Stem Cell Research 17 (2016) 122–129

Contents lists available at ScienceDirect

Stem Cell Research

journal homepage: www.elsevier.com/locate/scr

Large-scale time-lapse microscopy of Oct4 expression in human embryonic stem cell colonies

Kiran Bhadriraju [a], Michael Halter [b,*], Julien Amelot [c], Peter Bajcsy [c], Joe Chalfoun [c], Antoine Vandecreme [c], Barbara S. Mallon [d], Kye-yoon Park [d], Subhash Sista [b], John T. Elliott [b], Anne L. Plant [b]

[a] Department of Bioengineering, University of Maryland, College Park, MD 20742, USA
[b] Biosystems and Biomaterials Division, Materials Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA
[c] Software Systems Division, Information Technology Laboratory, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA
[d] The NIH Stem Cell Unit, Division of Intramural Research, National Institute of Neurological Disorders and Stroke, NIH, U.S. Department of Health and Human Services, Bethesda, MD, USA

ARTICLE INFO

Article history:
Received 17 November 2015
Received in revised form 15 April 2016
Accepted 20 May 2016
Available online 22 May 2016

Keywords:
Fluorescence microscopy
Stem cells
Live cell imaging
Cell therapy
Pluripotency

ABSTRACT

Identification and quantification of the characteristics of stem cell preparations is critical for understanding stem cell biology and for the development and manufacturing of stem cell based therapies. We have developed image analysis and visualization software that allows effective use of time-lapse microscopy to provide spatial and dynamic information from large numbers of human embryonic stem cell colonies. To achieve statistically relevant sampling, we examined >680 colonies from 3 different preparations of cells over 5 days each, generating a total experimental dataset of 0.9 terabyte (TB). The 0.5 Giga-pixel images at each time point were represented by multi-resolution pyramids and visualized using the Deep Zoom Javascript library extended to support viewing Giga-pixel images over time and extracting data on individual colonies. We present a methodology that enables quantification of variations in nominally-identical preparations and between colonies, correlation of colony characteristics with Oct4 expression, and identification of rare events.

1. Introduction

Better understanding of stem cell biology will be aided by measurement methods that allow validation of assumptions about gene expression and morphological characteristics as criteria for assessing cell state and biological activity. Such methods will also support the development of cell therapy products, which requires methods for quantitatively assessing the quality and consistency of cells and colonies (Fink, 2009). The defining characteristics of desirable cells is often unclear (Baker, 2012), and this lack of knowledge, and the lack of robust measurement methods, complicates decision-making about starting materials, processes, and product quality. Having quantitative and relevant cell and colony characterization criteria is necessary for determining the consistency of preparations, assuring that culture processes are robust, and achieving a reliable, safe and effective product. While flow cytometry or genomics measurements provide useful data about some population characteristics at a point in time, they cannot provide spatial and dynamic information from individual cells and colonies. Tracking the relationship between cellular characteristics at a specific time and the fate of those cells in the future provides a means of determining what characteristics are meaningful for evaluating preparations and are predictive of the future response of cells (Filipczyk et al., 2015). To facilitate these goals, we have developed image analysis and visualization software that allows effective use of time-lapse microscopy to quantify spatial and dynamic differences in gene activity in a large number of human embryonic stem cell (hESC) colonies over several days under conditions designed to maintain pluripotency. We used the H9 hESC line which was modified by homologous recombination to include the gene for enhanced green fluorescent protein (EGFP) downstream of the endogenous Octamer binding transcription factor 4 (Oct4) gene according to the method of Zwaka and Thomson (2003).

We have examined Oct4 as a marker for this study because it is well known as a critical factor for maintaining pluripotency, and its expression is lost in differentiated cells (van den Berg et al., 2010; Niwa et al., 2000). The relationship between Oct4 expression and pluripotency is not a simple one and is not fully understood. While loss of pluripotency is often accompanied by Oct4 down-regulation (Pan and Thomson, 2007; Nichols et al., 1998), other factors are required (Niwa, 2007; Boyer et al., 2005).Thomson et al. (2011) demonstrated the dynamic response of Oct4 and Sox2 to differentiation factors in mouse ESCs. They showed with time-lapse imaging that Oct4 and Sox2 levels increased or decreased according to the lineage to which those cells were committing. Nanog, Sox2 and Oct4 have a complex relationship (Boyer et al., 2006; Mitsui et al., 2003) in which an assembly of Oct4

Abbreviations: hESC, human embryonic stem cell; EGFP, enhanced green fluorescent protein; FOV, field of view; TB, terabyte; GB, gigabyte.

* Corresponding author.

http://dx.doi.org/10.1016/j.scr.2016.05.012

and Sox2 affect their own promoter activities and the activity of the Nanog promoter. Together, Oct4 and Sox2 promote self-renewal of ESCs by preventing differentiation (Ambrosetti et al., 2000; Chew et al., 2005; Rodda et al., 2005) at least in part by control of Nanog. High levels of Nanog help maintain ESC self-renewal (Mitsui et al., 2003; Chambers et al., 2003). At the same time, it has been observed that overexpression of Oct4 can repress its own promoter as well as that of Nanog (Pan et al., 2006) and induce differentiation (Karwacki-Neisius et al., 2013). Engineering cells for low Oct4 production stabilizes their pluripotent state and reduces their efficiency of differentiation (Karwacki-Neisius et al., 2013; Radzisheuskaya et al., 2013). These observations point to a complex interplay of pluripotency factors, and dynamic regulation of them.

Determining how to accurately interpret the presence, absence and levels of Oct4 and other factors requires the ability to quantify their dynamic responses over long times in individual cells (Filipczyk et al., 2015; Sokolik et al., 2015; Ochiai et al., 2014). Direct observation of the history of a cell or colony by tracking forward and backward in time will further our understanding of markers of cell state and our ability to correctly predict future states (Singer et al., 2014; Bajcsy et al., 2016). For the data to be meaningful, it is necessary to sample a relevant number of colonies at an appropriate level of spatial resolution with sufficient temporal interrogation without compromising the health of the cells. Handling and analyzing the very large resulting dataset requires the development and application of advanced data manipulation and visualization methods.

Here, we present imaging and analysis methods that make it possible to quantify the dynamic and spatial behavior of a reporter of stem cell pluripotency in large numbers of hESC colonies over extended periods of time. To demonstrate these methods, we observe and quantify Oct4 expression under culture conditions that are expected to retain pluripotency. The image analysis, visualization software, and analytical pipeline developed for this study allow spatial and dynamic characterization of the growth and expression of Oct4 in a large number of hESC colonies. These methods for acquisition, image analysis, and visualization allow us to quantify differences in 3 preparations of pluripotent colonies, and to identify rare behaviors in colonies. The 3 preparations were all nominally identical in the sense that there were no controlled or systematic differences between them.

2. Results

2.1. Imaging and visualization

Colonies were imaged with a 10× objective at 45 min intervals per field of view (FOV) over an area of approximately 4 cm^2 comprising hundreds of FOVs. Three preparations were observed for approximately 120 h each, producing a total experimental data set with a storage size of 0.9 TB. Because of the size and movement of colonies, tracking and quantifying large numbers of them requires stitching together adjacent FOV into a single composite image in which colonies can span across FOV. The size of just a single composite image within these datasets (one time point ~1 gigabyte (GB) of data) precluded viewing on a typical desktop computer, in part, because of the lag time associated with unloading and loading sequential images in a time-lapse series. The data were hence visualized by a multi-resolution pyramid representation using the Deep Zoom Javascript library. Fig. 1A shows a schematic of the computational steps in the image analysis pipeline and representative images. The Deep Zoom software and data analysis pipeline provided the means of examining the entire data set through time at the level of the mosaic images down to full resolution images of each individual colony or groups of colonies in phase contrast and/or fluorescence mode over time (Fig. 1B), extracting data from the images for visualization of information, and performing quantitative evaluation (Fig. 1C).

Given the challenges associated with maintaining stem cells in a healthy state over 5 days of imaging on a microscope incubation chamber, we first examined if cells on the microscope exhibited growth characteristics that are consistent with cells maintained in a jacketed incubator. Colonies were segmented in phase contrast, and the mask used to measure the overall culture area and growth rate. Colony growth rates on the microscope stage-top incubator were similar to that of cells maintained in a water-jacketed incubator (Supplemental Fig. 4). While this exposure regime was designed to minimize damage to cells, it resulted in a relatively low signal to noise ratio in the GFP fluorescence channel.

We then performed analysis of each of 3 separate preparations of cells at the level of individual colonies, tracking colonies through time. At each time point, individual colonies were segmented, labeled with an identifying number, and evaluated for area, average GFP intensity, and other features (a total of 68 features, which are available to the user for manual browsing by clicking on the image of the colony in the Deep Zoom tool, as described in the Methods). Colonies were assigned a unique id when they merged with other colonies as they expanded. The change in area with time for hundreds of individual colonies is plotted for each of the 3 stem cell preparations (Fig. 2A–C). The integrated intensity from the GFP signal in each colony was computed at each time point and represented using an indicating color for each trajectory. During this period, colony areas increased approximately over 4 orders of magnitude, as colonies merged and the cultures became increasingly confluent (Fig. 2A–C). Further examination of panels A–C suggested that smaller, dimmer colonies that are slow growing are present in all three preparations (blue colored traces). These and additional differences and similarities were quantified through the data analysis pipeline.

Accurate analysis of colony intensities is complicated in these live cell imaging experiments because of relatively significant and uneven background fluorescence generated from the components in the media including riboflavins and residual phenol red and by uneven illumination and photobleaching. Image data were subjected to an optimized background correction procedure that involved a number of functions applied at the mosaic and sub-mosaic levels, and which was evaluated by a minimum RMS error after correction (Chalfoun et al., 2015). The background levels were not identical in the three preparations. To be sure that it was valid to quantitatively compare the 3 preparations to one another, we determined for each preparation a threshold for reliable detection of a dim fluorescent colony, and the sensitivity to detecting changes in the fluorescence intensity of a colony in the form of a signal to noise ratio (SNRcolony) (refer to Supplemental Information for a complete description). The SNRcolony and threshold for reliable detection were determined from the stitched fluorescence fields of view in the mosaic image after flatfield correction and background subtraction, thus accounting for any bias or uncertainty that may have been introduced during these processing steps. The detection threshold is the value above which GFP intensity can be reliably measured from a colony because it is unlikely to be due to random fluctuations in the background signal. The detection threshold took into account both spatial and temporal variations in background intensities and was set as 3× the standard deviation of the background intensity. The standard deviation of the background intensity was calculated from 70 × 70-pixel (1930.6 μm^2) areas of background regions positioned some distance away from a colony. This area was chosen to approximate the size of a small colony, and because averaging over a smaller number of pixels is more likely to give a larger standard deviation than averaging over a larger number of pixels, this provided a sensitive estimate of background standard deviation. For each preparation the intensities of background areas were collected over 20 time frames, for a total of 100 areas used in the calculation of the mean and standard deviation in background pixel intensities. The colony GFP fluorescence signal was computed for each preparation using the mean GFP intensity of representative colonies from that preparation that appeared to be uniformly expressing GFP. The GFP signal divided by standard deviation