

## Automated Cell Counting in a High Density, Polymer-Coated, Live Single Cell Sandwich Microarray

Jordan R. Yaron<sup>1,2</sup>, Jieying Pan<sup>1</sup>, Tejas Borkar<sup>3</sup>, Kristen B. Lee<sup>1</sup>, Kuo-Chen Wang<sup>1,3</sup>, Clifford L. Anderson<sup>1</sup>, Honor L. Glenn<sup>1</sup>, Deirdre R. Meldrum<sup>1</sup>

<sup>1</sup> Center for Biosignatures Discovery Automation, The Biodesign Institute, Arizona State University, Tempe, AZ, USA

<sup>2</sup> Biological Design Graduate Program, Arizona State University, Tempe, AZ, USA

<sup>3</sup> School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ, USA

Single cell analysis is essential for elucidating the contribution of rare cellular events or sub-populations in the onset and progression of disease pathology. Despite the importance of identifying cellular heterogeneity, most of the current understanding of disease is based on bulk population measurements. To address this need, recent advances in single cell technology and associated methods are gaining popularity and are resulting in the identification of critical features of disease.

We have developed a high-throughput, single cell sandwich microarray, called the Cellarium, for isolation of live single cells in hermetically sealed microchambers followed by real-time measurement of oxygen consumption rate (OCR) and extracellular acidification rate (ECAR, or pH rate of change) within each microchamber. An important step in analysing data collected on such high density live cell microarrays is the identification of wells containing only a single cell since the desired results are OCR and ECAR per cell and since multiple cell occupancy can result in cell-to-cell interaction effects [1]. High accuracy in identification of single cell occupancy is needed to achieve confidence in outlier response. This can be challenging as cells are both motile and capable of dividing within the duration of an experiment. Also, high density arrays are tedious to count manually and resultant counts are susceptible to human error and assessment bias. To facilitate well occupancy assessment at large scale, we developed an automated pipeline for cell counting based on binding and fluorescence of the cell-permeable DNA-binding dye Hoechst 33342. We verified our pipeline by comparing to a manual counting method for 3 individual data sets.

We identified CellProfiler ([www.cellprofiler.org](http://www.cellprofiler.org)) as an ideal environment to develop our pipeline because of its multi-platform compatibility, broad accessibility and existing awareness among the biological research community. CellProfiler is an open-source image analysis suite with pre-built modules for performing parameterized segmentation, measurement and modification routines on batch image data sets [2].

Special technical challenges for performing automated cell counting on the Cellarium platform relate to distinguishing each nucleus from 1) adjacent nuclei, 2) the side-wall of the microchamber and 3) background fluorescence of the polymer layer [3].

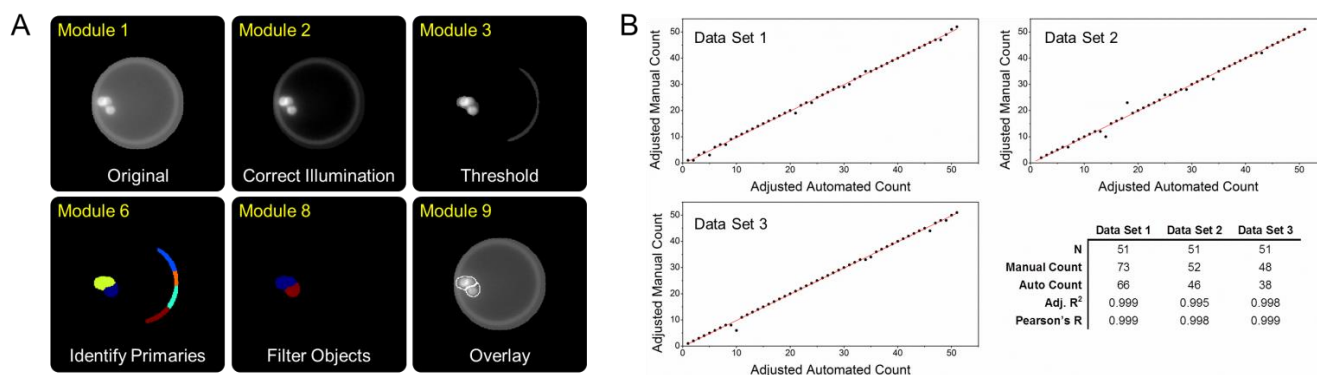
Figure 1A depicts major stages in the workflow for the presented pipeline. Briefly, data is collected on a high content imager and a software journal is run to map the x- and y-location of each well. The montaged image and a map describing well location is processed by a script that generates an individual image for each well of the array (we use MATLAB, but this code could easily be ported to other languages). Individual well images are then processed through the following CellProfiler pipeline: 1)

*LoadImages* opens all images with a given series filename format; 2) *CorrectIlluminationCalculate* removes the majority of the background associated with polymer autofluorescence; 3) *ApplyThreshold* performs a finer cut-off of polymer-associated background fluorescence while maintaining positive signal from Hoechst 33342 via an adaptive strategy background threshold; 4) *Crop* cuts off a large portion of the brightly fluorescent polymer layer on the microwell lip while preserving cells which may be adhered to the sidewall of the well; 5) *EnhanceOrSuppressFeatures* suppresses aberrant sections of the microwell lip while maintaining the positive signal from cell nuclei; 6) *IdentifyPrimaryObjects* identifies individual nuclei within a defined diameter range and distinguishes clumped neighbours via a Laplacian of Gaussian routine wherein objects are identified based on a bright center point; 7) *MeasureObjectSizeShape* defines the character of each identified nucleus; 8) *FilterObjects* filters out leftover well lip fluorescence via an eccentricity cutoff by favouring more round objects; 9) *OverlayOutlines* produces a quality control image illustrating which features in a well were identified as cell nuclei; 10) *ExportToSpreadsheet* produces a Microsoft Excel file for downstream data analysis; 11) *SaveImages* exports files depicting final cell count boundaries over the initial well view for post-processing quality control assessment. Results of three independent data sets are shown in Figure 1B. Less than 1% variance (as determined by adjusted  $R^2$  and Pearson's R) between counting methods was achieved with the presented pipeline.

The pipeline described here can be modified for a variety of input data, but is especially useful when complicated background signal (e.g. autofluorescent wells, as in the present data) shares spectral space with experimental signal. Additionally, because the pipeline is developed for use within the CellProfiler environment, additional fluorescence signal (such as from an intracellular sensor or dye) can be multiplexed with cell count, extracellular sensor data and DNA content.

#### References:

- [1] L. Kelbauskas *et al*, *Journal of Biomedical Optics* **17** (2012) p. 037008.  
 [2] L. Kamensky *et al*, *Bioinformatics* **27** (2011) p. 1179-1180.  
 [3] Y.S. Shin *et al*, *Journal of Micromechanics and Microengineering* **13** (2003) p. 768-774.  
 [4] This work is supported by the NIH Common Fund LINCS (Library of Integrated Network-Based Cellular Signatures) program grant: U01CA164250 (to D.R.M.).



**Figure 1.** (A) Summary and illustration of select, major module stages in the automated cell counting pipeline. Detailed descriptions of all module processes can be found in the text. (B) Variance between manual and automated count represented as  $N_{\text{method}} + (\text{Count}_{\text{auto}} - \text{Count}_{\text{manual}})$ . Linear regression is shown in red.