High-throughput approaches for the analysis of extrinsic regulators of stem cell fate
Adrian Ranga and Matthias P Lutolf

The complexity of stem cell niches poses a tremendous challenge to understanding mechanisms of extrinsic regulation of stem cell fate. In order to better understand niche signaling and its effect on stem cell fate choices, in vitro systems are being engineered which recapitulate, in a simplistic but increasingly sophisticated manner, native stem cell niches. New technologies or new combinations of existing technologies allow more systematic ways to probe niche signaling in high-throughput. Systems biology approaches in experimental design, data acquisition and analysis will be necessary to tackle the challenges that lie ahead.

Address
School of Life Sciences (SV), Institute of Bioengineering (IBI) and Laboratory of Stem Cell Bioengineering (LSCB), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

Corresponding author: Lutolf, Matthias P (matthias.lutolf@epfl.ch)

Introduction
One of the central questions in stem cell biology focuses on understanding the role of extrinsic cues in directing stem cell fate. Indeed, cell fate choices are made in response to a complex symphony of regulatory signals, and are uniquely regulated in microenvironments termed stem cell niches [1]. These signals include local cell–cell interactions with neighboring cells as well as soluble factors and physiochemical modulation from the extracellular matrix (ECM) (Figure 1a). Together, these signals determine stem cell fate to ensure developmental robustness and, under normal circumstances, lifelong maintenance of function. However, when these signaling processes and fate choices become imbalanced, tissue homeostasis is lost, leading to diseases such as cancer.

The remarkable combinatorial complexity of stem cell niches and the context-dependent cellular responses that arise from these microenvironments pose a formidable challenge to understanding the underlying mechanisms of stem cell regulation. In order to deconstruct the role of each of these signals as well as their cross-talk there is a need to recreate stem cell niches in vitro [2–4] (Figure 1b). Though simplified, such niche models can recapitulate the essential molecular interactions between the cell and its microenvironment.

While many in vitro experiments in stem cell research are currently performed in low-throughput and often in a trial-and-error fashion, a consensus is emerging that more quantitative, systematic and integrative studies must be performed to capture a more global picture of extrinsic control of stem cell fate. In the last decade, various technologies have emerged which have allowed us to interrogate stem cell fate decisions in more systematic ways (Figure 1c). In this review we attempt to give a general overview of the state-of-the-art of high-throughput approaches for the analysis of extrinsic regulators of stem cell fate, focusing specifically on ways in which new technologies allowing a systematic deconstruction and reconstruction of niche signaling have been deployed to interrogate stem cell fate decisions. We then proceed to survey practical and theoretical approaches which have been somewhat underexplored in the context of stem cell research but which could significantly enhance current methodologies, providing a roadmap for future work.

Cellular microarrays and microfluidics in 2D
In the last decade, the physical miniaturization of experiments, from the scale of a Petri dish down to the single cell, has permitted the concurrent parallelization of experimental conditions. Technological advances have permitted thousands of experimental conditions to be performed on a single microscope slide, an experimental format that can be termed ‘cellular microarrays’ [5]. To begin to dissect and interrogate the function of various proteins involved in stem cell niche signaling, ECM microarrays have been created using robotic spotting techniques and equipment borrowed from DNA microarray fabrication field. In such experiments, purified or recombinant proteins found in the ECM, involved in cell–cell communication or present in the extracellular milieu as soluble factors have been immobilized in defined spots, allowing cells to selectively attach and respond to individual or combinatorial cues. Various stem cell systems have been explored using such systems in combination with end-point immunocytochemistry, and in-depth analysis has been able to reconstruct some of the
relationships between specific microenvironments and cell fate (e.g. [6–8]).

Despite the significant insights they have provided, ECM microarrays, due to their reliance on hard non-physiological glass or plastic supports, have not been able to capture the essential physical characteristics of a stem cell niche. As such, hydrogel-based arrays have offered an alternate solution: by patterning microwells into a soft, physiologically relevant hydrogel layer, followed by cell seeding, these platforms have allowed the capture of single cells into well-defined environments of controllable stiffness (e.g. [9,10*]). However, microwell arrays have not allowed for the combinatorial complexity of patterned signals afforded by ECM microarrays.
A recent advance by our group has combined elements of these two approaches [11**] (Figure 2): instead of spotting a combinatorial array of proteins onto a glass slide, proteins were spotted onto the microfabricated pillars of a silicon wafer, which was then pressed against a hydrogel layer of tunable stiffness (Figure 2a). Thus, with this platform, it was possible to benefit both from the modularity of the protein-spotting approach (single or double protein combination, adjustable concentration), as well as the topography and stiffness modulation of the microwell system (Figure 2b). To demonstrate the possibilities of this platform, novel modulators of mesenchymal and neural stem cell fate were uncovered, showing in particular a synergistic interplay between biochemical and biophysical cues.

At the same time as research into protein modulators of stem cell fate, the development of purely synthetic substrates for guiding stem cell fate have represented an important goal in the field. The vast combinatorial possibilities available from chemical synthesis have given hope that screening approaches could yield fully synthetic instructive cues to direct stem cell function, a therapeutically highly relevant approach. Polymer microarrays have notably been carried out using contact spotting and inkjet approaches (e.g. [12*,13,14]), and have begun to relate how stiffness or other physical properties such as surface roughness and wettability could be modulated to understand stem cell response to biophysical cues.

Despite the flexibility afforded by such cellular microarrays, once cells are immobilized on a spot or in a microwell it is difficult to move and manipulate them. Moreover, the described approaches are all static, limiting the possibility to manipulate cells dynamically, for example, to expose them to certain chemical stimuli.

---

**Figure 2**

Example of a state-of-the-art combinatorial biomaterial microarray to probe biochemical and biophysical niche effectors. (a) A DNA spotter with solid pins can be used to spot different protein solutions on micropillars of a microfabricated silicon stamp. The printed stamps can be pressed against a thin (partially crosslinked) hydrogel layer. Finally, the stamp can be demolded to obtain an artificial niche microarray suitable for both adherent or nonadherent stem cells. (b) Representative example of two full arrays (combined in mosaic fashion from individual images) spotted with two fluorescently labeled model proteins. Six concentrations of FITC-BSA (green) and six concentrations of rhodamine-BSA (magenta) were printed either in 12 × 12 random motifs (left) or as overlapping gradients (right), all in the context of a topographically patterned gel substrate. Scale bars, 2 mm. A three-dimensional reconstruction of confocal stacks showing microwells with a diameter of 450 μm (bottom right). Scale bar, 500 μm. Reprinted with permission from [11**].
Example of a state-of-the-art microfluidic stem cell culture array to dynamically manipulate and probe stem cell fate. (a) Schematic of the microfluidic chip with micrographs as insets. The cell culture layer contains 1600 chambers (pink) connected by flow channels (gray). Hydration lines are located on each side of the array to minimize edge effects. The control lines (blue) consist of an isolation valve and a peristaltic pump to control cell loading and perfusion rates. Arrows point at single cells. Scale bars, 1 mm (left) and 100 μm (right). (b) Culture of primary hematopoietic stem cells (HSCs) under dynamic conditions. HSCs were exposed to the cytokine interleukin-11 (IL-11) plus the indicated amounts of stem cell factor (SF) for the durations shown. (c) Differences in cell survival during microfluidic culture in the indicated conditions compared to the high [SF] condition. Cells were imaged every 12 min. (d) Cumulative division kinetics of primary HSCs that are cycling in the indicated in vitro conditions for the first and second divisions. Reprinted with permission from [17].
during an experiment. As such, microfluidic approaches have proven to be particularly appropriate for programming complex flow regimes [15], as well as in manipulating cells towards downstream evaluations such as on-chip flow cytometry [16]. As well, microfluidic approaches have been of particular value in situations where fluid flow is an important physiological consideration, and have proven to be extremely versatile, for example, in the context of reconstructing 'lineage trees' for single-cell fate analysis of hematopoietic stem cells [17] (Figure 3), as well as in designing spatial configurations of complex organ-like geometries, for example, in creating lung-like microenvironments [18].

Adding the third dimension in high-throughput: towards 3D cellular microarrays
While much can be learned from cellular microarrays and chip-based microfluidic assays, these two-dimensional (2D) cell culture systems do not capture the three-dimensionality of the in vivo niche. It has been shown that this essential component of the microenvironment allows for gene expression patterns and cellular phenotypes that more closely resemble those found in vivo [19,20]. A cornerstone in moving towards the third dimension is a 3D scaffold, preferably based on a soft hydrogel network structure, functioning as an analog of the instructive natural microenvironment. By breaking down the associated physiological complexity into a smaller number of distinct interactions, such ECM analogs are ultimately expected to bridge the gap between traditional 2D cell culture methods (i.e. plastic culture dishes) and labor-intensive animal models. The application of such models has been reviewed in the context of human pluripotent stem cells [21].

Most array-based work to date with 3D hydrogels has been carried out in bulk gels using hand-made fabrication techniques, including a first study looking at combinatorial modulation of ECM components for human ES cell differentiation [22]. A few groups have recently attempted to recreate biomimetic niches in high-throughput using 3D hydrogels. These techniques have borrowed both microarray spotting techniques to deposit drops of cell-containing alginate solutions [23] as well as inkjet printing approaches to deposit material in the nanoliter range [24]. Although both approaches have claimed to operate in the miniaturized high-throughput space, they have in fact mainly achieved the patterning of multiple, repeated samples in an array, without combinatorial complexity.

Others have tackled the challenge of generating high-throughput 3D constructs via microfluidic approaches. Indeed, using elegant microfabrication techniques, it has been possible to create cell-encapsulating beads of varying stiffness by droplet microfluidics [25] as well as to encapsulate co-cultures [26], and to perform FACS-like analysis on these constructs [27]. However, as with spotted or inkjet-patterned approaches, while the number of 3D samples can be tremendous, these technology platforms do not yet address the need for true combinatorial modulation of the extracellular microenvironment in three dimensions.

Indeed, the creation of a combinatorial cellular microarray in 3D remains a major challenge in the field. It requires as a starting point an artificial ECM with sufficient flexibility to independently modulate biophysical properties such as stiffness and degradability as well as allowing the possibility to combine both tethered and soluble biochemical signaling cues (Figure 1a). One of the key challenges in engineering functional artificial ECMS has been the design of molecular strategies to render matrices permissive for cell migration, cell proliferation, and ultimately morphogenetic multicellular processes. Our group has developed one family of synthetic hydrogel networks, whose backbone is biologically inert poly(ethylene glycol) (PEG), that contain some of the essential biochemical signals to mimic the biological character of natural ECMS (e.g. [28]). The susceptibility to proteolytic degradation of these synthetic hydrogels enables the establishment of a bi-directional cell-matrix crosstalk and thus permits cell-autonomous processes to occur within the synthetic matrix. While these PEG-based artificial extracellular matrices allow these properties to be tuned to a large extent independently of each other, it remains a significant technical challenge to array such hydrogels in a spatially addressable format, and to image cells in such three-dimensional constructs in a systematic and high-throughput manner [29]. One way to implement such an approach would be through a combinatorial mixing of liquid hydrogel precursors solutions with a protein library using automated liquid handling robotics, followed by dispensing (either on a slide or in a microwell plate) and in situ gelation.

Challenges and future directions: understanding the dynamics of stem cell fate
Dynamic materials: changing niche signaling on demand
The relationship between stem cells and their microenvironment has an important temporal component, which is, to some extent, reflected in current multi-step in vitro differentiation protocols. To recreate the changing stem cell microenvironment in a 3D biomimetic context, new strategies in material design now permit the selective modulation of hydrogel biophysical properties in space and time: in addition to network degradation in response to cell-demanded protease secretion [28], hydrogels can be softened [30] or stiffened [31] by the incorporation of selectively photo-inducible or photo-degradable cross-linkers into the network. Other strategies for dynamic tuning of biomaterials have focused on drug [32] or light-inducible [33] release or patterning of bioactive proteins;
of course, other payloads such as small molecules can be considered using analogous strategies.

**Reading stem cell fate in real time using reporter libraries and novel imaging techniques**

As a direct complement to dynamic temporal modulation of experimental conditions, there is a need to understand how cells behave not only at specified end-points or even at discrete time points, but rather in real or near-real time. Furthermore, it would be valuable to monitor multiple gene activities within the same experiment. Gene reporter libraries associated with imaging could be a key enabling technology to achieve this. Indeed, while the notion of using large small molecule or siRNA libraries to modulate input conditions in a given cellular model system is now commonplace, the outputs from such screens have often been of far reduced dimensionality, primarily making use of end-point immunocytochemistry or single-gene reporters. To obtain time-course results on multiple stem cell fate decisions, a large number of gene promoters (or a restricted set of interest) could each be fused to a fluorescent reporter protein and inserted into an ES cell line using lentiviral vectors [34]. Thus, sectors of a cellular microarray could have repeated instances of a microenvironment, and the cells in each spot could each express a different reporter construct. Other variants on such approaches include the generation of cell lines engineered to provide multicolor reporting [35], as well as inducible overexpression and knockdown, possible even in certain primary stem cell cultures [36].

Advances in imaging techniques could complement the temporal modulation of experimental condition by analyzing these cellular processes in real or near-real time. Indeed new lens-free microscopy techniques such as the ePetri dish [37**], could allow just such monitoring. By incorporating relatively inexpensive commercial CMOS sensors to image large areas at sufficient resolution, this approach could be parallelized to observe large areas, directly in the incubator without disturbing samples. A further extension to this approach could be to modulate materials and signaling cues in direct feedback response to dynamically observed phenotypes. As an example, a complex multi-lineage differentiation protocol could be optimized based on the temporal appearance of genes, monitored by reporter constructs.

**Dealing with data complexity: smarter experimental designs and more powerful statistical approaches**

By deploying high-throughput techniques to array dynamically cell-responsive and tunable materials, it is possible to imagine an almost complete spatiotemporal control over the instructive *in vitro* environment required for stem cell fate specification. Combining these inputs with equally flexible read-outs (outputs), notably using reporter libraries and novel imaging techniques possibly coupled to feedback control, the range of possible experiments becomes nearly infinite. A fundamental question is how this unprecedented and multifactorial control can be harnessed to address biologically relevant questions. More specifically, it becomes important to address ways in which experiments can be set up in a more efficient way and how statistical methods can best be deployed to extract the most relevant information from the mass of collected data.

One way to address the need to establish more systematic approaches to experimental setup is to adopt elements of Design of Experiments (DOE) methodologies. In multifactorial studies, DOE methods provide a rationale and statistical framework for significantly reducing the number of experiments required to arrive at meaningful results: instead of conducting experiments in a traditional way where factors are varied one by one and independently, multiple factors are changed at once in a controlled manner. Such methodologies have been widely adopted in industries such as bioprocess optimization as well as, to some extent, in pharmaceutical drug screening [38], but far less frequently in academic stem cell research — a notable recent example has used DOE approaches to quantitatively define and optimize *in vitro* erythropoiesis from human cord blood cells to enucleation, where DOE approaches were used to define a cocktail of cytokines which led to greatly enhanced expansion and accelerated erythroid maturation [39]. An alternative to carrying out high-throughput experiments in a more systematic way has been based on Feedback System Control (FSC). This approach, used successfully to find a cocktail of minimal compounds enabling hESC self-renewal in feeder-free conditions [40*](Figure 4), involves performing sequential experiments evaluated by a mathematical learning algorithm, which iteratively suggests modifications to the input conditions to arrive towards an optimal desired biological response [41].

Another key issue for high-throughput stem cell biology experiments has been the biological interpretation of the relationships between inputs and outputs. One of the key advantages of multifactorial high-throughput experiments is that they can reveal interactions between factors, which can sometimes be subtle or difficult to disentangle from main factor effects and experimental noise. Mathematical models can help in understanding the role each of these factors plays in contributing to the observed cell fate. Multivariate statistics, including generalized linear models (useful for mixed continuous and categorical factors) or higher-order modeling strategies [42] can provide a look not only at the contribution of single factors but also on combinations of factors. Thus, in some cases a combination of two or more factors present together can have an effect which could not be predicted by looking at the contribution of each factor individually: in biological terms these are synergistic or antagonistic effects [43]. High-throughput
methods also provide ways to simultaneously quantify multiple measures of cell fate, not only using such measures as cell or colony counts and size, but also in the case of image-based analyses on a variety of more esoteric image-based metrics. Principal component analysis (PCA) is a method which can be used to reduce the dimensionality of the outputs (readouts); thus, if an image analysis software outputs tens or hundreds of measures, PCA could potentially find which combination of readouts is best explained by the inputs. Other mathematical techniques which can also be used to better understand the generated data include hidden Markov models, neural networks, singular value decomposition and Fisher discriminant analysis [44].

Conclusion

Technological developments in carrying out and analyzing combinatorial experiments are beginning to allow for a systematic understanding of stem cell biology. By using canonical model systems such as embryonic stem cells for many of such studies, the proliferation of new techniques can be benchmarked and compared. New advances in materials design allow us to reproduce increasingly biomimetic microenvironments, notably in 3D. Thus far, even the most sophisticated analyses have mostly been confined to descriptions relating extrinsic cues and their phenotypic manifestation. Underlying these cause and effect relationships lie intracellular signaling pathways, which have begun to be linked to the extracellular signaling inputs [45–47,48**]. In silico models linking these various datasets [49*,50,51] will, ultimately, allow us to come closer to a completely deterministic view of biological systems [52,53]. Translating such fundamental understanding of the regulatory machinery will allow us to begin to link single stem cell fate decisions to morphogenetic processes involved in arranging cells into organized tissue structures.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest


Using a hydrogel microwell platform, substrate elasticity was shown to be an important determinant of muscle stem cell fate: by matching hydrogel stiffness to that of the in vivo muscle tissue, it was possible to achieve previously impossible stem cell self-renewal in vitro, proven by subsequent potency in an in vitro transplantation assay.


In this study the advantages of microcontact printing and protein spotting have been combined, thereby enabling a truly multifactorial analysis of stem cell fate. This platform was versatile enough to be used both with non-adherent (neural) and adherent (mesenchymal) stem cells, and deployed sophisticated statistical methods to decouple the role of various microenvironmental cues.


While a number of papers have shown the possibility of on-side combinatorial material synthesis, this paper demonstrated that the properties of the synthesized materials can be assessed in high-throughput and can be directly related to human embryonic stem cell fate. As such, washability, surface topography, surface chemistry and indentation elastic modulus of the substrates were measured, and suggested that biophysical cues mediated via a purely synthetic substrate could be sufficient to maintain a particular stem cell fate.


Three-dimensional constructs are created using microfluidics, and a system of fluorescent cell reporters is used for multiplexed analysis and sorting. This platform is shown to be useful in discerning gene interactions on cells in 3D, and is demonstrated to be amenable to in vivo transplantation.


This study presents a technique to dynamically modify gel properties using light, in mild conditions which do not affect cell survival. Such manipulation allows the arbitrary creation of three-dimensional channels or changes in the biochemical composition of the gel which is shown to direct chondrogenic differentiation of mesenchymal stem cells.


The ePetri dish is an example of a technological advance that has the potential to transform high-throughput imaging-based studies. By repurposing a relatively inexpensive existing technology (high performance image sensors), it is possible to imagine that this type of approach could be parallelized and widely adopted as a potential replacement to high-end scanning microscopy.
40. Tsutsui H, Valamehr B, Hindoyan A, Qiao R, Ding X, Guo S,
   A methodology based on a feedback control scheme is introduced to rapidly and efficiently perform experimental iterations leading to the identification of an optimal set of small molecule inhibitor types and concentrations required to maintain undifferentiated hESCs.
48. Lu R, Markowitz F, Unwin RD, Leek JT, Airoldi EM, MacArthur BD,
   In this study, multiple layers of the regulatory machinery of ES cells were probed to understand their co-ordinated response to downregulation of Nanog, a gene known to be important for self-renewal. Here, epigenetic, transcriptional and translational mechanisms were simultaneously assessed over 5 days of differentiation, showing for the first time a true systems-level analysis of ES cell fate and how various components of gene expression are regulated in time.
   Computational approaches such as the one presented here can serve as a valuable addition to existing tools used to predict stem cell fate. In this case, by identifying key image-based features of retinal progenitor cells, it was possible to determine with high accuracy key phenotypical features leading to specific self-renewal or differentiation fates.