BMC Bioinformatics

Methodology article

Local cell metrics: a novel method for analysis of cell-cell interactions

Jing Su[1], Pedro J Zapata[2], Chien-Chiang Chen[2] and J Carson Meredith*[2]

Address: [1]Coulter School of Biomedical Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA and [2]School of Chemical & Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

Email: Jing Su - jing.su.66@gmail.com; Pedro J Zapata - pjzapatab@gatech.edu; Chien-Chiang Chen - chen.iverson@gmail.com; J Carson Meredith* - carson.meredith@chbe.gatech.edu

* Corresponding author

Published: 23 October 2009

BMC Bioinformatics 2009, **10**:350 doi:10.1186/1471-2105-10-350

Received: 6 March 2009
Accepted: 23 October 2009

This article is available from: http://www.biomedcentral.com/1471-2105/10/350

Abstract

Background: The regulation of many cell functions is inherently linked to cell-cell contact interactions. However, effects of contact interactions among adherent cells can be difficult to detect with global summary statistics due to the localized nature and noise inherent to cell-cell interactions. The lack of informatics approaches specific for detecting cell-cell interactions is a limitation in the analysis of large sets of cell image data, including traditional and combinatorial or high-throughput studies. Here we introduce a novel histogram-based data analysis strategy, termed local cell metrics (LCMs), which addresses this shortcoming.

Results: The new LCM method is demonstrated via a study of contact inhibition of proliferation of MC3T3-E1 osteoblasts. We describe how LCMs can be used to quantify the local environment of cells and how LCMs are decomposed mathematically into metrics specific to each cell type in a culture, e.g., differently-labelled cells in fluorescence imaging. Using this approach, a quantitative, probabilistic description of the contact inhibition effects in MC3T3-E1 cultures has been achieved. We also show how LCMs are related to the naïve Bayes model. Namely, LCMs are Bayes class-conditional probability functions, suggesting their use for data mining and classification.

Conclusion: LCMs are successful in robust detection of cell contact inhibition in situations where conventional global statistics fail to do so. The noise due to the random features of cell behavior was suppressed significantly as a result of the focus on local distances, providing sensitive detection of cell-cell contact effects. The methodology can be extended to any quantifiable feature that can be obtained from imaging of cell cultures or tissue samples, including optical, fluorescent, and confocal microscopy. This approach may prove useful in interpreting culture and histological data in fields where cell-cell interactions play a critical role in determining cell fate, e.g., cancer, developmental biology, and tissue regeneration.

Background

Cell-cell recognition is critical to a wide range of problems in biology and medicine [1-16]. The development of experimental approaches associated with cell-cell recognition has promoted advances in understanding these effects, e.g., biochemical assays for protein binding and transcription,. However, less attention has been focused on developing algorithms for the detection of cell-cell rec-

ognition from the structure and spatial distribution of cells. Such methods would offer complimentary benefits to biochemical assays, due to the relative ease of collecting microscopy data from cell cultures and tissues. This would be useful also in combinatorial and high-throughput screening of cell-cell and cell-material interactions [17-23]. In adhesion dependent cells, cell-cell recognition is known to be a crucial step in initiating contact inhibition (CI) of cell migration[24,25] and proliferation[14]. CI plays an important role in the proliferation, invasion, and metastasis of cancer cells, [26-28] cardiovascular tissue homeostasis and development, [5,29] embryonic development, [1] and wound healing, [16] among many other biological phenomena. Conversely, it has also been shown that under certain conditions cell-cell contact can promote cell proliferation, known as "density-dependent" contact stimulation of cell proliferation[12,30-32]. The investigation of mechanisms relating cell contact, cell proliferation, migration, and differentiation, in which cell adhesion molecules play a major role, is a rich area of research. Cadherins, e.g., VE-cadherin, [3] E-cadherin[13] and N-cadherin, [8,9,15,33] and notch proteins[1] have been shown to mediate contact-dependent phenomena in a wide variety of cell types.

Here, CI of proliferation, a known cell-cell recognition phenomenon, is used as a model system for developing algorithms for the analysis of cell-cell recognition from microscopy data. Usually, the effects of cell density on proliferation are studied as relationships between *global descriptions*, such as average cell density, proliferation rate, and protein expression level[2,11,12,14]. However, as we demonstrate below, these summary-statistic descriptions are only sensitive to the effects of very large changes in cell density. As a result, global metrics do not illuminate all of the information available from image data for cell contact phenomena. This is because cell-cell contacts are *local interactions* and are very sensitive to short-range cell-to-cell distance. When global metrics are used, such as cell density in a set of images, 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, which is observed as noise in the data. Furthermore, the stochastic characteristics of cell behavior add to the noise so that only major trends of the responses of cells to neighbors is distinguishable over very large changes in cell density[15]. To address this "dilution" effect, Nelson and Chen studied contact stimulation effects on the growth of a single pair of cells by using specially-designed surfaces to decouple the effects of cell-cell contact from others[12].

In this paper, we introduce a complementary approach that allows focused analysis on nearest-neighbor cells, but permits sampling from cultures with high cell densities and use of any type of surface. We had previously applied this technique to screen large image databases from cell cultures on combinatorial libraries of biomaterials[22,23]. Here, we outline the details of this method and discuss it's generalization to Bayesian modelling. The method is based upon defining local cell metrics (LCMs), which are histograms of cell properties. The use of these local variables expands the sample space considerably and allows separation of arbitrarily-defined short- and long-range effects. We show how the local cell metrics 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 osteoblast proliferation.

Methods

Experimental

Surface Preparation

Poly (DL-lactic-glycolic acid) (PLGA, block copolymer, 50:50 ratio of PGA and PLA, 40,000~75,000 Da) and poly (ε-caprolactone) (PCL, 114,000 Da, M_w/M_n = 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 (22 × 22 mm). To provide adhesion of these polymers to the silicon during cell culture, the silicon was pretreated with a Piranha etch (70% H_2SO_4/21% H_2O/9% H_2O_2 at 80°C for 1 h) followed by 1 min in a HF acid bath and a final rinse in DI water (filtered at 0.2 μm).

Cell Culture

Established from newborn mouse calvaria, [34] the MC3T3-E1 cell line has been shown capable of differentiating into osteoblast and osteocytes *in vitro*[35]. MC3T3-E1 cells have been shown to exhibit specific bone related protein expression patterns, under different developmental stages, similar to primary mouse calvaria cells[35,36]. 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[37]. Since the original MC3T3-E1 cell line has been found phenotypically heterogeneous with regard to cell differentiation, more homogeneous subclones have been established[38]. In this work, MC3T3-E1 subclone 4 (from ATCC, VA), which shows homogenous capabilities of osteogenesis both *in vitro* and *in vivo*, [38] was chosen in order to minimize variations due to phenotypical heterogeneities.

Cell proliferation was assayed by BrdU 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, 30 min),

MC3T3-E1 cells (passage 6) were seeded onto the coated wafers at 4×10^4 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. After seeding, microplates were shaken for 10 min on a shaker (Instrument model, operation frequency) 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. At 5 h post 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. At 18 h post seeding, 2 mM BrdU (5-bromo-2-ϕ-deoxyuridine, Sigma, MO) in PBS was added to the culture medium to reach a final concentration of 20 μM. After 6 h of 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).

Low calcium concentration suppresses contact inhibition of cell growth by deactivating calcium-dependent cadherins[7,39]. This phenomenon was used in this study to validate the local cell metrics, and at the same time the dependency of contact inhibition on calcium was quantitatively studied. In order to investigate the role of Ca^{++} on cell spreading and proliferation, BrdU incorporation experiments in low Ca^{++} medium were performed on TCPS surfaces. Fifteen minutes before the introduction of BrdU, cells were rinsed twice with DPBS (without Ca^{++} and Mg^{++}) and afterword cultured in the low Ca^{++} medium (0.5% FBS in Ca^{++} and Mg^{++} free DPBS)[39]. The rest of the protocol was the same as previously described.

Image Acquisition

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™ monochromic digital camera (SKU S99826, Optronics, CA). The image locations were fixed on a 16 × 20 grid with horizontal and vertical spacing of 1280 μm and vertical spacing of 960 μm. For each location a 1189 × 892 μm^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® 10 g (Oracle, CA) database for further image processing and data analysis.

Image Processing

The Image Processing Toolbox of Matlab™ R14 (MathWorks, MA) was employed for image processing. Due to the volume of image data dynamic, self-adapting algorithms were developed for automated image processing. Binary images of both surface lateral patterns of cell nuclei counter staining were obtained from raw grayscale microscopic images by a *variation-adjusted iterative selection method* (*VAIS*), which was modified from the original iterative selection method [40-45] (details in the BrdU thresholding part below).

Binary images of cell nuclei were segmented by the *marker-controlled watershed* method[42] to separate images of closely-spaced cell nuclei. This process was critical because the nuclei of a pair of recently-proliferated cells were often too close to be distinguished with thresholding alone. Resultant black-and-white cell nuclei images were used as masks by overlaying them with corresponding BrdU staining images to determine the fluorescence intensity of incorporated BrdU. The histogram of BrdU staining intensity per nucleus [see Additional File 1: Figure S.1] is composed of two major peaks: the low intensity peak (background) represents cells at rest, while the high intensity peak (foreground) indicates proliferating cells. The optimal threshold between these two peaks was determined automatically by *VAIS*. Briefly, starting at an initial threshold $T_i = 0.5$, the histogram was divided into resting (background) and proliferating (foreground) parts. Means and standard deviations of the foreground and background, respectively denoted as M_{bi}, M_{fi}, σ_{bi}, and σ_{fi}, were determined by fitting each peak to a Gaussian curve. A new threshold was calculated as $T_{i+1} = (\sigma_{fi}M_{fi} + \sigma_{bi}M_{bi})/(\sigma_{fi} + \sigma_{bi})$ and was repeated until convergence on a stable threshold. Compared with more common iterative selection methods, which use a simple mean intensity, the modified VAIS procedure is more robust when background and foreground intensities have different variances. Indeed, the variance of the BrdU signal intensity from non-proliferating cells was significantly greater than that of the proliferating cells [see Additional File 1: Figure S.1]. During image processing, data washing in the form of median filtering was performed to remove noise below a certain threshold. Image processing was supervised in order to assure the performance of self-adaptive algorithms and images of poor quality not permitting quantification were occasionally discarded. Proliferation behaviors were determined for every cell and stored in the database along with the cell location on the surfaces.

Methodologies of Data Analysis

Global Metrics

Cell density and proliferation were described with summary statistics such as number of resting and proliferated cells computed for each image. This provides a set of glo-

bal metrics for features in each image. As indicated in Figure 1a, 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. The ability to detect contact inhibition of cell proliferation, a known phenomenon, was used as an indicator of the effectiveness of the global metrics cell density and proliferation averages.

Definition of Local Cell Metrics

Source codes that implement the algorithms presented in this section have been made available by the authors. [see Additional File 2] Consider that the collection of all cells (A) is composed of either proliferated (P) or resting cells (R), such that A = P + R. The symbol A represents any cell chosen at random, regardless of proliferative status. The proliferating-resting cell distance, *PR*, is used here to illustrate the definition and properties of local cell metrics, as indicated in Figure 1b. The definitions below are generalizable to any type of cell-cell distance, or any other spatial or temporal metric of cells. 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. In the k^{th} image, the set of all such distances, PR_k is defined as

$$PR_k = \left\{ PR_{ijk} \mid i = 1, 2, \cdots n_{Pk}; j = 1, 2, \cdots n_{Rk} \right\}.$$

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

$$PR = \bigcup_k PR_k.$$

A set of N+1 distance bins is defined as

$$bin_{dist} = \left\{ [d_0, d_1), [d_1, d_2), \cdots, [d_{N-1}, d_N] \right\}$$

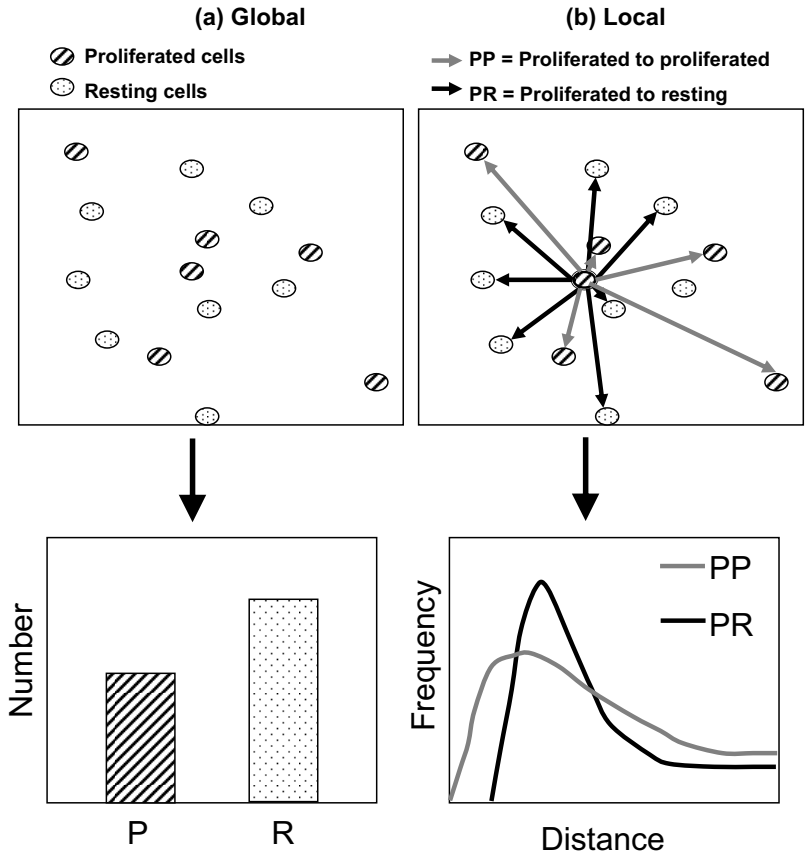

Figure 1

Schematic comparing global versus local metrics. Schematic comparing global versus local metrics. (a) global point of view leading to summary statistics, illustrated by bar graph, (b) local or individual-cell point of view leading to histograms, illustrated by frequency plot.

where d_0, d_1, ..., d_N is a user-defined distance scale over which analysis is to be performed. The centroid of each interval in bin_{dist} is defined as

$$\hat{d}_i = \frac{d_{i-1}+d_i}{2} \quad (\text{if linearly spaced})$$

or

$$\hat{d}_i = \sqrt{d_{i-1} \cdot d_i} \quad (\text{if logarithmically spaced}),$$

and the resultant centroid set for bin_{dist} is

$$\hat{d} = \{\hat{d}_1, \hat{d}_2, \cdots, \hat{d}_N\}.$$

The bin_{dist} is used to sort set *PR* into an *N*-bin histogram

$$N_{PR}(d) = \{N_{PR}(d_i)\}, \; i = 1, N.$$

where $N_{PR}(d_i)$ is the number of *PR* distances that fall in the interval $[d_{i-1}, d_i)$, which is centered at $\hat{d}_i$.

The total number of elements in set *PR* is

$$n_{PR} = \sum_k (n_{Pk} \cdot n_{Rk}).$$

After normalizing by n_{PR}, the frequency function LCM is

$$f_{PR}(\hat{d}) = \frac{N_{PR}(\hat{d})}{n_{PR}}.$$

and $f_{PR}(\hat{d})$ represents the R-type cell environment of the P-type cells observed over distances $\hat{d}$. Frequency functions, denoted as $f_{PP}(\hat{d})$, $f_{AA}(\hat{d})$, $f_{RR}(\hat{d})$, and $f_{PA}(\hat{d})$, may also calculated for cell-to-cell distances *PP*, *AA*, *RR*, and *PA* in similar manner.

LCM Normalization

Normalization is necessary to interpret LCMs in a meaningful manner and to compare the probability of cell responses under different cell environments. One method of normalization is to relate observed occurrences to random occurrences. Given the finite image size and generally non-overlapping nature of cultured cells, the distribution of random cell occurrences is not Gaussian. The random distribution for cell-cell distance, f_{std}, was calculated as the any cell-any cell distribution (f_{AA}) of 1×10^{10} randomly-chosen nuclei positions on a simulated image 1600 pixels by 1200 pixels. The normalized LCM $\tilde{f}_{PR}(\hat{d})$ is

$$\tilde{f}_{PR} = \frac{f_{PR}}{f_{std}}$$

Other LCMs (f_{AA}, f_{PA}, f_{RR}) are normalized similarly, which allows direct comparisons of different types of cell distances on different surfaces.

In addition to normalizing by the standard distribution, f_{std}, direct ratios between LCMs are used also in our analysis, in which case f_{std} cancels, as indicated in the next equation.

$$r_{PR|PA}(d) = \frac{\tilde{f}_{PR}(d)}{\tilde{f}_{PA}(d)} = \frac{f_{PR}/f_{std}}{f_{PA}/f_{std}} = \frac{f_{PR}}{f_{PA}}$$

The ratio $r_{PR|PA}$ highlights the specific effects of non-proliferated cells on the central proliferating cell relative to the effects of 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 investigation of the contribution of each subset to the overall effect. Therefore, ratios of cell backgrounds may be constructed and used as classifiers for screening and identifying significant cell environment patterns. These ratios also define *posterior odds* (PO) of observing certain proliferation behaviors. For example, consider $r_{PA|AA}(d_i)$, and using the subscript *i* to signify the evaluation at a certain distance $\hat{d}_i$, the ratio is calculated as

$$r_{PA|AA,i} = \frac{f_{PA,i}}{f_{AA,i}} = \frac{N_{PA,i}/n_{PA}}{N_{AA,i}/n_{AA}} = \frac{N_{PA,i}/N_{AA,i}}{n_{PA}/n_{AA}}$$

where

$$\frac{n_{PA}}{n_{AA}} = \frac{n_P n_A}{n_A n_A} = \frac{n_P}{n_A} = \langle PO \rangle.$$

$$\frac{N_{PA,i}}{N_{AA,i}} = \frac{N_{P,i} N_{A,i}}{N_{A,i} N_{A,i}} = PO_i = PO(d_i).$$

Applying equations (12) and (13) to equation (11),

$$r_{PA|AA,i} = \frac{PO(d_i)}{\langle PO \rangle}.$$

Thus, $r_{PA|AA}(d_i)$ is a posterior odds that quantifies how the probability of cell proliferation is changed by the presence of a second cell located at distance $\hat{d}_i$, relative to the average proliferation for all cell-cell distances. Computationally, to promote the efficiency of the codes, we defined $A_k = \{P_k, R_k\} = \{P_{1k}, P_{2k}, ..., P_{NPk}, R_{1k}, R_{2k}, ..., R_{NRk}\}$, and removed self-to-self cell pairs (PA_{ijk} where $i = j$) and identical cell pairs (PA_{ijk} where $i > j$) from PA.

LCM Decomposition

Furthermore, each set of cell-to-cell distances can be decomposed into subsets, which allows isolation of each subset's contribution. For example, consider $r_{PR|PA}$ defined above. As described graphically in Figure 2, since the denominator $PA = PP \cup PR$ (the union of distance sets PP and PR) and $PP \cap PR = \varnothing$ ($\varnothing$ = the empty set), one may remove the PR component from PA, and the denominator becomes PP. The result is that the ratio $r_{PR|PA}$ is transformed into $r_{PR|PP}$. By removing the shared, or overlapping, component PR from the denominator, $r_{PR|PP}$ has higher "contrast" for observing effects of R-cells on P-cells than $r_{PR|PA}$.

Connection of LCMs to Bayesian Statistics

Local cell metrics are naturally connected to Bayesian analysis, which is a powerful statistical method used for classification[46,47]. Specifically for the PR distance, the Bayesian approach allows one to quantify the local environment of P cells, as the conditional probability of finding an R cell a certain distance PR from a P cell. Based on the definition of f_{PR} in equation (8), a naïve Bayes model can be established as follows. Consider a "test" cell chosen at random. It is desired to predict the possibility this cell will be in proliferating status, based upon the local environment of non-proliferating cells, which is given by the following conditional probability function

$$p\left(prolif \mid R_{d1}, R_{d2}, ..., R_{dN}\right)$$

where $p(R_{di})$ represents the probability of finding a non-proliferating cell at a distance of $\hat{d}_i$ from the central, randomly chosen, test cell. Using Bayes's theorem,

$$p\left(prolif \mid R_{d1}, R_{d2}, ..., R_{dN}\right) = \frac{p(prolif) \cdot p(R_{d1}, R_{d2}, ..., R_{dN} \mid prolif)}{p(R_{d1}, R_{d2}, ..., R_{dN})}$$

In the above function, the components $p(R_{\hat{d}1}, R_{\hat{d}2}, ..., R_{\hat{d}N})$ and $p(prolif)$ are constants that can be determined from their frequency in the data. The only non-constant component, the class-conditional probability, is given by

$$\begin{aligned}
&p\left(R_{\hat{d}1}, R_{\hat{d}2}, ..., R_{\hat{d}N} \mid prolif\right) = \\
&= p(R_{\hat{d}1} \mid prolif) \cdot p\left(R_{\hat{d}2}, \cdots, R_{\hat{d}N} \mid prolif, R_{\hat{d}1}\right) \\
&= p(R_{\hat{d}1} \mid prolif) \cdot p\left(R_{\hat{d}2} \mid prolif\right) \cdot p\left(R_{\hat{d}3}, \cdots, R_{\hat{d}N} \mid prolif, R_{\hat{d}1}, R_{\hat{d}2}\right) \\
&\vdots \\
&= p(R_{\hat{d}1} \mid prolif) \cdot p\left(R_{\hat{d}2} \mid prolif\right) \cdots p\left(R_{\hat{d}N} \mid prolif, R_{\hat{d}1}, R_{\hat{d}2}, \cdots, R_{\hat{d}N-1}\right)
\end{aligned}$$

Assuming the occurrence probabilities around the non-proliferating cell distances $R_{\hat{d}1}, R_{\hat{d}2}, ..., R_{\hat{d}N}$ are conditionally independent (uncorrelated), then $p(R_{di} \mid prolif) = p(R_{di} \mid prolif, R_{dj})$ where $i \neq j$ (the naïve Bayes assumption). Hence, equation (17) reduces to

$$p\left(R_{\hat{d}1}, R_{\hat{d}2}, \cdots, R_{\hat{d}N} \mid prolif\right) = \prod_{i=1}^{N} p(R_{\hat{d}i} \mid prolif)$$

A key development is to notice that $p(R_{di} \mid prolif) = f_{PR}(d_i)$, which means that under the naïve Bayes assumption the LCM is in fact a class-conditional probability function. The term $p(R_{\hat{d}1}, R_{\hat{d}2}, ..., R_{\hat{d}N})$ represents the probability of locating the R cells at dis-

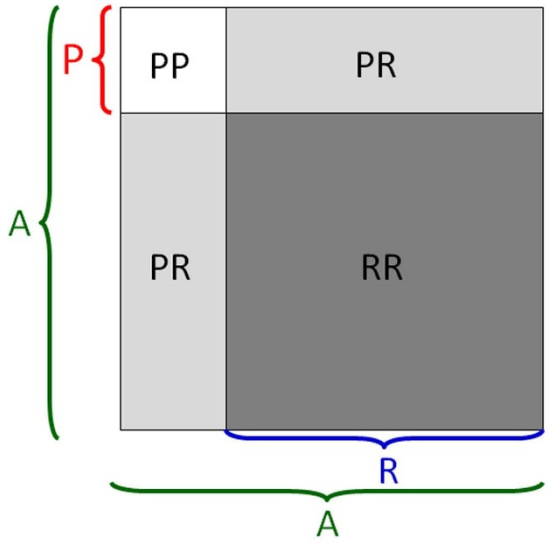

Figure 2
Schematic indicating the decomposition of local metrics. Schematic indicating the decomposition of local metrics into groups of cell-cell distances based on cell proliferative status. A = any cell, P = proliferated cell, R = resting cell (nonproliferated).

tances ($\hat{d}_1, \hat{d}_2, ..., \hat{d}_N$) from any cell, which is $f_{RA}\left(\hat{d}\right)$. Hence, the Bayes conditional probability function from equation (16) becomes

$$p\left(prolif \mid R_{d1}, R_{d2}, \cdots, R_{dN}\right) = p(prolif) \cdot \prod_{i=1}^{N} \frac{f_{PR}(d_i)}{f_{RA}(d_i)}$$

The naïve Bayes model allows prediction of the probability of proliferation as a function of the LCMs, $f_{PR}\left(\hat{d}\right)$ and $f_{RA}\left(\hat{d}\right)$, which are easily computed from a training data set, as is *p*(*prolif*). The evaluation of this modeling approach will be the subject of forthcoming work.

Results and Discussion

Traditional 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 3 and Table 1. Although a trend towards lower proliferation at higher cell density is somewhat apparent in Figure 3, global analysis, does not allow for quantitative detection of CI effects on cell proliferation at a statistically-significant level. The linear regression (Table 1) yields in an adjusted R^2 of 0.128 (on PLGA) or 0.109 (on PCL), indicating that the contact inhibition effect masked by "noise" in the data. Furthermore, it is obvious that in Figure 3, no regression function can be fit satisfactorily to the global statistics, since the noise is too high relative to the CI effects. The use of a larger range may allow the global analysis to distinguish contact effects from natural variance in cell properties. This has been demonstrated in repeated experiments on PLGA, PCL, and TCPS surfaces. [see Additional File 1: Figures S.2, S.3, S.4]. 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.

Local Cell Based Metrics

The noise level inherent to proliferation measurements, which are normally carried out over a small seeding density range, make contact inhibition a robust test-case for comparing local *vs*. 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) provides the random noise reference. For local metrics, the random cell-cell distance frequency distribution was calculated using a Monte-Carlo approach, termed the standard frequency distribution, f_{std}. The reference f_{std} is shown in Figure 4 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 due to the non-overlapping nature of the nuclei centers at close distances. The computed f_{std} distribution is nearly identical to the

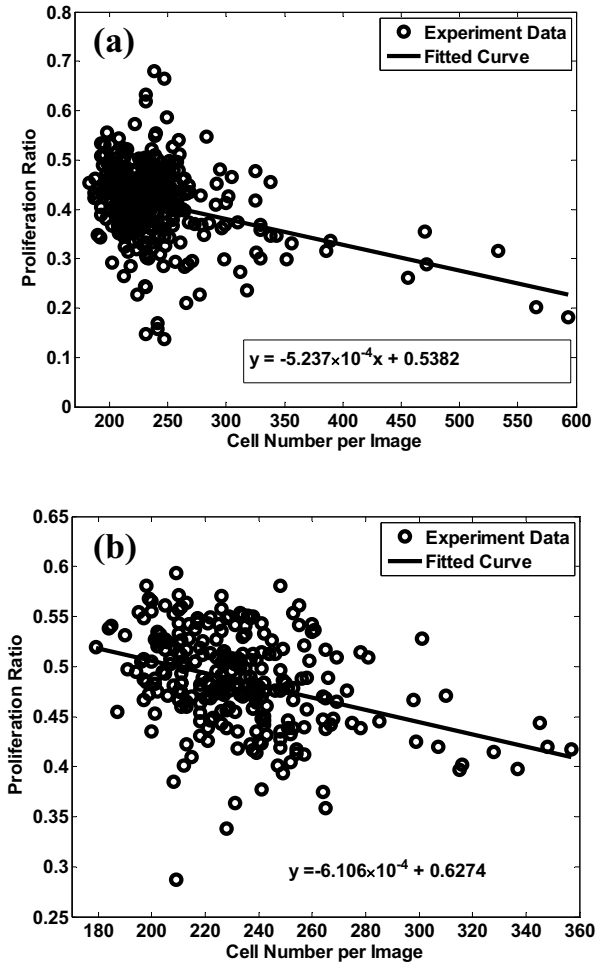

Figure 3
Effects of global cell density on global cell proliferation. Effects of global cell density on global cell proliferation ratio for MC3T3-E1 cultured on (a) PLGA and (b) PCL. Number of images used was (a) 353 and (b) 288.

Table 1: Linear Regression from Global Analysis Results from Figure 3

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

*SSE = Sum of squared error, RMSE = root mean squared error, R^2 = linear correlation coefficients

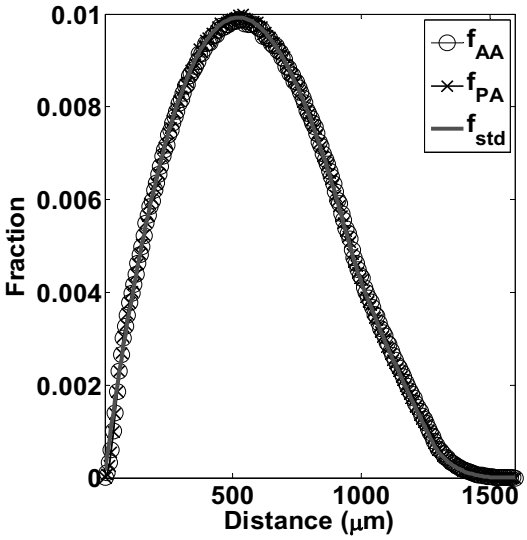

Figure 4

Comparison of experimental and calculated distributions. Comparison of experimentally determined f_{AA} and f_{PA} for MC3T3-E1 on a PLGA surface and the computed standard curve, f_{std}. These data represent the frequency at which any cell is located a certain distance from any other cell (f_{AA}) or from a proliferated cell (f_{PA}) in the experiments. The standard curve was computed from a Monte Carlo simulation and represents a uniformly random probability of locating any cell a given distance from any other cell on the same sized area as our microscope images. Number of images used was 353.

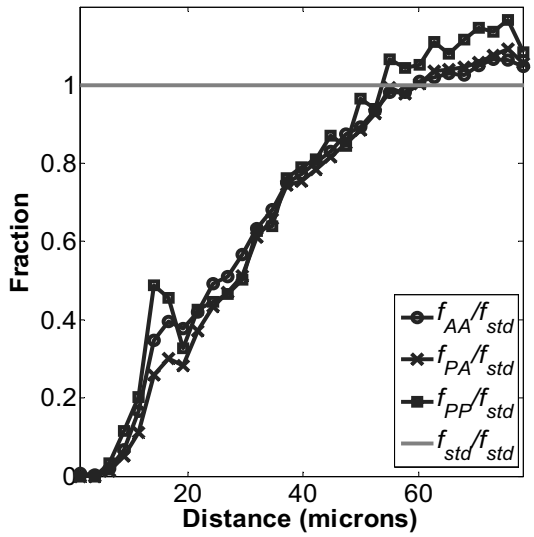

Figure 5

Normalized distributions for cells on PLGA surfaces. Normalized distributions for any cell-any cell (AA), proliferating-any cell (PA) and proliferating-proliferating cell (PP) on a PLGA surface: $\tilde{f}_{AA}(d)$, $\tilde{f}_{PA}(d)$, and $\tilde{f}_{PP}(d)$. Normalization was performed by dividing the experimental frequency distribution by f_{std}, the random cell-cell distribution determined from Monte Carlo simulation. Number of images used was 353.

experimental f_{AA} distribution at large distances (> 100 μm). 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 4 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 proliferative status then f_{PA} and f_{AA} should differ from one another and from f_{std}, but only at close distances where cell-cell contact is likely to occur.

Figures 5, 6, and 7, which show the normalized distributions f_{PA}/f_{std}, f_{AA}/f_{std} and f_{PP}/f_{std} at close distances, indicate the non-random effects of contact inhibition when the values become less than one. Specifically, CI occurs when the distance between cell nuclei becomes less than about 50 μm. The typical mean cell area was around 2500 μm^2, resulting in a mean diameter of 56 μm, which corresponds closely to the onset of CI. Representative images of cultured MC3T3-E1 cells on these surfaces have been presented in previous work[22,23]. In addition, local fine structure in the contact inhibition region is observed as a local maximum peak between 10 and 20 μm. This peak indicates enhanced local proliferation at very close distance, even when overall proliferation is being inhibited. Interestingly the local peak magnitudes at 10 to 20 μm 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, which have not had 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 was observed comparing Figures 5, 6, and 7 (f_{PP}) to Figures 8, 9, 10, and 11 (f_{PR}). In addition, in the Monte Carlo simulation of random cell positions (f_{std}), with no proliferation, this local peak is absent.)

Direct ratios between experimentally-determined distributions can be chosen specifically to illuminate the CI phenomena of interest. Specifically, common components in the numerator and denominator not related to CI phenomena may be removed, thus isolating the phenomena of interest. This process is examined in Figures 8, 9, 10, and 11. Figure 8 shows the $r_{PA|AA}$ profile, which is classified into two regions: the *proliferation suppressed region* (0

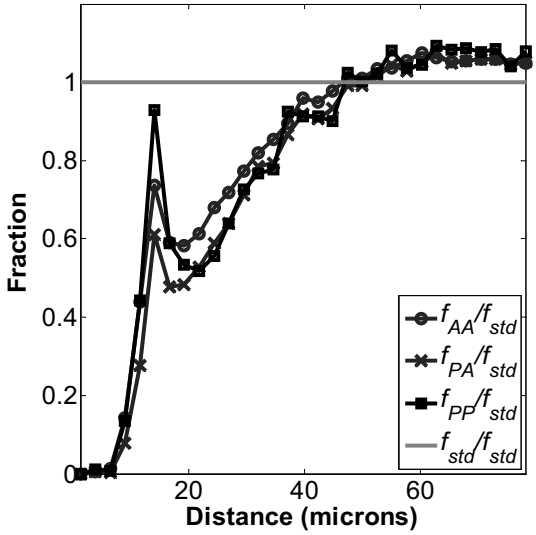

Figure 6

Normalized distributions for cells on PCL surfaces. Normalized distributions as described in Figure 5 on a PCL surface: $\tilde{f}_{AA}(d)$, $\tilde{f}_{PA}(d)$, and $\tilde{f}_{PP}(d)$. Number of images used was 288.

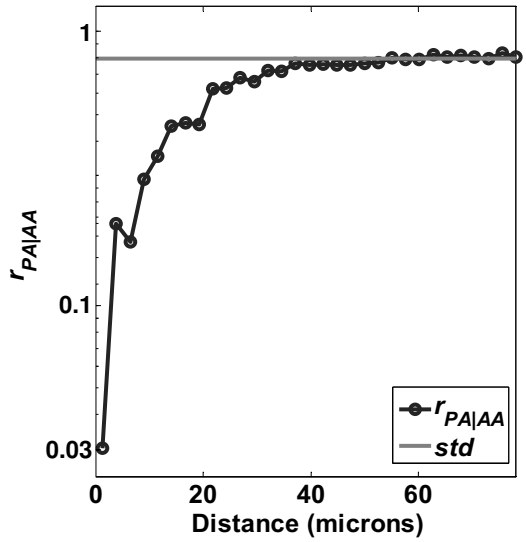

Figure 8

Ratio of distributions PA/RR on PLGA surfaces. Ratio of distributions $r_{PA|RR} = \tilde{f}_{PA}(d) / \tilde{f}_{RR}(d)$ on a PLGA surface. Number of images used was 353.

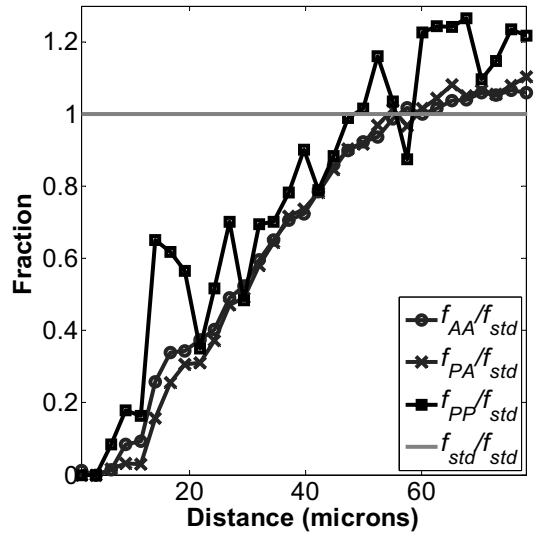

Figure 7

Normalized distributions for cells on TCPS surfaces. Normalized distributions as described in Figure 5 on a TCPS surface: $\tilde{f}_{AA}(d)$, $\tilde{f}_{PA}(d)$, and $\tilde{f}_{PP}(d)$. Number of images used was 291.

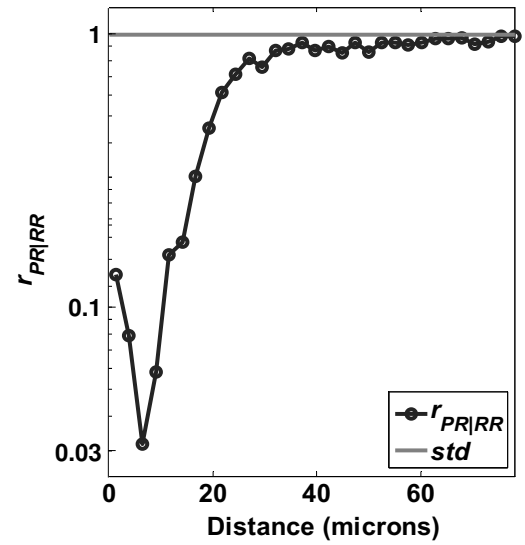

Figure 9

Ratio of distributions PR/RR on PLGA surfaces. Ratio of distributions $r_{PR|RR} = \tilde{f}_{PR}(d) / \tilde{f}_{RR}(d)$ on a PLGA Surface. Number of images used was 353.

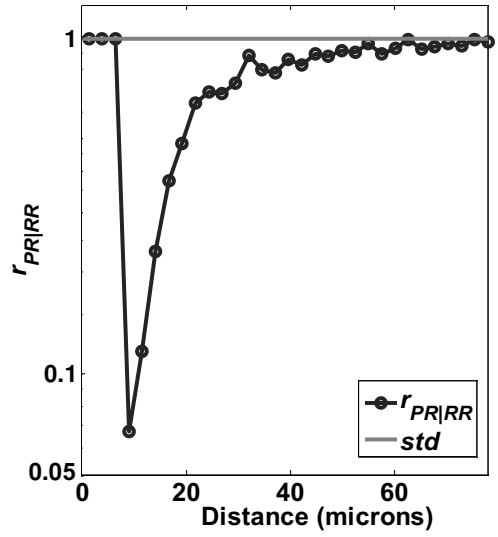

Figure 10
Ratio of distributions PR/RR on PCL surfaces. Ratio of distributions $r_{PR|RR}$ on a PCL surface. Number of images used was 288.

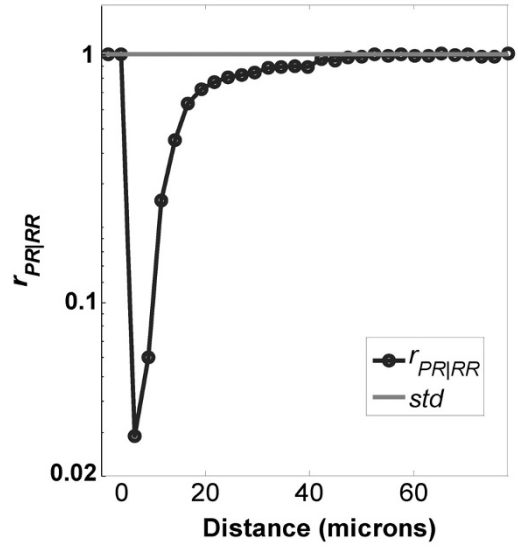

Figure 11
Ratio of distributions PR/RR on TCPS surfaces. Ratio of distributions $r_{PR|RR}$ on a TCPS surface. Number of images used was 291.

to 40 μm) where cell proliferation was suppressed up to 6 fold ($r_{PA|AA}$ falls to ~1/6) and the *null region* (beyond 40 μm) where cell proliferation was not noticeably affected by the contact of other cells. Based on the discussion for Figure 7, 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 $PP \subset PA$. We illustrate how separation of the *PP* and *PR* components enhances the detection of contact inhibition of proliferation. By definition the various distances are related as follows

$$PA = PP \bigcup PR$$
$$RA = RR \bigcup PR$$
$$AA = (PP \bigcup RR) \bigcup PR$$

with

$$PR \bigcap RR = \varnothing$$
$$PP \bigcap PR = \varnothing$$

where ∅ is the empty set. The two shared components of *AA* and *PA* are *PP* and *PR*. The *PP* distance component represents distances between cells that have both proliferated, i.e., *proliferation in those pairs was not contact inhibited*. Removal of the common *PP* component from the numerator and denominator of $r_{PA|AA}$ leads to $r_{PR|RR}$, shown in Figures 9, 10, and 11 for the PLGA, PCL, and TCPS surfaces, respectively. By definition $r_{PR|RR}$ should be more sensitive to CI of proliferation, because non contact inhibited cell pairs have been removed. In Figure 9, the $r_{PR|RR}$ ratio is classified into two regions: the *NaN region* (below 5 μm) where few pairs occur, and the *contact inhibition region* (5 to 40 μm). 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 μm, where contact inhibition effects are maximized. To our knowledge, this is the *first time both the magnitude and the range of contact inhibition of cell proliferation* have been determined quantitatively in a single function.

The physical meaning of the LCM ratio $r_{PR|RR}$ can be can be illustrated by recognizing that it is the posterior odds (PO) of proliferation as a function of cell-cell distance. Consider two cells that are well-separated at 40 μm, and another two cells that are at a close distance of 8 μm, where the extreme in contact inhibition behavior is found (minimum $r_{PR|RR}$ in Figure 9). The PO that one of the closely-spaced cells has proliferated is $PO_{PR|RR} = r_{PR|RR}(8)/r_{PR|RR}(40) = 1/32$. This means there is a 32 fold lower

Table 2: Minima in $r_{PR|RR}$ Curve Indicating Maximum Contact Inhibition

Surface	$1/r_{PR\|RR,\ min}$	d_{min}(μm)
PLGA	31.6	8
PCL	15.8	9
TCPS	35.5	6

chance of proliferation at 8 μm than at a distance of 40 μm.

The profiles of $r_{PR/RR}$ from the other polymer surfaces are shown in Figure 10 (PCL surface) and Figure 11 (TCPS surface). The ratios are similar in shape but have different magnitudes for the minimum point as a function of the surface. Table 2 summarizes the variation of $r_{PR|RR,\ min}$ and d_{min} on the different surfaces. The different location and strength of contact inhibition might be due to surface features such as roughness, crystallinity, hydrophobicity, surface charge, or protein adsorption, factors which are known to influence osteoblast proliferation[48]. For example, the surface roughness increases in the order TCPS < PLGA < PCL, and at the same time the $PO_{PP/PR}$ is decreasing, and the d_{min} is increasing. We illustrate this point, however, not to make a definite mechanistic argument about surface effects on proliferation, which is certainly more complicated than roughness alone. Rather, the point made is that the LCM method is capable of sensitive detection of differences in proliferation for cells cultured on different surfaces.

The effect of calcium depletion on LCMs is presented in Figure 12. The $\tilde{f}_{AA}(d)$ distribution of the control (calcium +) was significantly lower than the low calcium (calcium -) case. In addition, the 'calcium -' curve stays close to unity except at very close distances, less than 20 μm, whereas the 'calcium +' curve falls below unity at 50 μm. Hence, the LCM $\tilde{f}_{AA}(d)$ detects the expected result: that low calcium should inhibit the cell-cell self-avoidance and contact inhibition[7,39]. This effect is seen more clearly in examining the ratio $r_{PR/RR}$, in Figure 13. Contact inhibition was very strong within a cell-to-cell distance of 30 μm when cells were cultured under physiological calcium concentration (calcium +). However, contact inhibition disappeared when calcium was depleted (calcium -).

Conclusion

We have shown that global summary statistics are not adequate metrics for detecting local cell interactions, due to noise and non-local effects inherent to cell-cell contact phenomena. A novel data analysis strategy, *local cell metrics*, has been introduced in this paper. LCMs, which are cell-cell distance histograms, describe cell environments from the "point of view" of individual cells. These metrics allow focusing of analysis onto arbitrarily-defined close distances. In addition, LCMs can be 'tuned' to be sensitive to specific contact phenomena by decomposing the distributions into specific cell-types (proliferating vs. nonproliferating) and removing unwanted components. Local metrics as defined herein are not limited to proliferation analysis, nor to cell-cell interactions alone. The metrics are generic and can be, in principle, applied to any type of

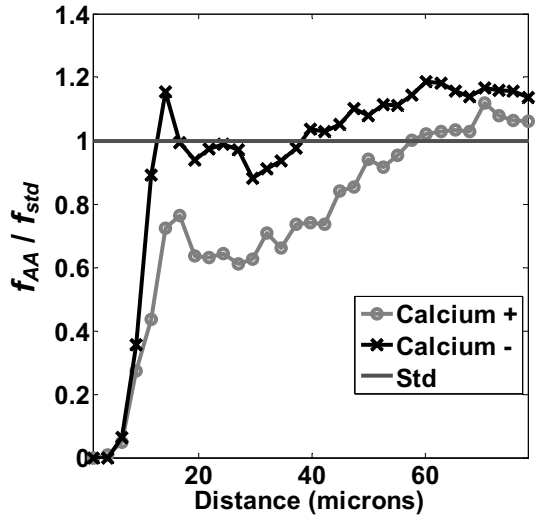

Figure 12

Effects of calcium depletion: loss of contact inhibition. Normalized distribution $\tilde{f}_{AA}(d)$ examining effects of calcium depletion on proliferation of MC3T3-E1 cultured on TCPS surfaces. Number of images used was 198.

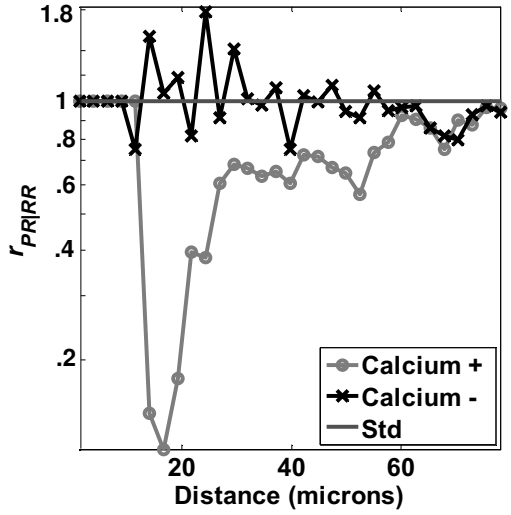

Figure 13

Effects of calcium depletion: loss of contact inhibition. Ratio of distributions $r_{PR|RR}$ examining effects of calcium depletion on proliferation of MC3T3-E1 cultured on TCPS surfaces. Number of images used was 204.

quantifiable cell assay, and can be applied to cell-biomaterial and cell-tissue interactions as well. We have also shown how LCMs are related to the naïve Bayes model, which makes them useful for data mining and classification (the subject of forthcoming work.)

We have demonstrated the new local metrics by considering the contact inhibition of proliferation of the osteoblast cell line MC3T3-E1. A quantitative and probabilistic description of the contact inhibition effect as a function of cell-cell distance has been achieved. In fact, the probability of proliferation is shown to be strongly dependent on the distance to, and proliferative state of, neighboring cells. The LCMs were also sensitive to effects of the culture surface, and of calcium composition in the culture media, on proliferation.

Authors' contributions

JCM directed the experimental design, data analysis method development, and writing of the manuscript. JS performed most cell culture experiments and implemented the LCM method, including programming, and wrote the manuscript. PJZ helped in data interpretation and statistical analysis. CCC performed the calcium-depletion experiments and analysis. All authors have read and approved this manuscript.

Additional material

Additional file 1

Supplementary materials. *Supplementary theoretical and calculation details, as well as additional data on image analysis and global analysis are provided.*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2105-10-350-S1.DOC]

Additional file 2

Source codes for local cell metrics. *This .zip archive folder contains sources codes for implementing local cell metrics provided under terms of the GNU open source license.*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2105-10-350-S2.ZIP]

Acknowledgements

We gratefully acknowledge support from NIH Grant Numbers RR17425 and HK072039.

References

1. Artavanis-Tsakonas S, Matsuno K, Fortini M: **Notch signaling.** *Science* 1995, **5208:**225-232.
2. Castor LN: **Contact Inhibitions of Cell Division and Cell Movement.** *J Investig Dermatol* 1972, **59(1):**27-32.
3. Caveda L, Martin-Padura I, Navarro P, Breviario F, Corada M, Gulino D, Lampugnani MG, Dejana E: **Inhibition of Cultured Cell Growth by Vascular Endothelial Cadherin (Cadherin-5/VE-Cadherin).** *Journal of Clinical Investigation* 1996, **98(4):**886-893.
4. Charrasse S, Comunale F, Gilbert E, Delattre O, Gauthier-Rouviere C: **Variation in cadherins and catenins expression is linked to both proliferation and transformation of Rhabdomyosarcoma.** *Oncogene* 2004, **23(13):**2420-2430.
5. Grego-Bessa L, Luna-Zurita G, del Monte V, Bolós P, Melgar A, Arandilla A, Garratt H, Zang Y, Mukouyama HC: **Notch Signaling Is Essential for Ventricular Chamber Development.** *Developmental Cell* 2007, **12:**415-429.
6. Hirano S, Nose A, Hatta K, Kawakami A, Takeichi M: **Calcium-dependent cell-cell adhesion molecules (cadherins): subclass specificities and possible involvement of actin bundles.** *J Cell Biol* 1987, **105(6):**2501-2510.
7. Kandikonda S, Oda D, Niederman R, Sorkin BC: **Cadherin-Mediated Adhesion Is Required for Normal Growth Regulation of Human Gingival Epithelial Cells.** *Cell Communication & Adhesion* 1996, **4(1):**13-24.
8. Levenberg S, Katz BZ, Yamada KM, Geiger B: **Long-range and selective autoregulation of cell-cell or cell-matrix adhesions by cadherin or integrin ligands.** *Journal of Cell Science* 1998, **111(3):**347-357.
9. Levenberg S, Sadot E, Goichberg P, Geiger B: **Cadherin-mediated transmembrane interactions.** *Cell Adhes Commun* 1998, **6(2-3):**161-170.
10. Mary S, Charrasse S, Meriane M, Comunale F, Travo P, Blangy A, Gauthier-Rouviere C: **Biogenesis of N-Cadherin-dependent Cell-Cell Contacts in Living Fibroblasts Is a Microtubule-dependent Kinesin-driven Mechanism.** *Molecular Biology of the Cell* 2002, **13(1):**285-301.
11. Nakatsuji Y, Miller RH: **Density dependent modulation of cell cycle protein expression in astrocytes.** *Journal of Neuroscience Research* 2001, **66(3):**487-496.
12. Nelson CM, Chen CS: **Cell-cell signaling by direct contact increases cell proliferation via a PI3K-dependent signal.** *FEBS Letters* 2002, **514:**238-242.
13. St Croix B, Sheehan C, Rak JW, Florenes VA, Slingerland JM, Kerbel RS: **E-Cadherin-dependent Growth Suppression is Mediated by the Cyclin-dependent Kinase Inhibitor p27KIP1.** *J Cell Biol* 1998, **142(2):**557-571.
14. Stoker MGP, Rubin H: **Density Dependent Inhibition of Cell Growth in Culture.** *Nature* 1967, **215(5097):**171-172.
15. Warchol ME: **Cell Density and N-Cadherin Interactions Regulate Cell Proliferation in the Sensory Epithelia of the Inner Ear.** *Journal of Neuroscience* 2002, **22(7):**2607-2616.
16. Zegers MMP, Forget M-A, Chernoff J, Mostov KE, Beest MBAt, Hansen SH: **Pak1 and PIX regulate contact inhibition during epithelial wound healing.** *The EMBO Journal* 2003, **22:**4155-4165.
17. Broderick S, Suh C, Nowers J, Vogel B, Mallapragada S, Narasimhan B, Rajan K: **Informatics for combinatorial materials science.** *Journal of the Minerals, Metals and Materials Society* 2008, **60:**56-59.
18. Kennedy S, Washburn N, Simon C, Amis E: **Combinatorial screen of the effect of surface energy on fibronectin-mediated osteoblast adhesion, spreading and proliferation.** *Biomaterials* 2006, **27:**3817-3824.
19. Kohn J: **New approaches to biomaterials design.** *Nature Materials* 2004, **3:**745-747.
20. Meredith JC: **Advances in Combinatorial and High-Throughput Screening of Biofunctional Polymers for Gene Delivery, Tissue Engineering and Anti-Fouling Coatings.** *Journal of Materials Chemistry* 2009, **19:**34-45.
21. Meredith JC, Sormana JL, Keselowsky BG, Garcia AJ, Tona A, Karim A, Amis EJ: **Combinatorial characterization of cell interactions with polymer surfaces.** *J Biomed Mater Res A* 2003, **66(3):**483-490.
22. Su J, Meredith JC: **Local Histogram Analysis: Detecting Cell-Microstructure Interactions on Combinatorial Biomaterial Libraries.** *Combinatorial Chemistry and High-Throughput Screening* 2009, **12:**626-633.
23. Su J, Zapata P, Meredith JC: **Knowledge Discovery Applications in High-Throughput Polymer Characterization.** In *Combinatorial Methods and Informatics in Materials Science* Edited by: Fasolka M. Warrendale, PA: Materials Research Society; 2006. 0894-LL0805-0804

24. Abercrombie M, Ambrose EJ: **Interference microscope studies of cell contacts in tissue culture.** *Experimental Cell Research* 1958, **15(2):**332-345.
25. Abercrombie M, Heaysman JEM: **Observations on the social behaviour of cells in tissue culture: II. "Monolayering" of fibroblasts.** *Experimental Cell Research* 1954, **6(2):**293-306.
26. Abercrombie M, Heaysman JEM, Karthauser HM: **Social behaviour of cells in tissue culture III. Mutual influence of sarcoma cells and fibroblasts.** *Experimental Cell Research* 1957, **13(2):**.
27. Bremnes RM, Veve R, Hirsch FR, Franklin WA: **The E-cadherin cell-cell adhesion complex and lung cancer invasion, metastasis, and prognosis.** *Lung Cancer* 2002, **36(2):**115-124.
28. Sharma VM, Draheim KM, Kelliher MA: **The Notch1/c-Myc pathway in T cell leukemia.** *Cell Cycle* 2007, **6:**927-930.
29. Liu Z-J, Shirakawa T, L Y, Soma A, Oka M, Dotto GP, Fairman RM, Velazquez OC, Herlyn M: **Regulation of Notch1 and Dll4 by Vascular Endothelial Growth Factor in Arterial Endothelial Cells: Implications for Modulating Arteriogenesis and Angiogenesis.** *Mol Cell Biol* 2003, **23:**14-25.
30. Burke JM: **Cell-cell contact promotes DNA synthesis in retinal glia but not in fibroblasts.** *Exp Cell Res* 1983, **146(1):**204-206.
31. Burke JM: **Growth in retinal glial cells in vitro is affected differentially by two types of cell contact-mediated interactions.** *Exp Cell Res* 1989, **180(1):**13-19.
32. Gaunt SJ, Subak-Sharpe JH: **Cell cycle variation associated with feeder effects in cultures of Chinese hamster fibroblasts.** *Exp Cell Res* 1977, **109(2):**341-348.
33. Polyak K, Kato JY, Solomon MJ, Sherr CJ, Massague J, Roberts JM, Koff A: **p27Kip1, a cyclin-Cdk inhibitor, links transforming growth factor-beta and contact inhibition to cell cycle arrest.** *Genes Dev* 1994, **8(1):**9-22.
34. Kodama H, Amagai Y, Sudo H, Kasai S, Yamamoto S: **Establishment of a clonal osteogenic cell line from newborn mouse calvaria.** *Japanese Journal of Oral Biology* 1981, **23:**899-901.
35. Sudo H, Kodama HA, Amagai Y, Yamamoto S, Kasai S: **In vitro differentiation and calcification in a new clonal osteogenic cell line derived from newborn mouse calvaria.** *The Journal of Cell Biology* 1983, **96(1):**191-198.
36. Choi JY, Lee BH, Song KB, Park RW, Kim IS, Sohn KY, Jo JS, Ryoo HM: **Expression patterns of bone-related proteins during osteoblastic differentiation in MC3T3-E1 cells.** *Journal Of Cellular Biochemistry* 1996, **61(4):**609-618.
37. Sodek J, Berkman FA: **Bone cell cultures.** *Methods Enzymol* 1987, **145:**303-324.
38. Wang D, Christensen K, Chawla K, Xiao G, Krebsbach PH, Franceschi RT: **Isolation and characterization of MC3T3-E1 preosteoblast subclones with distinct in vitro and in vivo differentiation/mineralization potential.** *J Bone Miner Res* 1999, **14(6):**893-903.
39. Qian X, Karpova T, Sheppard AM, McNally J, Lowy DR: **E-cadherin-mediated adhesion inhibits ligand-dependent activation of diverse receptor tyrosine kinases.** *EMBO J* 2004, **23(8):**1739-1748.
40. Magid A, Rotman SR, Weiss AM: **Comments on Picture thresholding using an iterative selectionmethod.** *ystems, Man and Cybernetics, IEEE Transactions on* 1990, **20(5):**1238-1239.
41. Meyer F, Beucher S: **Morphological segmentation.** *Journal of Visual Communication and Image Representation* 1990, **1(1):**21-46.
42. Ridler TW, Calvard S: **Picture thresholding using an iterative selection method.** *IEEE Transactions on Systems, Man and Cybernetics* 1978, **SMC-8(8):**630-632.
43. Sezgin M, Sankur B: **Survey over image thresholding techniques and quantitative performance evaluation.** *Journal of Electronic Imaging* 2004, **13(1):**146-168.
44. Trussell HJ: **Comments on 'Picture thresholding using an iterative selection method'.** *IEEE Transactions on Systems, Man and Cybernetics* 1979, **SMC-9(5):**311.
45. Yanni MK, Horne E: **A new approach to dynamic thresholding.** *Eur Assoc Signal Process: Edinburgh, UK* 1994:34-41.
46. Carlin BP, Louis TA: **Bayes and Empirical Bayes methods for data analysis.** *Volume 47*. Taylor & Francis CRC Press; 2000.
47. French S, Smith JQ: **The practice of Bayesian analysis.** New York: J. Wiley & Sons; 1997.
48. Martin JY, Schwartz Z, Hummert TW, Schraub DM, Simpson J, Lankford J, Dean DD, Cochran DL, Boyan BD: **Effect of Titanium Surface-Roughness on Proliferation, Differentiation, and Protein-Synthesis of Human Osteoblast-Like Cells (Mg63).** *Journal of Biomedical Materials Research* 1995, **29(3):**389-401.