Published in final edited form as:

Biotechnol J. 2013 April ; 8(4): 397–407. doi:10.1002/biot.201200201.

The Role of Single-Cell Analyses in Understanding Cell Lineage Commitment

Tyler M. Gibson¹ and Charles A. Gersbach^{1,2,3,*}

¹Department of Biomedical Engineering, Duke University, Durham, North Carolina, United States, 27708

²Institute for Genome Science and Policy, Duke University Medical Center, Durham, North Carolina, United States, 27708

³Department of Orthopaedic Surgery, Duke University Medical Center, Durham, North Carolina, United States, 27710

Abstract

The study of cell lineage commitment is critical to improving our understanding of tissue development and regeneration and to enhancing stem cell-based therapies and engineered tissue replacements. Recently, the discovery of an unanticipated degree of variability in fundamental biological processes, including divergent responses of genetically identical cells to various stimuli, has provided mechanistic insight into cellular decision making and the collective behavior of cell populations. Therefore the study of lineage commitment with single-cell resolution could provide greater knowledge of cellular differentiation mechanisms and the influence of noise on cell processes. This will require the adoption of new technologies for single-cell analysis, in contrast to traditional methods that typically measure average values of bulk population behavior. This review discusses the recent development of methods for analyzing the behavior of individual cells and how these approaches are leading to deeper understanding and better control of cellular decision making.

Keywords

single-cell; stochastic; single-molecule; lineage commitment; cellular heterogeneity; noise; gene expression; cell differentiation

Introduction

The goal of regenerative medicine is to develop cell-based approaches to restoring the function of diseased or damaged tissues [1]. This unique approach to repairing tissues is accompanied by a diverse set of challenges. Regenerative medicine strategies often rely on specific lineage-committed cell types that perform all the appropriate functions for target tissue regeneration. Alternatively, progenitor cells, such as stem cells, may be used such that native or external cues guide their differentiation into the appropriate cell type. For both approaches, it is essential to have a thorough understanding of cell lineage commitment and precise control over this process for these cell-based therapies to be safe and effective.

*Address for correspondence: Charles A. Gersbach, Ph.D., Department of Biomedical Engineering, Room 136 Hudson Hall, Box 90281, Duke University, Durham, NC 27708-0281, Phone: 919-613-2147, Fax: 919-668-0795, charles.gersbach@duke.edu.

The authors declare no commercial or financial conflict of interest.

A large toolkit exists for directed cell lineage commitment in regenerative applications. Soluble growth factors can be used to maintain pluripotency or direct stem cells to defined lineages by activating natural signaling pathways [2, 3]. Additionally, the composition and properties of the extracellular matrix and intercellular contact can control these pathways [4, 5]. Various biomaterials have been engineered to serve as scaffolds that guide cell fate for both *in vitro* and *in vivo* applications [6, 7]. New therapies being released to market show the promise of regenerative medicine using techniques such as these [8]. The field is being further refined by the development of gene therapies and genetic reprogramming, as discussed in more detail below. An increased understanding of cell lineage commitment has the potential to catalyze advances in all of these areas.

Long-term changes in cell behavior, including cell lineage commitment, are almost exclusively guided by changes in gene expression. Transcription factors are the main components of the cellular machinery that interact with DNA and modulate gene expression. The delivery of specific factors associated with particular cell states can reprogram the cell by activating the corresponding gene networks [9-13]. The prototypical example of transcription factor-driven differentiation in mammalian cells is the induction of myogenesis by the muscle-specific transcription factor MyoD [14, 15]. Forced expression of MyoD robustly converts various cell types to a skeletal myoblast-like phenotype [16, 17]. Master transcription factors that induce several other cell lineages have also been identified. For example, Runx2 drives osteoblast differentiation and skeletogenesis [18-22], Sox9 regulates cartilage development and chondrogenic gene expression [23-25], and Ascl1 in conjunction with other factors induces the development of a neuronal phenotype [26-30]. Furthermore, the delivery of Pdx1 transdifferentiates liver and exocrine cells into an insulin-producing phenotype similar to pancreatic beta-islet cells [31-35] and GATA4 with a cocktail of other factors can drive cells to become functionally similar to cardiomyocytes both *in vitro* [36] and *in vivo* [37, 38]. These are only a few examples of the different factors found to induce transdifferentiation. The landmark discovery that the transcription factors Oct4, Sox2, Klf4, and c-Myc can create a pluripotent state in terminally-differentiated adult cells [39-41] has created numerous possibilities for directing cells towards a desired phenotype for applications in regenerative medicine [13].

Importantly, all of these examples of transcription factor-driven genetic reprogramming are inefficient processes. Production of induced pluripotent stem cells (iPSCs) results in reprogramming frequencies that range from 0.002-2% of cells [42]. Early iterations of iPSC production methods were unable to meet some hallmarks of pluripotency such as chimera generation or germline-competency [39, 43]. These results suggested that cells can exist in a partially reprogrammed state. In this state, cells are not able to revert to their original phenotype but also are not completely reprogrammed to the intended phenotype [44]. Similarly, individual cells show variable responses to the same reprogramming stimuli, possibly because of stochastic variability in the population [45]. Furthermore, reprogrammed iPSCs that have not differentiated are capable of forming tumors after implantation, and therefore it must be ensured that all cells used therapeutically have been completely directed to a nontumorigenic phenotype. A thorough understanding of decision making at the single-cell level is necessary to address these issues. Additionally, the observation of single-cell behavior and heterogeneity within a cell population can provide deeper insight into the mechanisms of natural differentiation and lineage commitment. This review focuses on cellular heterogeneity in the context of cell differentiation and genetic reprogramming and discusses methods for analyzing single-cell behavior that can expand our understanding of cellular lineage commitment.

Origins of Heterogeneity in Cell Populations

The value of a biochemical measurement averaged across a large cell population does not necessarily describe the value for any one cell within that population (Fig. 1). This misrepresentation is exacerbated in data sets that consist of dissimilar binary states, such as distinct cell phenotypes. In these systems, the average does not accurately represent either state. Because traditional biochemical assays of cell activity, such as Western blot and RT-PCR, make bulk measurements of the aggregate cellular population, there is a clear opportunity to more accurately describe cell behavior with assays that quantitatively describe single cells.

Genetically identical cell lines can display a wide disparity of cell morphologies and behaviors, often due to random fluctuations, or noise, in gene expression [46]. Noise in cellular gene expression has been categorized into two classes. The stochasticity that results directly from the biochemical events of gene expression is the “intrinsic noise” [47]. This heterogeneity originates from the nature of intracellular chemistry. Macroscopic chemical models assume that the reactants are present in large quantities in dilute, well-mixed environments. However, the intracellular environment is nearly the opposite: the cell is crowded with macromolecules that can be locally sequestered [48]. As the number of molecules within a system decreases, the observed noise resulting from stochastic fluctuations increases. Because diploid cells contain only two copies of each autosomal gene locus, gene regulation within a cell is particularly sensitive to perturbations in the states of individual molecules, which fluctuate stochastically as a result of Brownian motion [49, 50]. In contrast, “extrinsic noise” is the variation in the concentrations of molecular components between cells that can globally affect gene expression [47]. Examples of extrinsic noise include changes in cellular microenvironments or different levels of a common upstream transcription factor [51].

The observation of large degrees of stochastic behavior in cells has challenged conventional wisdom of the cell as a precise and finely tuned machine [52]. In particular, how are biological processes such as organism development and regeneration deterministic and reproducible when the underlying chemical interactions are stochastic at the individual cell level? A variety of investigations of how cells sense and react to the intrinsically noisy intracellular environment suggest the answer can be found in the complex topology of the network motifs that govern gene regulation. For example, a series of genes that activate one another in a sequential cascade can act as a low-pass filter that only responds to a stimulus of sufficient strength and duration, and therefore is resistant to small noisy fluctuations [53]. Similarly, mutual inhibition of genes can create a bistable system where cells stably exist in one of two attractor states to which the system will return following transient noisy signals [54]. Noise can also influence cell decision making by producing a bimodal response to a signaling pathway that does not contain bistable regulatory networks [55]. In other networks, such as the stress response pathway, motifs have been selected to respond more sensitively to noisy signals, and thus increase population-wide variability. This can be an advantageous by ensuring some cells in the population survive under stress [56]. Therefore nature has adapted to the noisy microenvironment of the cell and regulatory pathways have evolved to control the sensitivity of cellular responses to noisy stimuli.

Methods for Analysis of Single Cells

Immunofluorescence staining

Heterogeneity in a cell population is most easily observed with a microscope. Immunofluorescence staining facilitates the observation of biochemical information within individual cells (Fig. 2A). However, immunofluorescence staining is a destructive assay and

therefore only discrete time points can be assessed. This method also typically does not report quantitative levels of protein expression, as appropriate quantification and validation requires an extensive set of standards and controls[57]. Several alternative methods have been developed to overcome the shortcomings of these techniques. In particular, fluorescent reporter genes in live cells enable continuous, non-destructive observation of gene expression and single-molecule microscopy provides definitive quantification of molecular copy number.

Fluorescent reporter genes in live cells

The most common method for developing a reporter of gene expression is to genetically encode fluorescent proteins under the control of specified promoter elements (Fig. 2B). Fluorescence is then used as a surrogate for expression levels of the gene of interest. Many fluorescent proteins have been modified from their original forms found in nature [58, 59] to encompass a veritable rainbow of excitation and emission wavelengths that range from the violet to the infrared spectrum [60, 61]. The wide variety of distinct spectra enables researchers to track multiple targets at once.

Regulation of the promoter sequence of a gene of interest can be used to infer individual gene expression. This sequence contains binding sites for transcription factors that recruit transcriptional machinery to control mRNA polymerization. In measurements of bulk cell populations, promoter sequences are often used to drive the expression of reporter genes such as luciferase. However, the advent of bright, robust, and genetically-encodable fluorescent proteins has enabled the detection of expression levels in individual cells [47, 62]. This allows for live-cell tracking of cell behavior.

A limitation of the use of engineered promoter-reporter constructs is their inability to reflect the full regulatory environment of the endogenous locus. Although promoter sequences contain many of the necessary transcription factor binding sites, they do not contain distal enhancers or reproduce the full chromatin structure at the locus of the gene being tracked. To more fully reflect gene regulation, “mini-loci” have been created in transgenic mice. For example, by including sequences from a 70 kb region extending through multiple DNase I hypersensitive sites on both the 5′ and 3′ of the human β -globin gene, transgenic mice expressed levels of human globin that were independent of the genomic locus of the transgene [63]. Similarly, the mapping of DNase-sensitive regions flanking the IL-4 gene enabled the construction of an IL-4 mini-locus reporter that recapitulated the natural regulation of this gene, in contrast to the isolated promoter [64]. The importance of the control regions distal to the promoter required to generate these mini-loci exemplifies the regulatory complexity of the genome, and the perpetual challenge of assuring that reporter behavior is an accurate surrogate for expression levels of the target gene. The effect of distal regulatory elements and chromatin structure on gene expression levels may not be measurable by reporters of insufficient size and complexity. Therefore it is imperative to independently verify all reporters correlate to endogenous expression levels.

Exogenous promoter-reporter constructs have been used extensively to track individual cell differentiation and reprogramming. Synthetic constructs combining binding sites for Oct4 and Sox2 driving GFP were virally delivered into both embryonic stem cells and iPS cells as a marker for pluripotency [65]. An alternative system to track reprogramming used GFP targeted to the Nanog gene in a bacterial artificial chromosome to identify reprogrammed iPS cells [41]. A two-color system was used to track different levels of osteoblast differentiation [66]. In this system, the promoter for the Osteocalcin/*Bglap1* gene and the promoter for collagen type 1a1 each drove distinct isomers of GFP in transgenic mice and indicated significant heterogeneity of expression levels among mature osteoblasts. In addition to these systems, fluorescence reporter systems have been used to monitor

chondrogenesis [67], cardiogenesis [68], and neurogenesis [69], to name just a few examples.

An alternative to delivering exogenous promoter-reporter constructs is to directly integrate a reporter into a targeted genomic locus. Homologous recombination is a method of gene targeting that is commonly used to achieve site-specific recombination of genomic DNA [70, 71]. Targeted nucleases, including zinc finger nucleases [72] and TALE nucleases [73], can be used to increase the efficiency of homologous recombination and promote targeted integration of DNA constructs [74]. The recent development of efficient methods to custom-design TALE nucleases [75-78] offers a promising tool for gene targeting with minimal toxicity [79]. Engineered nucleases can be used to precisely and reproducibly integrate reporter constructs into “safe harbor” sites in mammalian genomes that have been well-characterized [55, 80-82]. Additionally, this approach of targeted genome editing can insert a reporter gene directly into the locus of the gene of interest. For example, targeted homologous recombination was used to replace the *Oct4* gene with an eGFP reporter [83]. Because replacing an endogenous gene may not always be the desired outcome, this method has also been used to fuse the endogenous *Oct4* gene with eGFP, which can also be linked by a 2A ‘skipping’ peptide [84]. These targeted reporters have the advantage of being fully integrated into the genomic locus with fewer regulatory interactions disrupted.

To measure reporter activity at the single-cell level, either fluorescence microscopy or flow cytometry can be used (Fig. 2B). Flow cytometry confers the advantage of measuring hundreds to thousands of cells per second, rapidly gathering a comprehensive look at the population distribution. This method also enables the use of fluorescence-activated cell sorting (FACS) to sort and recover cells by fluorescence intensity for further study. Therefore FACS offers a powerful method to isolate distinct cell populations with various levels of biochemical markers.

Single-molecule microscopy

Recent advances in single-molecule detection strategies offer a quantitative alternative to the use of fluorescent reporters. The visualization of individual molecules gives definite and absolute counts of molecular copy number on a per-cell basis. Furthermore, these single-molecule technologies enable researchers to observe the spatial distribution of molecules within the cell.

The MS2-GFP system has been developed to track single molecules in living cells [85]. This system takes advantage of a strong binding affinity between the coat protein of the MS2 bacteriophage and a 19-nucleotide stem-loop RNA sequence. By genetically fusing a fluorescent protein, such as GFP, to the MS2 coat protein, the fluorescent fusion protein will specifically co-localize with target mRNAs that have been modified to contain the 19 bp sequence [86]. This method only requires delivery of the encoding DNA to the cell. Although it requires the genetically engineering the cells, in contrast to IF or FISH, this approach allows the detection of mRNA transcripts in living cells. However, a disadvantage of this method is that the MS2-GFP complexes may aggregate to one another, potentially damaging the cell and complicating the mRNA quantification [85].

Single-molecule fluorescence *in situ* hybridization (smFISH) is a method that enables the direct fluorescent labeling of mRNA molecules inside a cell (Fig. 2C) [87]. It has the advantage of not requiring any genetic modification of the target gene or other genetic engineering of the cell. Unlike conventional FISH, smFISH amplifies the fluorescence signal from each mRNA molecule by using multiple complementary probes. The probes are “tiled” along the mRNA molecule, creating a bright fluorescent spot at the location of the single mRNA molecule. The size of the spots are limited by diffraction [87]. The spots are

typically bright enough to be observed using conventional widefield fluorescence microscopy and provide quantitative information regarding the number of mRNA molecules present in a cell at a given time [88].

A limitation of smFISH is that only “snapshots” of cell behavior are acquired because cells must be fixed for analysis, similar to immunofluorescence staining. Therefore it is not possible to track the dynamic behavior of individual cells. Additionally, only a limited number of genes can be tracked simultaneously, making it difficult to perform a systems-level analysis of many interacting genes. New advances in microscopy and image analysis are beginning to enable the study of multiple genes at once. smFISH has been adapted to monitor multiple genes simultaneously by combinatorial labeling—attaching probes with different fluorophores in pre-determined sequences. Super-resolution microscopy facilitates the identification of the combinatorial label localized to each spot and the ability to dramatically increase the number of measurable genes [89]. As fluorescence microscopy and image processing technologies continue to improve, the precise number, location, and identity of macromolecules inside a cell will become easier to determine.

Single-cell gene expression analysis

While several methods have been developed to enhance flow cytometry and microscopy for single-cell and single-molecule experiments, there have also been advances in conventional biochemical assays to enable measurements in single cells. For example, quantitative reverse transcriptase PCR (qRT-PCR) has been modified to enable quantification of gene expression levels in single-cell samples (Fig. 2D) [90]. Advancements in microfluidics have further optimized this process and allowed for smaller-scale reactions. There are now commercially available microfluidic systems and chips that can run high-throughput qRT-PCR reactions. For example, the commercial microfluidic systems allows researchers to load up to 96 wells of single-cell lysate and 96 wells of the qRT-PCR reaction mix and primers targeting desired genes. The system will then analyze each cell for expression of each gene, and has been used to develop extensive transcription profiles of many cell types [91]. These assays offer the benefit of allowing researchers to measure more genes simultaneously than fluorescence-based methods.

For a more comprehensive analysis of gene expression, next-generation sequencing methods such as RNA-Seq offer a powerful tool for analyzing the entire transcriptome of cell populations in a variety of studies. Early iterations of RNA-Seq required far more source mRNA than could be obtained from a single cell. More recently, the sensitivity of RNA-Seq has been improved to the point that the transcriptome of individual cells can be analyzed (Fig. 2E) [92]. This method originally relied on PCR amplification of the starting mRNA material, and later iterations of this technology focused on improving the length of the transcript reads and decreasing the bias intrinsic to the PCR steps. Because PCR amplifies its product exponentially, less favorable amplicons can be depleted exponentially [93]. More recently, this method was adapted by adding *in vitro* transcription (IVT) steps. The incorporation of IVT in single-cell RNA-Seq, termed CEL-Seq, differs from PCR amplification of the starting material because it facilitates linear amplification of the material. This has led to an improvement in sensitivity and fidelity [94]. The application of high-throughput whole-transcriptome sequencing to single-cell studies enables the study of gene expression in a cell with unprecedented detail.

Gene Expression Noise in Cellular Fate Decisions

Gene expression noise can lead to substantial phenotypic differences in cells. Stochastic gene expression is a fundamental part of cellular decision making in biological processes across a variety of systems, ranging from bacterial stress response to stem cell lineage

commitment [95]. As described above, cell control systems have developed regulatory mechanisms to sense and modulate the population-wide response to noise [53-56]. By studying cell responses at the individual cell level, mechanistic insights can reveal potential means to control and modify cell behavior during differentiation. Several landmark examples of this approach are reviewed below.

Hematopoietic stem cell differentiation

Noisy gene expression modulates cellular lineage choices. For example, expression levels of the stem cell marker Sca-1 randomly fluctuate in hematopoietic stem cells. Cells isolated from a bulk population with transiently higher levels of Sca-1 produce myeloid cells after differentiation with a significantly greater rate than cells with lower levels of Sca-1 [96]. Conversely, cells that are low in Sca-1 produce erythroid cells at a higher rate. However, if either the Sca-1 low or high population is isolated and cultured without being induced to differentiate, they will recapitulate the original distribution of Sca-1 expression levels over time (Fig. 3A). This is evidence of random, stochastic fluctuations within a normally distributed unimodal population that may be priming cells for different states. This finding also provides a novel mechanism for guiding cell differentiation for therapeutic applications.

Adipocyte differentiation

Once a stem cell has committed to a specific lineage, there are still several steps to becoming a terminally differentiated cell. For example, the transition from proliferating adipocytes to non-dividing adipocytes is marked by a clear change in cell phenotype, but the gene network that regulates this decision is poorly understood. This process is important to maintaining healthy weight balance. To better understand the regulation of adipocyte differentiation, single cells were studied using immunofluorescence staining for two transcription factors, C/EBP α and C/EBP β , as well as the nuclear receptor PPAR γ [97]. To obtain sufficiently high cell counts to identify distinct populations, automated image analysis software was used to quantify the staining intensity levels for these proteins in each nucleus. At a bulk level, it appeared that differentiating adipocytes experienced a transient increase in C/EBP β expression, followed by upregulation of C/EBP α and PPAR γ as C/EBP β levels dropped off. However, by analyzing the behavior of the individual cells, a different picture emerged: After upregulation of C/EBP β , cells were making an all-or-nothing commitment to adipogenesis. Cells that became differentiated adipocytes maintained higher expression levels of C/EBP β while upregulating PPAR γ , but a separate population of cells emerged that did not appear to differentiate as indicated by low C/EBP β and PPAR γ expression (Fig. 3B). The authors determined that adipocyte differentiation took place via the activation of three sequential positive feedback loops. Differentiation started with positive feedback between C/EBP α and PPAR γ . A second loop between PPAR γ and C/EBP β would activate only after levels of PPAR γ reached a critical level. Finally, a late-acting positive feedback loop exists between PPAR γ and the insulin receptor that further boosts PPAR γ expression. This loop is temporally limited by the slower upregulation of the insulin receptor during differentiation. The observed delays appeared to be crucial to creating an irreversible commitment to adipogenesis. This observation was further supported by the development of a quantitative model of the serial positive feedback loops. The model, like the experimental data, showed an irreversible commitment and the emergence of a bimodal cell population[97]. Not only did this line of experiments demonstrate how single-cell measurements can show distinct trends from the bulk population, but they also led to the uncovering of a novel regulation pathway in lineage commitment.

Developmental lineage specification

The single-cell studies of hematopoietic stem cell differentiation and adipogenesis focused on lineages that were already partially characterized. Therefore the analysis focused on the

few genes known to regulate these processes. However, other studies have investigated developmental systems for which the critical regulatory genes are unknown. In these cases, the studies use high-throughput methods that enable a broader, systems biology approach to single-cell analysis.

These systems-level methods were used to study the developing zygote as it divides into a multicellular blastocyst [91]. As the monocellular zygote divides, three different cell types begin to develop: the trophectoderm (TE) appears first, followed by the epiblast (EPI) and primitive endoderm layers (PE) that emerge from the inner cell mass. To determine the genes associated with lineage commitment in the different layers, a microarray analysis was performed to measure gene expression within the whole blastocyst as well as in the isolated inner cell mass. 48 genes were then selected for single-cell analysis. A commercially available array chip and microfluidic system were used to analyze the 48 genes by quantitative RT-PCR in 48 individual cells. The gene expression profiles enabled a clear categorization of the 48 cells into the TE, PE, or EPI layers. Following these findings, the investigators identified Sox2 upregulation as one of the first identifiers separating the inner cell mass from the TE. They also identified differential regulation of the signaling molecule Fgf4 and its receptor Fgfr2 as one of the early differences between the PE and EPI layers within the inner cell mass. None of these experiments involved blastocysts of more than 64 cells, demonstrating the need for single-cell analysis to understand the transcription programs that drive early blastocyst differentiation.

Genetic reprogramming

Cellular heterogeneity and stochastic phenotype switching are not confined to natural developmental and regulatory pathways. After the factors Oct4, Sox2, Klf4, and c-myc were discovered to be capable of inducing pluripotency in somatic cells [39], possible explanations for the extremely low efficiency of reprogramming were widely postulated [98]. A series of experiments demonstrated that genetic reprogramming to a pluripotent state was stochastic at the single-cell level, and following sufficient duration of reprogramming factor expression most cells could be reprogrammed (Fig. 4A) [45]. Furthermore, increasing the cell division rate by delivering Lin28 showed that the reprogramming process could be accelerated. A separate study used live-cell tracking to identify cells that would eventually form iPS cell colonies [99]. These observations showed that an average of 3% of fibroblasts converted to iPS cells. Interestingly, these originating fibroblasts underwent an early transition to a faster-dividing and smaller-sized phenotype that did not occur in fibroblasts that did not reprogram. These studies demonstrate findings at the single-cell level that offer potential to improve cellular reprogramming.

Buganim *et al.* used a combination of single-molecule mRNA FISH and single-cell quantitative RT-PCR to analyze the conversion process from fibroblasts to iPSCs [100]. The quantitative RT-PCR results identified genes that predict whether a cell will eventually reprogram into a pluripotent state. Strikingly, these genes appeared to be more stringent predictors of successfully reprogrammed cells than the previously accepted reprogramming markers. A search for late markers of pluripotency indicated that endogenous Sox2 expression was an effective marker for successful reprogramming. This work also revealed a sequential activation of other pluripotency-related genes beginning with Sox2, suggesting a specific hierarchy of genes activated during reprogramming (Fig. 4B). Identification of this hierarchical pathway led to the substitution of the conventional “Yamanaka” cocktail of Oct4, Sox2, Klf4, and c-Myc with downstream regulatory factors. iPSCs were successfully produced by this approach, although often with lower efficiencies of reprogramming. The activation of Sox2 upstream of this well-ordered hierarchy of reprogramming events appears to occur stochastically, and therefore this new model is consistent with the previous studies indicating that reprogramming is a stochastic process [45]. The single-cell and single-

molecule experiments in this study facilitated the dissection of the significant heterogeneity during genetic reprogramming, leading to new pathways of generating iPSCs.

Conclusions and Future Directions

Although the cell is a complex entity assembled from many interacting components, science is moving towards the ultimate ability to simulate and predict a full cell's behavior *in silico*. Recently, a “first draft” simulation was developed to simulate the behavior of *Mycoplasma genitalium*, a bacterium with a genome consisting of less than six hundred thousand base pairs [101]. For comparison, the human genome is roughly three billion base pairs in size. Even though mammalian cells are orders of magnitude more complex than *Mycoplasma*, it is likely that combinations of experimental and computational studies will lead to useful models of mammalian cell lineage commitment in the near future. A thorough comprehension of cellular decision making at the single-cell level will clearly be a vital component of these efforts. As regenerative medicine, gene therapy, and tissue engineering develop into viable medical therapies, it will be increasingly important to understand the behavior of biological therapeutics and grafts—an understanding that must range from the scale of the entire tissue down to the scale of the single cell.

Acknowledgments

This work was supported by an NIH Director's New Innovator Award (DP2OD008586), NIH R03AR061042, NSF Faculty Early Career Development (CAREER) Award (CBET-1151035), a Scientist Development Grant from the American Heart Association (10SDG3060033), and the Duke Center for Systems Biology. T.M.G. was supported by an NIH Training Grant (T32GM008555).

Biography

Charles A. Gersbach is Assistant Professor in the Departments of Biomedical Engineering and Orthopaedic Surgery and the Institute for Genome Science and Policy at Duke University. He received a Bachelor of Science degree in chemical engineering and a Ph.D. in biomedical engineering from the Georgia Institute of Technology, followed by postdoctoral training in molecular and chemical biology at The Scripps Research Institute. His research is focused on cellular and molecular engineering applied to regenerative medicine, gene therapy, and synthetic biology. Dr. Gersbach is a recipient of the NIH Director's New Innovator Award, an NSF CAREER Award, and a Hartwell Foundation Individual Biomedical Research Award.

Tyler M. Gibson is pursuing a Ph.D. in Biomedical Engineering at Duke University, thanks in part to the support of the Duke Center for Biomolecular and Tissue Engineering. He holds a B.S. degree in Chemical Engineering from the University of Oklahoma, where he was awarded a Barry Goldwater Scholarship to conduct tissue engineering research. He has also earned an M.S. degree in Biomedical Engineering from Duke University. His doctoral research focuses on understanding and modeling molecular processes involved in cell fate decisions made during progenitor cell lineage commitment and reprogramming, as well as the use of synthetic biology techniques to influence cellular differentiation.

Abbreviations

DNA	deoxyribonucleic acid
iPSC or iPS cell	induced pluripotent stem cell
PCR	polymerase chain reaction

RT-PCR	reverse transcription - polymerase chain reaction
mRNA	messenger ribonucleic acid
kb	kilobase
GFP	green fluorescent protein
TALE	transcription activator-like effector
IF	immunofluorescence
FISH	fluorescence <i>in situ</i> hybridization
smFISH	single-molecule fluorescence <i>in situ</i> hybridization
qRT-PCR	quantitative reverse transcription - polymerase chain reaction
IVT	<i>in vitro</i> transcription
TE	trophectoderm
EPI	epiblast
PE	primitive endoderm

References

- [1]. Orlando G, et al. Regenerative medicine as applied to general surgery. *Ann Surg.* 2012; 255:867–80. [PubMed: 22330032]
- [2]. McDevitt TC, Palecek SP. Innovation in the culture and derivation of pluripotent human stem cells. *Curr Opin Biotechnol.* 2008; 19:527–33. [PubMed: 18760357]
- [3]. Murry CE, Keller G. Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. *Cell.* 2008; 132:661–80. [PubMed: 18295582]
- [4]. Discher DE, Mooney DJ, Zandstra PW. Growth factors, matrices, and forces combine and control stem cells. *Science.* 2009; 324:1673–7. [PubMed: 19556500]
- [5]. Lutolf MP, Hubbell JA. Synthetic biomaterials as instructive extracellular microenvironments for morphogenesis in tissue engineering. *Nat Biotechnol.* 2005; 23:47–55. [PubMed: 15637621]
- [6]. Chai C, Leong KW. Biomaterials Approach to Expand and Direct Differentiation of Stem Cells. *Mol Ther.* 2007; 15:467–480. [PubMed: 17264853]
- [7]. Lutolf MP, Gilbert PM, Blau HM. Designing materials to direct stem-cell fate. *Nature.* 2009; 462:433–41. [PubMed: 19940913]
- [8]. Nerem RM. Regenerative medicine: the emergence of an industry. *J R Soc Interface.* 2010; 7(Suppl 6):S771–5. [PubMed: 20843840]
- [9]. Gurdon JB, Melton DA. Nuclear reprogramming in cells. *Science.* 2008; 322:1811–5. [PubMed: 19095934]
- [10]. Jaenisch R, Young R. Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell.* 2008; 132:567–82. [PubMed: 18295576]
- [11]. Yamanaka S, Blau HM. Nuclear reprogramming to a pluripotent state by three approaches. *Nature.* 2010; 465:704–12. [PubMed: 20535199]
- [12]. Graf T, Enver T. Forcing cells to change lineages. *Nature.* 2009; 462:587–94. [PubMed: 19956253]
- [13]. Cherry AB, Daley GQ. Reprogramming cellular identity for regenerative medicine. *Cell.* 2012; 148:1110–22. [PubMed: 22424223]
- [14]. Lassar AB, Paterson BM, Weintraub H. Transfection of a DNA locus that mediates the conversion of 10T1/2 fibroblasts to myoblasts. *Cell.* 1986; 47:649–56. [PubMed: 2430720]
- [15]. Davis RL, Weintraub H, Lassar AB. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell.* 1987; 51:987–1000. [PubMed: 3690668]

- [16]. Weintraub H, et al. Activation of muscle-specific genes in pigment, nerve, fat, liver, and fibroblast cell lines by forced expression of MyoD. *Proc Natl Acad Sci U S A*. 1989; 86:5434–8. [PubMed: 2748593]
- [17]. Weintraub H, et al. Muscle-specific transcriptional activation by MyoD. *Genes & Development*. 1991; 5:1377–1386. [PubMed: 1651276]
- [18]. Ducy P, et al. Osf2/Cbfa1: a transcriptional activator of osteoblast differentiation. *Cell*. 1997; 89:747–54. [PubMed: 9182762]
- [19]. Ducy P, et al. A Cbfa1-dependent genetic pathway controls bone formation beyond embryonic development. *Genes & Development*. 1999; 13:1025–1036. [PubMed: 10215629]
- [20]. Byers BA, et al. Cell-type-dependent up-regulation of in vitro mineralization after overexpression of the osteoblast-specific transcription factor Runx2/Cbfa1. *J Bone Miner Res*. 2002; 17:1931–44. [PubMed: 12412799]
- [21]. Gersbach CA, et al. Runx2/Cbfa1 stimulates transdifferentiation of primary skeletal myoblasts into a mineralizing osteoblastic phenotype. *Exp Cell Res*. 2004; 300:406–17. [PubMed: 15475005]
- [22]. Gersbach CA, Guldberg RE, Garcia AJ. In vitro and in vivo osteoblastic differentiation of BMP-2- and Runx2-engineered skeletal myoblasts. *J Cell Biochem*. 2007; 100:1324–36. [PubMed: 17131362]
- [23]. Bi W, et al. Sox9 is required for cartilage formation. *Nat Genet*. 1999; 22:85–9. [PubMed: 10319868]
- [24]. Henry SP, et al. The postnatal role of Sox9 in cartilage. *J Bone Miner Res*. 2012
- [25]. Dy P, et al. Sox9 directs hypertrophic maturation and blocks osteoblast differentiation of growth plate chondrocytes. *Dev Cell*. 2012; 22:597–609. [PubMed: 22421045]
- [26]. Vierbuchen T, et al. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature*. 2010; 463:1035–41. [PubMed: 20107439]
- [27]. Caiazzo M, et al. Direct generation of functional dopaminergic neurons from mouse and human fibroblasts. *Nature*. 2011; 476:224–7. [PubMed: 21725324]
- [28]. Pang ZP, et al. Induction of human neuronal cells by defined transcription factors. *Nature*. 2011; 476:220–3. [PubMed: 21617644]
- [29]. Pfisterer U, et al. Direct conversion of human fibroblasts to dopaminergic neurons. *Proc Natl Acad Sci U S A*. 2011; 108:10343–8. [PubMed: 21646515]
- [30]. Adler AF, et al. Non-viral direct conversion of primary mouse embryonic fibroblasts to neuronal cells. *Molecular Therapy - Nucleic Acids*. 2012; 1:e32. [PubMed: 23344148]
- [31]. Serup P, et al. Induction of insulin and islet amyloid polypeptide production in pancreatic islet glucagonoma cells by insulin promoter factor 1. *Proceedings of the National Academy of Sciences*. 1996; 93:9015–9020.
- [32]. Sapir T, et al. Cell-replacement therapy for diabetes: Generating functional insulin-producing tissue from adult human liver cells. *Proc Natl Acad Sci U S A*. 2005; 102:7964–9. [PubMed: 15899968]
- [33]. Horb ME, et al. Experimental conversion of liver to pancreas. *Curr Biol*. 2003; 13:105–15. [PubMed: 12546783]
- [34]. Koya V, et al. Reversal of streptozotocin-induced diabetes in mice by cellular transduction with recombinant pancreatic transcription factor pancreatic duodenal homeobox-1: a novel protein transduction domain-based therapy. *Diabetes*. 2008; 57:757–69. [PubMed: 18086901]
- [35]. Zhou Q, et al. In vivo reprogramming of adult pancreatic exocrine cells to beta-cells. *Nature*. 2008; 455:627–32. [PubMed: 18754011]
- [36]. Ieda M, et al. Direct Reprogramming of Fibroblasts into Functional Cardiomyocytes by Defined Factors. *Cell*. 2010; 142:375–386. [PubMed: 20691899]
- [37]. Qian L, et al. In vivo reprogramming of murine cardiac fibroblasts into induced cardiomyocytes. *Nature*. 2012; 485:593–8. [PubMed: 22522929]
- [38]. Song K, et al. Heart repair by reprogramming non-myocytes with cardiac transcription factors. *Nature*. 2012; 485:599–604. [PubMed: 22660318]

- [39]. Takahashi K, Yamanaka S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*. 2006; 126:663–676. [PubMed: 16904174]
- [40]. Wernig M, et al. In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature*. 2007; 448:318–24. [PubMed: 17554336]
- [41]. Okita K, Ichisaka T, Yamanaka S. Generation of germline-competent induced pluripotent stem cells. *Nature*. 2007; 448:313–7. [PubMed: 17554338]
- [42]. Maherali N, et al. A High-Efficiency System for the Generation and Study of Human Induced Pluripotent Stem Cells. *Cell Stem Cell*. 2008; 3:340–345. [PubMed: 18786420]
- [43]. Stadtfeld M, Hochedlinger K. Induced pluripotency: history, mechanisms, and applications. *Genes & Development*. 2010; 24:2239–2263. [PubMed: 20952534]
- [44]. Nagy A, Nagy K. The mysteries of induced pluripotency: where will they lead? *Nat Meth*. 2010; 7:22–24.
- [45]. Hanna J, et al. Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature*. 2009; 462:595–601. [PubMed: 19898493]
- [46]. Spudich JL, Koshland DE. Non-genetic individuality: chance in the single cell. *Nature*. 1976; 262:467–471. [PubMed: 958399]
- [47]. Elowitz MB, et al. Stochastic gene expression in a single cell. *Science*. 2002; 297:1183–6. [PubMed: 12183631]
- [48]. McGuffee SR, Elcock AH. Diffusion, crowding & protein stability in a dynamic molecular model of the bacterial cytoplasm. *PLoS Comput Biol*. 2010; 6:e1000694. [PubMed: 20221255]
- [49]. Kaern M, et al. Stochasticity in gene expression: from theories to phenotypes. *Nat Rev Genet*. 2005; 6:451–64. [PubMed: 15883588]
- [50]. Raser JM, O'Shea EK. Noise in gene expression: origins, consequences, and control. *Science*. 2005; 309:2010–3. [PubMed: 16179466]
- [51]. Volfson D, et al. Origins of extrinsic variability in eukaryotic gene expression. *Nature*. 2006; 439:861–864. [PubMed: 16372021]
- [52]. Raj A, van Oudenaarden A. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell*. 2008; 135:216–26. [PubMed: 18957198]
- [53]. Hooshangi S, Thiberge S, Weiss R. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102:3581–3586. [PubMed: 15738412]
- [54]. Gardner TS, Cantor CR, Collins JJ. Construction of a genetic toggle switch in *Escherichia coli*. *Nature*. 2000; 403:339–42. [PubMed: 10659857]
- [55]. DeKolver RC, et al. Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome. *Genome Res*. 2010; 20:1133–42. [PubMed: 20508142]
- [56]. Kittisopikul M, Süel GM. Biological role of noise encoded in a genetic network motif. *Proceedings of the National Academy of Sciences*. 2010; 107:13300–13305.
- [57]. Bhadriraju K, et al. Quantifying myosin light chain phosphorylation in single adherent cells with automated fluorescence microscopy. *BMC Cell Biol*. 2007; 8:43. [PubMed: 17941977]
- [58]. Heim R, Tsien RY. Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr Biol*. 1996; 6:178–82. [PubMed: 8673464]
- [59]. Baird GS, Zacharias DA, Tsien RY. Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral. *Proc Natl Acad Sci U S A*. 2000; 97:11984–9. [PubMed: 11050229]
- [60]. Shu X, et al. Mammalian expression of infrared fluorescent proteins engineered from a bacterial phytochrome. *Science*. 2009; 324:804–7. [PubMed: 19423828]
- [61]. Shaner NC, Steinbach PA, Tsien RY. A guide to choosing fluorescent proteins. *Nat Methods*. 2005; 2:905–9. [PubMed: 16299475]
- [62]. Yao G, et al. A bistable Rb-E2F switch underlies the restriction point. *Nat Cell Biol*. 2008; 10:476–482. [PubMed: 18364697]

- [63]. Grosveld F, et al. Position-independent, high-level expression of the human β -globin gene in transgenic mice. *Cell*. 1987; 51:975–985. [PubMed: 3690667]
- [64]. Lee GR, Fields PE, Flavell RA. Regulation of IL-4 Gene Expression by Distal Regulatory Elements and GATA-3 at the Chromatin Level. *Immunity*. 2001; 14:447–459. [PubMed: 11336690]
- [65]. Hotta A, et al. Isolation of human iPS cells using EOS lentiviral vectors to select for pluripotency. *Nat Meth*. 2009; 6:370–376.
- [66]. Bilic-Curcic I, et al. Visualizing levels of osteoblast differentiation by a two-color promoter-GFP strategy: Type I collagen-GFPcyan and osteocalcin-GFPtpz. *genesis*. 2005; 43:87–98. [PubMed: 16149065]
- [67]. Grant TD, et al. Col2-GFP reporter marks chondrocyte lineage and chondrogenesis during mouse skeletal development. *Developmental Dynamics*. 2000; 218:394–400. [PubMed: 10842365]
- [68]. Domian IJ, et al. Generation of Functional Ventricular Heart Muscle from Mouse Ventricular Progenitor Cells. *Science*. 2009; 326:426–429. [PubMed: 19833966]
- [69]. Wernig M, et al. Tau EGFP embryonic stem cells: An efficient tool for neuronal lineage selection and transplantation. *Journal of Neuroscience Research*. 2002; 69:918–924. [PubMed: 12205684]
- [70]. Smithies O, et al. Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. *Nature*. 1985; 317:230–4. [PubMed: 2995814]
- [71]. Thomas KR, Folger KR, Capecchi MR. High frequency targeting of genes to specific sites in the mammalian genome. *Cell*. 1986; 44:419–428. [PubMed: 3002636]
- [72]. Urnov FD, et al. Genome editing with engineered zinc finger nucleases. *Nat Rev Genet*. 2010; 11:636–46. [PubMed: 20717154]
- [73]. Bogdanove AJ, Voytas DF. TAL effectors: customizable proteins for DNA targeting. *Science*. 2011; 333:1843–6. [PubMed: 21960622]
- [74]. Perez-Pinera P, Ousterout DG, Gersbach CA. Advances in targeted genome editing. *Curr Opin Chem Biol*. 2012; 16:268–77. [PubMed: 22819644]
- [75]. Miller JC, et al. A TALE nuclease architecture for efficient genome editing. *Nat Biotech*. 2011; 29:143–148.
- [76]. Cermak T, et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res*. 2011; 39:e82. [PubMed: 21493687]
- [77]. Reyon D, et al. FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotechnol*. 2012; 30:460–5. [PubMed: 22484455]
- [78]. Zhang F, et al. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol*. 2011; 29:149–53. [PubMed: 21248753]
- [79]. Mussolino C, et al. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Research*. 2011; 39:9283–9293. [PubMed: 21813459]
- [80]. Perez-Pinera P, et al. Gene targeting to the ROSA26 locus directed by engineered zinc finger nucleases. *Nucleic Acids Res*. 2012; 40:3741–52. [PubMed: 22169954]
- [81]. van Rensburg R, et al. Chromatin structure of two genomic sites for targeted transgene integration in induced pluripotent stem cells and hematopoietic stem cells. *Gene Ther*. 2012
- [82]. Lombardo A, et al. Site-specific integration and tailoring of cassette design for sustainable gene transfer. *Nat Methods*. 2011; 8:861–9. [PubMed: 21857672]
- [83]. Hockemeyer D, et al. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat Biotech*. 2009; 27:851–857.
- [84]. Hockemeyer D, et al. Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotech*. 2011; 29:731–734.
- [85]. Bertrand E, et al. Localization of ASH1 mRNA Particles in Living Yeast. *Molecular Cell*. 1998; 2:437–445. [PubMed: 9809065]
- [86]. Querido, E.; Chartrand, P. Using Fluorescent Proteins to Study mRNA Trafficking in Living Cells. In: Kevin, FS., editor. *Methods in Cell Biology*. Academic Press; 2008. p. 273-292.
- [87]. Raj A, et al. Imaging individual mRNA molecules using multiple singly labeled probes. *Nat Methods*. 2008; 5:877–9. [PubMed: 18806792]

- [88]. Raj A, Tyagi S. Detection of individual endogenous RNA transcripts in situ using multiple singly labeled probes. *Methods Enzymol.* 2010; 472:365–86. [PubMed: 20580972]
- [89]. Lubeck E, Cai L. Single-cell systems biology by super-resolution imaging and combinatorial labeling. *Nat Meth.* 2012; 9:743–748.
- [90]. Peixoto A, et al. Quantification of Multiple Gene Expression in Individual Cells. *Genome Research.* 2004; 14:1938–1947. [PubMed: 15466292]
- [91]. Guo G, et al. Resolution of Cell Fate Decisions Revealed by Single-Cell Gene Expression Analysis from Zygote to Blastocyst. *Developmental Cell.* 2010; 18:675–685. [PubMed: 20412781]
- [92]. Tang F, et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Meth.* 2009; 6:377–382.
- [93]. Hebenstreit D. Methods, Challenges and Potentials of Single Cell RNA-seq. *Biology.* 2012; 1:658–667.
- [94]. Hashimshony T, et al. CEL-Seq: Single-Cell RNA-Seq by Multiplexed Linear Amplification. *Cell Reports.* 2012; 2:666–673. [PubMed: 22939981]
- [95]. Eldar A, Elowitz MB. Functional roles for noise in genetic circuits. *Nature.* 2010; 467:167–173. [PubMed: 20829787]
- [96]. Chang HH, et al. Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature.* 2008; 453:544–7. [PubMed: 18497826]
- [97]. Park, Byung O, Ahrends R, Teruel, Mary N. Consecutive Positive Feedback Loops Create a Bistable Switch that Controls Preadipocyte-to-Adipocyte Conversion. *Cell Reports.* 2012; 2:976–990. [PubMed: 23063366]
- [98]. Yamanaka S. Elite and stochastic models for induced pluripotent stem cell generation. *Nature.* 2009; 460:49–52. [PubMed: 19571877]
- [99]. Smith ZD, et al. Dynamic single-cell imaging of direct reprogramming reveals an early specifying event. *Nat Biotech.* 2010; 28:521–526.
- [100]. Buganim Y, et al. Single-Cell Expression Analyses during Cellular Reprogramming Reveal an Early Stochastic and a Late Hierarchic Phase. *Cell.* 2012; 150:1209–1222. [PubMed: 22980981]
- [101]. Karr, Jonathan R, et al. A Whole-Cell Computational Model Predicts Phenotype from Genotype. *Cell.* 2012; 150:389–401. [PubMed: 22817898]

Figure 1.

Bulk population measurements do not always reflect the behavioral dynamics of individual cells. **(A)** An example of misleading bulk population measurements when cells exhibit two distinct states, such as a basal state and a state of highly upregulated gene expression. The average value accounts for both states and does not capture the distinct populations. **(B)** A schematic representation of a graded and a bimodal response to reprogramming induction. The black dots indicate hypothetical measurements of individual cells, while the grey line is the average that represents a bulk measurement. Individual cell behavior is different in the two cases, but the bulk average behavior is indistinguishable.

Figure 2.

Methods for single-cell analysis. **(A)** Immunofluorescence uses antibodies tagged with fluorophores to bind the target protein of interest. The proteins can be visualized under a fluorescent microscope or using a flow cytometer. Scale bar = 100 μm . **(B)** Schematic of an example of a genetic system consisting of an inducible reprogramming factor tracked by a red fluorescent protein (RFP) and a promoter-GFP construct. Fluorescence may be tracked using flow cytometry to measure fluorescence levels over a variety of induction levels, as seen in the representative result. The RFP level increases with the induction level of the reprogramming factor. Cells transition into a GFP-high state as the reprogramming increases. Black: unmodified cells; grey: reporter-only control. **(C)** Schematic of single-molecule FISH probes tiled along an mRNA molecule. A representative image of single-molecule mRNA FISH shows the locations of mRNA molecules for two different genes involved in cell differentiation. Spot-counting algorithms can quantify the number of mRNA molecules in each imaged cell. Scale bar = 10 μm . **(D)** Workflow for microfluidic-assisted, single-cell qRT-PCR, where cells are manually isolated and RNA is extracted prior to running the qRT-PCR. The results from this assay are similar to those of conventional qRT-PCR, although often with lower $C(t)$ values. **(E)** Single-molecule RNA-Seq involves isolating the target cells manually before processing them for high-throughput analysis. The output sequence fragments are computationally reconstructed to give a comprehensive list of the copy number of each member of the cellular transcriptome.

Figure 3.

Single-cell analysis of cell differentiation. **(A)** The population-wide expression of Sca-1 in hematopoietic stem cells is an approximately normal, unimodal distribution. However, when cells are sorted based on Sca-1 expression levels into low (dark grey), middle (white), or high (light grey) populations, they favor differentiation towards different downstream lineages. However, if the sorted cells are maintained in culture they reconstitute the full population, indicating that changes in Sca-1 levels are stochastic. **(B)** Single-cell analysis of adipogenesis showed that the bulk measurements of C/EBP α , C/EBP β , and PPAR γ do not completely describe adipocyte differentiation. Although measurements of the bulk population showed a moderate decrease in C/EBP β levels following differentiation, single cells actually either make an all-or-nothing commitment to terminal differentiation and maintain elevated C/EBP β expression levels, or revert to proliferating pre-adipocytes and dramatically decrease C/EBP β expression.

Figure 4.

Single-cell analysis of genetic reprogramming. **(A)** Reprogramming somatic cells to pluripotency is a stochastic process. As the time of induction increases, a larger proportion of the population express signs of pluripotency, indicating that the reprogramming process may be stochastic with variable latency time between cells. **(B)** Single-cell analysis of somatic cells induced to become iPSCs confirmed that the initial steps of reprogramming are stochastic, leading to endogenous Sox2 upregulation. However, the study also revealed that the steps following Sox2 activation follow an ordered hierarchy of gene upregulation leading to complete reprogramming.