearly identical, only one of them can be used for screening. In this work $PerS$ is used.)

Based on a significant level $\alpha = 0.05$, $\bar{L}$ was screened out as the only significant factor among all selected candidates. Linear regression results of model $Prol \sim \bar{L}$ are listed in Table 6. Therefore, PCL island size is the only major factor among candidates that significantly affects cell proliferation. The result is consistent with that from screening by localized data analysis, shown in Section 7.2. However, traditional factor importance analysis methods based on global descriptions (such as $\bar{L}$) could only provide a general guide and do not illuminate specific relationships.

7.1.2 Principal Component Analysis

PCA was performed on candidate global descriptors N_{PCL}, $SumS$, $\bar{S}$, $\bar{S}_S$, $SumL$, $\bar{L}$, and

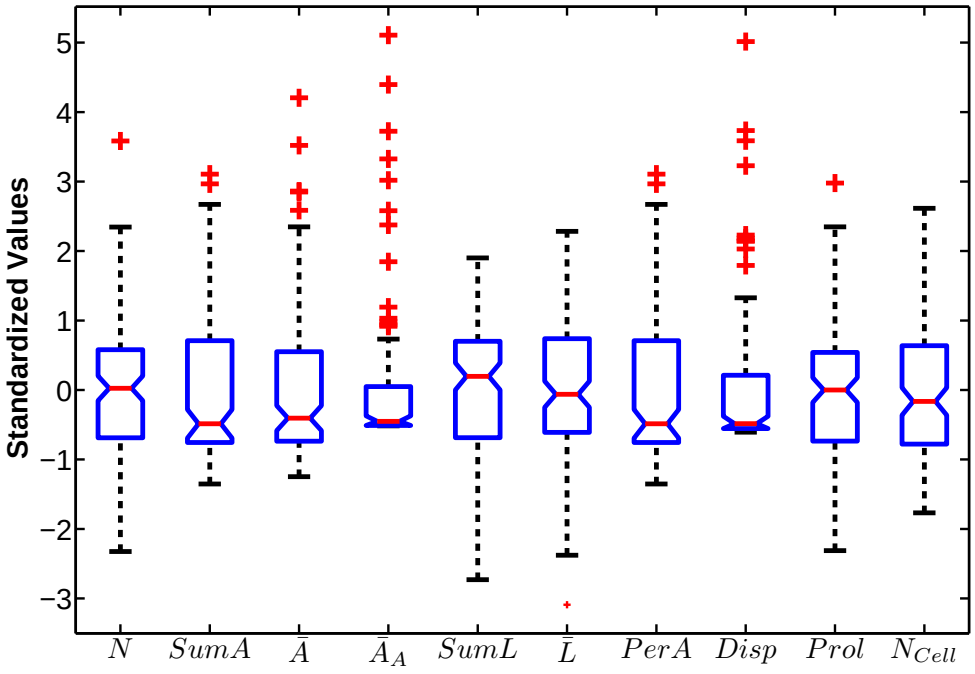


Figure 26: Boxplot of traditional summary metrics as applied to the library, including surface descriptors, cell proliferation ratio, and cell density at 126 locations. Values have been standardized by centering and normalizing the range by high and low values.

Table 5: Screening of candidate traditional surface feature descriptors

Descriptor	N_{PCL}	$\bar{S}$	$\bar{S}_S$	$SumL$	$\bar{L}$ *	$PerS$	$DispS$
p-value	0.564	0.999	0.561	0.669	0.0015	0.411	0.929

Table 6: Linear regression results for $Prol \sim \bar{L}$

Parameter	$Coefficient$	R^2	$RMSE$	$Adj\,R^2$	F-value	p-value
Value	0.278	0.0783	0.964	0.0635	10.54	0.0015

DispS, and the results are shown in Table 7. No candidate descriptor significantly contributes to the primary component. Though $\bar{L}$ contributes most to the secondary component, it is not distinguishable from other descriptors.

The first two principal components (*PC*s) explain 84.5% variation observed from the original descriptor set (Table 8 and Figure 27) and thus provide an acceptable representation of the original data set of candidate global descriptors for further analysis.

However, cell proliferation behaviors show no noticeable correlation to the first two principal components (Figure 28). Attempts of clustering against the first two or three principal components yield no meaningful result at all.

Based on global descriptor sets, *PCA*, as a typical and common tool in exploratory data analysis and data mining, was not effective in this study for knowledge discovery and modeling on combinatorial libraries. Clustering, as another typical approaches for data mining, also does not yield impressive result.

7.2 Local Feature Analysis

The local cell-feature analysis results are discussed in this section. The enhancement of cell proliferation, represented by $\Delta_{PC|AC}$, were screened among combinations of *PCL* island size and cell-to-*PCL* distance and favorable combinations were highlighted in Figure 29. Different from previous definitions of distance, the results were modified by converting the centroid-to-centroid distance (dis_{PC} and dis_{AC}) to the centroid-to-edge distance (dis'_{PC} and dis'_{AC}) between cell nuclei and *PCL* islands,

$$dis'_{PC}\left(grid_{\hat{d},\widehat{dia}}\right) = dis_{PC}\left(grid_{\hat{d},\widehat{dia}}\right) - \frac{\widehat{dia}}{2} \tag{65}$$

$$dis'_{AC}\left(grid_{\hat{d},\widehat{dia}}\right) = dis_{AC}\left(grid_{\hat{d},\widehat{dia}}\right) - \frac{\widehat{dia}}{2} \tag{66}$$

as the modified one is independent to *PCL* island size and provides direct visualization of cell-*PCL*-island interactions. After the conversion, for the cases that the cell-to-*PCL* distance is negative, cell nuclei are on top of *PCL* islands.

Table 7: Results of *PCA*

Symbol	Principal component						
	1^{st} PC	2^{nd} PC	3^{rd} PC	4^{th} PC	5^{th} PC	6^{th} PC	7^{th} PC
N_{PCL}	0.3072	-0.5525	0.1598	-0.0896	0.1171	-0.6618	-0.3393
$SumS$	-0.4281	-0.2086	0.4000	-0.5392	-0.2743	0.2998	-0.3967
$\bar{S}$	-0.4821	0.0792	0.1996	-0.3257	0.3598	-0.4194	0.5568
$\bar{S}_S$	-0.4412	-0.2830	-0.0213	0.4683	0.6110	0.2195	-0.2897
$SumL$	0.3396	-0.1789	0.7639	0.2314	0.0741	0.2922	0.3531
$\bar{L}$	-0.1793	0.5969	0.4309	0.3677	-0.1563	-0.3714	-0.3582
$DispS$	-0.3845	-0.4202	-0.0699	0.4326	-0.6151	-0.1574	0.2838

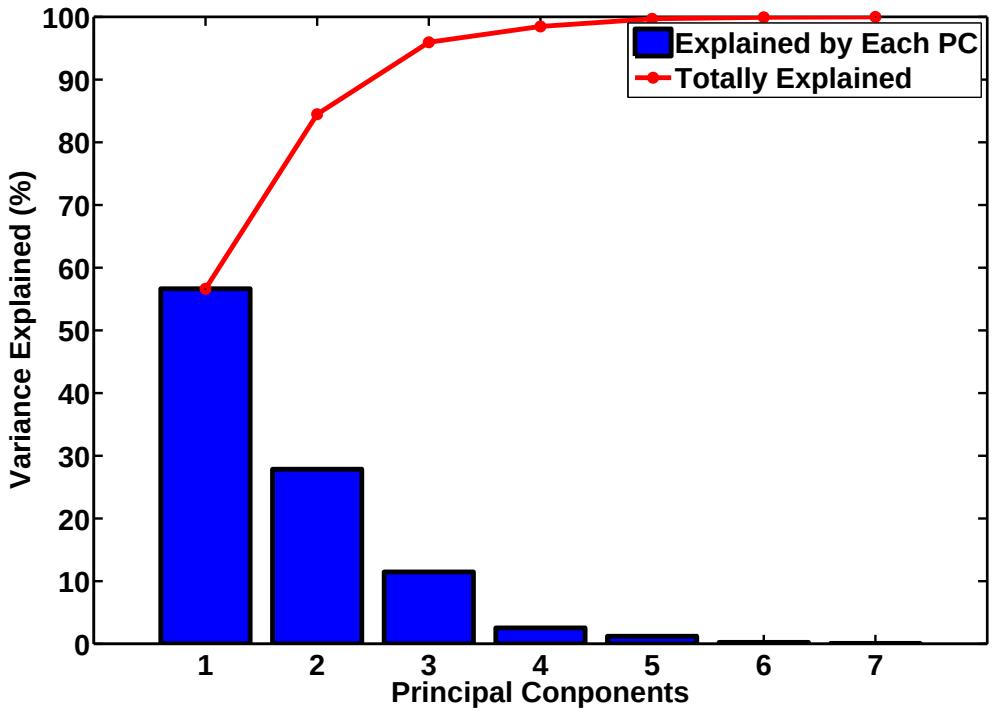


Figure 27: PCA: percentage of data explained by principal components

Table 8: *PCA*: data explained

	N_{PCL}	$SumS$	$\bar{S}$	$\bar{S}_S$	$SumL$	$\bar{L}$	$DispS$
Explained (%)	56.63	27.84	11.47	2.536	1.210	0.2348	0.07975
Cumulatively Explained (%)	56.63	84.47	95.94	98.48	99.69	99.92	100.0

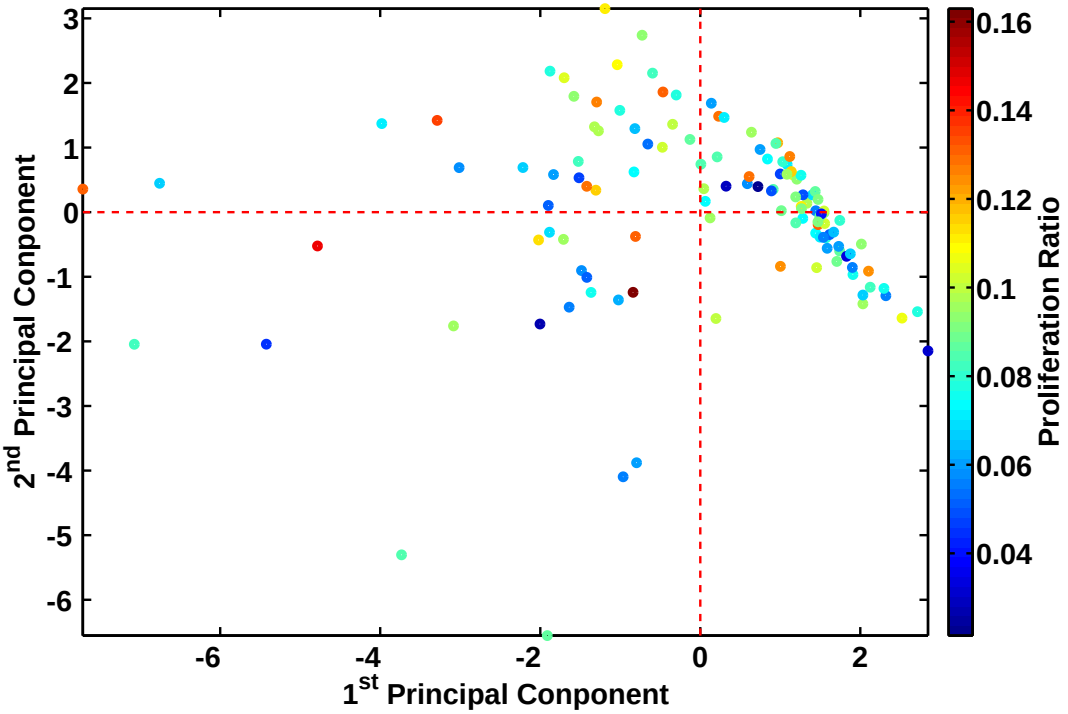


Figure 28: PCA: re-mapping *Prol* according to the first two principal components

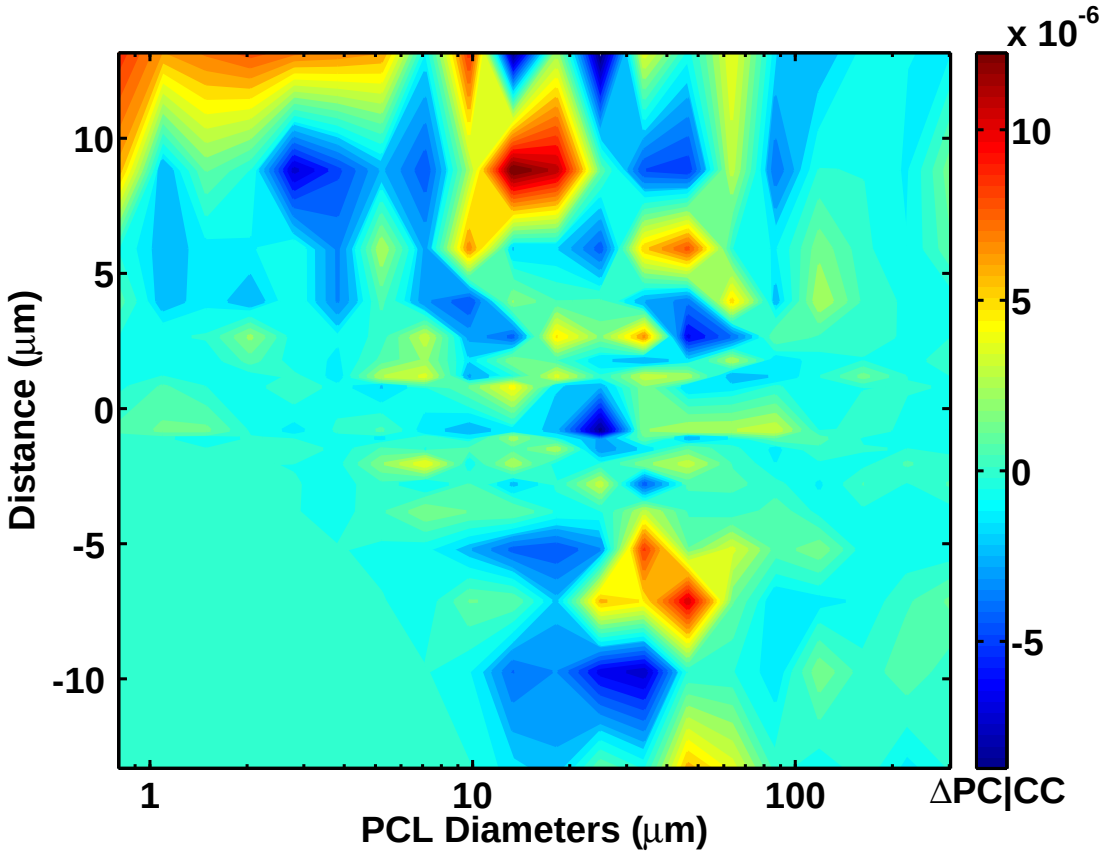


Figure 29: Screening of enhancement of cell proliferation $\Delta_{PC|AC}$ as a function of cell-to-*PCL* lateral distance and *PCL* diameter.

7.3 "Holder" and "Shaper" Model

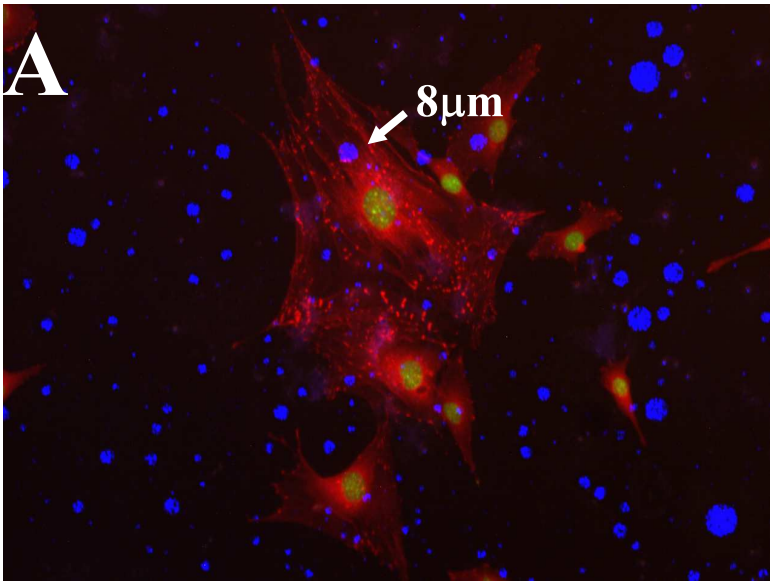
The data analysis results shown in Figure 29 suggest a *shaper-holder model*. Several local "hot spots" have been found from the two dimensional profile. According to the cell-*PCL* island distance and *PCL* island size, the surface lateral patterns represented by these hot spots can be classified as *holder*s, which are about $10\mu m$ away or farther from cell nuclei, and *shaper*s, which have a large size and are very close to or even overlapped with cell nuclei.

In the holder-shaper model, holders provide anchor sites for cell attachment, while shapers affect cell shapes by physically hindering cell spreading or increasing cellular internal tension. Each type of holder or shaper is characterized by its size range and cell-to-*PCL* distance range. The combinations of holders and shapers, which are defined as *holder-shaper pattern*s, are used as descriptors of surface features. The comprehensive effects of both holders and shapers of a specific pattern may affect cell proliferation.

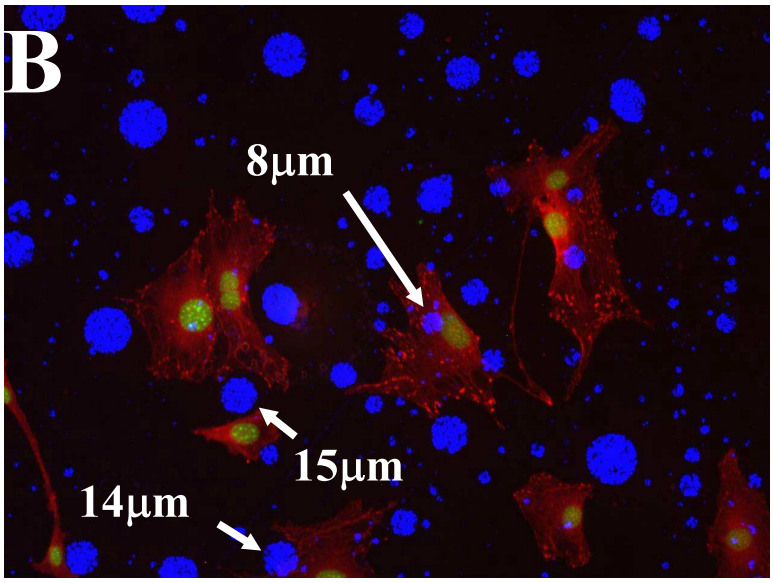
$\Delta_{PC|AC}$ was used to fast screen the combinatorial libraries for apparent holders and shapers and specific patterns of holder-shaper combinations as candidates. Such candidate patterns can be further investigated under well-controlled conditions.

Two candidate holders (in the region $\left(\hat{d} \in [11, 12]\mu m),\ \widehat{dia} < 3\mu m\right)$ and the region $\left(\hat{d} \in [7, 13]\mu m),\ \widehat{dia} \in [10, 22]\mu m\right)$) and two shapers (in the region $\left(\hat{d} \in [-8, -4]\mu m),\ \widehat{dia} \in [25, 60]\mu m\right)$ and the region $\left(\hat{d} \in [-11, -14]\mu m),\ \widehat{dia} \in [40, 70]\mu m\right)$ are recognized from the profile.

This quantitative model and the discovery of holders and shapers suggests that not only the average size of *PCL* islands, $\bar{L}$, but also the combinations of different sizes of islands, is important for cell proliferation. Being incapable of recognizing distribution patters, traditional data analysis methods, as shown in the $Prol \sim \bar{L}$ regression, would have suggested that larger *PCL* islands promoted cell proliferation, while smaller ones suppressed it. However, such conclusion would against the observations shown in Figure 29 that holders smaller then $3\mu m$ may also play an important role for promoting cell proliferation.

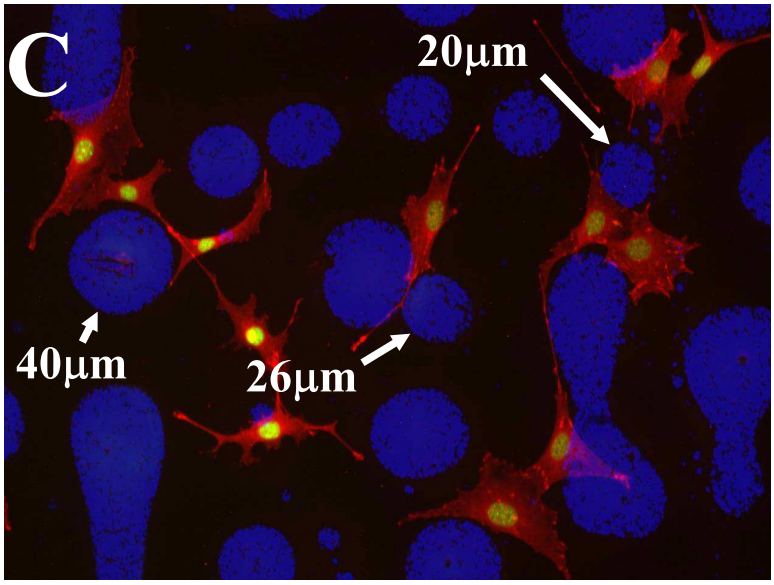


(a) Cells spread freely on small *PCL* islands



(b) Cells begin to recognize medium *PCL* islands

Figure 30: Micrographs of cells overlaid on *PCL* microstructures images from three select locations from the library, illustrating the effect of *PCL* size on cell shapes. Red: vinculin (over-exposed to show cell shapes); green: cell nuclei; blue: *PCL* islands



(c) Cells avoid large *PCL* islands

Figure 30 (continued): Micrographs of cells overlaid on *PCL* microstructures

The *holder-shaper model* casts new light on a comprehensive understanding of cell-surface interactions on multiple scales, and is of potentials to be applied in tissue engineering, medical care, and cellular biology research. Although the effectiveness of this model on other material/cell behavior systems needs to be validated, the methodologies demonstrated in these work can be applied to most studies of cell behaviors.

Three locations on the combinatorial libraries featured with *PCL* islands of different dominating sizes were shown in Fig. (6). While spreading freely on small islands (Figure 30(a)), cells began to recognize medium size islands (Figure 30(b)), and obviously avoided big islands (Figure 30(c)). When *PCL* size increased, cell spreading was constrained, and cells were reshaped by big *PCL* islands. The observations of cell shape and *PCL* island correlation were direct evidence of the shaper-holder model.

CHAPTER VIII

CONCLUSIONS

This work was motivated by developing a systematic, comprehensive methodology to address the difficulty of analyzing short-distance sensitive cell behaviors. A primary application is the analysis of large databases resulting from high-throughput cell culture, in order to discover cell-cell and cell-substrate interactions. A novel point of view, *the point of view from individual cell*, was used for distinguishing the near neighborhood environment of cells from the global average ones. Contact inhibition of cell proliferation, a typical short-scale cell-to-cell interaction, was used as the prototype for establishing the methodology, developing novel data analysis methods from it, and validating the effectiveness of the methodology and related new statistical tools developed in this study. A contact inhibition cell model based on *a priori* hypothesis was proposed. Based on this cell model, *Monte Carlo* simulation was performed to

- validate the *individual cell-based data analysis methods*
- validate the *individual cell viewpoint-based methodology* for investigating short-distance sensitive cell behaviors, and
- validate the contact inhibition model
- develop novel describing methods for the characteristics of biomaterial that affects cell behaviors, and
- provide novel computational tools for investigating cell-to-cell and cell-to-surface interactions.

The proposed methodology and related data analysis methods, after being established and tested, was extended into the high-dimensional cell-to-surface-pattern interactions, and a

novel high-throughput screening tool for material informatics on combinatorial libraries was developed and used for knowledge discovery and pattern screening.

It has been shown that describing the interactions of adherent cells with global (summary) statistics can be problematic due to the noise intrinsic to cell-cell interactions, and the inherent local nature of these interactions. The novel data analysis strategy, *individual cell-based local metrics*, has been introduced in this work. Local metrics such as cell-cell distance histograms describe cell environments from the "*point of view*" of individual cells. These metrics allow facile "fine tuning" to search a variety of phenomena by decomposing the distributions into specific cell-types (proliferating *vs.* non-proliferating). However, local metrics as defined herein are not limited to proliferation analysis, nor to cell-cell interactions alone. The metrics are generic and can be applied to any type of quantifiable cell assay, and can be applied to cell-biomaterial and cell-tissue interactions as well. We show how the method of local metrics is related to the naïve Bayes model, which makes them useful for data mining and classification.

This work has demonstrated the new local metrics by considering the contact inhibition of cell proliferation on three types of polymeric surfaces. A quantitative and probabilistic description of contact inhibition effect has been achieved for the first time. The results of localized, individual-cell-based analysis of cell-to-cell interaction on three different surfaces strongly support that the analysis method itself is repeatable and reliable.

A comprehensive model on contact inhibition was proposed, and *Monte Carlo* simulation approaches were designed and performed based on the results of *individual cell based local data analysis*. The simulation brought new insight into the mechanism of cell contact inhibition, suggested new descriptions of cell-behavior based biomaterial characteristics, and showed that the contact inhibition model and the simulation method can be potentially extended to other adhesion-dependent cell lines and other biomaterials.

Furthermore, in this work the novel high-throughput lateral pattern screening strategy, *local cell-feature analysis*, which integrated individual cell based descriptions, localized

data analysis, Bayes models, and combinatorial design of libraries of polymer of individual cells, was successfully developed. The screening results were consistent with traditional factor importance analysis and with observations of cell-pattern interactions.

This *local cell-feature analysis* strategy, established from the concepts of the *individual cell viewpoint-based methodology* and directly extended from the *individual cell-based data analysis methods for contact inhibition*, is promising in screening out candidate surface lateral patterns for further study, discovering new knowledge of cell-surface interactions, and exploring new materials for tissue engineering and medical applications.

The data scarcity for higher dimensional cases (such as the two dimensional case in the section of high throughput screening of surface microstructural patterns (surface microstructure pattern)s) is one of the major challenges of this method. For example, in order to achieve the same sensitivity and statistical confidence, if 1,000 data elements is required in a one dimensional case, 1×10^6 data elements are required for a two dimensional system, and 1×10^9 data elements are required for a three dimensional system. The current full orthogonal method used in this work is sufficient for study the two dimensional case of *PLGA*/*PCL* libraries (where we investigated two parameters interacting cell-*PCL* distance and *PCL* size). We would like to suggest two strategies for higher dimensional cases.

1. Such challenge can be addressed by a better analysis and sampling design that employs an orthogonal factorial method. That is, data elements are not simply binned into full orthogonal grids, but orthogonally randomized and binned into a much finer factorial grid to achieve much higher resolution.

2. Another work-around is using a so-called "bootstrap" strategy and randomly resampling the original data for much higher efficiency of usage of experimental data.

Another major limitation of this work is that all cell and surface information must be measured *in situ*, which is a challenge in some fields. Consider, for example, studying the distance effects on mRNA transcription. One would need to establish a quantitative

fluorescent *in situ* hybridization (FISH) method for the mRNA of interest before using our methodology.

This work offers new insight into related fields.

- For contact inhibition phenomena, the method allowed us to measure experimentally a "contact-inhibition profile" (e.g., $f_{PR}(\hat{d})$) for the first time. Such functions quantitatively describes the probability of cell proliferation as a function of distance to other cells. It is unique because it is measured from a set of static (all cells are fixed) immuno-stained microscope images, and yet it gives us a quantifiable measure of the dynamics of cell-cell interactions relevant to proliferation. For example, we now know that at a distance of 9 microns, MC3T3-E1 cells show a 15 fold reduction in probability of proliferating (in term of proliferation odd) on PCL surface. We also learned that this statistic is about the same on $PLGA$ and $TCPS$, but is significantly different on PCL. We know of no other report of such information, especially not coming from a static set of images. Therefore, specifically to the cell-cell, we quantitatively visualized and simulated, in a distance-based way, the contact inhibition phenomenon, which provides powerful tools to study underlying mechanisms, to disentangle such effects from other cell behaviors, and to engineer devices and protocols to clinically diagnose of related diseases, such as early screening of cancers based on the lost of capability to inhibit cell growth.

- For the micro-structured surfaces ($PLGA/PCL$ system), we have discovered that cell-surface interactions are only "active" to effect a proliferation in certain size ranges, and then only when cells are at a certain distance from the appropriately-sized islands. We didn't know this before and the global methods did not allow us to detect it. Why is this important? For one, it leads one to propose phenomenological models (such as the cell model of contact inhibition, the "critical-size" model, and the "shaper-and-holder" model) for further investigation. These may, in turn, lead to concrete hypotheses about how cells sense the surface patterns. Secondly, we now

know how to design, quantitatively, an experiment to test these ideas. We know what size *PCL* islands to make and what cells to look in detail at based on how far they are from those islands – in subsequent hypothesis testing. In an empirically-driven field this is tremendously important information that can save time and lead to more quantitative testing.

- The method also provides promising material informatics tools for high throughput screening of candidate biomaterials based on specific cell behaviors, and can be used to develop bio-MEMS to screen biomaterials.

- It is important that the field of this work is methodology. The methodology proposed and developed in this work provides a quantitative approach to conveniently studying the distance sensitive phenomena in all related fields, and may have fundamental impact to not only cell-cell and cell-surface interactions, but also other similar systems such as polymer physics, nano-particles, and so on.

In conclusion, this work focused on the point of view of individual cells, established a systematic and comprehensive methodology from that viewpoint and developed novel data analysis methods. We showed that local metrics successfully recognized contact inhibition of cell proliferation from experimental data of MC3T3-E1 cells on *PLGA*, *PCL* and $TCPS$ surfaces. A theoretical model for contact inhibition was proposed and used to validate the local metric analysis method. The local cell-behavior based characteristics can be viewed as quantitative descriptors for biomaterials, which extend the field of high-throughput screening and knowledge discovery of biomaterials.

CHAPTER IX

FUTURE DIRECTIONS

This work addressed the difficulty of analyzing short-distance sensitive cell behaviors by traditional global and summary metrics, established the novel "individual cell based localized data analysis method", proposed a cell model for contact inhibition, performed *Monte Carlo* simulation for the cell model, and developed screening methods of material informatics.

Because this work focused on methodology, the data methods as well as statistical tools established in this work are potential to fundamentally impact related fields. Some promising directions for future study based on this work are listed bellow.

Apply the contact inhibition model to other cell lines Contact inhibiton effect is common in most adhesion dependent cells. The effectiveness of the proposed cell model can be tested on other cell lines, and the characteristics of these cell lines, described by parameters used in the model, is potentially useful for the mechanism study of contact inhibition and comparison of these cell lines.

Apply the contact inhibition model to other surfaces The results of this study have shown that, fot the same cell line, cells behave differently on different surfaces, and such difference can be described by parameters in the cell model. Other materails and surfaces can be characterized by such parameters for their effects on contact inhibition of proliferation. That is, the new describing parameters can be used for the screening and characterization of biomaterials.

Study other cell-to-cell interactions As a general method for studying short-range cell-to-cell interactions, the local metrics can be used for other cell-to-cell inteactions,

such as cell recognition, differentiation, migration, stem cell and cancer study.

Screening and material informatics As mentioned in Chapter 7.1, a novel high-throughput screening method based on local metrics and combinatorial libraries has been developed and tested. This screening method can be used for other cell lines, other biomaterials and surfaces, and other short-range cell-surface interactions.

"Shaper and holder" model Discovered in *PLGA*/*PCL* combinatorial library screening for the effects of surface macrostructures of PCL islands on cell proliferation, the "shaper and holder" model should be further studied under well-controlled conditions, so do the underlying mechanisms and potential application of this discovery in bioscience, tissue engineering, and medicine.

APPENDIX A

SELECTED CODES

Some M codes used in this work are listed in this chapter for illustration of the algorithms and strategies.

A.1 Variation-adjusted Iterative Selection

The original iterative selection algorithm, known as file "isodata.m" (listed in Listing 1) which is available from Matlab™ Central File Exchange provided by Dhanesh Ramachandram, was modified to address the influence of the variations of the foreground and the background. The modified code can be found in Listing 2

```
function level=isodata(I)
%    ISODATA Compute global image threshold using iterative isodata method.
%    LEVEL = ISODATA(I) computes a global threshold (LEVEL) that can be
%    used to convert an intensity image to a binary image with IM2BW. LEVEL
%    is a normalized intensity value that lies in the range [0, 1].
%    This iterative technique for choosing a threshold was developed by Ridler and ↓
→       Calvard .
%    The histogram is initially segmented into two parts using a starting threshold ↓
→       value such as 0 = 2B-1,
%    half the maximum dynamic range.
%    The sample mean (mf,0) of the gray values associated with the foreground pixels↓
→       and the sample mean (mb,0)
%    of the gray values associated with the background pixels are computed. A new ↓
→       threshold value 1 is now computed
%    as the average of these two sample means. The process is repeated, based upon ↓
→       the new threshold,
%    until the threshold value does not change any more.

%
%    Class Support
%    -------------
%    The input image I can be of class uint8, uint16, or double and it
%    must be nonsparse.  LEVEL is a double scalar.
```

```
%
%    Example
%    -------
%          I = imread('blood1.tif');
%          level = graythresh(I);
%          BW = im2bw(I,level);
%          imshow(BW)
%
%    See also IM2BW.
%
% Reference :T.W. Ridler, S. Calvard, Picture thresholding using an iterative ↓
→      selection method,
%                 IEEE Trans. System, Man and Cybernetics, SMC-8 (1978) 630-632.

% Convert all N-D arrays into a single column.  Convert to uint8 for
% fastest histogram computation.
I = im2uint8(I(:));

% STEP 1: Compute mean intensity of image from histogram, set T=mean(I)
[counts,N]=imhist(I);
i=1;
mu=cumsum(counts);
T(i)=(sum(N.*counts))/mu(end);
T(i)=round(T(i));

% STEP 2: compute Mean above T (MAT) and Mean below T (MBT) using T from
% step 1
mu2=cumsum(counts(1:T(i)));
MBT=sum(N(1:T(i)).*counts(1:T(i)))/mu2(end);

mu3=cumsum(counts(T(i):end));
MAT=sum(N(T(i):end).*counts(T(i):end))/mu3(end);
i=i+1;
% new T = (MAT+MBT)/2
T(i)=round((MAT+MBT)/2);

% STEP 3 to n: repeat step 2 if T(i)~=T(i-1)
while abs(T(i)-T(i-1))>=1
    mu2=cumsum(counts(1:T(i)));
    MBT=sum(N(1:T(i)).*counts(1:T(i)))/mu2(end);

    mu3=cumsum(counts(T(i):end));
    MAT=sum(N(T(i):end).*counts(T(i):end))/mu3(end);

    i=i+1;
```

```
    T(i)=round((MAT+MBT)/2);
    Threshold=T(i);
end

 % Normalize the threshold to the range [i, 1].
level = (Threshold - 1) / (N(end) - 1);
```

Listing 1: (isodata.m) The Original iterative selection algorithm in M

```
function level=isohist(I)
%    Jing Su modified the original 'ISODATA' function to 'ISOHIST'.
%    Improvement:
%    1. determine a threshold of any histograph.
%    2. use median instead of mean, which is more robust
%    3. introduce the std's of the two groups for justification of peak
%          distributions.
%
%    ISODATA Compute global image threshold using iterative isodata method.
%    LEVEL = ISODATA(I) computes a global threshold (LEVEL) that can be
%    used to convert an intensity image to a binary image with IM2BW. LEVEL
%    is a normalized intensity value that lies in the range [0, 1].
%    This iterative technique for choosing a threshold was developed by Ridler and ↓
→       Calvard .
%    The histogram is initially segmented into two parts using a starting threshold ↓
→       value such as 0 = 2B-1,
%    half the maximum dynamic range.
%    The sample mean (mf,0) of the gray values associated with the foreground pixels↓
→       and the sample mean (mb,0)
%    of the gray values associated with the background pixels are computed. A new ↓
→       threshold value 1 is now computed
%    as the average of these two sample means. The process is repeated, based upon ↓
→       the new threshold,
%    until the threshold value does not change any more.


%
%    Class Support
%    -------------
%    The input image I can be of class uint8, uint16, or double and it
%    must be nonsparse.  LEVEL is a double scalar.
%
%    Example
%    -------
%          I = imread('blood1.tif');
%          level = graythresh(I);
```

```
%          BW = im2bw(I,level);
%          imshow(BW)
%
%    See also IM2BW.
%
% Reference :T.W. Ridler, S. Calvard, Picture thresholding using an iterative ↓
→       selection method,
%                 IEEE Trans. System, Man and Cybernetics, SMC-8 (1978) 630-632.

% Convert all N-D arrays into a single column.  Convert to uint8 for
% fastest histogram computation.

I = I(:);
I = sort(I);

% STEP 1: Compute mean intensity of image from histogram, set T=mean(I)

i = 1;
T(i) = median(I);

% STEP 2: compute Mean above T (MAT) and Mean below T (MBT) using T from
% step 1

MA = I(find(I <= T(i)));
MB = I(find(I >  T(i)));

MAT = median(MA);
MBT = median(MB);
i=i+1;

% new T = (MAT*stdA + MBT*stdB)/(stdA + stdB)

T(i) = round( (MAT*std(MA)+MBT*std(MB))/(std(MA)+std(MB)↓
→     ));

Threshold = T(i);

% STEP 3 to n: repeat step 2 if T(i)~=T(i-1)
while abs(T(i)-T(i-1)) >= mean(I)/40 & i <= 100
      %%mu2=cumsum(counts(1:T(i)));
      %%MBT=sum(N(1:T(i)).*counts(1:T(i)))/mu2(end);
      %%
      %%      mu3=cumsum(counts(T(i):end));
      %%      MAT=sum(N(T(i):end).*counts(T(i):end))/mu3(end);
      %%
```

```
76      MA = I(find(I <= T(i)));
        MB = I(find(I >  T(i)));

        MAT = median(MA);
        MBT = median(MB);
81      i = i+1;

        T(i) = round( (MAT*std(MA)+MBT*std(MB))/(std(MA)+std↓
→          (MB)));
    end

86   % Normalize the threshold to the range [i, 1].
    if i > 100,
        if abs(T(i)-T(i-2)) < mean(I)/40 & ...
            abs(T(i-1)-T(i-3)) < mean(I)/40,
            level = (T(i)+T(i-1))/2;
91      else,
            level = -1;   %Warning signal for error
        end;
    else,
        level = Threshold;
96  end;
```

Listing 2: (isohist.m) The Variation-adjusted iterative selection algorithm in M

A.2 Marker Controlled Watershed Segmentation

The *MCWS* algorithm is shown in Listing 3.

```
1   % Marker-controlled watershed segmentation
    % Modified from Matlab News & Notes & Demo
    % Final version: 102606
4   %    by Jing Su
    %
    % i_o: original raw image
    % i_info: metadata of the raw image abtained
    %————————→by function iminfo
9   % i_MCWS: output image
    %
```

```
function i_MCWS = MCWS(i_o, i_info)
    if ~strcmp(i_info.ColorType, 'grayscale'),
        i_o = rgb2gray(i_o);
    end;

    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    %               Leveling
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    bg = imopen(i_o, strel('disk', 15));
    i_o = imsubtract(i_o,bg);

    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    %           Openning and filling
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    SE3 = strel('disk',3);
    i_bw_fill = imfill(imopen(i_bw, SE3), 'holes');

    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    %           Watershed
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    D_QE = bwdist(~i_bw_fill, 'quasi-euclidean');
    D_QE = mat2gray(D_QE);
    f_a = fspecial('disk', 5);
    D_QE5 = imfilter(D_QE, f_a, 'replicate');
    D_QE5 = -D_QE5;
    D_QE5(~i_bw_fill) = -Inf;
    L_QE5 = watershed(D_QE5);
    i_MCWS = i_bw_fill .* L_QE5;
    %Remove fragmental particles
    [lf,nf] = bwlabel(i_MCWS);
    stats = regionprops(lf,'all');
    idx = find([stats.Area] <= 20);
    BW2 = ismember(lf,idx);
    i_MCWS = i_MCWS .* (~BW2);
```

Listing 3: (MCWS.m) The Marker controlled watershed segmentation algorithm in M

REFERENCES

[1] A , M. and A , E. J., "Interference microscope studies of cell contacts in tissue culture," *Experimental Cell Research*, vol. 15, no. 2, pp. 332–345, 1958.

[2] A , M., H , J. E. M., and K , H. M., "Social behaviour of cells in tissue culture iii. mutual influence of sarcoma cells and fibroblasts," *Experimental Cell Research*, vol. 13, no. 2, pp. 279–291, 1957.

[3] A , M. and H , J. E. M., "Observations on the social behaviour of cells in tissue culture : Ii. "monolayering" of fibroblasts," *Experimental Cell Research*, vol. 6, no. 2, pp. 293–306, 1954.

[4] A , A., L , D., A , D., and L , R., "Parallel synthesis and biophysical characterization of a degradable polymer library for gene delivery," *Journal of the American Chemical Society*, vol. 125, pp. 5316–5323, 2003.

[5] A , C. F. and H , P. A., "Adaptation of plastic surfaces for tissue culture by glow discharge.," *Journal of Clinical Microbiology*, vol. 2, no. 1, pp. 46–54, 1975.

[6] A , D. G., P , D., L , E. B., M , T. A., and L , R., "Biomaterial microarrays: rapid, microscale screening of polymer-cell interaction," *Biomaterials*, vol. 26, pp. 4892–4897, 2005.

[7] A , K. and B , M., "Modelling approach in cell/material interactions studies," *Biomaterials*, vol. 27, pp. 1187–1199, Mar 2006.

[8] B , C. T. H., B , G. A., P , C. A. H., and R , F. A., "Modelling and analysis of time-lags in some basic patterns of cell proliferation," *Journal of Mathematical Biology*, vol. 37, pp. 341–371, Oct. 1998.

[9] B , C. T. H., B , G. A., P , C. A. H., and R , F. A., "Modelling and analysis of time-lags in cell proliferation," Tech. Rep. No. 313, Manchester, England, 1997.

[10] B , S. and L , C., "Use of watersheds in contour detection," in *International Workshop on image processing, real-time edge and motion detection/estimation, Rennes, France*, 1979.

[11] B , S., M , H., H , D., and P , A. S., "Quantification of cell turnover kinetics using 5-bromo-2'-deoxyuridine," *Journal of Immunology*, vol. 164, no. 10, pp. 5049–5054, 2000.

[12] B , B. D., B , L. F., P , E. P., L , C. H., R , J., C , D. L., D , D. D., S , Z., and B , A. L., "Osteoblast-mediated mineral deposition in culture is dependent on surface microtopography," *Calcified Tissue International*, vol. 71, no. 6, pp. 519–529, 2002.

[13] B , R. M., V , R., H , F. R., and F , W. A., "The e-cadherin cell-cell adhesion complex and lung cancer invasion, metastasis, and prognosis," *Lung Cancer*, vol. 36, no. 2, pp. 115–124, 2002.

[14] B , S., J , K., T , V., and K , J., "A combinatorial approach to polymer design," *Journal of the American Chemical Society*, vol. 119, pp. 45–53, 1997.

[15] B , S., J , K., T , V., and K , J., "Structure-property correlations in a combinatorial library of degradable biomaterials," *Journal of Biomedical Materials Research*, vol. 42, no. 1, pp. 66–75, 1998.

[16] B , M. M., "Proteolytic enzymes initiating cell division and escape from contact inhibition of growth," *Nature*, vol. 227, no. 5254, pp. 170–171, 1970.

[17] B , J. M., "Cell-cell contact promotes dna synthesis in retinal glia but not in fibroblasts," *Experimental Cell Research*, vol. 146, no. 1, pp. 204–206, 1983.

[18] B , J. M., "Growth in retinal glial cells in vitro is affected differentially by two types of cell contact-mediated interactions," *Experimental Cell Research*, vol. 180, no. 1, pp. 13–19, 1989.

[19] C , L. N., "Contact inhibition of cell division and cell movement," *Journal of Investigative Dermatology*, vol. 59, no. 1, pp. 27–32, 1972.

[20] C , L., M -P , I., N , P., B , F., C , M., G , D., L , M. G., and D , E., "Inhibition of cultured cell growth by vascular endothelial cadherin (cadherin-5/ve-cadherin)," *Journal of Clinical Investigation*, vol. 98, no. 4, pp. 886–893, 1996.

[21] C , J. L., B , L. E., G ´ , A. J., and K , W. P., "Hot embossing for micropatterned cell substrates.," *Biomaterials*, vol. 25, pp. 4767–4775, Aug 2004.

[22] C , J. L., E , M. T., G ´ , A. J., and K , W. P., "Combined microscale mechanical topography and chemical patterns on polymer cell culture substrates.," *Biomaterials*, vol. 27, pp. 2487–2494, Apr 2006.

[23] C , J. L., G ´ , A. J., and K , W. P., "Myoblast alignment and differentiation on cell culture substrates with microscale topography and model chemistries.," *Biomaterials*, vol. 28, pp. 2202–2210, Apr 2007.

[24] C , S., C , F., G , E., D , O., and G -R , C., "Variation in cadherins and catenins expression is linked to both proliferation and transformation of rhabdomyosarcoma," *Oncogene*, vol. 23, no. 13, pp. 2420–2430, 2004.

[25] C , S., C , F., F , M., P -C , E., D , A., and G -R , C., "M-cadherin activates rac1 gtpase through the rho-gef trio during myoblast fusion," *Molecular Biology of the Cell*, vol. 18, no. 5, pp. 1734–1743, 2007.

[26] C , S., C , F., G , Y., P , F., B , A., and G -R , C., "Rhoa gtpase regulates m-cadherin activity and myoblast fusion," *Molecular Biology of the Cell*, vol. 17, no. 2, pp. 749–759, 2006.

[27] C , C. S., M , M., H , S., W , G. M., and I , D. E., "Geometric control of cell life and death," *Science*, vol. 276, no. 5317, pp. 1425–1428, 1997.

[28] C , M. E., S , R. H., M , C. J., and OI , D. E., "Integrin binding and mechanical tension induce movement of mrna and ribosomes to focal adhesions," *Nature*, vol. 392, pp. 730–733, 1998.

[29] D , M. J., C , S., R , M. O., J , H. J., A , S., and C - , A. S., "Fibroblast reaction to island topography: changes in cytoskeleton and morphology with time," *Biomaterials*, vol. 24, pp. 927–935, 2003.

[30] P , J. J., L , M., S , J. I., and S , U. W., "Continuum-configurational-bias Monte Carlo simulations of long-chain alkanes," *Molecular Physics*, vol. 80, pp. 55–63, 1993.

[31] P , J. J. and E , F. A., "Monte carlo methods for polymeric systems," in *Advances in Chemical Physics* (I. P , S. A. R., ed.), pp. 337–367, 2007.

[32] D , D. D., B , B. D., B , L. F., P , E. P., L , C. H., R , J., C , D. L., S , Z., and B , A. L., "Bmp-2 regulation of calcification is modulated by substrate microtopography," *Journal of Bone and Mineral Research*, vol. 17, pp. S399–S399, 2002.

[33] D , O., G , M., H , H., K , M., M , J. A., and T , H. R., "On the formulation and analysis of general deterministic structured population models. ii. nonlinear theory.," *J Math Biol*, vol. 43, pp. 157–189, Aug 2001.

[34] D , O., G , M., and M , J. A. J., "Steady-state analysis of structured population models.," *Theor Popul Biol*, vol. 63, pp. 309–338, Jun 2003.

[35] D , O., H , J. A., and M , J. A., "On the definition and the computation of the basic reproduction ratio r0 in models for infectious diseases in heterogeneous populations.," *J Math Biol*, vol. 28, no. 4, pp. 365–382, 1990.

[36] D , L. E., C , C. S., M , M., T , J., W , G. M., and I , D. E., "Geometric control of switching between growth, apoptosis, and differentiation during angiogenesis using micropatterned substrates," *In Vitro Cell Dev Biol Anim*, vol. 35, pp. 441–448, 1999.

[37] D , D. and H , S., "Individual-based approaches to birth and death in avascu1ar tumors," *Mathematical and Computer Modelling*, vol. 37, pp. 1163–1175, June 2003.

[38] D , D., K , R., and M C , J. S., "Monte carlo approach to tissue-cell populations," *Physical Review E*, vol. 52, pp. 6635–6657, Dec 1995.

[39] D , D. and L , M., "Individual-based models to growth and folding in one-layered tissues: intestinal crypts and early development," *Nonlinear Analysis*, vol. 47, pp. 245–256, 2001.

[40] D , D., *On selected individual based approaches to the dynamics in multicellular systems*. Leipzig: Max-Planck-Inst. Mathematik in den Naturwiss., 2002.

[41] D , D. and H¨ , S., "A single-cell-based model of tumor growth in vitro: monolayers and spheroids.," *Physical Biology*, vol. 2, pp. 133–147, Sep 2005.

[42] E , E., M , J., N , W., and B , J., "Influence of the surface structure of titanium materials on the adhesion of fibroblasts.," *Biomaterials*, vol. 17, pp. 1399–1403, Jul 1996.

[43] E , E., V , D., and B , J., "Biomimetic implant coatings.," *Biomol Eng*, vol. 24, pp. 27–32, Feb 2007.

[44] F , J. and M , A., "Role of cell shape in growth control," *Nature*, vol. 273, pp. 345–349, 1978.

[45] G , N. D., C , J. L., K , W. P., and G ´ , A. J., "Micro- and nano-patterned substrates to manipulate cell adhesion.," *J Nanosci Nanotechnol*, vol. 7, pp. 803–807, Mar 2007.

[46] G , J., L , M., and D , D., "Modeling the effect of deregulated proliferation and apoptosis on the growth dynamics of epithelial cell populations in vitro," *Biophysical Journal*, vol. 88, no. 1, pp. 62–75, 2005.

[47] G ´ , A. J., "Get a grip: integrins in cell-biomaterial interactions.," *Biomaterials*, vol. 26, pp. 7525–7529, Dec 2005.

[48] G , S. J. and S -S , J. H., "Cell cycle variation associated with feeder effects in cultures of chinese hamster fibroblasts," *Experimental Cell Research*, vol. 109, no. 2, pp. 341–348, 1977.

[49] G -B , J., P , D., and B , D., "Effects of substratum surface topography on the organization of cells and collagen fibers in collagen gel cultures," *Journal of Biomedical Materials Research*, vol. 61, no. 4, pp. 608–618, 2002.

[50] H , S., C , S. C., and I , D. E., "Cell-shape-dependent control of p27kip and cell cycle progression in human capillary endothelial cells," *Molecular Biology of the Cell*, vol. 9, pp. 3179–3193, 1998.

[51] H , S. and I , D. E., "Shape-dependent control of cell growth, differentiation, and apoptosis: Switching between attractors in cell regulatory networks," *Experimental Cell Research*, vol. 261, pp. 91–103, 2000.

[52] K , S., W , N., S , A., K , A., and A , E., "Combinatorial assay of cells on topographically patterned surfaces," *Abstracts Of Papers of The American Chemical Society*, vol. 224, p. U237, 2002.

[53] K , B. G., C , D. M., and G ´ , A. J., "Surface chemistry modulates fibronectin conformation and directs integrin binding and specificity to control cell adhesion.," *J Biomed Mater Res A*, vol. 66, pp. 247–259, Aug 2003.

[54] K , B. G., C , D. M., and G ´ , A. J., "Surface chemistry modulates focal adhesion composition and signaling through changes in integrin binding.," *Biomaterials*, vol. 25, pp. 5947–5954, Dec 2004.

[55] K , B. G., C , D. M., and G ´ , A. J., "Integrin binding specificity regulates biomaterial surface chemistry effects on cell differentiation.," *Proc Natl Acad Sci U S A*, vol. 102, pp. 5953–5957, Apr 2005.

[56] K , V., S , J. R., K , D., A , S., K , J., and W , W. J., "Accurate predictions of cellular response using qspr: a feasibility test of rational design of polymeric biomaterials," *Polymer*, vol. 45, no. 22, pp. 7367–79, 2004.

[57] K , J., "New approaches to biomaterials design," *Nature Materials*, vol. 3, pp. 745–747, Nov. 2004.

[58] K , T., H , H., S , W. S., T , T., F , S., A , T., K , N., I , K., and T - , T., "Cell density and growth-dependent down-regulation of both intracellular calcium responses to agonist stimuli and expression of smooth-surfaced endoplasmic reticulum in mc3t3-e1 osteoblast-like cells," *J. Biol. Chem.*, vol. 278, no. 8, pp. 6433–6439, 2003.

[59] L , M. A., G , C. A., M , K. E., K , B. G., and G ´ , A. J., "Myoblast proliferation and differentiation on fibronectin-coated self assembled monolayers presenting different surface chemistries.," *Biomaterials*, vol. 26, pp. 4523–4531, Aug 2005.

[60] L , A., L , K., and M , M. C., "Stability properties of proliferatively coupled cell replication models.," *Acta Biotheor*, vol. 39, pp. 1–14, Mar 1991.

[61] L , S., K , B. Z., Y , K. M., and G , B., "Long-range and selective autoregulation of cell-cell or cell-matrix adhesions by cadherin or integrin ligands," *Journal of Cell Science*, vol. 111, no. 3, pp. 347–357, 1998.

[62] L , S., S , E., G , P., and G , B., "Cadherin-mediated transmembrane interactions," *Cell Adhes Commun*, vol. 6, no. 2-3, pp. 161–170, 1998.

[63] L , S., Y , A., K , Z., and G , B., "p27 is involved in n-cadherin-mediated contact inhibition of cell growth and s-phase entry.," *Oncogene*, vol. 18, pp. 869–876, Jan 1999.

[64] L , M., P , C. S., and W , H. E., "Epidermal cell proliferation. ii. a comprehensive mathematical model of cell proliferation and migration in the basal layer predicts some unusual properties of epidermal stem cells.," *Virchows Archiv. B, Cell pathology including molecular pathology*, vol. 53, pp. 286–300, 1987.

[65] L , C. H., T , E. M., S , V. L., H -V , A. K., C , D. L., D , D. D., B , B. D., and S , Z., "Response of normal female human osteoblasts (nhost) to 17 beta-estradiol is modulated by implant surface morphology," *Journal of Biomedical Materials Research*, vol. 62, no. 2, pp. 204–213, 2002.

[66] L , S., S , Z., W , L., L , C. H., T , J. D., W , M., C , D. L., and B , B. D., "Microrough implant surface topographies increase osteogenesis by reducing osteoclast formation and activity," *Journal of Biomedical Materials Research Part A*, vol. 70A, no. 3, pp. 361–369, 2004.

[67] M , M. C. and M , J. G., "A deterministic approach to survival statistics.," *Journal of Mathematical Biology*, vol. 28, no. 1, pp. 33–48, 1990.

[68] M , A., R , S., and W , A., "Comments on picture thresholding using an iterative selectionmethod," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 20, no. 5, pp. 1238–1239, 1990.

[69] M , S., C , S., M , M., C , F., T , P., B , A., and G -R , C., "Biogenesis of n-cadherin-dependent cell-cell contacts in living fibroblasts is a microtubule-dependent kinesin-driven mechanism," *Molecular Biology of the Cell*, vol. 13, no. 1, pp. 285–301, 2002.

[70] M , J. C. and A , E., "Lcst phase separation in biodegradable polymer blends: poly (d,l-lactide) and poly (ε-caprolactone)," *Macromolecular chemistry and physicss*, vol. 200, pp. 733–739, 2000.

[71] M , J. C., S , J. L., K , B. G., G , A. J., T , A., K , A., and A , E. J., "Combinatorial characterization of cell interactions with polymer surfaces," *Journal of Biomedical Materials Research Part A*, vol. 66, no. 3, pp. 483–490, 2003.

[72] M , J. A. J. and D , O., *The dynamics of physiologically structured populations*, vol. 68. Springer, Berlin, ALLEMAGNE (1974- 1994), 1986.

[73] M , F. and B , S., "Morphological segmentation," *Journal of Visual Communication and Image Representation*, vol. 1, no. 1, pp. 21–46, 1990.

[74] M , J. G., L , A., B , A., M , M. C., and G , L., "Complex dynamics and bifurcations in neurology.," *Journal of Theoretical Biology*, vol. 138, pp. 129–147, May 1989.

[75] M , J. G., H , U., L , A., and M , M. C., "Complex dynamics and noise in simple neural networks with delayed mixed feedback.," *Biomedica Biochimica Acta*, vol. 49, no. 8-9, pp. 697–707, 1990.

[76] N , Y. and M , R. H., "Density dependent modulation of cell cycle protein expression in astrocytes," *Journal of Neuroscience Research*, vol. 66, no. 3, pp. 487–496, 2001.

[77] N , C. M. and C , C. S., "Cell-cell signaling by direct contact increases cell proliferation via a pi3k-dependent signal," *FEBS Letters*, vol. 514, pp. 238–242, 2002.

[78] R , W., H , W., N , E., and B , N., "Surface treatments and cell attachment," *In Vitro Cellular & Developmental Biology - Plant*, vol. 20, pp. 802–808, Oct. 1984.

[79] R , C. S. and M , P. V., "Polymer substrate topography actively regulates the multicellular organization and liver-specific functions of cultured hepatocytes," *Tissue Engineering*, vol. 5, pp. 407–420, 1999.

[80] R , P., L , C. H., T , J., W , L., P , N., B , C., B , B. D., and S , Z., "1 alpha,25(oh)(2)d-3 regulation of integrin expression is substrate dependent," *Journal of Biomedical Materials Research Part A*, vol. 71A, no. 2, pp. 217–225, 2004.

[81] R , T. W. and C , S., "Picture thresholding using an iterative selection method," *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-8, no. 8, pp. 630–632, 1978.

[82] S , J. R., S , A., W , N., K , D., A , S., and K , J., "Integration of combinatorial synthesis, rapid screening, and computational modeling in biomaterials development," *Macromolecular Rapid Communications*, vol. 25, no. 1, pp. 127–140, 2004.

[83] S , J., K , D., K , J., R , K., W , N., K , V., and WJ, W., "Using surrogate modeling in the prediction of fibrinogen adsorption onto polymer surfaces," *Journal of Chemical Information and Modeling*, vol. 44, pp. 1088–1097, 2004.

[84] S , M. G. P. and R , H., "Density dependent inhibition of cell growth in culture," *Nature*, vol. 215, no. 5097, pp. 171–172, 1967.

[85] S , J., Z , P., G , A. J., and M , J. C., "Knowledge discovery applications in high-throughput polymer characterization," *MRS Proceedings*, vol. 2005 Fall Meeting, pp. 147–152, 2005.

[86] S , H., S , J., B , J., R , B., M , J., and G , Z., "The use of temperature/composition combinatorial libraries to study the effects of biodegradable polymer blend surfaces on vascular cells," *Biomaterials*, vol. 26, pp. 4557–4567, 2005.

[87] T , H. J., "Comments on 'picture thresholding using an iterative selection method'," *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-9, no. 5, p. 311, 1979.

[88] T , E., K , J., and M , P. V., "Peg-variant biomaterials as selectivly adhesive protein templates: model surfaces for controlled cell adhesion and migration," *Biomaterials*, vol. 21, pp. 511–520, 2000.

[89] W , L., O -N , R., B , B. F., Z , G., W , M., C , D. L., B , B. D., and S , Z., "Integrin beta-1 silencing alters osteoblast response to substrate microtopography and 1,25(oh)2d3," *Journal of Bone and Mineral Research*, vol. 20, no. 9, pp. S363–S363, 2005.

[90] W , M. E., "Cell density and n-cadherin interactions regulate cell proliferation in the sensory epithelia of the inner ear," *Journal of Neuroscience*, vol. 22, no. 7, pp. 2607–2616, 2002.

[91] W , N., Y , K., S , C., K , S., and A , E., "High-throughput investigation of osteoblast response to polymer crystallinity: influence of nanometer-scale roughness on proliferation," *Biomaterials*, vol. 25, pp. 1215–1224, 2004.

[92] W , N., B , D., B , S., and K , J., "Small changes in the polymer structure influence the adsorption behavior of fibrinogen on polymer surfaces: Validation of a new rapid screening technique," *Journal of Biomedical Materials Research*, vol. 68A, pp. 496–503, 2004.

[93] W , M., C , B., T , M., and B , D. M., "Use of ti-coated replicas to investigate the effects on fibroblast shape of surfaces with varying roughness and constant chemical composition," *Journal of Biomedical Materials Research*, vol. 60, no. 3, pp. 434–444, 2002.

[94] Z , G., Z , O., S , Z., W , M., L , D., and B , B. D., "Osteoblast-like cells are sensitive to submicron-scale surface structure," *Clinical Oral Implants Research*, vol. 17, no. 3, pp. 258–264, 2006.