Biotechnology Advances xxx (2013) xxx–xxx

Contents lists available at ScienceDirect

Biotechnology Advances

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

Deconstructing stem cell population heterogeneity: Single-cell analysis and modeling approaches

Jincheng Wu [a], Emmanuel S. Tzanakakis [a,b,c,d,*]

[a] Department of Chemical and Biological Engineering, State University of New York at Buffalo, Buffalo, NY 14260, USA
[b] Department of Biomedical Engineering, State University of New York at Buffalo, Buffalo, NY 14260, USA
[c] New York State Center of Excellence in Bioinformatics and Life Sciences, Buffalo, NY 14203, USA
[d] Western New York Stem Cell Culture and Analysis Center, State University of New York at Buffalo, Buffalo, NY 14214, USA

ARTICLE INFO

Available online xxxx

Keywords:
Human embryonic stem cells
Induced pluripotent stem cells
Heterogeneity
Single-cell analysis
Time-lapse microscopy
Flow cytometry
Multiple displacement amplification
Mass cytometry
Stochastic multiscale model
Gene expression noise

ABSTRACT

Isogenic stem cell populations display cell-to-cell variations in a multitude of attributes including gene or protein expression, epigenetic state, morphology, proliferation and proclivity for differentiation. The origins of the observed heterogeneity and its roles in the maintenance of pluripotency and the lineage specification of stem cells remain unclear. Addressing pertinent questions will require the employment of single-cell analysis methods as traditional cell biochemical and biomolecular assays yield mostly population-average data. In addition to time-lapse microscopy and flow cytometry, recent advances in single-cell genomic, transcriptomic and proteomic profiling are reviewed. The application of multiple displacement amplification, next generation sequencing, mass cytometry and spectrometry to stem cell systems is expected to provide a wealth of information affording unprecedented levels of multiparametric characterization of cell ensembles under defined conditions promoting pluripotency or commitment. Establishing connections between single-cell analysis information and the observed phenotypes will also require suitable mathematical models. Stem cell self-renewal and differentiation are orchestrated by the coordinated regulation of subcellular, intercellular and niche-wide processes spanning multiple time scales. Here, we discuss different modeling approaches and challenges arising from their application to stem cell populations. Integrating single-cell analysis with computational methods will fill gaps in our knowledge about the functions of heterogeneity in stem cell physiology. This combination will also aid the rational design of efficient differentiation and reprogramming strategies as well as bioprocesses for the production of clinically valuable stem cell derivatives.

1. Introduction

Phenotypic diversity is an intrinsic feature observed in isogenic populations of prokaryotic and eukaryotic cells. This diversity is also observed in stem cells including embryonic stem cells (ESCs), induced pluripotent stem cells (iPSCs) and mesenchymal stem cells (MSCs) (Enver et al., 2009; Graf and Stadtfeld, 2008; Narsinh et al., 2011; Phinney, 2012; Young et al., 2012). It is becoming increasingly clear that population variation contributes substantially to the variability observed in stem cell responses to their microenvironment and factors inducing their self-renewal or lineage commitment (Losick and Desplan, 2008). Within individual stem cell lines, heterogeneity can arise from several sources including stochastic fluctuations in gene regulatory networks (GRNs) (Arias and Hayward, 2006; Kaern et al., 2005), the kinetics of protein synthesis and degradation, partitioning of cellular material during division (Huh and Paulsson, 2011; Wu and Tzanakakis, 2012), asynchronous or asymmetric cell proliferation, allelic regulation of gene expression (Miyanari and Torres-Padilla, 2012), and spatial gradients of soluble cues and matrix factors in the extracellular millieu (Park et al., 2009; Parmar et al., 2007; Suslov et al., 2002).

Although heterogeneity is commonly observed in stem/progenitor cell ensembles, its sources and roles in stem cell biology and engineering have only recently attracted greater attention. Such heterogeneity can be manifested at the population level with the existence of phenotypically distinct cells (e.g. pluripotent and committed cells) and is termed 'macro-heterogeneity' (Huang, 2009). In contrast, 'micro-heterogeneity' refers to variations, for example, in gene or protein expression displayed by a particular subpopulation. Additional concepts related to the variations within stem cell populations can be found in Table 1. The discrimination of these two types of inhomogeneity is important because each has different biological implications. Macro-heterogeneity is typically easier to observe and may suggest a bistability/multistability feature of the system while micro-heterogeneity simply refers to the dispersion in the distribution of a trait (e.g. mRNA or protein). Furthermore, the mechanisms cells employ to maintain macro-heterogeneity or transition between multiple stable states remain to be elucidated.

* Corresponding author at: Department of Chemical and Biological Engineering, State University of New York at Buffalo, 907 Furnas Hall, Buffalo, NY 14260, USA. Tel.: +1 716 645 1201; fax: +1 716 645 3822.
E-mail addresses: jincheng@buffalo.edu (J. Wu), emtzan@buffalo.edu (E.S. Tzanakakis).

0734-9750/$ – see front matter © 2013 Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.biotechadv.2013.09.001

Table 1
Terms used in the analysis of stem cell population heterogeneity.

Gene expression heterogeneity	Cell-to-cell variation in the expression level of a gene or genes in an isogenic cell population
Micro-heterogeneity	The dispersion in the distribution of a trait (e.g. specific protein content) within a state (attractor)
Macro-heterogeneity	The multimodal distribution of a trait
Cell-fate heterogeneity	Varying lineage specification potential among cells exposed to the same environmental condition(s)
Gene expression noise	Fluctuations in gene expression due to stochasticity in pertinent biochemical reactions, e.g. random promoter activation, transcription bursts, and mRNA/protein degradation
Extrinsic noise	Noise contributed from sources affecting cell properties globally, e.g. asynchronous proliferation and asymmetric division

It is becoming increasingly clear that studies on the diversity of stem/progenitor cell populations will require methodologies with single-cell and even single-molecule resolutions. Routine biochemical assays such as reverse transcription-polymerase chain reaction (RT-PCR), and western blotting provide population-level information with property averages (mainly gene or protein expression levels) masking micro-/macro-heterogeneity. This makes imperative the development of novel methods for the high-throughput analysis of single cells. In this review, we discuss the adaptation of existing methods as well as emerging technologies for the genomic, transcriptomic and proteomic profiling of single stem cells within populations.

The high content and complexity of information generated by single-cell analytical methods necessitates synergistic efforts in parallel with mathematical and computational modeling. Stem cell specification is an intricate process entailing multilevel interactions among extensive GRNs, extracellular signals and intercellular cross-talk. Deeper understanding of these interactions is often confounded by the multiple sources of 'noise' present in stem cell processes. Equally important is the fact that those processes span multiple temporal and physical scales. For example, DNA transcription and translation transpires in seconds to minutes or faster whereas cell division occurs every 10–30 h. Diversity in cell populations is also affected by subcellular actions (e.g. by transcription factor networks or signal transduction), cell–cell and cell-substrate processes (e.g. paracrine effects, extracellular matrix components), and population-wide regulation (e.g. interactions among phenotypically dissimilar subpopulations). Here, we review quantitative frameworks for the analysis of stem cells. Models commonly fall within two categories: those describing (temporally evolving) traits within an individual cell (or groups of identical cells), and those simulating whole ensembles of cells (even with single cell resolution). The latter provide a vista of heterogeneity at the level of the population. Approaches based on the landscape model of cell states introduced by Waddington (1957) are also presented.

2. Experimental methods for the analysis of stem cell population heterogeneity

Traditional experimental methods such as western blotting and quantitative PCR (qPCR) commonly employed for cell analysis yield end-point, population-average information. As such their utility is limited for addressing questions about the origins and role of heterogeneity on the evolving properties and fate specification of stem/progenitor cell ensembles. Instead, real-time analysis at single-cell resolution is necessary to investigate cell-to-cell variability. Techniques routinely used for analyzing single cells include mainly flow cytometry and fluorescence-activated cell sorting (FACS) (Chang et al., 2006, 2008; Hayashi et al., 2008; Kalmar et al., 2009a), and fluorescence microscopy (Davidson et al., 2012; Kalmar et al., 2009a) including time-lapse microscopy (Eden et al., 2011; Smith et al., 2010). Yet, recent advances reviewed below have brought forward powerful new methods for high-throughput analysis of the genome, transcriptome and proteome of single cells (Fig. 1). The obtained data allow for the visualization of cell properties not just as mean values but as probability density distributions providing researchers with new insights about stem cell heterogeneity and associated mechanisms.

Various methods in this review are available for single cell analysis depending on the desired type of profiling (e.g. genomic) as shown in Fig. 1. The typical workflow entails the culture of cells in monolayers (e.g. single cells or colonies) or three-dimensional (3D) structures (e.g. aggregates), subsequent dispersion into single cells and selection followed by acquisition of pertinent data. Cell isolation is accomplished by FACS, micromanipulation or the use of microfluidic devices in addition to laser capture microdissection, which can be used for the isolation of cells from tissues as well. Some methods such as live-cell time-lapse microscopy, which provides real-time information for individual cells, require the genetic modification of cells with reporter transgenes (Giudice and Trounson, 2008) under particular promoters. Furthermore, cells are maintained in chambers or microfluidic devices with controlled conditions (e.g. humidity, gas, pH, temperature) allowing continuous microscopic observation. Besides the relatively high capital costs associated with these setups, cells can only be maintained for short periods (typically a few hours) compared to the duration of differentiation experiments (days). Computational image processing (e.g. segmentation algorithms (Chirieleison et al., 2011; Fero and Pogliano, 2010)) increases the throughput by reducing the time for analysis and eliminating user bias. However, other methods may be more straightforward technically but at the expense of providing only static rather than dynamic information. Flow cytometry and fluorescence microscopy of fixed cells are implemented to obtain 'snapshots' of populations with single-cell resolution at different time points. Flow cytometry analysis can be performed on live cells expressing a transgene or incubated with a non-cytotoxic dye (e.g. carboxylfluorescein diacetate succinimidyl ester (CFSE)), and on fixed cells after immunostaining. Live-cell flow cytometry can be combined with FACS for narrowing down to subpopulations of interest for the acquisition of data over time. Here, cell manipulations including repetitive sorting may have adverse effects on stem cell viability and differentiation propensity. Moreover, both flow cytometry and fluorescence microscopy require labeling cellular components with high specificity. Therefore, label-free methods based on differential characteristics (e.g. morphology, adhesion (Singh et al., 2013)) for the analysis of single cells are highly desirable. To that end, microfluidic platforms present an appealing technology for developing tools for isolating and analyzing single cells under defined microenvironments ('niches') (Chung et al., 2005; Zare and Kim, 2010; Zhong et al., 2008).

2.1. Single-cell analysis by time-lapse microscopy

Time-lapse microscopy has been utilized to monitor protein expression dynamics in live single cells within populations (Fig. 2). Typically, cells are transfected or incubated with a fluorescent reporter (e.g. GFP, CFSE) prior to their observation. Previously, time-resolved microscopy was applied to bacteria and yeast carrying two reporters (e.g. YFP, CFP) flanked by identical promoters to demonstrate stochasticity in gene expression (Fig. 3). Elowitz et al. (2002) utilized this dual reporter assay to discern noise (Table 1) due to the low copy number of transcripts and random promoter activation (intrinsic noise) from other sources including asynchronicity among cells (extrinsic noise) in *E. coli* bacterial populations. Detailed mathematical analysis of the intrinsic and extrinsic noise concepts can be found elsewhere (Swain et al., 2002).

The division kinetics and cell cycle distribution of rhesus monkey ESCs has also been studied by time-lapse microscopy (Fluckiger et al., 2006). Cells were transduced with a lentiviral vector carrying the gene for the enhanced GFP (eGFP) downstream of the human elongation factor 1α (EF1α) promoter. The results showed a variable cycle duration

Please cite this article as: Wu J, Tzanakakis ES, Deconstructing stem cell population heterogeneity: Single-cell analysis and modeling approaches, Biotechnol Adv (2013), http://dx.doi.org/10.1016/j.biotechadv.2013.09.001