Abstract # 2209 / L118

Teachable Tool for Standardization of Human Induced Pluripotent Stem Cell Colony Selection from Live Cell Microscopy Image Sequences

Hoyin Lai[1], William T. Hendriks[2,3], Trupti V. Wadadekar[1], Samuel V. Alworth[1], Laurence Daheron[3], Chieko Nakada[4], Yasujiro Kiyota[4], Chad A. Cowan[2,3], Lee L. Rubin[2,3] and James S.J. Lee[1]

[1]*DRVision Technologies, LLC. Bellevue, WA 98008;* [2]*Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138;* [3]*Harvard Stem Cell Institute, Harvard University, Cambridge, MA 02138,* [4]*Nikon Corporation, 471 Nagaodai, Sakae, Yokohama, Kanagawa, 244-8533 Japan*

Abstract

The ability to reprogram somatic cells to an embryonic stem cell-like state using four transcription factors has had landmark impact on basic biological research, drug screening, and drug discovery. However, iPS colony selection remains technically challenging, while stemness characterization is costly and time consuming. Tools that can reliably classify iPS colonies in real time would be extremely useful to standardize selection, and drive down costs through the selection and expansion of fewer colonies. We developed a teachable image recognition tool that can be used for the detection and classification of human induced pluripotent stem cell (iPS) colonies. Whole well scans of human iPS colonies were recorded for four weeks after transduction. Colony reprogramming status was assessed based on morphology and virus silencing at the end of four weeks (Day 0) to establish ground truth. Our automated image recognition analytic achieved iPS colony classification accuracy of 99.83% (7 false positives from 4077 candidates). We found that using measurements from multiple time points improved the classification performance.

Methods

iPS culture and image acquisition

Human fibroblast cell lines are provided by the Harvard Stem Cell Institute. The cells are plated (100,000 cells into one well of a 6-well plate) and transduced for two consecutive days with four retroviruses (1ml MIG-Oct4, 1ml MIG-Klf4, 1ml MIG-Sox2 and 200ul MIG c-Myc). Five days later, the cells are trypsinized and replated on feeder (irradiated mouse embryonic fibroblasts) in one 10 cm dish. The cell culture dishes are placed in a BioStation CT(Nikon, Japan) and incubated over a period of four weeks. Whole well scans were obtained by 2x magnification in phase contrast and fluorescence (GFP) every 6 hours. At the end of Week 4, 27 iPS colonies are identified manually by morphological appearance and viral GFP-silencing to establish ground truth data for comparison.

Image Analysis

Whole well scans from Day 0, -1 week, -1.5 weeks, and -2 weeks were loaded into SVCell software (DRVision Technologies, Bellevue, WA) and stitched to create 20,000-by-20,000 pixel composite images for analysis. Automated colony detection is performed using a segmentation procedure incorporating Soft Matching™[1,2] technology, and measurements are made for each colony. A classification rule is taught using SVCell decision procedure incorporating regulation tree[3,4] technology. 179 colonies were taught (27 iPS, 152 other) using the ground truth. Terminal nodes were limited to a minimum of 10 samples to prevent over-training.

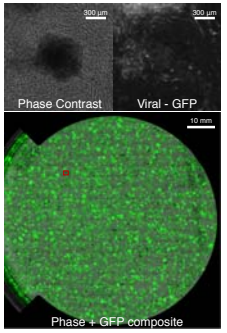

Fig. 1. *BioStation Images.* Above: Insets of phase and fluorescence images corresponding red ROI. Below: 2x composite scan of the 10cm dish

Colony Detection

An automated segmentation procedure is created using Soft Matching (Fig. 2) and applied to the phase contrast channel of the Day 0 image. Basic measurements of compactness, standard deviation of phase channel intensity, and area were calculated in SVCell for each colony and used to filter obvious non-iPS colonies. The gating rule removed 3,871 non-iPS colonies, leaving 179 challenging cases for further analysis.

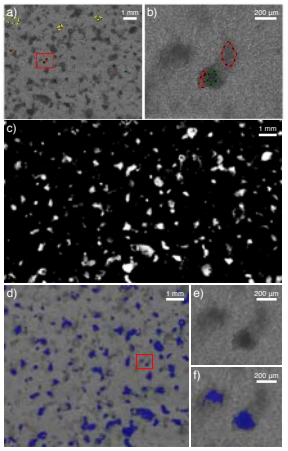

Fig. 2. *Colony detection by Soft Matching in SVCell.* Images shown on the right and in Figures 3 and 4 are downsampled by 50%. a) Regions of interests are drawn on the image and assigned labels for "Enhance" (green), "Suppress" (red), and "Background" (yellow). b) Enlarged image of the Enhance and Suppress regions. The user teaching is used to create an image transformation template that generates a c) "confidence image" that can be thresholded and manipulated using traditional image processing techniques to create the final detection mask d). Note that only the iPS colony - like portions of the colonies are detected; thus Soft Matching can discard extraneous image information and reduce the requirement for the classification step.

Colony Classification

Phase contrast channel measurements on the whole colony, its center and boundary, are calculated using the Week -1, Week -1.5 and Week -2 images. Three classification trees were created using only Week -1 data, Weeks -1 and -1.5 data, and lastly Weeks -1, -1.5, and -2 data (Fig. 3). In all three trees, the primary rule (first decision node) uses the mean intensity measured in the colony border region (Fig. 4). iPS colonies have been noted to have a more textured and lighter border. Additionally, as the Week -1 measure is made using the Day 0 mask, it could indicate faster growth for iPS colonies where borders no longer cover the colonies.

Fig. 3. *Classification Rule.* Screen shot of the regulated decision tree used for classification.

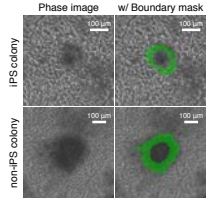

Fig. 4. *Key Measures.* Illustration of colony boundary region (Week -1 image).

Results

Table 1. Decision tree error rate.

	Missed iPS detections	False iPS detections	Total Error
Week -1 only	0/27(0%)	25/52 (48.08%)	25/179 (13.97%)
Week -1, -1.5	1/27 (3.7%)	7/34 (20.6%)	8/179 (4.47%)
Week -1, -1.5, -2	0/27 (0%)	7/34 (20.6%)	7/179 (3.91%)

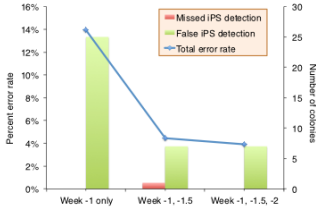

Fig. 5. *Temporal measures improve classification.* Error rates are shown for the three decision trees. Total error rate is shown as the blue line corresponding to the axis on left. Number of missed or false iPS colonies are shown as columns corresponding to the axis on right.

Of the 179 colonies analyzed, 27 were identified as true iPS colonies using the Day 0 image. Fig. 5 and Table 1 plot the classification error rates for the three decision trees. By using all three data sets for prediction, only 7 colonies out of 179 (3.91%) were misclassified, of which no iPS colonies were missed.

Introducing additional time points reduces the total error rate and the number of false iPS detections. In addition the 7 false detections were reviewed and appear reasonable.

Overall performance is excellent with just 7 false detections and no misses from 4,077 candidate colonies (0.17% error).

Conclusion

The selection of iPS colonies for expansion is a critical task that is difficult to standardize. This study validates SVCell as a useful tool for colony classification. We will continue to use the tool to assess classification of iPS colonies using many more images as test data. We will look to determine the earliest time point from which iPS colony formation can be reliably predicted. It is hoped that SVCell could enable reliable and automated selection of iPS colonies in the future.

References

1. **Alworth SV, Watanabe H, Lee JSJ** (2010). Teachable, high content analytics for live cell, phase contrast movies. J Biomol Screening.
2. **Lee JSJ, Oh S** (2007). Learnable object segmentation. US Patent # 7203360.
3. **Lee JSJ** (2006). Regulation of hierarchic decisions in intelligent systems. US Patent # 7031948.
4. **Lee JSJ** (2007). Information integration method for decision regulation in hierarchic decision systems. US Patent # 7293000.

Acknowledgement

This study is funded in part by NIH grant #R44HL106863 from the National Heart Lung and Blood Institute.