Machine Learning Analysis for Identification of Cell Shape Metrics Associated with Stem Cell Differentiation in Nanofiber Scaffolds

Desu Chen[1], Sumona Sarkar[2], Julian Candia[1], Stephen J. Florczyk[2], Subhadip Bodhak[2], Meghan K. Driscoll[1], Carl G. Simon, Jr.[2], Joy P. Dunkers[2], Wolfgang Losert[1]

[1]University of Maryland, College Park, MD; [2]Biosystems & Biomaterials Division, National Institute of Standards & Technology, Gaithersburg, MD

Statement of Purpose: Cell shape has been demonstrated to be closely related to cell function and may be an important predictor of cell fate. Many metrics are available to describe cell shape, however the relationship of these metrics to cell fate are not well understood. Specifically, nanofiber scaffold structures have been demonstrated to uniquely induce osteogenic differentiation of human bone marrow stromal cells (hBMSCs) and alter cell shape, similarly to chemically induced differentiation [1]. This phenomenon occurs after only 1 day of culture, and may allow for early prediction of stem cell fate. In this study, we aim to classify cells on nanofiber scaffolds versus flat film substrates based on their shape metrics and correlate these metrics to functional outcomes of the cells. We have developed computational tools based on Support Vector Machines (SVMs) to identify cell morphological features associated with nanofiber induced differentiation of hBMSCs. To accurately classify cells based on shape and to determine the most significant cell shape metrics we have utilized the supercell method which accounts for cell shape heterogeneity, as well as a jackknife method to test the robustness of our classifications [2].

Methods: Poly(ε-caprolactone) (PCL) films (SC) were generated with spin-coating and PCL nanofiber scaffolds (NF) were fabricated by electrospinning onto tissue culture polystyrene discs. hBMSCs were seeded and cultured on the PCL films and PCL nanofiber scaffolds for 24 hours (37° C, 5% CO_2), fixed with 3.7% formaldehyde and permeabilized with 0.1% Triton-X, then stained with Alexa Fluor 546 phalloidin (0.33μM) for actin and 4',6-diamidino-2-phenylindole (DAPI, 0.03mM) for nucleus. High-resolution 3-D z-stack images of hBMSCs were taken with a confocal microscope (Leica SP5) with 63x water immersion objective. A total of 121 hBMSCs in NF and 114 hBMSCs on SC were imaged. Max projections of the z stacks were processed with snake algorithm to define cell outlines with sub-pixel resolution in MATLAB [3] (Fig. 1 top). Shape metrics were extracted from these outlines for each hBMSC. A Support Vector Machine tool, perceptron, associated with the supercell paradigm was then implemented to classify hBMSCs in different culturing conditions [2]. Supercells were generated by averaging the shape metrics over a certain number of randomly picked original cells to reduce single cell heterogeneity in cell shape. All possible combinations of uncorrelated metrics (correlation coefficient < 0.5) were considered to identify the optimal combination of metrics and proper supercell size.

Results: 14 metrics were extracted from shape analysis as candidate shape indicators including 6 metrics showing significant difference (p-value < 0.001) in the comparison of NF versus SC. After applying perceptron analysis to data of varied sizes of supercell to all possible combinations of the uncorrelated featured metrics, we tested our results with a jackknife method. We found that with the combination of the following 3 metrics ("circularity, minor axis length and boundary tortuosity" or "solidity, minor axis length and boundary tortuosity"), we achieved the optimal classification of hBMSCs on NF and SC. Averaging over every 5 cells randomly (i.e. supercell size = 5) yields 99% correctness and 95% robustness in the jackknife test (Fig. 1 bottom). The analysis also yields the ranking of metrics by importance [2]. We found the following rankings of ability to properly classify the cells: "circularity > minor axis length > boundary tortuosity" or "minor axis length > solidity > boundary tortuosity" respectively.

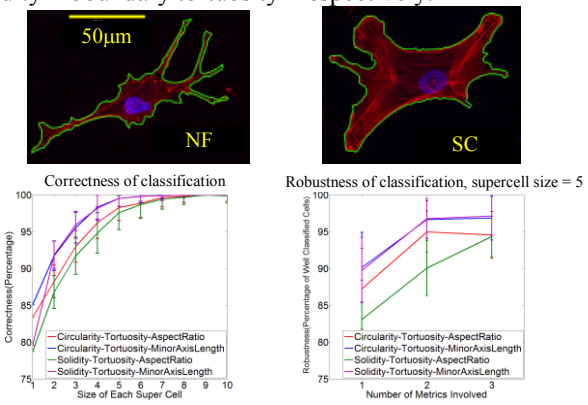

Figure. 1 (Top) Maximum intensity projections of the image z-stacks for hBMSCs on NF (left) and SC (right) with snake outlines in green (nucleus = blue; actin = red). (Bottom left) Correctness of classification with several shape metric combinations. (Bottom right) Results of jackknife test for the robustness of classification

Conclusions: The robust classification achieved with the supercell/SVM method successfully confirms and quantifies the morphological difference of hBMSCs on NF and SC on day 1 based on shape metrics. The resulting optimal linear classifiers will also enable predictions of future data. Several 2-D cellular morphological properties were included in the optimal classifier and emerged as important shape indicators for this difference. These shape indicators may be related to not only mechanical response of cells to the geometry of nanofiber environment but also cellular behavior of osteogenic differentiation at early stage. This approach is easily extendable to image stacks and 3D shape metrics.

References:

(1) Kumar G, et al. Biomaterials 32, 9188, 2011.
(2) Candia J, et al, PLOS Comp Biol 9: e1003215 , 2013
(3) Driscoll MK, et al. Aging Feb;4(2):119-32, 2012