ANALYSIS OF SPATIAL VARIATION OF NUCLEAR MORPHOLOGY IN TISSUE MICROENVIRONMENTS

S. Singh1, S. Raman1, E. Caserta2, G. Leone2, M. Ostrowski2, J. Rittscher3, R. Machiraju1

1Dept. of Computer Science and Engineering, The Ohio State University, Columbus, OH, U.S.A.
2Dept. of Molecular Genetics, The Ohio State University, Columbus, Ohio, U.S.A.
3General Electric Global Research Center, Niskayuna, NY, U.S.A.

ABSTRACT

We present a study of the spatial variation of nuclear morphology of stromal and cancer-associated fibroblasts in the mouse mammary gland. The work is part of a framework being developed for the analysis of the tumor microenvironment in breast cancer. Recent research has uncovered the role of stromal cells in promoting tumor growth and progression. In specific, studies have indicated that stromal fibroblasts - formerly considered to be passive entities in the extra-cellular matrix - play an active role in the progression of tumor in mammary tissue. We have focused on the analysis of the nuclear morphology of fibroblasts, which several studies have shown to be a critical phenotype in cancer. An essential component of our approach is that the nuclear morphology is studied within the 3D spatial context of the tissue, thus enabling us to pose questions about how the locus of a cell relates to its morphology, and possibly to its function. In order to make quantitative comparisons between nuclear populations, we build statistical shape models of cell populations and infer differences between the populations through these models. We present our observation on both normal and tumor tissues from the mouse mammary gland.

Index Terms— Tumor Microenvironment, Nuclear Morphometry, Statistical Shape Analysis, Spatial Statistics, Cancer Biology

1. INTRODUCTION

In recent years, the paradigm in cancer research has progressively moved away from viewing carcinogenesis as a cell autonomous, cancer cell-centered condition, and rather towards viewing it as a disease involving complex multicellular interactions within the cancer tissue [1]. Taking this approach, several studies of the tumor microenvironment have revealed that cells in tumor stroma play an active role in promoting tumor growth and progression [2, 3]. Rapid advances in microscopy have enabled the imaging of the tumor microenvironment with the complete 3D context. Further, recombinant DNA technology is used to create transgenic strains that have cell-type specific fluorescence, enabling the identification of various cell types in the microenvironment. Thus, it is now possible to study the biology of the tumor microenvironment with a spatial context - a context that is key to understanding the interactions between tumor cells and the surrounding (stromal) cells.

Fibroblasts, stromal cells that are the principal constituent of the extra-cellular matrix, have been discovered to play an active role in the initiation of cancer [3]. In our studies, which are focussed on deriving insights into human breast cancer through sophisticated mouse models, we seek to explore the mechanisms of interactions between these cells and the tumor they surround. A key phenotypic characteristic we seek to understand is the variation in nuclear morphology of the stromal cells. This is motivated by the fact that while significant progress has been made in understanding the molecular changes in cancer cells, cancer diagnosis remains largely guided by assessing the histological and cytological characteristics of tumor tissue. In particular, advances in the understanding of nuclear structure of cancer cells have thrown light on the process of tumor development [4, 5]. Along similar lines, the nuclear morphology of stromal cells is an essential phenotype, the study of which could lead to important insights into the complex process of tumor growth.

In this paper, we present an investigation of the morphological variations in the nuclei of stromal fibroblast cells in both normal and tumor tissue in the mouse mammary gland. The tissues are imaged in 3D using confocal fluorescence microscopy. A salient aspect of our work is that we conduct this morphometric study with reference to the 3D spatial context of the fibroblasts. Mammary ducts are an important structural landmarks in mammary tissue and serve as a frame of reference for these cells. More crucially, in breast cancer, tumor is initiated in the epithelial cells that surround the mammary ducts. Thus, the ducts are structurally as well as functionally salient landmarks in the mammary tissue. In our study, we use the distance of fibroblast cells from ducts to quantify the spatial characteristic of the cells. The morphological characteristics are derived from the shape of the cell nuclei as
explained in detail in Section 2.3.

Recent work in quantifying nuclear morphology in 3D has focussed on the analysis of individual nuclei [6, 7]. Rohde et. al. [7] propose the use of large deformation metric mapping for studying nuclear morphology. Gladilin et. al. [6] present a physically-based technique to map the nucleus onto a sphere in order have a common frame of reference to compare nuclear architecture. These studies were conducted on individual cell nuclei in isolation, without an emphasis on the spatial context of the cells. Our work is guided by the goal of uncovering the dynamics of the tumor microenvironment, and thus our morphometric studies have been conducted with an emphasis on the spatial context of cells. We note, however, that in [8], the organism being studied (larval L1 stage C. elegans) is several orders of magnitude smaller than mice, where it is impossible to construct such an atlas. In contrast, we have focussed on a salient structural and functional landmark - the mammary ducts - which serves as an atlas for our studies. The rest of the paper is organized as follows. Section 2 describes the methods used in our studies. Section 3 provides the results of our study followed by concluding remarks in Section 4.

2. METHODOLOGY

2.1. Sample Preparation and Imaging

The mice used in the study were from two transgenic strains. Both strains selectively expressed YFP in fibroblasts. One strain was a normal control and the other had Neu selectively expressed in epithelial cells, resulting in the development of tumor at 2 months. Both populations of mice were 6 months of age. Tissue from mouse mammary gland was sectioned into 80μm sections. The sections were stained with DRAQ5, a fluorescent marker for DNA, to identify cell nuclei. The tissue sections were imaged with an Olympus FV1000 confocal microscope with an objective of 40x/1.3NA. The images were acquired at an in-plane resolution 0.31μm and axial resolution of 0.53μm. The images have a field of view 317μm × 317μm and depth of 50μm. Samples images from the study are shown in Fig. 1.

2.2. Image Processing Pipeline

An overview of the image processing pipeline is shown in Fig. 2. The acquired image stacks have two channels - one identifying the cell nuclei (DRAQ5) of all cells in the tissue and the other identifying fibroblast cells (YFP). There are two entities in the image that we are interested in extracting - the fibroblast nuclei and the mammary ducts. First, a cell nuclei segmentation pipeline is used to obtain the nuclei in DRAQ5 image. The pipeline consists of smoothing using anisotropic diffusion, followed by a sequence of morphological image processing operations, followed by connected components to identify individual blobs. The segmentation pipeline was run over several parameter configurations and the best segmentation was selected using a validation framework proposed in [9]. Second, the foreground in the YFP channel was extracted by median filtering followed by adaptive thresholding. YFP in fibroblasts is expressed in the cytoplasm, thus surrounding the nucleus. To extract the fibroblast nuclei, an “expansion” of each segmented nucleus was computed using morphological dilation and the extent of overlap between the expanded version of the nucleus and the foreground mask of the YFP channel was used to identify the fibroblast nuclei. The segmented fibroblast nuclei were manually corrected for segmentation errors. To identify the mammary ducts, the raw DRAQ5 image volume was first volume rendered to visualize the ductal regions. Regions corresponding to mammary duct were manually identified. The segmented nuclei that fell within the identified region were extracted to give a more accurate estimate of the mammary duct. Fig. 3(b) shows an example of the extracted foreground region corresponding to the mammary duct. The distance transform of the mammary duct foreground image was computed, which was then used to compute the distance of the fibroblast nuclei from the mammary ducts.

Fig. 1. Confocal images of mammary tissue (a) 2-channel image with cell nuclei identified using Draq5 (yellow channel). Through transgenic modifications, YFP (blue channel) is selectively expressed in fibroblasts. (b) Volume rendering of nuclei channel

Fig. 2. Image Processing Pipeline
2.3. Shape Model

Statistical analysis of nuclear morphology requires the selection of an appropriate shape representation for the cell nucleus. We considered several approaches from shape analysis literature and found the point distribution model (PDM) [10] to be the most appropriate for our application. PDM’s provide a vector space which allows easy computation of statistical models. PDM’s are effectively obtained using a parameterization of the surface [11], which is the approach we have taken in our study. Brechbuhler et. al. had proposed the use of spherical harmonics to represent simply connected 3D shapes [12]. In this approach, the surface of the object is mapped on to a sphere through a transformation that minimizes area distortion. The spherical mapping results in a representation of the original surface through a (truncated) set of spherical harmonic coefficients. Each function maps a point on the sphere (given by $\phi$ and $\theta$) to the corresponding coordinate on the surface. A representation of the functions is obtained in terms of their projection onto spherical harmonic basis functions. Spherical harmonics are the equivalent of the Fourier basis for functions on a sphere ($S^2$). Each of the three functions is thus represented through a (truncated) set of spherical harmonic coefficients. For instance, $X(\phi, \theta)$ is given by

$$x(\phi, \theta) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c_{l,m} Y_l(m, \phi, \theta)$$ (1)

where $Y_l(m, \phi, \theta)$ are the spherical harmonic basis functions of degree $l$ and order $m$ [12]. This parameterization enables a dense resampling (known as reparameterization) of the original surface by sampling the spherical parameterization uniformly on the sphere. An effective of generating such a sampling is through a linear, uniform icosahedron subdivision [11] which provides a good approximation of a homogeneous sampling of the spherical parameter space. Correspondence between points across surfaces is established through a sequence of two operations. First, the surface is rotated to a canonical frame of reference using the first order ellipsoid of the surface. Second, a template image is used to further align the points using the Procrustes algorithm. The resulting set of points are the point distribution model of the original surface. The model is denoted by the vector $V = (x_1, y_1, z_1, x_2, y_2, z_2, \ldots, x_n, y_n, z_n)$

where $(x_i, y_i, z_i)$ are the coordinates of the $i^{th}$ sampled point and $n$ is the total number of points sampled. The pipeline for building this model has been implemented Styner et. al. [11]. We have used this implementation of the pipeline for our studies. In order to evaluate differences between two populations of nuclei, we use the statistical testing framework presented in [11] to measure local and global shape differences.

3. RESULTS

A set of 20 images of dimensions $317 \mu m \times 317 \mu m \times 50 \mu m$ ($1024 \times 1024 \times 96$ voxels) for each of the two transgenic mouse strains - normal control (NOR) and with the Neu oncogene (NEU) were acquired. A random subset of 3 images each were selected for analysis. The images were run through the image processing pipeline to extract fibroblast nuclei and the mammary ducts, where present. Fig. 3(a) shows a representative set of fibroblast nuclei. Fig. 3(b) shows an example of mammary ducts extracted from the image. 156 fibroblasts were obtained for NEU and 75 were obtained for NOR. The datasets were further divided on the distance of the fibroblasts to the closest point on a duct. The distribution of the distances was analyzed and two groups CLOSE and FAR were created based a distance threshold of $50 \mu m$. This resulted in four in groups in all - NEU-CLOSE (85 nuclei), NEU-FAR (71 nuclei), NOR-CLOSE (36 nuclei), NOR-FAR (39 nuclei). The following analysis was performed using these four groups.

The PDM for the fibroblasts in all the groups was obtained using by sampling 1002 points on a $5^{th}$ order spherical harmonics approximation of the surface. The principal modes of shape variation in the dataset was obtained on the pooled sample set of all the four datasets using PCA. Samples from each dataset were then projected onto the first two principal components. Fig. 4 shows the projection of the fibroblasts onto the top two principal components. The plot shows an interesting result - the morphologies of fibroblasts appear to be well separated across CLOSE and FAR. To analyze this trend, cluster analysis was performed using density-based EM clustering with 10-fold cross-validation. Clustering resulted in two clusters which were evaluated using the class information (CLOSE or FAR). 86% of the fibroblasts were correctly classified by this method, thereby suggesting that the spatial proximity of the fibroblasts to the ducts may be factor in determining the morphology of their nuclei.

The trend was further analyzed by visualizing the group differences of the local surface point distributions using the
Fig. 4. Projection of the fibroblasts onto the first two eigen-vectors

(a) (b) (c)

Fig. 5. (a) Mean shape of group CLOSE (b) Mean shape of group FAR (c) Mean difference vector field

framework in [11]. The means of each group were computed and the mean difference vectors were plotted. Fig. 5 shows the result of the mean difference tests. The vector field indicates that fibroblasts in the FAR group are larger that those in the CLOSE group in some regions of the nucleus.

4. DISCUSSION

We presented a study on the variation of nuclear morphology of fibroblasts in the spatial context defined by their tissue microenvironment. Focussing on a pertinent biological problem of analyzing the tumor microenvironment in breast cancer, we studied the shape variation of nuclei of stromal fibroblasts as a function of their distance from the mammary ducts. Fibroblasts from both normal and tumor tissues were divided into two groups based on their proximity to ducts. By building statistical shape models of the nuclei, we observed a difference in the nuclear morphologies between these two groups. Principal components analysis showed a significant separation between classes when projected onto the first two eigenvectors. The observations were verified using cluster analysis as well as visualizing the local variation of shape.

The processing pipeline for our studies required minimal manual intervention. The computations were performed in parallel on a computing cluster of 32 nodes. Thus, our framework can be used for the analysis of large amounts of image data and be used in both an exploratory and hypothesis testing framework.

5. REFERENCES