High-throughput Analysis and Control of the Pluripotent Stem Cell Microenvironment to Interrogate and Direct Cell Fate

by

Emanuel Nazareth

A thesis submitted in conformity with the requirements for the degree of Doctorate of Philosophy
Institute of Biomaterials & Biomedical Engineering
University of Toronto

© Copyright by Emanuel Nazareth 2015
Abstract

Human pluripotent stem cells (hPSCs) are heralded to transform regenerative medicine and drug discovery and development, as they can potentially supply a scalable source of clinically relevant cell populations. Numerous exogenous and endogenous factors interplay to control stem cell fate; however, these are often unrecognized and unreported, undermining reproducibility and precluding the systematic analysis of hPSC response to molecular and chemical perturbations. We have developed a high-throughput (HTP) platform to screen hPSCs in configurable micro-environments, in which we optimized parameters including colony size and cell density to achieve rapid and robust cell fate responses to cues. We used this platform to perform single-cell protein expression profiling, revealing that Oct4 and Sox2 co-staining discriminate pluripotent, neuroectoderm, primitive streak, and extraembryonic cell fates. We applied this code to analyze dose responses of 27 developmental factors to delineate lineage-specific concentration optima and to quantify cell-line endogenous signaling pathway activation and differentiation bias. We additionally demonstrate that response profiles can classify cell lines by similarity, are stable through passages, predict definitive endoderm induction efficiency and can be used to rescue differentiation of cell-lines reticent to cardiac
induction. We next applied this platform to screen 400 small molecule kinase inhibitors, measuring resultant yield and purity of multiple lineages on a per colony basis. Enrichment analysis revealed mTOR inhibitors as strong inducers of mesendoderm cells. This mesendoderm enhancing effect of mTOR inhibitors is blocked by TGFβ or BMP receptor inhibition, and knockdown of Raptor, a mTORC1 component, via siRNA phenocopied the effects of small molecule mTOR inhibitors such as rapamycin. When rapamycin was added to a blood differentiation protocol, CD34+VECAD+ hemogenic endothelium purity and yield were increased 3 fold, with a concomitant enhancement of blood colony forming cells (CFCs). Lastly, to aid translation of these approaches, we validate a protocol of our 96-well hPSC assay in multiple international laboratories via the International Stem Cell Initiative (ISCI) consortium and provide open-source GUI-based software for analysis of spatial context in HTP studies. Ultimately, these technologies and applications enable a more comprehensive elucidation of the regulatory networks linking cell-extrinsic factors, signal transduction, transcription factor regulation, and cell fate decisions.
Acknowledgments

It is a pleasure to express my gratitude to all those who supported my throughout the course of this Ph.D. thesis. This work would not have been possible without the efforts of numerous colleagues, friends, and family.

First, I would like to thank my supervisor Dr. Peter Zandstra for his support and guidance. Throughout my thesis, he has provided encouragement, sound advice, good teaching, and has been an incredible mentor. I also give many thanks to my supervisory committee, Dr. Derek van der Kooy, Dr. Gary Bader, and Dr. Jeff Wrana, who generously gave their time to offer me valuable comments toward improving my work.

I would also like to acknowledge my numerous co-authors including Joel Ostblom, Mukul Tewary, Nafees Rahman, Petra Lücker, Shreya Shukla, Manuel Alvarez, Ting Yin, Michael Prakesh, Rima Al-Awar, Orla O'Shea, Shalinee Khadun, Lyn Healy, Glyn Stacey, Dominika Dziedzicka, Mieke Geens for their contributions to my project.

I would like to thank all members of the Zandstra lab for providing a stimulating and fun environment in which to learn and grow. A special thanks to Ting Yin who, in addition to cells, has tirelessly provided an endless supply of support and positivity. The friendship and camaraderie of many lab mates and floor mates has made coming into the lab, even on weekends, always a pleasure. I consider myself blessed for having the opportunity to work with such an exceptional group of people.

I would like to express my heart-felt gratitude to my family, to whom I dedicate this thesis. I have been lucky to grow up with the incredible support of my brother Savio, aunts, uncles, cousins, and recently an incredible group of in-laws who have all been a constant source of love, concern, support and strength. I would also like to thank my wife Laura for her understanding and love, and for making each and every moment brighter. Lastly, I wish to thank my and parents, Juliet and Paul, who receive my deepest gratitude for their dedication and decades of unconditional love and support. Thank you.
Table of Contents

ACKNOWLEDGMENTS ........................................................................................................................................ IV

TABLE OF CONTENTS ......................................................................................................................................... V

LIST OF TABLES ............................................................................................................................................... X

LIST OF FIGURES ............................................................................................................................................. XI

LIST OF ABBREVIATIONS ................................................................................................................................. XIII

1 INTRODUCTION ........................................................................................................................................... 2

1.1 Controlling Niche Factors for Robust Biochemical Assays ........................................................................... 2
  1.1.1 Abstract .................................................................................................................................................... 2
  1.1.2 Introduction ........................................................................................................................................... 2
  1.1.3 HPSC Cell Fate Regulation .................................................................................................................... 4
    1.1.3.1 TGFβ family ................................................................................................................................... 7
    1.1.3.2 FGF ................................................................................................................................................ 8
    1.1.3.3 FGF inhibition of BMP signaling ..................................................................................................... 10
    1.1.3.4 Insulin ............................................................................................................................................ 11
    1.1.3.5 Wnt ............................................................................................................................................... 11
    1.1.3.6 Extra Cellular Matrix Proteins ....................................................................................................... 12
  1.1.4 Sources of Culture Heterogeneity .......................................................................................................... 14
    1.1.4.1 Endogenous Cytokines .................................................................................................................... 15
    1.1.4.2 Mechanical forces and cell-cell contact ............................................................................................ 15
    1.1.4.3 Heterogeneous Subpopulations ...................................................................................................... 16
  1.1.5 Localized cell density – Master Regulator of Culture Heterogeneity ....................................................... 17
  1.1.6 Discussion ............................................................................................................................................. 19

1.2 Human pluripotent stem cells & drug screening ......................................................................................... 21
  1.2.1 Abstract ................................................................................................................................................ 21
  1.2.2 Introduction ......................................................................................................................................... 22
  1.2.3 Pluripotent stem cells ............................................................................................................................ 24
  1.2.4 Stem cells in lead discovery and disease modeling ............................................................................... 26
  1.2.5 Stem cells in safety pharmacology ...................................................................................................... 28
1.3 Conclusions and future perspectives ................................................................. 43

2 HIGH-THROUGHPUT FINGERPRINTING OF HUMAN PLURIPOTENT STEM
CELL CUE-FATE RESPONSES AND LINEAGE BIAS .......................................................... 46

2.1 Abstract ........................................................................................................... 46

2.2 Introduction ..................................................................................................... 46

2.3 Results ............................................................................................................. 47
  2.3.1 Microenvironmental control for hPSC screening ........................................ 47
  2.3.2 Oct4 and Sox2 expression discriminate early cell fates ............................... 52
  2.3.3 Analysis of early hPSC cell fate responses .................................................. 56
  2.3.4 Quantification of cell-line differentiation propensities ............................... 61
  2.3.5 Quantification of cell line endogenous signaling ....................................... 66

2.4 Discussion ....................................................................................................... 67

2.5 Methods .......................................................................................................... 68
  2.5.1 Cell culture .................................................................................................. 68
  2.5.2 Microcontact printing of substrate into 96-well plates ................................ 69
  2.5.3 Seeding hiPSCs onto patterned substrates .................................................. 69
  2.5.4 Immunocytochemistry and image analysis .................................................. 70
  2.5.5 hiPSC derivation .......................................................................................... 71
  2.5.6 hiPSC cardiomyogenic and endoderm induction ....................................... 71
  2.5.7 Flow cytometry ........................................................................................... 72
  2.5.8 Estimation of endogenous pathway activation .......................................... 72
  2.5.9 Statistical analysis ....................................................................................... 72

2.6 Acknowledgements ......................................................................................... 72

2.7 Author Contributions ....................................................................................... 73
3 A MULTI-LINEAGE SCREEN REVEALS MTORC1 INHIBITION ENHANCES HUMAN PLURIPOTENT STEM CELL MESENDOERM AND BLOOD PROGENITOR INDUCTION ................................................................. 75

3.1 Abstract ................................................................................................................................. 75

3.2 Introduction ............................................................................................................................ 75

3.3 Results ................................................................................................................................. 77
    3.3.1 A kinase inhibitor screen of hPSCs reveals lineage-specific regulators ......................... 77
    3.3.2 Target enrichment analysis identified kinase inhibitor targets for NE and PS responders..... 81
    3.3.3 ERK/MEK inhibition induces hPSC differentiation towards NE in a BMP-sensitive manner .... 86
    3.3.4 mTOR inhibitors induce mesendoderm in hPSCs ............................................................. 87
    3.3.5 mTOR inhibition induction of mesendoderm is BMP receptor and TGFβ receptor dependent ... 91
    3.3.6 Knockdown of Raptor phenocopies the mesendoderm inducing effect of rapamycin .......... 95
    3.3.7 mTOR inhibition enhances hPSC differentiation towards mesoderm and not endoderm derived tissues 96

3.4 Discussion ............................................................................................................................. 99

3.5 Methods ............................................................................................................................... 101
    3.5.1 Cell culture .................................................................................................................... 101
    3.5.2 Microcontact printing of substrate into 96-well plates ....................................................... 102
    3.5.3 Seeding hPSCs onto 96-well plates and application of small molecules or controls .......... 102
    3.5.4 Immunocytochemistry and high-content image analysis ...................................................... 103
    3.5.5 96-well siRNA transfection ............................................................................................... 103
    3.5.6 Viability analysis .............................................................................................................. 104
    3.5.7 Differentiation of hPSCs towards blood and definitive endoderm ........................................ 104
    3.5.8 Colony forming cell (CFC) assay ....................................................................................... 104
    3.5.9 Statistical analysis .......................................................................................................... 105

3.6 Acknowledgements ............................................................................................................... 105

3.7 Author Contributions .......................................................................................................... 105

4 TOOLS FOR HIGH-THROUGHPUT SCREENING HUMAN PLURIPOTENT STEM CELLS IN CONFIGURABLE MICROENVIRONMENTS ................................................................. 107
4.1 A multi-lab validated protocol for simultaneous and reproducible quantification of human pluripotent stem cell pluripotent, neuroectoderm, mesendoderm, and extra-embryonic cell fate responses to cues 107
4.1.1 Abstract .................................................................................................................. 107
4.1.2 Introduction .......................................................................................................... 107
4.1.3 Materials .............................................................................................................. 115
   4.1.3.1 Reagents .......................................................................................................... 115
   4.1.3.2 Equipment ....................................................................................................... 117
   4.1.3.3 Reagent setup .................................................................................................. 118
4.1.4 Procedure ............................................................................................................ 119
4.1.5 Author Contribution ............................................................................................ 128
4.1.6 Acknowledgements .............................................................................................. 128

4.2 ContextExplorer: Open-source software for spatial analysis in high-throughput screens ........ 129
4.2.1 Abstract .............................................................................................................. 129
4.2.2 Introduction .......................................................................................................... 130
4.2.3 Methods ............................................................................................................... 132
   4.2.3.1 Cell culture ....................................................................................................... 132
   4.2.3.2 Microcontact printing of substrate into 96-well plates ....................................... 132
   4.2.3.3 Seeding hPSCs onto patterned substrates ......................................................... 132
   4.2.3.4 Immunocytochemistry and high-content image analysis .................................. 133
   4.2.3.5 Statistical analysis ........................................................................................ 133
   4.2.3.6 Radial analysis .............................................................................................. 133
4.2.4 Results .................................................................................................................. 134
   4.2.4.1 ContextExplorer enables automated colony Identification ............................... 134
   4.2.4.2 Comparison of cells within colonies and not in colonies ................................. 136
   4.2.4.3 Analysis of transcription factor intra-colony variation .................................... 139
   4.2.4.4 Quality control of patterns ............................................................................. 142
4.2.5 Conclusions .......................................................................................................... 144
4.2.6 Author contributions ............................................................................................ 145
4.2.7 Acknowledgements .............................................................................................. 145

5 DISCUSSION AND FUTURE WORK .............................................................................. 147

5.1 Contributions .......................................................................................................... 147
   5.1.1 Understanding and controlling cell-to-cell variability in a HTP platform ........................ 147
   5.1.2 A novel method for patterning proteins into standard 96-well plates ........................ 148
   5.1.3 Understanding and controlling cell-to-cell variability at the bench scale .................... 149
   5.1.4 Oct4 and Sox2 based discrimination of early development subpopulations .................... 150
List of Tables

Chapter 1

Table 1-1 Human Pluripotent Stem Cell Fate Regulation – Conflicting Observations in the Literature.................................................................5
Table 1-2 Sources of Culture Heterogeneity........................................13

Chapter 2

Table 2-1 96µCP assay configuration.............................................52
Table 2-2 Factor Characterization..................................................60

Chapter 4

Table 4-1 Deep-well plate layout.....................................................124
Table 4-2 Layout of control media applied to patterned hPSCs........125
List of Figures

Chapter 1

Figure 1-1 The Stem Cell Microenvironment .................................................................13

Chapter 2

Figure 2-1 Generation of PDMS stamps for 96-well plates ...........................................49
Figure 2-2 Development of defined and fully disclosed media and substrate .................50
Figure 2-3 The 96µCP platform .......................................................................................51
Figure 2-4 Single cell protein profiling reveals Oct4 and Sox2 mark early cell fates ........55
Figure 2-5 Characterization of factors modulating cell fate choices .........................59
Figure 2-6 Dose curves of non-responsive factors ........................................................59
Figure 2-7 Quantitative assessment of cell line-specific endogenous signaling and
differentiation ..............................................................................................................63
Figure 2-8 Differentiation of H7, H9, HES2, H7*, ZAN11, and ZAN3 cell lines ..........64
Figure 2-9 Definitive endoderm induction across a panel of cell lines .........................65
Figure 2-10 Quantitative comparison of pathway activation across hiPSC lines ............67

Chapter 3

Figure 3-1 High-throughput screening of small molecule regulators of hPSC pluripotency,
primitive streak, neuroectoderm, and extraembryonic/other cell fate decisions ..........79
Figure 3-2 Response of control conditions, distributed across 10 96-well plates ............79
Figure 3-3 Responding compounds ..............................................................................81
Figure 3-4 Target enrichment analysis ........................................................................83
Figure 3-5 Target enrichment in clusters ......................................................................85
Figure 3-6 MEK inhibitor PD0325901 induced neuroectoderm is sensitive to BMP
inhibition but not Activin A inhibition ..........................................................................87
Figure 3-7 mTOR inhibitors in defined media upregulate mesendoderm via TGFβ and
BMP pathways ...........................................................................................................89
Figure 3- 8 SB-431542 and LDN-193189 antagonize the mesendoderm enhancing effect of rapamycin .................................................................91

Figure 3- 9 mTOR inhibition synergizes with BMP4 and activin in inducing Brachyury+ mesendoderm ........................................................................................................94

Figure 3- 10 mTOR inhibitors enhance Brachyury+ and Oct4+Sox2- induction in a dose dependent manner in non-patterned hPSCs ..............................................................................................94

Figure 3- 11 Rapamycin enhances blood progenitor induction ....................................97

Figure 3- 12 Differentiation of hPSC cell lines towards hemogenic endothelium is enhanced by rapamycin ........................................................................................................97

Figure 3- 13 Differentiation of HES2 cell line towards hemogenic endothelium is enhanced by rapamycin ........................................................................................................98

**Chapter 4**

Figure 4- 1 Protocol overview ...............................................................................................109

Figure 4- 2 Sample x-y- scatters of all cells in 96-well plates ..................................110

Figure 4- 3 Sample images of colonies patterned in 96-well plates, showing differential expression of Oct4 and Sox2 in response to control conditions ......................................................112

Figure 4- 4 Subpopulation frequency in response to control conditions .........................113

Figure 4- 5 Mesendoderm predictions for VUB01, VUB02, VUB07, and VUB14 ..........114

Figure 4- 6 Experimental and analytical workflow and colony identification ...............136

Figure 4- 7 Differences in fluorescent protein intensities between cells inside and outside colonies ..................................................................................................................................138

Figure 4- 8 Superimposition of colonies enables statistical analysis of radial trends .......141

Figure 4- 9 Automatic colony detection facilitates patterning fidelity analyses ..........143
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>96µCP</td>
<td>96-well micro-contact printing assay</td>
</tr>
<tr>
<td>ADRs</td>
<td>adverse drug reactions</td>
</tr>
<tr>
<td>ANOVA</td>
<td>analysis of variance</td>
</tr>
<tr>
<td>B</td>
<td>BMP4</td>
</tr>
<tr>
<td>BA</td>
<td>BMP4A with Activin A</td>
</tr>
<tr>
<td>CCLE</td>
<td>Cancer Cell line Encyclopedia</td>
</tr>
<tr>
<td>CDM</td>
<td>chemically defined media</td>
</tr>
<tr>
<td>CNO</td>
<td>CellNetOptimizer</td>
</tr>
<tr>
<td>CFCs</td>
<td>colony forming cells</td>
</tr>
<tr>
<td>CGP</td>
<td>Cancer Genome Project</td>
</tr>
<tr>
<td>CIHR</td>
<td>Canadian Institutes of Health Research</td>
</tr>
<tr>
<td>CM</td>
<td>MEF conditioned medium</td>
</tr>
<tr>
<td>CNC</td>
<td>computer numerical control</td>
</tr>
<tr>
<td>CSCs</td>
<td>cancer stem cells</td>
</tr>
<tr>
<td>CSV</td>
<td>comma-separated values</td>
</tr>
<tr>
<td>DCM</td>
<td>dilated cardiomyopathy</td>
</tr>
<tr>
<td>ddH2O</td>
<td>double distilled water</td>
</tr>
<tr>
<td>DE</td>
<td>definitive endoderm</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco’s modified Eagle’s medium</td>
</tr>
<tr>
<td>ECM</td>
<td>extra-cellular matrix</td>
</tr>
<tr>
<td>ECMPs</td>
<td>extra-cellular matrix proteins</td>
</tr>
<tr>
<td>EGFRs</td>
<td>epidermal growth factor receptors</td>
</tr>
<tr>
<td>ESCs</td>
<td>embryonic stem cells</td>
</tr>
<tr>
<td>FGFR</td>
<td>fibroblast growth factor receptor tyrosine kinase</td>
</tr>
<tr>
<td>GSK3β</td>
<td>glycogen synthase kinase 3 beta</td>
</tr>
<tr>
<td>GUI</td>
<td>graphical user interface</td>
</tr>
<tr>
<td>HAF</td>
<td>heregulin-1β with activin A and FGF2</td>
</tr>
<tr>
<td>HCA</td>
<td>high content analysis</td>
</tr>
<tr>
<td>hdF</td>
<td>autologously derived fibroblast-like cells</td>
</tr>
<tr>
<td>HE</td>
<td>hemogenic endothelium</td>
</tr>
<tr>
<td>hESCs</td>
<td>human Embryonic stem cells</td>
</tr>
<tr>
<td>hiPSC</td>
<td>human induced pluripotent stem cell</td>
</tr>
<tr>
<td>HLCs</td>
<td>hepatocyte-like cells</td>
</tr>
<tr>
<td>hPSCs</td>
<td>Human pluripotent stem cells</td>
</tr>
<tr>
<td>HTP</td>
<td>high-throughput</td>
</tr>
<tr>
<td>HTS</td>
<td>high-throughput screening</td>
</tr>
<tr>
<td>ICM</td>
<td>inner cell mass</td>
</tr>
<tr>
<td>iPSCs</td>
<td>induced pluripotent stem cells</td>
</tr>
<tr>
<td>ISCI</td>
<td>the International Stem Cell Initiative</td>
</tr>
<tr>
<td>KO-DMEM</td>
<td>DMEM with 20% KO-serum replacement</td>
</tr>
<tr>
<td>KOSR</td>
<td>Knockout Serum Replacement</td>
</tr>
<tr>
<td>LIF</td>
<td>leukemia inhibitory factor</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>MEFs</td>
<td>mouse embryonic fibroblasts</td>
</tr>
<tr>
<td>MEK</td>
<td>mitogen-activated protein kinase</td>
</tr>
<tr>
<td>MERFISH</td>
<td>multiplexed error robust FISH</td>
</tr>
<tr>
<td>mEpiSCs</td>
<td>Murine epiblast stem cells</td>
</tr>
<tr>
<td>mESCs</td>
<td>mouse ESCs</td>
</tr>
<tr>
<td>MG</td>
<td>Matrigel</td>
</tr>
<tr>
<td>MxIF</td>
<td>multiplexed fluorescence microscopy method</td>
</tr>
<tr>
<td>mPSCs</td>
<td>mouse pluripotent stem cells</td>
</tr>
<tr>
<td>MPSS</td>
<td>massively parallel signature sequencing</td>
</tr>
<tr>
<td>NEAA</td>
<td>MEM minimum non-essential amino acids solution</td>
</tr>
<tr>
<td>OICR</td>
<td>Ontario Institute of Cancer Research</td>
</tr>
<tr>
<td>PCB</td>
<td>printed circuit board</td>
</tr>
<tr>
<td>PDGFRs</td>
<td>platelet-derived growth factors receptors</td>
</tr>
<tr>
<td>PDMS</td>
<td>polydimethylsiloxane</td>
</tr>
<tr>
<td>PEDF</td>
<td>pigment epithelium-derived factor</td>
</tr>
<tr>
<td>PSCs</td>
<td>pluripotent stem cells</td>
</tr>
<tr>
<td>rsPSCs</td>
<td>region-selective pluripotent stem cells</td>
</tr>
<tr>
<td>RTKs</td>
<td>receptor tyrosine kinase receptors</td>
</tr>
<tr>
<td>siRNA</td>
<td>siRNA Small-interfering ribonucleic acid</td>
</tr>
<tr>
<td>TGFβ1</td>
<td>transforming growth factor β1</td>
</tr>
<tr>
<td>TRKRs</td>
<td>tropomyosin-related kinase receptor</td>
</tr>
</tbody>
</table>
Chapter 1
Introduction

Portions of this chapter have been prepared for submission for publication. Co-authors for this chapter include Emanuel J. P. Nazareth and Peter W. Zandstra. Authorization to reproduce this work has been obtained from all co-authors.
1 Introduction

1.1 Controlling Niche Factors for Robust Biochemical Assays

1.1.1 Abstract

Human pluripotent stem cells (hPSCs) offer a multitude of opportunities for both clinical cell-based therapies and for understanding cell fate decisions and lineage commitment. In particular, after the advent of human induced pluripotent stem cell (hiPSC) technology there has been a tremendous burst of interest in applying hPSCs and derivatives towards understanding development, generating relevant differentiated populations, disease modeling and drug screening, and even using hPSCs as a model of cancer stem cells. However, multi-lab comparisons of hPSC work show a concerning discordance between the reproducibility of hPSC based assays. Numerous factors, both exogenous and endogenous, interplay to control stem cell fate, and these factors are all too often unrecognized and not reported, ultimately undermining reproducibility and relevancy of assays. In this perspective, using hPSCs as a model screening system we first summarize studies of the major signaling pathways implicated in hPSC fate, highlighting open questions and contradictory reports in the literature. We then review factors that frequently confound assays, such as spatial heterogeneity, endogenous factors secreted by cells, mechanical forces, cell-cell contact, and the induction of subpopulations of cells. We conclude by suggesting several paths towards increasing drug discovery efficacy, including engineering the microenvironment to mitigate assay variance, reporting and varying cell density, and utilizing multivariate single cell assays to track heterogeneous populations.

1.1.2 Introduction

Human pluripotent stem cells (hPSCs) have opened up many avenues in drug discovery, as witnessed by the proliferation of hPSCs and induced pluripotent stem cells (iPSCs) based applications in disease modeling and drug discovery, studies to more efficiently generate target differentiated cells, and even as a model of cancer stem cells (CSCs) (Sachlos, Risueno et al. 2012). Although great progress has been made in understanding the factors that guide stem cells towards clinically relevant cell types it is becoming increasingly clear that numerous factors, both exogenous and endogenous, interplay to control stem cell fate, and these factors
are all too often unrecognized and not reported, ultimately undermining reproducibility and relevancy of drug discovery and development campaigns. The surprising discordance between laboratories attempting to reproduce the same results is most clearly illustrated in a study by the International Stem Cell Initiative (ISCI), which concluded that out of seven published serum-free chemically defined media tested, only the two commercial media were successful in long-term culture of hPSCs in a multi-laboratory comparative study (Akopian, Andrews et al. 2010). These findings mirror the result of recent comparisons of two large-scale pharmacogenomic studies on immortalized cancer lines which found that although gene expression is well correlated between the studies, the drug response data is highly discordant, critically limiting the application of either of these data sets to assess gene-drug associations, determine mechanism-of-action, and select potential cancer drugs (Haibe-Kains, El-Hachem et al. 2013). As increased efforts are poured into stem cell based drug discovery programs, these studies underscore the urgent need to understand and address the sources of these assay response differences.

In this perspective, using hPSCs as a model drug discovery system, we first summarize investigations of the major signaling pathways implicated in human pluripotent stem cell fate, including fibroblast growth factor (FGF), Activin/Nodal, insulin, and Wnt signaling, highlighting open questions and contradictory reports in the literature. We then review factors that introduce spatial heterogeneity and likely run-to-run and lab-to-lab variance into assays, such as endogenous factors secreted by cells, mechanical forces, cell-cell contact, and the generation of often poorly characterized subpopulations. We propose that local cell density (for a single cell heuristically defined as the number of cells within a 200 µm radius (Peerani, Rao et al. 2007)) is often a ‘master regulator’ of these factors, and hence is an excellent parameter to measure and/or control in order to dramatically increase assay reproducibility. By striving to account for these obfuscating factors, for example by engineering the microenvironment (e.g. controlling niche factors such as colony size to minimize endogenous factors) and employing methods to assay multiple phenotypes at the single cell level to probe multiple population subtypes while maintaining spatial information (e.g. ‘high content analysis’, HCA), stem cell drug discovery campaigns will be much more effective in producing a mechanistic understanding of cell fate decisions and ultimately generating clinically relevant therapeutics.
1.1.3 HPSC Cell Fate Regulation

Human embryonic stem cells (hPSCs) were first isolated from the inner cell mass of blastocyst-stage embryos in 1998 (Thomson, Itskovitz-Eldor et al. 1998). These pluripotent cells can be differentiated into all somatic cell types, and thus in theory can be differentiated into all tissues and organs of the human body. It has been demonstrated that somatic cells can be reprogrammed into ‘induced pluripotent stem cells’ (iPSCs) by ectopic induction of four or fewer transcription factors (Takahashi and Yamanaka 2006, Yu, Vodyanik et al. 2007), thus offering an additional source of pluripotent stem cells (PSCs). HPSCs were initially cultured on feeder layers of mouse embryonic fibroblasts (MEFs), and later on cultured in feeder-free conditions using MEF conditioned medium (CM) and Matrigel™, fibronectin, and additional substrates (Xu, Inokuma et al. 2001, Amit, Shariki et al. 2004, Amit and Itskovitz-Eldor 2006). These are still widely used today, and both conditions employ serum, which exhibits significant batch-to-batch variation, and is undefined. In addition, Matrigel™ is mouse sarcoma-derived matrix and is also undefined. In order to diminish variability and increase hPSC culture robustness, several groups have established feeder-free and defined culture conditions (Beattie, Lopez et al. 2005, Li, Powell et al. 2005, Pebay, Wong et al. 2005, Vallier, Alexander et al. 2005, Wang, Zhang et al. 2005, Xu, Rosler et al. 2005, Levenstein, Ludwig et al. 2006, Liu, Song et al. 2006, Lu, Hou et al. 2006, Xiao, Yuan et al. 2006, Yao, Chen et al. 2006) including commercial formulations such as mTeSR™1 (Ludwig, Bergendahl et al. 2006, Ludwig, Levenstein et al. 2006), StemPro® (Wang, Schulz et al. 2007), and TeSR™-E8™ (Chen, Gulbranson et al. 2011).

The defined media developed for the extended culture of hPSCs typically consist of DMEM/F12 based media supplemented with albumin, transferrin, vitamins, antioxidants, trace minerals, specific lipids and select growth factors. The growth factor combinations employed to date in such chemically defined media include high FGF2 (Li, Powell et al. 2005, Levenstein, Ludwig et al. 2006, Liu, Song et al. 2006, Yao, Chen et al. 2006), FGF2/transforming growth factor β1 (TGFβ1) (Chen, Gulbranson et al. 2011), FGF2/Insulin/Wnt3A/April or Baff (Lu, Hou et al. 2006), FGF2/Activin A/insulin (Vallier, Alexander et al. 2005), ActivinA/Keratinocyte Growth Factor (FGF7)/nicotinamide (Beattie, Lopez et al. 2005), FGF2/ TGFβ1/leukemia inhibitory factor (LIF) (Amit, Shariki et al. 2004), FGF2/Noggin (Xu, Peck et al. 2005), and FGF2/ActivinA/insulin-like growth factor 1
(IGF1)/heregulin-1β (HRG1β) (Wang, Schulz et al. 2007). From these studies it has emerged that FGFs, TGFβ/Activin, insulin, and extra-cellular matrix proteins (ECMPs) play an important role in hPSC propagation.

Although several such chemically defined media (CDM) for the long-term maintenance of hPSCs have been published, the majority of these media have failed in attempts to be reproduced in independent laboratories. In order to address this, ISCI recently completed an in-depth hPSC CDM comparative analysis where eight published CDM formulations were tested across a panel of ten hPSC lines in five independent laboratories (Akopian, Andrews et al. 2010). Surprisingly, out of the eight published formulations tested only the two commercial formulations, mTeSR®1 and STEMPRO, supported long-term maintenance of most of the hPSC lines tested. The reasons for the poor reproducibility of these hPSC CDM conditions are unknown, however the study indicated that the complexity of media preparations, non-optimal formulations, lack of rigorous quality control in media preparations, and fundamental differences between cell lines across different laboratories play an important role. It is clear that independent of the provided media and substrate, there are unappreciated factors that vary greatly between laboratories that exert a strong influence on stem cells. Elucidating these factors will aid in the refining of defined culture conditions and will also help reconcile seemingly contradictory findings in the hPSC literature (see Table 1-1). More importantly, reproducible assays are a prerequisite for truly elucidating the underlying regulatory networks that link cell-extrinsic factors, signal transduction, transcription factor regulation, and cell fate decisions.

**Table 1-1 Human Pluripotent Stem Cell Fate Regulation – Conflicting Observations in the Literature**

<table>
<thead>
<tr>
<th>Regulatory Molecule</th>
<th>Published results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Activin/TGFβ not required if 100 ng/ml FGF2 present (Xu, Peck et al. 2005) Activin/TGFβ only required to suppress BMP signaling (Avery, Zafarana et al. 2010) TGFβ1 does not delay differentiation of hPSCs in CDM (Vallier, Alexander et al. 2010)</td>
</tr>
<tr>
<td></td>
<td>FGFR Expression in hPSCs</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-----------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>Expressed in both undifferentiated and differentiated cells (Rosler, Fisk et al. 2004)</td>
</tr>
<tr>
<td></td>
<td>(Eiselleova, Matulka et al. 2009)</td>
</tr>
<tr>
<td></td>
<td>Not expressed in hPSCs, only expressed in differentiated cells (Bendall, Stewart et al. 2007)</td>
</tr>
<tr>
<td>Wnt/ β-catenin signaling</td>
<td>Results in differentiation (Sumi, Tsuneyoshi et al. 2008)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Heregulin-β1</td>
<td>Needed in CDM with FGF, Activin/TGF β, and insulin/IGF to support pluripotency (Wang, Schulz et al. 2007)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Spontaneously Differentiated Subpopulation</td>
<td>Support hPSCs, secrete IGFs and TGFβ (Bendall, Stewart et al. 2007)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>ROCK inhibitor</td>
<td>Decreases cell-cell contact, results in the disassociation of colonies (Harb, Archer et al. 2008)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>ROCK inhibitor, poly-D-lysine,</td>
<td>Can maintain hPSCs long-</td>
</tr>
</tbody>
</table>
and defined media | long-term (Mahlstedt, Anderson et al. 2010)
---|---
Vitronectin | Vitronectin, collagen IV, fibronectin, and laminin combination required for long-term maintenance of hPSCs in defined media (Ludwig, Levenstein et al. 2006)
Vitronectin alone capable of maintaining hPSCs long-term (Braam, Zeinstra et al. 2008)

### 1.1.3.1 TGFβ family

Several studies indicate that self-renewal of PSCs is associated with TGFβ/Activin/Nodal subfamily signaling and suppression of BMP subfamily signaling (Besser 2004, James, Levine et al. 2005, Vallier, Alexander et al. 2005, Xu, Peck et al. 2005). BMP4 initiates differentiation to trophoblast or primitive endoderm via SMAD 1/5/8 signaling (Xu, Chen et al. 2002, Pera, Andrade et al. 2004). It has been observed that in undifferentiated cells SMAD2/3 are activated while Smad1/5 are only active in mitotic cells, upon differentiation SMAD2/3 signaling decreases while SMAD1/5 signaling increases, and inhibiting the Activin/Nodal signaling pathway results in differentiation (James, Levine et al. 2005, Vallier, Alexander et al. 2005). Both TGFβ and BMP responsive SMADs bind and regulate the NANOG proximal promoter, and TGFβ/Activin signaling and FGF signaling enhance NANOG promoter activity, while BMP signaling suppresses NANOG promoter activity (Xu, Sampsell-Barron et al. 2008). Factors associated with TGFβ signal transduction have been found to be highly expressed in undifferentiated hPSCs, such as Nodal (Rosler, Fisk et al. 2004), Cripto, Lefty1 and Lefty2 (Sato, Sanjuan et al. 2003). Furthermore, Nodal, Lefty1 and Lefty2 are reduced upon differentiation (Besser 2004).

TAZ, a critical regulator of SMAD shuttling, has been found to complex with SMAD2, SMAD3 and SMAD4, and only weakly complex with SMAD1 (Varelas, Sakuma et al. 2008). TAZ knockdown disrupts TGFβ/Activin/SMAD2/3 signaling and induces hPSC specification into neuroectoderm, however TAZ knockdown does not disrupt BMP signaling. Although TAZ is considered to be a canonical effector of the Hippo pathway, it is presently unknown what signals regulate TAZ in hPSCs. Interestingly, it has been shown that TAZ also regulated Wnt/β-catenin signaling (Varelas, Miller et al. 2010), making TAZ an important site for crosstalk between these pathways.
SMAD4 binds to SMAD2/3 as well as SMAD1/5/8 and facilitates translocation of these SMADs into the nucleus. Short hairpin RNA interference mediated knockdown of SMAD4 has been observed to allow the maintenance of hPSCs, despite inhibiting both TGFβ signaling and BMP signaling (Avery, Zafarana et al. 2010). Intriguingly, the differentiation usually observed during inhibition of TGFβ/activin/nodal signaling is limited with SMAD4 knockdown, suggesting that SMAD2/3 signaling in hPSCs is only required to suppress SMAD1/5/8 signaling. Previously, the supportive effect of FGF2 on hPSC maintenance was shown to be dependent on ALK4/5/7, indicating that FGF2 activation of SMAD2/3 signaling, and subsequent repression of SMAD1/5/8 signaling, is a mechanism whereby FGF inhibits BMP signaling (Vallier, Alexander et al. 2005).

GDF3 is expressed in the mouse embryo, in the inner cell mass, and is expressed in mouse and human embryonic stem cells (Levine and Brivanlou 2006). Analysis of transcription profiles of 59 hPSC lines from 17 laboratories reveal that GDF3 is a robust marker of pluripotency (Adewumi, Aflatoonian et al. 2007). GDF3 is a BMP inhibitor, although at very high non-physiological concentrations it may be capable of activating Activin/Nodal signaling (Levine, Levine et al. 2009). HPSC culture typically consists of colonies of varying size, with local cell density dramatically varying in different regions. We have identified GDF3 as a major regulator of the microenvironment-dependent response of hPSCs, as cells in high local cell density express higher GDF3, likely due to increased paracrine accumulation, and this antagonizes BMP endogenously secreted by differentiated cells (Peerani, Rao et al. 2007).

1.1.3.2 FGF

Several population-based studies indicate the importance of FGF signaling in the self-renewal of PSCs. cDNA oligonucleotide microarray technologies have highlighted a role for components of the FGF, TGFβ/bone morphogenic protein (BMP), and Wnt pathways in the maintenance of pluripotency (Rao and Stice 2004). Among the most highly enriched genes in hPSCs include Oct3/4, FGF2/4/14, FGFR1/2/4, LeftyA, LeftyB, GDF3, Nodal, and Wnt1/5A. Additional array technologies have also indicated that FGFs and FGFRs are highly expressed in hPSCs (Sato, Sanjuan et al. 2003, Brandenberger, Kherbtukova et al. 2004). Using RT-PCR, FGFR expression in hPSCs has been reported as FGFR1>FGFR3>FGFR4>FGFR2 (Dvorak and Hampl 2005). Also, massively parallel signature sequencing (MPSS) profiling has also shown hPSCs express FGFR1, FGFR3, FGFR4, FGF2, and FRS2 (Brandenberger,
Khrebtukova et al. 2004). Loss of pluripotency correlates with downregulation of FGF2 and its receptors (Sato, Sanjuan et al. 2003, Brandenberger, Wei et al. 2004) and autocrine FGF2 signaling has been shown to be operative in hPSCs (Dvorak, Dvorakova et al. 2005), further highlighting the important role of FGF in maintenance. Several groups (Boyer, Lee et al. 2005) and references therein) have shown that OCT4, NANOG, and SOX2 form the core transcriptional switch regulating hPSC maintenance. These three core transcription factors are autoregulatory, and all co-occupy the FGFR2 and FGFR1 promoters (Boyer, Lee et al. 2005).

A critical disclaimer of this data is that the population-based studies summarized above do not discriminate between the hPSC culture subpopulations, hence the expression and behavior of undifferentiated cells may be obfuscated by subpopulations of differentiated progeny. Especially in cases where signaling or protein expression is bimodal, population-average measurements are poor indicators of single-cell behavior, necessitating the use of single-cell level measurements (Hoffman and Carpenter 2005). Based on fluorescent microscopy and anti-FGFR1 staining of hPSC cultures, Bendall et al. have reported that the differentiated cells in hPSC culture exclusively express FGFR1 and that this has long hidden the fact that undifferentiated cells do not express FGFR1, instead the hPSCs exclusively express IGFR1 (Bendall, Stewart et al. 2007). This study contradicts other single cell studies that show sorted SSEA3+ hPSCs express all four FGFRs (Eiselleova, Matulka et al. 2009). FACS costaining of SSEA4 and FGFR1 has also shown high expression of FGFR1 in SSEA4 high populations (Rosler, Fisk et al. 2004). The Bendall et al. study also found minimal FGFR3 and FGFR4 expression in the hPSC bulk population analyzed by QPCR, contradicting previous studies of bulk hPSC cultures (Dvash, Mayshar et al. 2004). FGF2 binding assays with hPSCs by two independent groups have indicated that hPSCs bind FGF2 (Levenstein, Berggren et al. 2008, Eiselleova, Matulka et al. 2009). Lastly, immunohistochemistry by Wang et al. have shown that IGFR1 is expressed both on hPSCs and differentiated derivatives (Wang, Schulz et al. 2007). One possible explanation of these findings is that receptor expression can vary dramatically between subpopulations and also between culture conditions, and the results of Bendall et al. may be due to assay sensitivity issues, with both hPSCs and hdFs expressing FGFRs but hdFs having increased FGFR expression (Eiselleova, Matulka et al. 2009).
Most recently, Eiselleova and colleagues have revealed that hPSCs have intracrine FGF2, and intriguingly, exogenous FGF2 is unable to rescue shRNA-mediated knockdown of FGF-2 in hPSCs (Eiselleova, Matulka et al. 2009). The authors propose a model where intrinsic FGF2 maintains hPSCs, exogenous FGF2 can enhance this but is mainly required to enhance adhesion and survival during passaging, which indirectly contributes to hPSC maintenance.

In addition to FGFRs, other receptor tyrosine kinase receptors (RTKs) such as epidermal growth factor receptors (EGFRs) (Wang, Schulz et al. 2007), platelet-derived growth factors receptors (PDGFRs) (Pebay, Wong et al. 2005) and tropomyosin-related kinase receptor (TRKRs) (Pyle, Lock et al. 2006) support hPSC maintenance. These RTKs have similar downstream targets implicated in maintenance, such as MAPK and PI3K, however an open question is what, if any, kinase targets are differentially activated by these receptors, and what sets these receptors apart from other RTKs that do not have a pro-maintenance effect.

1.1.3.3 FGF inhibition of BMP signaling

FGF2 stimulation appears to have an effect on TGFβ ligand expression in both MEFs and hPSCs. FGF2 treatment leads to BMP4 suppression and up-regulation of Gremlin, TGFB1, as well as Activin A transcripts (Greber, Lehrach et al. 2007). Xu and colleagues have found that 40 ng/ml bFGF did not reduce SMAD1 phosphorylation, however 100 ng/ml bFGF reduced pSMAD1 to levels comparable to CM (Xu, Peck et al. 2005). The effects of high bFGF could be recapitulated with 40 ng/ml bFGF and various BMP inhibitors such as noggin, gremlin, or soluble BMP receptor IA, indicating that high FGF may result in the secretion of BMP inhibitors. Interestingly, even 40 ng/ml bFGF reduces BMP reporter assay (luciferase reporter of ID1) activity, despite no change in total pSmad1 levels, so total pSmad1 alone may not accurately reflect BMP transduction. The authors note that bFGF has been reported to repress BMP signaling by preventing the nuclear translocation of pSmad1 (Pera, Ikeda et al. 2003) or by inhibiting Smad1 activity in the nucleus (Nakayama, Tamura et al. 2003).

One possible mechanism for this is that FGF may induce phosphorylation of the linker region of SMAD1 by MAPK, abolishing nuclear accumulation (Yamagata, Matsuzaki et al. 2005). To investigate this further, Peerani et al. pulsed hPSCs for 90 minutes with serum-free chemically defined media either with or without cytokines (Peerani, Rao et al. 2007). Using image analysis and pSmad1 and Smad1 (total Smad1) staining, in conjunction with Hoechst
staining to separate nuclear and cytoplasmic compartments, it was revealed that FGF2 alone decreases both cytoplasmic and nuclear pSmad1. Conversely, total Smad1 increased in the cytoplasm after FGF2 stimulation. It therefore appears that FGF2 prevents phosphorylation of Smad1, thus reducing translocation to the nucleus and increasing nonphosphorylated Smad1 in the cytoplasm. Furthermore, the rapid kinetics of this assay (90 minutes) indicates that the mechanism likely does not involve changes in the de novo synthesis of BMP antagonists or TGFβ agonists. This does not preclude additional effects of FGF2, including upregulation of BMP antagonists and TGFB agonists over longer time spans, these are additional mechanisms by which FGF2 inhibits BMP.

1.1.3.4 Insulin

Insulin or IGF is present in all PSC expansion media published to date, although it is most often present in unknown concentrations in media components such as fetal bovine serum or in undisclosed amounts in proprietary supplements such as Knock-Out Serum Replacement (Price, Goldsborough et al. 1998), N2 or B27 (Brewer, Torricelli et al. 1993). It has been reported that IGF1R inhibition results in increased differentiation (Wang, Schulz et al. 2007), however during EB differentiation it appears that insulin has no pro-maintenance effect and that it instead shifts differentiation towards neural and away from mesendoderm (Freund, Ward-van Oostwaard et al. 2008). At the high concentrations of insulin present in PSC formulations, insulin is known to activate IGF1R signaling (Freund, Ward-van Oostwaard et al. 2008), however the majority of studies to date do not discriminate between insulin receptor and IGF1R.

1.1.3.5 Wnt

In contrast to mESCs, hPSCs have upregulated mRNA expression of Wnt ligands and signal transduction components, indicating that they are fully capable of responding to Wnt (Wei, Miura et al. 2005, Okoye, Malbon et al. 2008). Initially, it was reported that Wnt pathway activation by a specific pharmacological inhibitor of glycogen synthase kinase-3 (GSK-3) by 6-bromoindirubin-3'-oxime (BIO) maintains pluripotency of both hPSCs and mESCs (Sato, Meijer et al. 2004). Longer term studies have failed to recapitulate these short-term studies and may indicate that Wnt3a has a partial and transient pro-maintenance and survival effect (Xiao, Yuan et al. 2006). Studies by Dravid et al. revealed that although Wnt3a
stimulates transient hPSC proliferation, it also increases differentiation (Dravid, Ye et al. 2005). Furthermore, inhibiting Wnt signaling, with sFRP2 or Dkk-1 for 4 days, did not reduce maintenance of hPSCs on supportive feeder cells. Importantly, canonical Wnt signaling, indicated by a functional β-catenin mediated transcriptional activation reporter, is minimal in pluripotent hPSCs but increased during differentiation. In fact, in direct contradiction with the initial BIO studies, activation of canonical Wnt/β-catenin signaling has been reported to result in rapid differentiation of hPSCs (Sumi, Tsuneyoshi et al. 2008), however this study used an inducible β-catenin system to test hPSC response, and did not actually use Wnt ligands.

1.1.3.6 Extra Cellular Matrix Proteins

With the development of defined media, the push for defined, cost-effective substrates has also accelerated, with several novel defined substrates recently being reported. HPSCs are most often cultured on gelatin and MEFs, or Matrigel™, a protein mixture secreted by Engelbreth-Holm-Swarm mouse sarcoma cells. HPSCs have receptors for all major ECM components (Braam, Zeinstra et al. 2008). HPSCs have been found to attach efficiently to vitronectin, fibronectin, and Matrigel™, and poorly to laminin, entactin, and collagen IV (Braam, Zeinstra et al. 2008). Vitronectin has been reported to be capable of maintaining hPSCs long-term similarly to Matrigel™ in defined media (Braam, Zeinstra et al. 2008), however other studies have found vitronectin alone to be insufficient, instead a combination of collagen IV, fibronectin, laminin and vitronectin is required for long-term maintenance of hPSCs in defined media (Ludwig, Levenstein et al. 2006). Similarly, ROCK inhibitor has been reported to allow the culture of HPSCs on poly-D-lysine coated surfaces in defined media (mTeSR®) (Harb, Archer et al. 2008), however other groups have reported this combination of ROCK inhibitor, poly-D-lysine, and mTeSR™1 do not maintain hPSCs (Mahlstedt, Anderson et al. 2010). Another synthetic matrix, oxygen plasma etched tissue culture polystyrene in conjunction with conditioned medium has also been reported to be capable of maintaining hPSCs long-term (Mahlstedt, Anderson et al. 2010). A synthetic polymer coating poly[2-(methacryloyloxy)ethyl dimethyl-(3-sulfopropyl)ammonium hydroxide] (PMEDSAH) was also reported to sustain H9s for ten passages in StemPro® hESC SFM defined media (Villa-Diaz, Nandivada et al. 2010). The surface was unable to sustain hPSCs grown in mTeSR™1 defined media, and in StemPro® the BG01 hPSC line could not be passaged long term. It is unclear why these media and cell line response differences exist, and it remains to be shown if passage
numbers much higher than ten can be supported. It is becoming increasingly clear that as with the case of defined media, these novel defined substrates will need to be validated in multiple independent laboratories before widespread acceptance in the hPSC field.

Figure 1- 1 The Stem Cell Microenvironment

The stem cell microenvironment consists of various interacting factors that regulate cell fate. These factors are difficult to measure, rarely reported, and typically vary spatially within a culture vessel and can dramatically vary between laboratories. Local cell density is often a master regulator of these factors. Depicted here, cells are distributed heterogeneously in a culture vessel. This results in different cell-cell contacts, endogenous ligands, mechanical forces, and ECM interactions signaling to cells in different microenvironments.

Table 1- 2 Sources of Culture Heterogeneity

<table>
<thead>
<tr>
<th>Source of Heterogeneity</th>
<th>References</th>
<th>Example</th>
<th>Control measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endogenous Factors</td>
<td>(Dvorak, Dvorakova et al. 2005, Janes, Gaudet et al. 2006, Xiao, Yuan et al. 2006, Adewumi, Aflatoonian et al. 2007, Peerani, Rao et al. 2007, Hsu and Sabatini 2008)</td>
<td>e.g. cytokines and extracellular matrix proteins secreted by cultured cells</td>
<td>Assess cell line endogenous signaling pathway activation via cell based assays (Nazareth, Ostblom et al. 2013) or PCR</td>
</tr>
<tr>
<td>Mechanical Forces and Cell-cell contact</td>
<td>(McBeath, Pirone et al. 2004, Watanabe, Ueno et al.</td>
<td>e.g. flow, shear stress, cell tension from cell or colony shape, packing density</td>
<td>Report images of cell distributions, perform experiment at multiple cell densities</td>
</tr>
</tbody>
</table>
2007, Harb, Archer et al.
2008, Li, Krawetz et al. 2009

| Heterogeneous Subpopulations | (Bendall, Stewart et al. 2007, Peerani, Rao et al. 2007, Hayashi, Lopes et al. 2008, Toyooka, Shimosato et al. 2008) | e.g. differentiated populations and heterogeneous pluripotent states | Assess single-cell expression of population markers (ex. Oct4 and Sox2 to delineate pluripotent, early neuroectoderm, mesendoderm, and extra-embryonic/other populations (Nazareth, Ostblom et al. 2013)) |

1.1.4 Sources of Culture Heterogeneity

A key difficulty in investigating hPSC signaling lies in the heterogeneities introduced by the stem cell microenvironment, or "niche", which has been shown to be spatially organized and have an important role in determining stem cell responsiveness to exogenous signals.
(Davey and Zandstra 2006, Stewart, Bosse et al. 2006, Singh and Terada 2007, Hayashi, Lopes et al. 2008, Stewart, Bendall et al. 2008) (see summary of these factors and suggested control measures in Table 1-2). In addition to the exogenous media and substrate, there are several important factors that affect stem cell fate decisions. Such factors include endogenous ligands and ECM proteins secreted by the cells, mechanical forces and cell-cell contact, and the existence of multiple sub-populations of cells (Figure 1-1). These factors not only vary greatly between different culture systems, but also are spatially heterogeneous within each culture system. These factors are difficult to measure routinely and are not reported in the majority of the stem cell literature, and evidence is mounting that these factors play a far greater role than initially suspected in regulating stem cell fate.

1.1.4.1 Endogenous Cytokines

In addition to exogenous factors present in culture conditions, such as the cytokines in the media and the matrix proteins in the substrate, the cells secrete numerous endogenous cytokines and matrix proteins. These factors are rarely characterized; however, it is becoming clear that endogenous factors play an important role in cell fate regulation. For example, in human epithelial cell culture systems it has been estimated that approximately half of the intracellular signaling is a response to induced autocrine factors, and that these induced autocrine loops perhaps aid the cells in detecting their local environment (Janes, Gaudet et al. 2006). In hPSC culture, which typically contains exogenous Activin A, Activin A has been shown to induce the expression of FGF2, FGF8, and Wnt3 (Xiao, Yuan et al. 2006). HPSCs secrete FGF2 into the media and express several FGF receptors, and this autocrine loop is thought to be critical in stem cell maintenance (Dvorak, Dvorakova et al. 2005). GDF3, a secreted factor antagonist of BMP signaling, is also secreted in the media (Peerani, Rao et al. 2007) and is in fact considered a marker of hPSCs (Adewumi, Aflatoonian et al. 2007). In addition to cytokines, hPSCs, like cancer cells, are thought to secrete high amounts of metabolites and the accumulation of these factors likely also regulates cell fate (Hsu and Sabatini 2008).

1.1.4.2 Mechanical forces and cell-cell contact

In addition to cytokine cues, the stem cell environment also includes mechanical cues. In particular, cell shape and resulting cytoskeletal tension can be transduced by the cell via RhoA...
signaling to regulate developmental decisions (McBeath, Pirone et al. 2004). HPSCs preferentially self-renew in colonies with tight cell-cell contact. A key difference between mouse and hPSCs is that hPSCs have high rates of apoptosis upon disassociation, with a survival rate of approximately 1% (Thomson, Itskovitz-Eldor et al. 1998, Amit, Carpenter et al. 2000, Reubinoff, Pera et al. 2000, Pyle, Lock et al. 2006). The specific mechanism for this reliance on cell-cell contact is unclear, however it has been demonstrated that survival after disassociation can be increased to approximately 27% by inhibiting ROCK, the major downstream effector of RhoA (Watanabe, Ueno et al. 2007). As a result, ROCK inhibitor has proven to be a valuable tool in applications of hPSCs that require single cell disassociation, such as sub-cloning after gene transfer and generation of homogeneous embryoid bodies (Ungrin, Joshi et al. 2008). Upon addition of ROCK inhibitor to hPSC colonies, the colonies disassemble and hPSCs take on a more mesenchymal appearance within 6 hours, thus implicating Rho-Rock signaling in cell-cell contact regulation (Harb, Archer et al. 2008).

ROCK inhibitor increases cell-cell interaction between hPSCs, as aggregate formation in single cell suspension culture is increased with Y-27632 (trans-4-[(1R)-1-Aminoethyl]-N-4-pyridinyl-cyclohexanecarboxamide dihydrochloride) (Li, Krawetz et al. 2009). In addition to cell-cell contact, ROCK inhibition also increase motility (Harb, Archer et al. 2008) and moderately increases cell growth (Watanabe, Ueno et al. 2007). The specific mechanism of the anti-apoptotic effect of ROCK inhibitors is unclear. One intriguing observation is that blocking cell-cell adhesion with EGTA abolishes the survival effect of Y-27632, indicating that ROCK inhibition decrease of apoptosis is indirect and mediated by increasing cell-cell contact (Li, Krawetz et al. 2009). Although the initial ROCK inhibitor study did perform clonal assays, where a single hPSC was plated per well in 96-well plates, this study was done in the presence of MEF feeder layers (Watanabe, Ueno et al. 2007), so increased cell-cell contact (between hPSCs and the feeders) cannot be ruled out as the mechanism behind the anti-apoptotic effect of ROCK inhibition. Repeating this study in feeder-free conditions at clonal densities will definitively address this issue.

1.1.4.3 Heterogeneous Subpopulations

PSC cultures consist of various subpopulations which exhibit different protein expression and signaling dynamics, which population-average measurements often obfuscate. As a prime example of this aspect of the hPSC culture, it has been reported that in a culture condition
consisting of hPSCs grown on autologously derived fibroblast-like cells (hdF) feeders, the exogenous FGF2 commonly added to medium does not in fact act on the undifferentiated hPSCs, FGF2 instead signals the feeder cells, which in turn release IGFII and other ligands which directly signal to undifferentiated cells (Bendall, Stewart et al. 2007). Previously, in MEF-feeder layer system, it has been noted that the response of hPSCs to BMP2 is inhibited by MEFs in a dose-dependent manner, as MEFs secrete the BMP antagonist gremlin (Peerani, Rao et al. 2007). However even in feeder-free chemically defined media, it has been observed that extra-embryonic endoderm (ExE) derived from the hPSCs secrete BMP2, and this is antagonized by hPSC-secreted GDF3, forming an intricate balance in the stem cell niche (Peerani, Rao et al. 2007). It has previously been demonstrated that in many mouse embryonic stem cell culture conditions, the Oct4+ population consists of two interconvertable populations, one similar to the in vivo inner cell mass (ICM) (Oct4+Rex1+) and another equivalent to the in vivo epiblast (Oct4+Rex1-) (Toyooka, Shimosato et al. 2008). These populations are difficult to distinguish, however they possess different differentiation behaviour and the relative contribution is dependent on the particular culture conditions (Hayashi, Lopes et al. 2008, Toyooka, Shimosato et al. 2008). These studies highlight the complex and often reciprocal signaling in the stem cell microenvironment between subpopulations. Although these studies focused on two major cell types in each case, in each culture condition there are likely numerous cell types all interacting with poorly understood dynamics and feedback loops. Future work characterizing these various cell types, what factors they secrete and respond to, and why they appear or do not appear in different culture systems will aid in developing more homogeneous expansion and directed differentiation.

1.1.5 Localized cell density – Master Regulator of Culture Heterogeneity

Forty seven years ago, Rein and Rubin observed that independent of the total number of cells in a culture vessel, phenotypes such as growth can be determined almost completely by local cell density, the number of cells in a given radius (Rein and Rubin 1968). In CDM, hPSCs have been shown to respond to the removal of supportive exogenous cytokines in a spatially heterogeneous and density dependent manner, where hPSCs in high local cell density areas largely maintain pluripotency and hPSCs in low local cell density niches rapidly differentiate (Peerani, Rao et al. 2007). Similar to the Rein and Rubin studies with chick
embryo explants, the density dependent effects in hPSC culture appear to be mediated by supportive factors, such as GDF3, which scale with localized cell density. In this culture system, GDF3 secreted by hPSCs antagonizes BMP2 secreted by extra-embryonic endoderm autologously derived from the hPSC population, with the balance between these factors being determined by local cell density. A similar phenomenon exists in mESC culture, with supportive Jak-Stat activation by endogenous factors scaling with local cell density (Davey, Onishi et al. 2007, Peerani, Onishi et al. 2009). Also, the results from these 2D studies of localized cell density can also be extrapolated to 3D. When mESCs are cultured in 3D alginate spots, spots seeded with a lower density are more sensitive to LIF removal (Fernandes, Kwon et al. 2009), this is presumably due to decreased endogenous Jak-Stat signaling. In all these cases, it remains unclear how local cell-density information is sensed by the cell – cytokines may be localized by immediate binding to the secreting cell, or alternatively, cytokines may diffuse in the media over short distances creating localized gradients.

For generating iPSC, it has long been observed that reprogramming efficiency is highly dependent on cell density. Although the mechanisms behind this are unclear, it is thought that fibroblasts at low densities may have increased senescence, resulting in lower reprogramming efficiency, and conversely at too high densities cells have hindered growth, especially after the long culture periods required for reprogramming (Maherali and Hochedlinger 2008). Indeed, in secondary systems with high reprogramming efficiency, at high cell densities colony formation frequency declines.

In addition to pluripotent stem cell reprogramming and maintenance, cell density also plays an important role in differentiation. Most notably, neural induction of mESCs in monolayers has been found to be robustly inhibited by even slight increases in density (Tropepe, Hitoshi et al. 2001, Ying and Smith 2003). In hPSCs, a homogeneous cell distribution, created by seeding hPSCs as single cells at high density in CM and ROCK inhibitor, instead of seeding whole colonies, has also been found to increase neural conversion efficiency (Chambers, Fasano et al. 2009). The high-density is necessitated by the high apoptosis rate in single-cell hPSCs. Also, in hPSC differentiation it has been found that definitive endoderm specification towards pancreatic fates, rather than liver, is highly dependent on seeding at low cell densities and disrupting cell-cell contact (Cai, Yu et al. 2009).
Local cell density dependence has largely been attributed to local concentration differences in signaling factors. However, localized cell density also effects mechanical forces, cell-cell contact, the distribution of subpopulations within the culture, the ratio of perimeter to internal cells, and cell adhesion and spreading against the substrate. For this reason, we propose that local cell density is an underappreciated master regulator of culture conditions. Furthermore, as localized cell density is difficult to quantify, and rarely reported, we also hypothesize that lab-to-lab variation in cell density distribution can explain many of the outstanding discrepancies in the stem cell literature.

1.1.6 Discussion

In recent years human pluripotent stem cell biology has experienced an explosion in research, with a great emphasis on discovering what factors can guide these stem cells towards clinically relevant cell types and applying hiPSCs for disease modeling and drug development. Much progress has been made in the field, in just over a decade culture conditions have evolved from feeder-based and serum-based to using feeder-free chemically defined media and substrates. However, numerous factors interplay to control stem cell fate, and these factors are all too often unrecognized and not reported, often undermining reproducibility of results.

Addressing these concerns about the reproducibility of hPSC assays should be an immediate priority. Towards this aim, we propose closer monitoring, reporting, and control of cell density distribution within assays. Local cell density has been demonstrated to be a dramatic regulator of pluripotent stem cell maintenance, differentiation, and for reprogramming and it serves as a master regulator of several key sources of culture heterogeneity, including endogenous factor secretion, mechanical forces, cell-cell contact, and subpopulations of cells. Endogenous secreted factors and cell density are expected to be particularly susceptible to variation between labs, however these factors are notoriously difficult to measure and manipulate. In addition to this measure, to address variability in endogenous secreted factors, assays could be performed under different background conditions, with exogenous factors added to mimic different endogenous conditions. For example, assaying in conditions with and without exogenous FGF2 and/or BMP4, and performing the assay at different cell densities. This may clarify under what conditions different labs see similar results, for example if variability is due to differential endogenous FGF2 or BMP4 in the culture system. As a simple measure, cell density and culture volumes
should also be more often reported, as endogenous factor concentration directly varies with these. For additional control, more advanced methods such as micro-contact printing can be employed to finely control cell spatial distribution. We have recently applied these concepts in a high-throughput platform for screening cells in microenvironments with configurable colony size and cell density and have found these parameters to be critical in designing robust and rapid hPSC screens (Nazareth, Ostblom et al. 2013). This assay additionally provides a standardized method for measuring cell-line specific endogenous pathway activation, and a rapid way to quantify cell-line specific differentiation propensity.

As an additional measure to regulate assay heterogeneity, multivariate analysis, particularly single cell microscopy based phenotypic analysis (such as HCA), would reap many benefits. HCA allows tracking of spatial heterogeneity, and the response of each cell can thus be linked to such parameters as local cell density (Peerani, Rao et al. 2007). By staining for cell markers, or using reporter cell lines, multiple cell lineages can be simultaneously tracked.

Murine epiblast stem cells (mEpiSCs), derived from postimplantation murine epiblast stage embryos, have been shown to have many similarities to hPSCs, suggesting that hPSCs are developmentally similar to postimplantation epiblast (Brons, Smithers et al. 2007, Tesar, Chenoweth et al. 2007). Both hPSCs and mEpiSCs are dependent on FGF and Activin signaling, have slower proliferation rates than mESCs, female lines have similar X-inactivation status, and both are typically dependent on cell-cell contact and require passaging as colonies, showing very low rates of clonal survival. mEpiSCs have been shown to also have similar growth factor requirements as hPSCs for differentiating into neur ectoderm, extra embryonic tissue, and mesendoderm (Vallier, Touboul et al. 2009). As such, the signaling discussed in this chapter and the caveats to interpreting signaling studies introduced by heterogeneous culture conditions also apply to mEpiSCs. Already, there are unresolved discrepancies in the mEpiSC literature. For example, it is well established that FGF inhibition reduces neuroectoderm induction in mEpiSCs and hPSCs (Vallier, Mendjan et al. 2009, Vallier, Touboul et al. 2009) as well as mESCs (Ying, Stavridis et al. 2003), however it has been reported that FGF inhibition actually increases neuroectoderm induction in mEpiSCs (Greber, Wu et al. 2010).
Several groups have reported the reversion of mouse EpiSCs (Bao, Tang et al. 2009, Guo, Yang et al. 2009, Hanna, Markoulaki et al. 2009, Onishi, Tonge et al. 2012) and hPSCs (Buecker, Chen et al. 2010, Hanna, Cheng et al. 2010, Xu, Zhu et al. 2010) from a “primed” pluripotent state into a mESC-like state (“naïve ground state” pluripotency). The requirements to stabilize naïve and primed pluripotent states appear to be related to genetic background, as different species, and genetically independent lines within species, have different requirements to stabilize in either pluripotent state (Hanna, Markoulaki et al. 2009, Hanna, Cheng et al. 2010). Determining how genetic backgrounds interplay with the microenvironment, and how the microenvironment can be manipulated to stabilize pluripotent states, especially without genetic manipulation, will enable a clearer mechanistic understanding of reprogramming.

Presently, efforts towards using hiPSCs and derivatives for in vitro disease modeling and drug discovery is rapidly scaling up, in particular for cardiac, neurological, hepatic, haematological, muscle, and eye disorders (reviewed in (Rajamohan, Matsa et al. 2013)). However, the discordance observed in large studies of drug response data of immortalized cancer lines (Haibe-Kains, El-Hachem et al. 2013), a much simpler model system, is troubling and underscores the importance of first resolving these issues and developing improved robust assays. By striving to recognize the various factors in play that dictate cell response, and by implementing strategies to control and/or monitor these factors, the reproducibility and relevancy of stem cell drug discovery and development programmes will be dramatically enhanced.

1.2 Human pluripotent stem cells & drug screening

1.2.1 Abstract

The convergence of stem cell biology and high-throughput (HTP) screening, two fields that have experienced tremendous growth in the past few years, has created exciting opportunities for regenerative medicine and disease modeling and therapy. Optimizing differentiation, maturation, scale-up, and improved characterization of hPSCs and derivatives are all major goals to be achieved in order to reap the rewards promised by hPSC technology in the clinic, and HTP screening may prove to be an invaluable tool towards these efforts. In this focused review, we first outline the new opportunities afforded by hPSC technology, providing a brief
history of human PSCs and outlining the applications of PSCs in lead discovery, disease modeling, and safety pharmacology. We next discuss screens performed directly with hPSCs, highlighting the successes and lessons learned, with a focus on technical aspects of translating PSCs in HTP screens. We conclude by concisely synthesizing these lessons that can be applied towards improving screens. To date PSC based HTP screens have been at a lower scale than typical pharmacological lead discovery programs, and studies have largely been proof of principle with little impact to patients - though this is changing rapidly. Learning from recent history and applying the lessons learned from the efforts at applying hPSCs and derivatives towards large scale screens promises to accelerate the translation of hPSC and derivative based lead discovery into the clinic.

1.2.2 Introduction

Human pluripotent stem cells (hPSCs) are heralded to transform regenerative medicine and drug discovery and development, as they potentially can supply a scalable source of clinically relevant cell populations. These cells can be used for cell therapies, toxicology and drug safety evaluation, lead discovery, and they afford unique opportunities to study diseases. Towards this end, techniques for differentiating hPSCs into desired cell populations and tissues has advanced rapidly. In parallel, high-throughput (HTP) screening in biomedical research has also grown tremendously, with 1 million+ compounds routinely screened in industrial settings. Human embryonic stem cells (hESCs) were derived in 1998 (Thomson, Itskovitz-Eldor et al. 1998), and 1997 was the first year that ten papers on PubMed with "high-throughput screening" in its title were published (Macarron, Banks et al. 2011). These two fields are slowly converging, with a shared vision of using hPSCs and derivatives to derive clinical therapeutics. HTP approaches are not only for lead discovery in terminally differentiated populations, but can be used to accelerate this process, namely by improving differentiation and maturation efficiencies, and finding cost-effective compounds to aid in scale-up of these processes; in this way HTP approaches can feedback to accelerate translation to the clinic.

Although several HTP screens to improve differentiation have been reported for hPSC differentiation towards cardiomyocyte (Spiering, Davidovics et al. 2015), pancreatic cells (Chen, Borowiak et al. 2009), and other cell types, the differentiation protocols widely employed have not been derived from HTP approaches, but rather from traditional hypothesis-driven approaches. The majority of PSC differentiation protocols have been made possible by
relying on a developmental biology perspective, leveraging findings from numerous other model systems to recapitulate key signaling events in the embryo to generate differentiated cell populations (Murry and Keller 2008). Many small molecules have been used to manipulate hPSC self-renewal and differentiation (more than 55 small molecule containing cocktails reviewed in (Atkinson, Lako et al. 2013)), these compounds are anticipated to be critical in developing cost-effective scale-up of pluripotent and differentiated cells. A few of these small molecules have gained wide spread use, such as Rho-associated kinase (ROCK) inhibitor Y-27632 to promote survival of disassociated hPSCs (Watanabe, Ueno et al. 2007), GSK3β inhibitor CHIR99021 and Wnt inhibitors IWP4 and IWP2 for differentiation towards cardiomyocytes (Lian, Hsiao et al. 2012), and ALK4/5/7 inhibitor SB431542 with a BMP inhibitor such as Noggin or LDN-193189 (Nazareth, Ostblom et al. 2013) for efficient neural induction (Chambers, Fasano et al. 2009). A completely small-molecule-driven protocol for generation of hPSC derived hepatocytes has recently been reported (Siller, Greenhough et al. 2015). It is notable that at present, the widespread small molecule applications have all been discovered through hypothesis-driven studies based on basic developmental biology, and not through the data-driven HTP screening approach.

In this review, we first provide a brief overview of the emergence of hESCs, hiPSCs, and their clinical relevance arising from their potential ability to generate differentiated cell types at scale. Next we review the role of hPSCs and derivatives in lead discovery and disease modeling. As examples, we examine select HTP hPSC disease modeling applications in detail including a HTP screen of hepatocyte-like cells (HLCs) differentiated from hiPSCs derived from patients with alpha-1 antitrypsin (AAT) deficiency (Choi, Kim et al. 2013), a HTP screen of neural crest precursors derived from hiPSCs generated from patients with familial dysautonomia (FD) (Lee, Ramirez et al. 2012), and a HTP screen of neoplastic hPSCs used as a disease model of cancer stem cells (CSCs) (Sachlos, Risueno et al. 2012). We next survey the opportunities of hPSC derivatives in safety pharmacology and toxicology screens, in which hPSC derived cardiomyocytes and HLCs have a prominent role. The bulk of this review is a comprehensive analysis of screens to date that have directly used hPSCs to screen for survival and self-renewal factors, with a focus on technical aspects of these screens. We conclude with a summary of specific recommendations derived from this analysis for future PSC based HTP programs. With further improvements, in particular by identifying limitations of screens
performed to date and how to address them, future screens of PSCs and derivatives promise to have a tremendous impact on the clinical outcome of patients.

1.2.3 Pluripotent stem cells

Stem cells have the ability to self-renew through mitosis and the ability to differentiate into a range of more specialized cell types. Stem cells can be broadly grouped as pluripotent, having the ability to differentiate into every somatic cell type, multipotent, having the ability to differentiate into several but not all somatic cell types, and unipotent, having the ability to differentiate into one terminal cell type. Human Embryonic stem cells (hESCs), which are isolated from the inner cell mass of the blastocyst, are pluripotent stem cells which can self-renew indefinitely in vitro (Thomson, Itskovitz-Eldor et al. 1998). In contrast, adult stem cells are multipotent and can typically only be expanded for a few passages before differentiation. Adult stem cells have been identified in diverse tissues and organs such as bone marrow (Pittenger, Mackay et al. 1999), adipose tissue (Zuk, Zhu et al. 2001), the brain (Song, Stevens et al. 2002), and muscle (Jankowski, Deasy et al. 2002). Due to their ability to generate large numbers of any cell of the human body, hESCs have opened the doors to many exciting applications in regenerative medicine and tissue engineering and are also increasingly recognized for their potential in pharmaceutical development.

Several pluripotent cell lines can be derived from the early embryo (reviewed in (Yu and Thomson 2008)). Predating hESC derivation by nearly two decades, mouse ESCs (mESCs) were first derived from the inner cell mass (ICM) of mouse blastocyst by Evans and Kaufman in 1981 (Evans and Kaufman 1981). These cells are capable of clonally differentiating into diverse cell types and when injected into mice they form teratocarcinomas (Martin 1981). mESCs can also be injected into the blastocyst and participate in embryogenesis similarly to host ICM cells (Gardner and Rossant 1979). Another pluripotent stem cell line, mouse epiblast stem cells (mEpiSCs), has been derived from the mouse epiblast (Brons, Smithers et al. 2007, Tesar, Chenoweth et al. 2007). These cells have many differences from mESCs but have many similarities to hESCs, raising the possibility that the differences between mouse and hESCs, previously attributed to species developmental differences, are in fact due to hESCs being developmentally equivalent to the epiblast stage, and not the ICM stage like mouse ESCs.
In 2006, Takahashi and Yamanaka demonstrated that pluripotent stem cells very similar to mESCs can be generated from mouse fibroblasts by introducing four factors, Oct4, Sox2, c-Myc, and KLF4 (Takahashi and Yamanaka 2006). These induced pluripotent stem cells (iPSCs) form tumors containing tissues from all three germ layers and contribute to mouse development when injected into blastocysts. Human iPSCs (hiPSCs) were also soon generated, using Oct4, Sox2, Nanog, and Lin28 (Yu, Vodyanik et al. 2007). Diverse alternative ways of generating iPSCs have since been developed (reviewed in (Gonzalez, Boue et al. 2011)). Much attention has been paid to the equivalency of hESCs and hiPSCs (review (Nazor, Loring et al. 2012)), due to their differences in methods of generation and origin. The evidence so far indicates these cells, are for the most part equivalent. Embryonic and induced origin pluripotent cells are now often spoken of with no distinction and referred to as mouse or human pluripotent stem cells (mPSCs, hPSCs, or collectively as PSCs). The ability to generate PSCs equivalent to embryonic stem cells from disease patients (patient derived iPSCs) enables the production of new disease models and promises to revolutionize the level of mechanistic understanding of diseases. Scalable generation of clinically relevant cell types, potentially autologously, also has immense potential to transform regenerative medicine and drug development.

The clinical relevance of PSCs mainly arises from their ability to differentiate into terminally differentiated cells; however, this is a complex multi-step process typically recapitulating in vivo development, with each step typically requiring different combinations of signaling molecules. For neuron differentiation, PSCs can be differentiated into neural precursors, then subtype-specific neuronal progenitor cells, and finally to mature functional neurons (Ying, Stavridis et al. 2003). For cardiac differentiation, PSCs are differentiated into mesendoderm, then mesoderm, then cardiovascular precursor cells, immature cardiac cell types, and then to more mature cardiomyocytes (Kattman, Huber et al. 2006, Laflamme, Chen et al. 2007). For endocrine cells, mesendoderm is directed towards definitive endoderm, then primitive gut tube, posterior foregut, pancreatic endoderm, endocrine precursor cells, and finally hormone-producing endocrine cells (D'Amour, Agulnick et al. 2005). Directed differentiation and maturation of PSCs into specific terminally differentiated cells is currently inefficient and current protocols often require co-culture with feeder cells, the presence of serum, or differentiation in 3D embryoid bodies which have high concentrations of endogenous
factors that drive differentiation. Addressing this bottleneck is a major effort in stem cell biology, and will be for the foreseeable future.

1.2.4 Stem cells in lead discovery and disease modeling

Drug discovery campaigns aim to efficiently identify compounds which will eventually lead into clinically relevant medicines. There are several aspects of stem cells which make them particularly advantageous in lead discovery. Stem cells can provide access to previously unavailable cell types, and they offer the chance to investigate cell expansion, differentiation, and regeneration. Although still in the early stages, there are several examples of discovery programs that combine the traditional pharmaceutical high-throughput screening approach with stem cells in order to find drugs to modify stem cell fate (McKernan, McNeish et al. 2010). Screens have been performed to find compounds to activate endogenous cardiac progenitors to treat congestive heart failure or myocardial infarction (Wu, Ding et al. 2004), expand pancreatic islet precursors for diabetes treatment (Chen, Borowiak et al. 2009), to activate adult progenitors to treat immune disorders (Flomenberg, Comenzo et al. 2010), and to identify drugs that selectively target cancer stem cells (Sachlos, Risueno et al. 2012). Perhaps the most exciting application of hPSCs to lead discovery, the advent of hiPSC technology has enabled the use of patient derived hiPSC in disease models. By deriving hiPSCs from a disease patient, it is possible to study the disease in vitro in differentiated cell types previously difficult to obtain, disease progression can be examined, and relevant cell types for drug screening may be generated at sufficiently large scale, opening up a new avenue to discover treatments of diseases. The hiPSC disease modeling field is flourishing, with numerous studies on cardiac, smooth and skeletal muscle, immune, imprinting, neurological, metabolic, haematological, eye, and multi-organ diseases (extensively reviewed in (Rajamohan, Matsa et al. 2013)).

A handful of HTP screens of disease model hiPSCs have also been reported. Choi et al. have generated hiPSC lines from patients with Alpha-1 antitrypsin (AAT) deficiency, and have differentiated these into hepatocyte-like cells (HLCs) to screen for compounds that reduce mutant AAT accumulation (Choi, Kim et al. 2013). HLCs were plated into 96-well plates, and then one of 3131 clinical compounds from the John Hopkins Drug Library was applied at 5 µM for 4 days to each well, and then immunostaining for AAT expression. Out of 262 initial hits (compounds that reduce AAT accumulation by > 50%), 5 hits were determined to reproducibly reduce AAT accumulation in 4 different AAT-deficient patient hiPSC lines. Mechanistic
follow-up studies indicate all 5 hits likely act via autophagy-mediated degradation of AAT, and not reduction of synthesised AAT. Carbamazepine was the control used to develop the assay and was previously known to enhance autophagy and decrease the hepatic load of AAT deficient mice (Hidvegi, Ewing et al. 2010). This study employed a blind screen, and carbamazepine turned up as a hit, demonstrating the proof of principle of HTP screens to detect therapeutics for AAT deficiency. However, no compound performed better across cell lines than carbamazepine. The possibility still remains that scaling up this screen would identify even more potent compounds.

Lee et al. have performed a HTP screen on neural crest precursors derived from hiPSCs generated from patients with FD (Lee, Ramirez et al. 2012). FD is a fatal genetic disorder affecting neural crest lineages caused by a point mutation in the I-κ-B kinase complex-associated protein (IKBKAP) resulting in reduced IKBKAP expression in neural crest cells. To search for potential therapeutics, Lee et al. first derived hiPSC cells from FD patients, generated neural crest cells (a 26-27 day process), FACS sorted these cells and expanded the neural crest cells for two weeks, and then froze down the neural crest cells, and thawed for subsequent use. Cells were seeded into 384-well plates (7500 cells per well), 24h later small molecules were added (to a final concentration of 10 µM), and 48 h later automated RNA extraction and qRT-PCR was performed, which is rarely used in HTP screens. 6912 small molecules were screened in triplicate. The Z' values for the control plates showed remarkable reproducibility, with a Z' score over 0.5 for all sixty 384-well plates. In addition to looking for hits using fold difference, the authors also used a rank-based method, which reduces effects from variability across plates. 43 hits were obtained, 35 of which were commercially available for secondary studies, and eight hits were pursued and determined to significantly increase WT-IKBKAP expression in FD neural crest cells. Notably, the authors found differential potency for IKBKAP induction in other cell types (FD-hiPSCS, FD-fibroblasts, and FD-lymphoblasts) confirming the need to do screens in disease-relevant cells. Testing across multiple clones from multiple patients indicated the compound effects were generally clone and patient-independent. The authors also determined that despite rescue of WT marker expression, the eight compounds could not rescue the reduced migratory propensity of FD neural crest precursors. In contrast to the Choi et al. screen on HLCs, where HLCs were differentiated in standard ways and then seeded into 384 well plates in a straight-forward
manner, Lee et al. employed cell sorting and freezing of batches, which will likely prove a critical strategy to translating other hiPSC derived cell types to HTP screens.

In addition to hiPSC based disease modeling, hPSCs have also been applied in HTP studies to screen for drugs that selectively target CSCs (CSCs are reviewed in (Dick 2009)). Sachlos et al., using a neoplastic hPSC line that was previously characterized as having features of CSCs (Werbowsk-Ogilvie, Bosse et al. 2009), performing a comparative screen of normal and neoplastic hPSC lines (Sachlos, Risueno et al. 2012). The authors reasoned that compounds that induce differentiation in the neoplastic line and have little effect on the normal hPSC line would be ideal clinical candidates as they selectively target CSCs and have little effect on non-cancerous endogenous stem cell populations. These selective agents may target pathways "usurped" in CSCs and not active in normal adult stem cell populations. For both normal and neoplastic hPSC lines, 10000 cells per well were seeded into 384-well plates for 24 h, then compounds were added for 120 h, after which Oct4 expression was measured. The authors identify thioridazine as a hit, which targets dopamine receptors, and dopamine receptor expression was found in patient samples to be unregulated in cancer cells. Furthermore, thioridazine inhibits human acute myeloid leukemia (AML) cell lines and augments AraC, the standard chemotherapeutic currently used in treatments. Based on this finding, Thioridazine is now in a Phase 1 clinical trial for AML (https://clinicaltrials.gov/ct2/show/NCT02096289). The authors note the exciting possibility that dopamine receptor upregulation may be a more general marker of tissue-specific CSCs. Indeed, thioridazine has previously been shown to inhibit human breast cancer growth (Strobl and Peterson 1992). It should also be noted that the criteria of hits selectively targeting the neoplastic hPSCs only is quite stringent, and may rule out many effective drugs. For example, the authors found that rapamycin downregulates pluripotency in both normal and neoplastic hPSCs, so does not meet their selectivity criteria, however rapamycin is known to be antileukemic for AML (Recher, Beyne-Rauzy et al. 2005) and is also part of a clinical trial for AML (https://clinicaltrials.gov/ct2/show/NCT01822015).

1.2.5 Stem cells in safety pharmacology

In addition to lead discovery, stem cells also are expected to have a critical role in safety pharmacology. Adverse drug reactions (ADRs) have a tremendous cost, perhaps best illustrated by the withdrawal of several drugs from the market due to unexpected induced cardiac arrhythmias, which resulted in many deaths (Redfern, Carlsson et al. 2003).
Discovering such off-target effects at an early stage of drug development is highly desirable to drug developers as this offers the opportunity to improve clinical safety and reduce development costs. In the past decades, drug safety screening has employed a host of animal and in vitro models for developmental toxicity, cardiotoxicity, hepatotoxicity, and neurotoxicity (reviewed in (Sison-Young, Kia et al. 2012)), however these tests have also been animal-intensive, expensive, and critically the prevalence of drug-induced toxicity and still result in high patient morbidity and drug development attrition costs (Sison-Young, Kia et al. 2012), demonstrating that the current models, which though useful, do not fully predict toxicity, thus driving the need for more predictive and clinically relevant toxicology models. Factors implicated in the current failure of toxicology models include patient genetic polymorphisms, age, and gender (Sison-Young, Kia et al. 2012), hPSC models offer an opportunity to address these. Additionally, iPSC based models also uniquely offer the potential for disease specific toxicity modeling. Driven by this potential of hPSCs, pharmaceutical companies are presently investing in and collaborating with academic institutes to develop hPSC in vitro toxicology and safety pharmacology models (Prescott 2011), a trend that may increase in the upcoming years.

As the heart appears to be particularly sensitive to these off-target effects, there is a pressing need for more reliable cardiac safety pharmacology assays (Braam, Tertoolen et al. 2010). During early drug development, presently a commonly used assay for QT prolongation employs aneuploid cell lines with over-expressed ion channels, however these do not accurately model functional cardiomyocytes (Pouton and Haynes 2007). There has not been a readily available source of human heart cells that can be used in large-scale screening studies, however stem-cell derived heart cells are now seen as a promising cell source for such studies. Braam et al. have demonstrated that field potential duration (FPD) of hESC derived cardiomyocytes (hESC-CM) responds dose dependently to 12 test drugs, and that the responses are predictive of clinical effects (Braam, Tertoolen et al. 2010). Generating large quantities of cardiomyocytes that are predictive in high-throughput screens is challenging, however the immense cost savings provided by such assays will continue to drive the development of this stem cell technology.

Drug induced liver injury is presently the leading cause of patient morbidity and mortality from ADRs (Halegoua-De Marzio and Navarro 2008), however hPSC derived hepatic model
systems offer the possibility to address this. Primary culture of freshly isolated human hepatocytes is presently the gold standard in assaying drug induced liver injury, but these cells are limited in availability, suffer from great donor variability, dedifferentiate from the hepatocyte phenotype in culture and lose cytochrome P450 enzyme expression, and cannot be used in long-term studies (Kia, Sison et al. 2013). Immortalized cell lines have also been used but have insufficient metabolic activity (Guguen-Guillouzo and Guillouzo 2010). There has been a burst of studies deriving HLCs from hPSCs (extensively reviewed in (Kia, Sison et al. 2013)). No hPSC derived cells have completely phenocopied primary hepatocyte cultures, in fact these cells most closely resemble foetal liver cells (Wobus and Loser 2011), hence the term HLC. Standardizing suitable characterization of HLCs, in particular their functional capabilities such as hepatic enzyme activity, uptake of low density lipoprotein, is another focus as these functions in HLCs are known to vary tremendously (Kia, Sison et al. 2013). For example, the activity of cytochrome P450, family 3, subfamily A (CYP3a) in HLCs has been reported to vary from 0% to 90% of primary hepatocyte controls, which has been attributed to differences in hPSC cell lines, differentiation protocols, the methods to measure CYP3a activity, the quality of reference primary hepatocytes, variability in laboratories, and inconsistent definitions of differentiation endpoints (Wobus and Loser 2011). Part of the observed variance arises from variation in metabolic enzymes in the population, sufficiently large libraries of hiPSC derived HLCs could in fact leverage this to be a potent tool in toxicology screening (Kia, Sison et al. 2013). Optimizing differentiation, maturation, and scale-up of these cells are all major goals to be achieved in order to effectively apply HLCs in pharmaceutical HTP applications.

1.2.6 High-throughput stem cell screens

High-throughput (HTP) screening is conventionally recognized as the parallel testing of large chemical libraries, with at least partially automated data capture and analysis. This method of scientific experimentation has been used extensively in both biology and chemistry. In biology, HTP screening has allowed for the rapid identification of agents that modulate particular biological processes. The field of HTP stem cell screening is still nascent, especially screens in human cells. Even so, important developments have been made and in this section key HTP screens looking at stem cell self-renewal and survival will be reviewed, with a
particular emphasis on what can be gleaned from these studies for the future development of HTP assays based on hPSCs and derivatives.

1.2.6.1 MESC self renewal

MESC were initially cultured in vitro using fibroblast feeder layers and serum containing media, the same conditions previously used for mouse embryonic carcinoma cell culture (Evans and Kaufman 1981). Studies of the fibroblast conditioned media eventually led to the discovery that leukemia inhibitory factor (LIF), which activates STAT3, was a key factor promoting pluripotency of mESCs (Smith, Heath et al. 1988, Williams, Hilton et al. 1988). Subsequent studies revealed that BMP is an additional critical factor in serum that in combination with LIF maintains the pluripotent state (Ying, Nichols et al. 2003). Nearly 3 decades of mESC signaling research led to LIF/STAT3 activation by being considered necessary for self-renewal, however in 2008 Ying et al. demonstrated inhibiting MAPK and GSK3b is sufficient to maintain the pluripotent state, even in mESCs devoid of STAT3 (Ying, Wray et al. 2008). This work indicates that extrinsic stimuli are in fact dispensable for mESC self-renewal, and all that is required is the inhibition of endogenous factors, such as FGF/MAPK signaling.

Prior to the Ying et al. studies, Chen et al. performed a HTP screen of 50000 small molecules to look for a compound that maintains Oct4 expression in a reporter Oct4-GFP mESC line (Chen, Do et al. 2006). The screen was conducted in the absence of LIF, serum, or other cytokines in order to find a small molecule that could support mESCs in defined conditions. After secondary confirmation assays and structure-activity-relationship studies, a novel compound SC1, later named pluripotin, was discovered which could maintain mESCs for more than 10 passages in defined conditions without LIF, BMP, or WNT. Pluripotin was shown to maintain mESCs even when STAT3 activation is inhibited by the JAK2 inhibitor AG490, and with treatment of siRNA against STAT3, indicating that pluripotin acts independently of STAT3. Additionally, RT-PCR of BMP induced ID gene expression indicated that pluripotin does not act through the BMP pathway, and the WNT activity reporter was not activated by pluripotin. Affinity pull-down experiments revealed that the pluripotin binds to RasGAP and ERK1. Additional studies revealed that although it was previously known that ERK inhibition can maintain mESCs for approximately 3 passages, dual inhibition of both ERK1 and RasGAP allowed for long-term maintenance. In confirmation of this,
mESCs stably expressing short hairpin RNA (shRNA) against RasGAP can be maintained long term with the MEK inhibitor PD098059. Ectopic expression of ERK1 induced differentiation, and this was partially rescued with pluripotin. The work of Chen et al. highlighted for the first time that mESCs may have the intrinsic ability to self-renew if differentiation factors are inhibited. Despite this, pluripotin was not adapted widely, unlike the “2i” MAPK inhibitor and GSK3b inhibitor cocktail of Ying et al. (Ying, Wray et al. 2008). This is perhaps reflective of the “black-box” nature of novel small molecules with insufficient mechanistic information, which perhaps are inadequately characterized for the stem cell community to adopt. This seminal work of Chen et al. also highlights the common necessity of modulating multiple targets in order to achieve a desired biological outcome.

1.2.6.2 HPSC survival

In contrast to mPSCs, hPSCs are passaged as colonies and after dissociation have a survival rate of less than 1% (Thomson, Itskovitz-Eldor et al. 1998, Reubinoff, Pera et al. 2000, Pyle, Lock et al. 2006). This limits the use of hPSCs in single cell culture, which is important for efficient expansion in suspension culture systems and also for subcloning after, for example, genetic transfection. To overcome this, Watanabe et al. searched for single-cell survival factors and found that the Rho-associated kinase (ROCK) inhibitor Y-27632 increases cloning survival to ~27% from ~ 1%, enabling subcloning and serum-free suspension culture (Watanabe, Ueno et al. 2007). Y-27632 was found by screening a small set of compounds, including caspase inhibitors, growth factors, trophic factors and kinase inhibitors (data on the initial library were not shown). The survival assay consisted of 1 hour pre-treatment with the test compound, used at the commonly used concentration of 10 µM, dissociating the hPSCs, and then plating the cells onto MEF feeder layers at 500 cells/well in 96-well plates in maintenance media containing the test compound. At day 6, cells were stained for alkaline phosphatase (ALP), and it was found that Y-27632 treated wells had many large colonies that were ALP+ (26.6%+/−4%) whereas untreated controls had few ALP colonies (1.0+/−0.4%). Further secondary assays confirmed that hESCs grown with Y-27632 were E-cadherin+, Oct4+, and SSEA4+. Survival was even increased in feeder-free conditions (MEF-CM and Matrigel, data not shown), however it is unclear if the survival increase is as pronounced as on feeders. Additionally, sorting single hESCs into each well of a 96-well plate, onto MEF feeders in the presence of Y-27632 for 7 d further confirmed the pro-survival effect of Y-
27632. It is still unclear if the pro-survival effect is by purely increased cell-cell contact via increased cell adhesion and mobility, as even in this clonal assay there are still MEFs present which may provide pro-survival cell-cell contact. Replicating this experiment in feeder-free conditions would address this issue.

Several groups have performed HTP screens of compounds in order to search for additional, and perhaps more effective, small molecules that can improve hPSC survival upon disassociation. Damoiseaux et al. performed a HTP screen of 1620 compounds from the Biomol enzyme and lipid libraries (http://www.biomol.com/1/) and the Prestwick library containing FDA approved drugs (Damoiseaux, Sherman et al. 2009). HESCs were disassociated and 5000 cells per well were seeded into 384 well plates containing MEF feeders, and then compounds were then added to the plates. After 4 days, cells were immunostained with Oct4 and Hoechst, and the number of Oct4+ cells were quantified using high-content analysis (HCA). A Z-score > .5 was achieved, and several compounds were found to improve hESC survival by more than 3 standard deviations from the control. Several small molecules that inhibit ROCK or Protein kinase C (PKC) were hits, and several compounds that modulate calcium flux were also detected. This study was one of the first HTP screens using hPSCs, however it is unclear if any of these compounds are an improvement over Y-27632, or if any can act additively with Y-27632 to further improve survival.

Barbaric et al. performed a similar hPSC survival screen, however multiple parameters were used to phenotype hPSCs including number of cells in a colony, colony area, colony area, nuclear intensity, number of total colonies per well, and the percentage of cells in the colony which express TRA-1-60, a marker of pluripotency (Barbaric, Gokhale et al. 2010). Compounds were screened from a Protein Kinase Inhibitor library (80 compounds, at 10 uM) and the Prestwick library (960 compounds, 2.5 ug/ml or ~5 uM). HPSCs were dissociated and 6000 cells per well were seeded into 96-well plates on feeders in the presence of compounds, and after 5 days the plates were immunostained for TRA-1-60 and Hoechst. 17 compounds were found to promote differentiation, and 5 compounds promoted survival. Pinacidil was a main hit of the assay, however this compound and the other survival compounds mainly targeted similar kinases as Y-27632, such as PRK2, ROCK, MNK1, RSK1 and MSK1 kinases. No compounds were found to result in as high survival as Y-27632.
Xu et al. also performed a hESC survival screen, testing an in-house library of 50,000 heterocycles (Xu, Zhu et al. 2010). Follow-up studies indicate that trypsin treatment results in hypercontraction and reduced cell attachment, however downregulating Rho-ROCK signaling can reduce this hypercontraction and increase cell-ECM adhesion. The trypsinable hESC line HUES9 was first plated at 4,000 cells per well into Matrigel coated 384-well plates, the cells were allowed to settle for an hour, and then compounds were added to each well at a 2 uM final concentration. After 6 days of incubation, including a media and compound exchange on day 3, cells were stained for alkaline phosphatase (ALP) and examined also for the characteristic compacted colony morphology. Follow-up studies were performed on two lead hits, a 2,4-disubstituted thiazole named Thiazovivin (Tzv) and a 2,4-disubstituted pyrimidine named Tyrntegin (Ptn), both of which enhance single cell survival of hESCs more than 30-fold.

These follow-up studies focused on the critical role of cell-cell adhesion and cell-ECM adhesion in hESC survival and maintenance. Disassociating hESCs disrupts E-cadherin signaling, and hESCs do not recover E-cadherin expression after trypsin treatment, in contrast to mESCs. By plating single hESCs onto plates coated with a E-cadherin-Fc chimera protein, it was revealed that hESC survival significantly increased. This assay was not clonal, so it is unclear if this increase is due mainly to increased cell-cell contacts. Interestingly, it appears that cell-cell contact increases cell-ECM adhesion, as cell plating efficiency with higher cell density was 5-fold higher than with low cell density. However, low density cells plated similarly as high density cells on the E-cadherin surface plates. Rho activity was also reduced on the E-cadherin plates, indicating that E-cadherin regulates Rho activity in hESCs. Tzv enhances cell adhesion by blocking E-cadherin internalization, likely via inhibition of endocytosis. Affinity pull-downs and in vitro Rho kinase assays revealed ROCK is a direct target of Tzv, and 2 µM of Tzv protects hESCs at a similar level as 10 uM Y-27632. Ptn has no effect on ROCK, and currently the mechanism is unknown. Lastly, Tzv and Ptn were shown to enhance phosphorylation of multiple growth factor receptors, such as FGFRs, IGFRs, and EGFRs, as well as downstream kinases such as PI3K, AKT and MAPK. Inhibition of these receptors reduces AKT phosphorylation and the survival effect of Ptn, indicating that integrin regulation is closely linked to signal transduction and both synergize to promote hPSC survival.

1.2.6.3 HPSC self renewal
HPSCs were initially cultured using feeder layers of mouse embryonic fibroblasts (MEFs) and exogenous FGF2 (Thomson, Itskovitz-Eldor et al. 1998), and subsequently cultured in feeder-free conditions using MEF conditioned medium (CM) and Matrigel or fibronectin substrates (Xu, Inokuma et al. 2001, Amit, Shariki et al. 2004, Amit and Itskovitz-Eldor 2006). The use of serum and feeders introduces great variability into the culture process, making large-scale long-term hESC culture challenging. Also, undefined components may complicate signaling studies and skew differentiation propensities. To address these concerns, several groups have established feeder-free and defined culture conditions (Beattie, Lopez et al. 2005, Li, Powell et al. 2005, Pebay, Wong et al. 2005, Vallier, Alexander et al. 2005, Wang, Zhang et al. 2005, Xu, Rosler et al. 2005, Levenstein, Ludvig et al. 2006, Liu, Song et al. 2006, Lu, Hou et al. 2006, Xiao, Yuan et al. 2006, Yao, Chen et al. 2006) and commercial formulations are available such as mTeSR™1 (Ludwig, Bergendahl et al. 2006, Ludwig, Levenstein et al. 2006), StemPro® hESC SFM (Wang, Schulz et al. 2007), and TeSR™-E8™ (Chen, Gulbranson et al. 2011). In order to scale up hPSC culture for clinical and screening applications, it is desirable to replace growth factors, which can be costly and have variable biological activity, with small molecules.

Desbordes et al. in 2008 performed the first HTP screen with hPSCs to identify small molecules regulating self-renewal and early differentiation (Desbordes, Placantonakis et al. 2008). HESCs were seeded onto Matrigel coated 384-well plates in CM (6000 per well), and on day 2 media was replaced with Knockout Serum Replacement (KOSR) based media containing the compounds at 10 µM, in 50 µl well volumes. On day 7 the plates were fixed and stained for Oct4 and Hoechst. Oct4 intensity, normalized to Hoechst, was used as the assay readout. FGF2 (200 ng/ml) and BMP4 (200 ng/ml) were used as pluripotent and differentiated controls, respectively, and no cytokine addition resulted in an intermediate state, allowing detection of compounds promoting both self-renewal and differentiation. The high FGF2 and BMP4 concentrations are likely necessary due to the long assay duration and cytokine degradation over time. Hits were defined as compounds which exhibit greater than 30% increase or decrease in normalized Oct4 intensity in at least two independent runs each performed in duplicate. Using this assay, 2880 small molecules were screened and out of 89 initial hits, secondary assays reduced this list to four factors promoting maintenance and ten compounds inducing differentiation. Microarray studies indicate that none of the
differentiation inducing factors result in lineage specific differentiation, so the utility of these compounds is uncertain. None of the four maintenance factors, or factor combinations, were able to prevent spontaneous differentiation after long-term culture in feeder-free conditions. Nonetheless, this watershed work demonstrated how hPSCs can be adapted for HTP and high-content screening, and the detailed protocol of this work (Desbordes and Studer 2013) is a valuable tool to translate these methods to the broader hPSC community.

The assay of Desbordes et al. was adapted by Gonzalez et al. to screen the mammalian secretome for hPSC regulators (Gonzalez, Jennings et al. 2010). Using a large-scale, HTP protein expression, purification, and screening platform, a library of 806 purified secreted proteins was screened for their ability to maintain Oct4 in hPSCs. HPSCs were plated onto Matrigel coated 384-well plates at 2000 cells per well, and were seeded in unconditioned media (UM) without cytokines in order to sensitize the assay and allow detection of pluripotency factors. After 1 day compounds were added to the wells, and after a 7-day incubation hESCs were fixed and stained for Oct4 and Hoechst. Three proteins were classified as hits, transmembrane emp24 protein-transport domain-containing 1, Epstein–Barr virus-induced 3, and pigment epithelium-derived factor (PEDF). Long term culture validation studies confirmed that PEDF maintains hESC pluripotency in the absence of FGF2, TGFβ, or Activin. PEDF appears to act through similar pathways as FGF2, it is unclear what the differences are, or what benefits PEDF has over FGF2. TMED1 and EBI3 were not able to maintain hPSCs long term, indicating a 66% false positive rate from the primary screen. TGFβ, heregulin, PDGF-A, and PDGF-B have previously been characterized as promoting pluripotency and were included in the library but were not detected as hits, indicating a high false negative rate. Unlike the Desbordes assay, differentiation inducing compounds could not be detected, as control conditions had very low %Oct4+. This difference is due to the reduced basal %Oct4+, caused by the reduced seeding density (one third of that used by Desbordes et al.) and seeding in UM, and this makes the assay more sensitive to detecting factors that support pluripotency, at the expense of detecting differentiation factors.

The above studies screen one factor at a time, and at one concentration. In order to achieve pure small molecule based hPSC expansion, Tsutsui et al. proposed that it is likely that several small molecules working in conjunction will be required, however these would need to be carefully dosed, creating a large experimental space to explore (Tsutsui, Valamehr et al. 2011).
To address this, Tsutsui et al. implemented a feedback control scheme, which they have previously used to search for optimal small molecule combinations to inhibit virus infection (Wong, Yu et al. 2008). First, five small molecules were chosen based on a literature search: ROCK inhibitor Y27632, glycogen synthase kinase 3 beta (GSK3β) inhibitor CHIR99021, mitogen-activated protein kinase (MEK) inhibitor PD0325901, FGF receptor tyrosine kinase (FGFR) inhibitor PD173074, and another MEK inhibitor PD98059. Each small molecule was considered at 6 concentrations, based on the "base" concentration found in the literature: 0, base/10, base/√10, base, base x √10, and base x 10. A differential evolution algorithm to optimize the concentrations as follow: N =32 combinations of small molecules were generated at random, and for each combination a "mutation" was also generated using a formula. The original and mutation combinations were then "crossed" to obtain N trial combinations. The original and trial combinations were then experimentally compared to each other by measuring the resulting number of alkaline phosphatase positive (AP+) colonies, and the combination with higher AP+ count was selected to carry over to the next generation. After the initial run, the authors performed three additional iterations of this process. After finding that the AP+ readout optimized for cell proliferation rather than maintenance, the authors repeated the experiment with percentage of OCT4-GFP expressing cells as the readout and N=16. Four combinations were selected for long-term trials, although all performed well at the first passages, none could maintain the number of Oct4+ cells at passage 3. The authors retested select combinations with added FGF2 and found that a combination of 100 µM ROCK inhibitor Y27632, 0.95 µM GSK3β inhibitor CHIR99021, and 0.4 µM MEK inhibitor PD0325901, with 10 ng per ml FGF2 maintains hPSCs (the H1 cell line) long-term on fibronectin. Interestingly 32 µM Y27632, 3 µM CHIR99021, and 0.4 µM PD0325901 were found to maintain the HSF1 cell line for 25 passages without FGF2, however H1 could not be supported in these conditions. This study demonstrates the integration of search algorithms with HTP screens, and highlights the importance of looking at combinations of compounds and the importance of proper dosing of these compounds. To achieve robust small molecule based expansion of hPSC lines using this approach a key barrier will be identification of the initial small set of small molecules. In the Tsutsui et al. study no modulators of TGFβ were included, and four out of the five compounds examined are highly associated with hPSC differentiation and not maintenance. Also, including two inhibitors of MEK and an FGFR inhibitor likely dramatically reduced the diversity of signaling perturbations in the experimental space. Why
FGF2 is needed in the presence of MEK inhibitor is unresolved, one possible explanation is that FGF2 is necessary for adhesion after passaging (Eiselleova, Matulka et al. 2009), independent of receptor activation.

In addition to optimizing the soluble factors needed to scale up hPSC culture, HTP approaches have also been used to optimize the substrate hPSCs adhere to. Mei et al. screened 496 different polymers, which were prepared by copolymerizing 16 "major" monomers and six "minor" monomers at different ratios (Mei, Saha et al. 2010). These polymers were spotted onto glass slides to create combinatorial polymer arrays, coated with FBS, and hPSCs with a GFP reporter for Oct4 were sorted onto the spots at low density in MEF-CM, with ROCK inhibitor added for the first 24h. Seven days after culture, the number of Oct4+ cells was quantified to obtain "hit" polymers that supported clonal expansion of hPSCs. Additionally, the polymer library was characterized (using HTP techniques) for surface roughness, indentation elastic modulus, and surface wettability, allowing links between these surface properties and hPSC clonal survival to be quantified. FBS or human serum was needed to support hPSCs on the polymers, and this is attributed to vitronectin which is present at high concentrations in serum. The clonal survival rates seen did not surpass that seen ROCK inhibitor is used in other culture conditions. A key finding from this study is that adsorbed proteins (from initial protein coatings, culture media, or secreted from hPSCs) requires specific polymer surface chemistry to promote hPSC clonal growth, thus modifying the surface chemistry of polymer surfaces is a desirable way to regulate hPSC expansion. The authors subsequently screened 1152 forms of poly (lactide-co-glycolide) (PLGA) and looked at biomaterial interactions with human mesenchymal stem cell (hMSC), a neural stem cell line (NSC), and primary articular chondrocytes, demonstrating that this approach can be generalized to diverse cell types (Anderson, Putnam et al. 2005).

Despite intense interest in performing siRNA screens with hPSCs, this has been technically challenging to achieve to date. Chia et al. have performed a genome-wide siRNA screen, targeting 21,121 genes, to identify regulators of self-renewal (Chia, Chan et al. 2010). First, siRNAs (siGenome, Dharmacon) were printed into Matrigel coated 384-well plates, and these plates were frozen before use. Typically, a pool of siRNAs were used to target the same gene, as this reduces off-target effects. Before transfection, Dharmafect1 (Dharmacon) and OptiMEM (Invitrogen) were added to each well and incubated for 20 minutes. Next, 3,000
cells in 40 ul of CM with ROCK inhibitor were seeded into each well. After 4 days of transfection, cells were fixed and stained with Hoechst. The H1 POU5F1–GFP reporter line was used, and the GFP fluorescence reduction (F_{av}) and nuclei number reduction were quantified for each well. Wells were screened in duplicate, and the average z-score of both runs was then used to rank each gene. A threshold of 2 was set on the F_{av} z-score to determined hits, resulting in 566 genes (the top 2.7%). Based on ranking and functional relevance, 200 of the 566 candidate genes were selected for validation in a battery of secondary screens, including screens based on Oct4-GFP and Oct4 and Nanog reduction as measured by immunocytochemistry. One hundred twenty-seven genes were validated in all three secondary screens, or 22.4% of the initial 566 hits. After further validation screens performed on HES2 and HES3 cell lines, ultimately 93 genes were found to downregulate Oct4 (16.4% of the 566 hits) and 54 genes (9.5% of the 566 hits) that downregulated NANOG across all cell lines. The lead hit from this screen, PRDM14 (the 10th ranked hit), directly regulates pluripotency genes and PRDM14 is integrated into the core Oct4/Sox2/Nanog transcription factor network. PRDM14 in conjunction with Oct4, Sox2, and KLF4 (OSK) was also shown to increase reprogramming efficiency of human fibroblasts by 6-fold relative to OSK alone, as shown by iPSC colony counts. In line with previous work, knockdown of Prdm14 in mouse ESCs results in no change in Oct4 and Sox2, and mouse EpiSCs are deficient in PRDM14, highlighting the importance of human based PSC. It is important to note that PRDM14 was previously identified as being essential for the maintenance of hESC self-renewal (Tsuneyoshi, Sumi et al. 2008). The z-score threshold for hits is important as too high or low a threshold can result in high false-negative and false-positive rates, respectively. Oct4 was ranked first, however using the z-score of greater than 2 threshold, Nanog (873rd rank) and Sox2 (6697th rank) were not detected as hits. Pathway analysis was performed using the 566 hits, to determine which pathways were enriched, however given the high false positive rate of this list the utility of such analysis remains to be seen. It is interesting that the lowest z-scores indicate genes which when silenced increase pluripotency, however the leads resulting from this readout were not discussed, and likely conducting the screen in highly supportive conditions (CM and Matrigel) likely preclude finding genes that even further enhance Oct4. Retesting the factors with lowest z-scores in less supportive conditions may reveal genes that when inhibited upregulate pluripotency.
1.2.6.4 Discussion

Induced and embryonic PSC technology has opened up exciting avenues for disease modeling, regenerative medicine, and drug development. Preliminary studies in the hiPSC disease modeling field have already shown the utility of hiPSC technology in modeling of numerous and diverse diseases, and it is anticipated that hPSCs will revolutionize lead discovery for disease treatments. Towards this goal, we have herein reviewed PSC based HTP studies, with an emphasis on extracting pertinent lessons learned from initial pioneering work that can be applied towards improving future HTP programs.

Beyond the need for more efficient means of scale-up of cell generation, the key limitation to date has been the inability to effectively mimic adult somatic cells, an obstacle perhaps best studied in HLCs. Research into maturation of hPSC derivatives is progressing, and more standardized and thorough characterization of generated cell types will enable better comparison of studies, especially as the differentiation protocols used are evolving rapidly. Additionally, although in some cases adapting differentiated cell types to HTP screens may be as simple as generating a single cell suspension and seeding into multi-well plates, implementing quality control steps (such as FACS analysis and/or sorting) and freezing large batches of cells may address the commonly observed run-run and plate-plate variance issues. Detailed follow-up studies are needed, ideally several in a scalable format, in particular to evaluate effects of lead compounds on bona fide adult somatic cells to characterize what (not if) differences exist between these and hiPSC disease model generated cell types. For secondary studies, in particular for iPSC disease model based screens evaluating leads across multiple patients and clones is needed, as high variance has been observed across lines. A common observation in long-term follow-up is that the observed effect of compound seen in routine culture diverges from the screen results after passaging, separate screens for effects on cells during passaging (adhesion effects, survival in suspension) and post-passaging may address this.

There has been a shortage of true leads discovered that are better than previously known drugs, scaling up of screens may address this. Industrial pharmaceutical screens can often assay 1 million compounds (Macarron, Banks et al. 2011), even the largest HTP hPSC based screens to date are in the range of ~10000's. In addition to screening large libraries of single factors there is an indication that screening combinations of fewer factors may be a powerful
approach, in particular if small-molecule based control of cell fate is desired. Screening combinations of soluble and extracellular matrix components may also be of utility, as these culture components are well known to interact (Flaim, Teng et al. 2008). Combinatorial screening in conjunction with feedback control schemes could be powerful, however presently the major bottleneck is identifying the select core small molecules to include in such an approach. IDE1 and IDE2, which activate TGFβ signaling (through unknown mechanisms) and improve definitive endoderm induction (Borowiak, Maehr et al. 2009), may be a useful compound in a small molecule cocktail to expand hPSCs. Similarly, Tzv and Ptn (Borowiak, Maehr et al. 2009), which enhance phosphorylation of FGFRs, IGFRs, EGFRs, PI3K, AKT and MAPK would also be intriguing to screen in combination with other small molecules.

As an alternative to larger and more elaborately designed screens, smaller more focused screens (for example with factors obtained from developmental biology literature) can be extremely effective. This approach is exemplified by Watanabe et al. (Watanabe, Ueno et al. 2007) where screening a select set of caspase inhibitors, growth factors, trophic factors and kinase inhibitors resulted in characterization of Y-27632, which is in widespread use today. Notably several subsequent approaches looking at much larger libraries (up to 50000 compounds (Xu, Zhu et al. 2010)) did not yield a lead to displace Y-27632 in routine single-cell hPSC applications. As stem cell based screens ramp up, keeping in mind the importance of library selection and leveraging prior knowledge will maximize the benefits obtained from such larger scales.

Studies on neoplastic hPSCs (Werbowetski-Ogilvie, Bosse et al. 2009) have demonstrated the power of comparative screens, standardized repositories of screening data would allow for a more wide-spread use of this approach. High-content approaches involve local storage of large image sets, sharing of these repositories could allow for independent post-publication analysis to extract new biological insights. For example, familial dilated cardiomyopathy (DCM) hiPSC-derived cardiomyocytes have a disease phenotype of disturbed sarcomeric organization, with punctate sarcomeric alpha-actinin staining and more scattered Z-lines (Sun, Yazawa et al. 2012); this disease phenotype has anecdotally been reported to be more frequent in cells at the edges of cardiomyocyte clusters than in the interior of such clusters. More broadly available high-content image sets would enable re-analysis to discover such spatial trends.
Given the large parameter space to optimize (seeding media, screen media, cell density, assay duration, assay outputs, input population, etc.) the use of more efficient optimization tools would be beneficial. For example, Design of Experiments was recently used to optimize hPSC culture media (Marinho, Chailangkarn et al. 2015), this approach may be extended to applications of larger screens (this initial study screened 32 media conditions at each of 3 rounds). Cell seeding density and seeding media are routinely optimized for screens, one underappreciated aspect is that this parameter can be fine-tuned based on the desired screen outcome. Desbordes et al. (Desbordes, Placantonakis et al. 2008) seeded 6000 hESCs per well in 384 well plates in MEF-CM, and targeted both postive and negative regulators of pluripotency, whereas Gonzalez et al. (Gonzalez, Jennings et al. 2010) seeded 2000 hESCs per well in 384 well plates in UM, and targeted only positive regulators of pluripotency.

The vast majority of screens, even the image-based screens, employ single assay output to classify compounds. Given the population heterogeneity and spatial complexity hPSCs and derivatives, multivariate outputs can be expected to improve classification of phenotypes resulting from compounds. For example Barbaric et al. hPSC survival screen looked at cell-level (nuclear intensity and expression of TRA-1-60), colony level (number of cells per colony, area, percentage of cells expressing TRA-1-60 in a colony), well-level (total colonies per well) outputs (Barbaric, Gokhale et al. 2010). Hit selection and statistical analysis of multivariate screen outputs also need further refinement. As an alternative to fold difference threshold, more elaborate hit criteria such as rank-based methods (Lee, Ramirez et al. 2012) appear to be beneficial in overcoming plate-plate variance in screens. A more widespread use of blind screens may also aid in analysis and interpretation of screen results, especially as the hit determination protocols used vary tremendously and are typically done post-hoc. Z’ score analysis is reported variably - usually only reported for the screen, occasionally reported per plate, and occasionally not reported at all. Plate level Z’ scores have emerged as being particularly variable in hPSC and derivative screens, increased reporting and optimizing of this value during screen design may aid in improving hit quality. Additionally, as there are rarely gold-standards to compare screen results with, to date reporting of false discovery rates has been non-existent in screens from hPSCs and derivatives, severely limiting evaluation of screen efficacy. As more published screens become available, more active comparison of studies may partially address this. Clearly a more standardized and less ad-hoc approach to
screen analysis is needed, and this may help address the poor translation of screen leads to routine stem cell culture and differentiation. Improved quality of screen readouts will also enable more systems biology approaches to analyzing screen data, for example pathway enrichment and integrating screen results with gene regulatory network data, which has been notably scarce in hPSC based screens to date. As stem cell biology and HTP screening increasingly converge, applying the lessons learned from the pioneering efforts at applying hPSCs and derivatives towards HTP screens promises to accelerate the delivery of the much heralded benefits of stem cell technology to the clinic.

1.3 Conclusions and future perspectives

In order to reap the immense rewards offered by PSCs there will need to be improved understanding and control of the complex molecular mechanisms underpinning PSC fate decisions. Numerous exogenous and endogenous factors interplay to control stem cell fate, and these factors can often obfuscate interpretation of studies. Engineering the stem cell microenvironment, in particular the control of cell spatial distribution, has emerged as a powerful strategy to mitigate these factors and improve assay variation and interpretability. Multivariate single cell assays, in particular image-based high-content analysis, is an attractive means to track the multiple emergent subpopulations inherent in stem cell based assays while maintaining spatial context of these cell populations. The burst of new hiPSC cell lines and disease models has also come with immense challenges to optimize differentiation, maturation, and scale up of countless clinically relevant cell types. To meet the scale of these complex challenges in stem cell biology, HTP screening is a powerful tool, but first present limitations must be understood and addressed.

In this study, we sought to use micro-technology approaches to engineer the stem cell niche in HTP applications, and demonstrate how this approach can be used to improve the understanding of mechanisms of stem cell fate regulation. In Chapter 2 we develop a 96-well plate microcontact printing based HTP platform (96µCP) to screen hPSCs in configurable microenvironments, which we have fine-tuned for robust assay response by controlling hPSC cell distribution, and in which we simultaneously measure pluripotent, neuroectoderm, primitive streak, and extraembryonic fated cell populations in response to applied cues. We apply this platform to characterize hPSC dose responses to 27 factors and we also develop an
assay to quantify ("fingerprint") cell line lineage induction bias, which we apply to 21 hPSC cell line samples. In Chapter 3 we apply the platform to screen a 400 compound kinase inhibitor library, and we discover mTOR inhibition enhances mesendoderm induction, and we apply this to improve purity and yield of hemogenic endothelium 3 fold, with a concomitant enhancement of blood colony forming cells (CFCs). In Chapter 4 we develop a protocol for the 96µCP assay and validate it in two additional member labs of the International Stem Cell Initiative (ISCI) consortium, and we also develop ContextExplorer, open-source GUI-based software for spatial analysis of HTP data. In Chapter 5 we discuss the insights provided by spatial control of hPSCs in HTP formats, and the translation of these tools to the broader stem cell community. We additionally discuss new approaches to interrogating and controlling stem cell fate. Overall, this body of work demonstrates the opportunities enabled by the convergence of micro-technologies, HTP screening, and stem cell biology. Striving to recognize the various factors in play that dictate cell response, and implementing strategies to control and monitor these factors promises to accelerate the delivery of the much heralded benefits of stem cell technologies to the clinic.
This chapter has been published in *Nature Methods* (Nazareth, Ostblom et al. 2013). The publication can be found at http://www.nature.com/nmth. Co-authors include Joel E.E. Ostblom, Petra B. Lücker, Shreya Shukla, Manuel M. Alvarez, Steve K.W. Oh, Ting Yin, and Peter W. Zandstra. Authorization to reproduce this work has been obtained from the publisher and all co-authors.
2 High-throughput fingerprinting of human pluripotent stem cell cue-fate responses and lineage bias

2.1 Abstract

Populations of cells create local environments that lead to emergent heterogeneity. This is particularly evident in hPSCs where microenvironmental heterogeneity limits cell fate control. We have developed a high-throughput platform to screen hPSCs in configurable microenvironments, in which we optimized colony size, cell density, and additional parameters to achieve rapid and robust cell fate responses to cues. We used this platform to perform single-cell protein expression profiling, revealing that Oct4 and Sox2 co-staining discriminate pluripotent, neuroectoderm, primitive streak, and extraembryonic cell fates. We applied this code to analyze dose responses of 27 developmental factors to delineate lineage-specific concentration optima and to quantify cell-line endogenous signaling pathway activation and differentiation bias. Finally, we demonstrate that short-term responses predict definitive endoderm induction efficiency and can be used to rescue differentiation of cell-lines reticent to cardiac induction. These findings facilitate high-throughput hPSC-based screening and quantification of lineage induction bias.

2.2 Introduction

Human pluripotent stem cells (hPSCs) offer opportunities for drug development, understanding mechanisms of human cell development and cell-based therapies, but this requires a predictive understanding of factors that control cell fate. Although much progress has been made, there is still a need for improved reproducibility and deeper biological insight. For example, fibroblast growth factor (FGF) (Levenstein, Ludwig et al. 2006, Bendall, Stewart et al. 2007), activin A (Vallier, Alexander et al. 2005, Avery, Zafarana et al. 2010), leukemia inhibitory factor (LIF) (Daheron, Opitz et al. 2004, Amit, Laevsky et al. 2011) and Wnt (Sato, Meijer et al. 2004, Dravid, Ye et al. 2005) signaling have all been reported to both maintain pluripotency of hPSCs and to have no effect on maintenance. The International Stem Cell Initiative (ISCI) recently reported the first multi-laboratory comparative study of published defined culture systems for hPSC expansion (Akopian, Andrews et al. 2010): only two out of
eight tested media reproducibly maintained hPSCs across laboratories and cell lines, highlighting the need to identify and control key confounding factors.

Population context can have dramatic consequences on stem cell maintenance, differentiation and reprogramming. HPSCs exist in complex microenvironments containing multiple factors that regulate cell fate, including endogenous ligands and extracellular matrix proteins (Peerani, Rao et al. 2007), mechanical forces and cell-cell contact (McBeath, Pirone et al. 2004), and cell subpopulations (Bendall, Stewart et al. 2007, Peerani, Rao et al. 2007). Critically, these factors are spatially heterogeneous (Peerani, Rao et al. 2007), introducing substantial variances in cell response. Indeed, reprogramming to pluripotency (Maherali and Hochedlinger 2008), differentiation towards neural (Chambers, Fasano et al. 2009), pancreatic (Cai, Yu et al. 2009) and cardiac (Hwang, Chung et al. 2009) cell types, and the disease phenotype of a familial dilated cardiomyopathy (DCM) induced hPSC model (Sun, Yazawa et al. 2012) all have specific organizational and density dependent optima.

We have developed a suite of tools to test hPSC responses to exogenous factors and to categorize cell lines based on differentiation bias. The platform consists of a method to pattern cells in 96-well plates, an optimized cell fate factor screening assay protocol, defined media and substrate, and a single-cell imaging and data analysis pipeline. We optimized colony size, cell density, media composition and substrate to allow robust cell fate responses to exogenous cues to be measured in 48 h. Characterization of single cell protein expression across diverse induction conditions revealed that Oct4 and Sox2 co-staining can discriminate pluripotent, neuroectoderm, primitive-streak, and extraembryonic subpopulations, enabling fingerprinting of cell line response to lineage induction stimuli. These responses were stable between passages, variable between cell lines, predictive of lineage induction efficiency, and can be used to improve differentiation of reticent cell lines. Additionally, we applied this platform to characterize 27 developmental signaling factors selected by the ISCI consortium across a wide range of doses, revealing new dose- and lineage-specific optima.

2.3 Results

2.3.1 Microenvironmental control for hPSC screening
We previously developed an assay in which hPSC distribution (colony size, shape, and spacing) can be controlled by micro-contact printing (µCP) substrates onto slides. This work indicated that colony size control could be an important parameter in stem cell screens (Peerani, Rao et al. 2007). To adapt this technique for high-throughput (HTP) studies, we developed a method of µCP in which substrates are directly printed into 96-well plates in user-defined patterns (Figure 2-1) in defined serum-free media (SF) and substrate (a fibronectin and gelatin mixture, “FnGel”) (Figure 2-2). We dispense single cell suspensions into the wells, allow the cells to settle and adhere to the patterned substrate for 6 h, wash away non-adherent cells, add test factors to the wells for 42 h, and then fix and analyze the plates (Figure 2-3a). Each well contains an array of hPSC colonies, and we use automated microscopy to obtain single cell data such as x- and y-coordinates and protein expression levels (Figure 2-3b).
Figure 2- 1 Generation of PDMS stamps for 96-well plates

(a) Liquid PDMS is poured into a teflon mould that is placed on top of the silicone master. The teflon mould allows control of the PDMS stamp shape, enabling the generation of PDMS stamps with micropatterned posts, with each post fitting into a well of a standard 96-well plate. (b) Image of the teflon mould that enables the generation of stamps with 24 posts (4x6). (c) Image of a PDMS stamp that fits into a 96-well plate. (d) Image of a PDMS stamp patterning 24 wells (one quadrant) of a 96-well plate.
Development of defined and fully disclosed media and substrate. Our previously hPSC µCP assay was based on growth factor reduced Matrigel™ (MG) substrate, which is an undefined sarcoma extract, and XVIVO10™ (XV) based media, which is defined but has an undisclosed formulation. We sought to replace MG with a defined substrate, thus we tested a combination of fibronectin and gelatin (FnGel) that we have previously validated with mESCs. We also developed our own DMEM-F12 based serum free chemically defined media (SF), which has a simple and fully disclosed formulation. Based on 200 µm diameter colonies with 500 µm pitch (centre-to-center distance) design (200-500), and seeding cells in SF with supportive cytokines (FGF2, Activin A, and ROCK inhibitor, “FARi”), we tested the pluripotency response to CM, and either XV or SF alone, or with BMP4 (B), or heregulin-1β, activin A, and FGF2 (HAF). We tested these 7 conditions on both MG and FnGel. On FnGel patterns, compared to SF media alone we found a decrease in pluripotency with BMP4 (p<.005) and an increase with HAF (p<.005) or CM (p<.004). In contrast, we found no significant increase in pluripotency when HAF was added to XV, compared to XV alone. Across all media conditions we generally found similar responses on FnGel and MG. We therefore used SF base media and FnGel for further experiments.
Figure 2-3 The 96µCP platform

(a) Schematic of the assay design. Cell substrates are patterned into standard 96-well plates using micro-contact printing. A single cell suspension is seeded, cells are allowed to settle and adhere for 6 h, plates are washed, and test media is added. Plates are fixed 48 h after seeding and then stained. (b) Image of 25 stitched fields (viewed at 20x) showing an array of colonies at the bottom of a well in a 96-well plate. DAPI staining is used to obtain nuclear masks which enable single cell analysis. The single cell features can be used to reconstruct vectorized figures of the wells. Using the x- and y-coordinates, cells can be clustered into colonies to enable colony-level analysis. (c) Colony size optimization. FnGel was patterned in different diameter islands as indicated. HPSCs (H9 line) were plated and the indicated marker response to six control conditions were tested: murine embryonic fibroblast conditioned media (CM), base media alone (SF), SF with FGF2 and SB431542 (FTi), SF with BMP4 (B), or SF with BMP4 and activin A (BA). Patterns tested were 150 µm diameter (500 µm pitch), 200 µm diameter (500 µm pitch), 400 µm diameter (640 µm pitch), and 1200 µm diameter (1920 µm pitch). Media conditions and colony diameter are significant sources of variation, with significant interaction (Two-way ANOVA \( p < 5 \times 10^{-5} \)). * ANOVA \( p < 0.01 \) and Tukey post-hoc \( p < 0.01 \). Error bars, s.d. (n = 3).
Based on our current understanding of the relation between colony size and endogenous signaling, we reasoned that calibration of colony size would allow robust detection of both positive and negative regulators of pluripotency. To test this, we assayed the response of hPSC (H9 line) colonies of different sizes to six control conditions: murine embryonic fibroblast-conditioned media (CM), SF base media alone, and SF with fibroblast growth factor 2 (FGF2) and the transforming growth factor beta (TGFβ) inhibitor SB431542 (FTi), SF with bone morphogenetic protein 4 (BMP4, B), or SF with BMP4 and activin A (BA) (Figure 2-3c). We observed that the frequencies of cells expressing the pluripotency factors Oct4 and Sox2 were colony-size dependent, with cells in SF alone maintaining expression in 1200 µm colonies and dramatically losing it in 150 µm colonies. Relative to baseline SF conditions, hPSCs in 200 µm colonies were found to respond with the greatest dynamic range to positive (CM, ANOVA $p = 1.17 \times 10^{-5}$) and negative (B, ANOVA $p = 2.74 \times 10^{-4}$) regulators, indicating that this colony size was ideal for our test criteria. Based on the above results and additional optimization we determined a final assay configuration which we term the “96µCP” platform and which we used for all subsequent 96-well assays (Table 2-1).

<table>
<thead>
<tr>
<th>Assay Parameter</th>
<th>Configuration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input Population</td>
<td>Single cell suspension of hPSCs off MEFs</td>
</tr>
<tr>
<td>Pattern</td>
<td>200 um diameter, 500 um pitch</td>
</tr>
<tr>
<td>Substrate</td>
<td>Defined, combination of fibronectin and gelatin</td>
</tr>
<tr>
<td>Base Media</td>
<td>Defined, DMEM/F12 based</td>
</tr>
<tr>
<td>Seeding Media</td>
<td>Base Media + FGF2 + Activin A + ROCK i</td>
</tr>
<tr>
<td>Assay Duration</td>
<td>48 hours (6h seeding + 42h with test compound)</td>
</tr>
<tr>
<td>Seeding Density</td>
<td>100,000 cells per well in 100ul Seeding Media</td>
</tr>
<tr>
<td>Analysis</td>
<td>Single cell, multiple fluorophores</td>
</tr>
</tbody>
</table>

### 2.3.2 Oct4 and Sox2 expression discriminate early cell fates
We next sought to determine if early lineage specification could be quantified using the 96µCP platform. In addition to the SF baseline condition, we chose five previously characterized conditions: CM (maintains pluripotency (Xu, Inokuma et al. 2001)), heregulin-1β with activin A and FGF2 (“HAF”) (maintains pluripotency (Wang, Schulz et al. 2007)), B (induces trophectoderm and primitive endoderm (Xu, Chen et al. 2002, Pera, Andrade et al. 2004, Vallier, Touboul et al. 2009)), BA (induces primitive streak (Vallier, Touboul et al. 2009)) and FTi (induces neuroectoderm (Vallier, Reynolds et al. 2004, Smith, Vallier et al. 2008)). For all conditions, we assessed single-cell protein expression of a panel of early developmental lineage markers including pluripotency (Oct4, Sox2, Nanog, Tra-1-60), primitive streak (Brachyury, Gata4, Snail), neural (Pax6, Sox1, Sox3), extraembryonic (Sox7, CDX2, Hand1), endoderm (CXCR4, Foxa2) and mesoderm markers (Brachyury, Gata4) (select images shown in Figure 2-4a). Several later stage markers, typically arising more than 6 days post-induction, including Pax6, Sox1, Sox3, Sox7, CDX2, Hand1, and CXCR4, showed no difference in expression, indicating that at this very early test point (42 h) these markers are not differentially expressed, as expected.

Two dimensional hierarchical clustering of marker expression levels across the six control conditions (Figure 2-4b) confirmed that different induction conditions result in distinct protein expression profiles, as expected. Note that thresholds for positive expression were determined for each protein based on differential expression across conditions (Figure 2-4c). SF and FTi conditions clustered together, indicating that under basal conditions the hPSCs are largely neuroectoderm fated, in line with previous observations (Paige, Osugi et al. 2010). Oct4+Sox2+, previously shown to mark committed neural precursors in human (Vallier, Mendjan et al. 2009) and mouse (Lowell, Benchoua et al. 2006) development, was expressed exclusively in the neuroectoderm-inducing conditions FTi and SF. Oct4+Sox2− expression clustered separately from the other groups, which we interpreted as marking a non-pluripotent, non-neural, non-primitive streak population, likely fated towards extraembryonic tissue (trophectoderm and primitive endoderm). Primitive streak markers also clustered together, and intriguingly Oct4+Sox2− expression clustered with this group. We observed that, across all control conditions, Oct4+Sox2− is exclusively found in the BA condition, which induces primitive streak (Figure 2-4d-e), and is exclusively associated with high expression levels of Snail, Brachyury, and Gata4 (Figure 2-4f).
This data supports the use of Oct4 and Sox2 as a binary code to discriminate four major early cell fates in human development: Oct4⁺Sox2⁺ for pluripotency, Oct4⁻Sox2⁺ for early neuroectoderm, Oct4⁺Sox2⁻ for early primitive streak, and Oct4⁻Sox2⁻ for early extraembryonic committed and other tissues (Figure 2-4g). This classification is congruent with previous reports of Oct4 and Sox2 expression in these lineages (Xu, Chen et al. 2002, Boyer, Lee et al. 2005, Mossman, Sourris et al. 2005, Chng, Teo et al. 2010).
Figure 2-4 Single cell protein profiling reveals Oct4 and Sox2 mark early cell fates
(a) Sample images of hPSC (H9 line) colonies stained for the indicated markers after treatment with the six control conditions (abbreviations as in Fig 1). Scale bar 100 µm. (b) 2D hierarchical clustering of protein expression (%positive) of markers and sets of markers across control conditions. Left panel displays sample similarity tree, with samples clustered using a distance threshold of 0.4. Clusters are indicated in red numerals. (c) Sample FACS plots of Oct4 and Sox2 intensity values from six controls. Thresholds can be based on these controls to classify cells as positive or negative for each marker. Color indicates cell density. (d) Quantification of Oct4 and Sox2 subpopulations in control conditions. Statistically compared to the value in SF alone, * ANOVA p < 0.01 and Tukey post-hoc p < 0.01, # ANOVA p < 0.005 and Tukey post-hoc p < 0.05. Error bars, s.d. (n = 3). (e) Sample composite images showing DAPI (blue), Oct4 (green), and Sox2 (red) expression in controls. Scale bar 100 µm. (f) Quantification of the indicated markers in control conditions. * Statistically compared to the value in SF alone, ANOVA p < 0.01 and Tukey post-hoc p < 0.01. Error bars, s.d. (n = 3). (g) Summary of Oct4 and Sox2 subpopulations.

2.3.3 Analysis of early hPSC cell fate responses

We used the 96µCP platform and the Oct4 Sox2 code to simultaneously characterize early cell fates in response to 27 developmental factors. Responding factors (Figure 2-5a, non-responding factors in Figure 2-6) were classified as promoting pluripotency, neuroectoderm, primitive streak, extraembryonic/other, or having a bimodal effect (inducing different subpopulations at different concentrations). We have summarized the classification of these factors as well as a recommended concentration for use in chemically defined media in
Table 2- 2.

<table>
<thead>
<tr>
<th>Subpopulation (%)</th>
<th>Concentration (ng mL⁻¹ or µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pluripotent</td>
<td>Neural ectoderm</td>
</tr>
<tr>
<td>Primitive streak</td>
<td>Extraembryonic/other</td>
</tr>
</tbody>
</table>

![Graphs and images showing pluripotent, neural ectoderm, primitive streak, and extraembryonic/other subpopulations across various concentrations of different molecules.]

![Bar charts showing the percentage of neural ectoderm and pluripotent subpopulations with and without LDN-193189 at different concentrations.]
Figure 2- 5 Characterization of factors modulating cell fate choices

(a) Dose-response curves of developmental signaling factors and their modulators. Title color indicates predominant effect of each factor: promoting pluripotency (blue), neuroectoderm (red), extraembryonic (orange), or primitive streak (purple). The plots show the percentages of these subpopulations as measured by the Oct4 Sox2 code. Ligand concentration is shown in ng/mL, small molecule concentration is shown in µM. Error bars, s.d. (n > 2).  
(b) Dose curve of Alk2/3/6 inhibitor LDN-193189 with and without 40ng/mL BMP4. Statistically compared to “0” control values, * ANOVA p < 0.01 and Tukey post-hoc p < 0.01. Error bars, s.d. (n = 3).  
(c) The micrographs show the effect of LDN addition during stage two of endoderm differentiation. HPSCs (H9 line) are stained with DAPI and for Pdx1. Scale bar 100 µm. The percentage of Pdx1+ cells is quantified in the plot (right). *ANOVA p < 0.0001. Error bars, s.d. (n = 3).  
(d) The plot shows subpopulation profiles under the indicated conditions, designed to test the effect of DKK1. * ANOVA p < 0.01 and Tukey post-hoc p < 0.01. Error bars, s.d. (n = 3).

Figure 2- 6 Dose curves of non-responsive factors.

Ligand concentration is shown in ng mL⁻¹, small molecule concentration is shown in µM. Error bars, s.d. (n>2).
Table 2- 2 Factor Characterization

<table>
<thead>
<tr>
<th>Promotes</th>
<th>Factor</th>
<th>Recommended Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pluripotency</td>
<td>Activin A</td>
<td>10 ng/mL (0.8 nM)</td>
</tr>
<tr>
<td></td>
<td>FGF2 (Fibroblast growth factor 2)</td>
<td>80 ng/mL (5 nM)</td>
</tr>
<tr>
<td></td>
<td>HRGβ1 (Heregulin-β1)</td>
<td>80 ng/mL (10.6 nM)</td>
</tr>
<tr>
<td></td>
<td>IGF1 (Insulin-like growth factor 1)</td>
<td>10 ng/mL (1.3 nM)</td>
</tr>
<tr>
<td></td>
<td>LDN-193189 (“LDN”, ALK2/3/6 inhibitor)</td>
<td>2.5 µM</td>
</tr>
<tr>
<td></td>
<td>Noggin</td>
<td>10 ng/mL (0.4 nM)</td>
</tr>
<tr>
<td></td>
<td>TGFβ1 (Transforming growth factor-β1)</td>
<td>1 ng/mL (0.08 nM)</td>
</tr>
<tr>
<td>Neuroectoderm</td>
<td>DKK1 (Dickkopf Homolog 1 Protein)</td>
<td>125 ng/mL (3.1 nM)</td>
</tr>
<tr>
<td></td>
<td>LDN-193189 (“LDN”, ALK2/3/6 inhibitor)</td>
<td>2.5 µM</td>
</tr>
<tr>
<td></td>
<td>PD0325901 (“PD”, MEK inhibitor)</td>
<td>1 µM</td>
</tr>
<tr>
<td></td>
<td>SB431542 (ALK4/5/7 inhibitor)</td>
<td>10 µM</td>
</tr>
<tr>
<td>Primitive Streak</td>
<td>Activin A</td>
<td>10-100 ng/mL (0.8 - 8 nM)</td>
</tr>
<tr>
<td></td>
<td>BMP4 (Bone morphogenetic protein 4)</td>
<td>10-100 ng/mL (0.8 - 8 nM)</td>
</tr>
<tr>
<td>Extraembryonic/Other</td>
<td>BMP4 (Bone morphogenetic protein 4)</td>
<td>40 ng/mL (3.2 nM)</td>
</tr>
<tr>
<td></td>
<td>CHIR99021 (GSK3β inhibitor)</td>
<td>6 µM</td>
</tr>
<tr>
<td></td>
<td>JAK Inhibitor I (JAK1/2/3 inhibitor)</td>
<td>10 µM</td>
</tr>
<tr>
<td></td>
<td>LY294002 (PI3K inhibitor)</td>
<td>10 µM</td>
</tr>
<tr>
<td></td>
<td>SB203580 (p38 MAPK inhibitor)</td>
<td>0.3 µM</td>
</tr>
<tr>
<td>No Observed Action</td>
<td>BAFF (B-cell activating factor)</td>
<td>[0.1-100 ng/mL]</td>
</tr>
<tr>
<td>[Range Tested]</td>
<td>Betacellulin</td>
<td>[0.1-100 ng/mL]</td>
</tr>
<tr>
<td></td>
<td>FGF4 (Fibroblast growth factor 4)</td>
<td>[0.1-100 ng/mL]</td>
</tr>
<tr>
<td></td>
<td>FLT3L (FMS-like tyrosine kinase 3 ligand)</td>
<td>[0.1-100 ng/mL]</td>
</tr>
<tr>
<td></td>
<td>GDF3 (Growth differentiation factor-3)</td>
<td>[0.1-200 ng/mL]</td>
</tr>
<tr>
<td></td>
<td>LIF (human Leukemia inhibitory factor)</td>
<td>[0.0001-200 ng/mL]</td>
</tr>
<tr>
<td></td>
<td>PDGF (platelet-derived growth factor-AB)</td>
<td>[0.01-100ng/mL]</td>
</tr>
<tr>
<td></td>
<td>SCF (Stem cell factor)</td>
<td>[0.01-100 ng/mL]</td>
</tr>
<tr>
<td></td>
<td>VEGF (Vascular endothelial growth factor)</td>
<td>[0.01-100ng/mL]</td>
</tr>
<tr>
<td></td>
<td>Wnt3a [0.01-100ng/mL]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wnt5a [0.1-200 ng/mL]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Y27632 (“ROCKi”, ROCK inhibitor)</td>
<td>[0.1-100 µM]</td>
</tr>
</tbody>
</table>

FGF2 (Vallier, Alexander et al. 2005), TGFβ1 and activin A (James, Levine et al. 2005), heregulin-β1 (Wang, Schulz et al. 2007), IGF1 (Wang, Schulz et al. 2007), and noggin (Wang, Zhang et al. 2005) all promoted pluripotency, as predicted from the literature. In contrast, a previously published 7 day non-patterned hPSC-based screen of 806 human extracellular factors assaying for pluripotency regulators found that only FGF2 and pigment epithelium-
derived factor maintained pluripotency, with TGFβ1, activin A, heregulin-β1, IGF-1, noggin, and all other factors tested resulting in differentiation (Gonzalez, Jennings et al. 2010), highlighting the potential benefits of microenvironmental optimization. Also in agreement with previous observations, the TGFβ inhibitor SB431542 (Vallier, Alexander et al. 2005) and the MEK inhibitor PD032590 (Greber, Wu et al. 2010) reduced pluripotency and increased neuroectoderm, while BMP4 suppressed pluripotency and neuroectoderm (Xu, Chen et al. 2002). Activin A plus 40 ng mL⁻¹ BMP4 resulted in a dose dependent increase in primitive streak, in contrast to activin A alone, which yielded no primitive streak (Vallier, Touboul et al. 2009). LDN-193189 (LDN), an inhibitor of BMP receptors Alk2/3/6 previously uncharacterized in hPSCs, increased neuroectoderm. To determine an optimal dose to inhibit BMP signaling, we performed the LDN dose curve in the presence of BMP4 (Figure 2-5b). LDN increased neuroectoderm both with and without BMP4 in a dose dependent manner up to 10 µM. Interestingly, LDN rescued the pluripotency-suppressing effect of BMP at concentrations up to 2.5 µM, however at higher concentrations pluripotency decreased. Nostro et al. have also shown that during endoderm induction from hPSCs there is a cell-line specific need to inhibit BMP signaling (Nostro, Sarangi et al. 2011). We therefore added 2.5 µM LDN during stage 2 of the Nostro et al. protocol and confirmed a significant induction of later-stage endodermal PDX1⁺ pancreatic progenitor cells (p = 0.0006) (Figure 2-5c). These results indicate that findings from our HTP assay are congruent with results from traditional assays and can yield predictive signaling insights about later stages of differentiation.

It has previously been reported that DKK1 has little or no effect on hPSC pluripotency. However, these studies were performed on hPSCs cultured on fibroblast feeders, measuring alkaline-phosphatase colony formation after 26 days (Dravid, Ye et al. 2005). In our assay, DKK1 had a surprising effect of inhibiting pluripotency and enhancing neuroectoderm. We speculate that endogenous supporting factors produced by the feeders mask this effect of DKK1. Indeed, when we tested the response of H9-hPSCs to DKK1 in the presence of additional factors (FGF, activin A or BMP), we found that DKK1 no longer induces neuroectoderm (Figure 2-5d).

2.3.4 Quantification of cell-line differentiation propensities

It has long been observed that hPSCs have varying differentiation propensities (Osafune, Caron et al. 2008). We tested whether the 96µCP platform would allow a rapid evaluation of
cell line-specific differentiation tendencies by measuring fate responses to six conditions (SF, CM, B, BA, TiF, and SF with PD0325901) for 21 cell samples. Our test panel consisted of 15 hESC and hiPSC lines, including two in-house generated human induced pluripotent stem cell (hiPSC) lines, ZAN3i-85UCBT (“ZAN3”) and ZAN11i-85UCBT (“ZAN11”), an in-house generated karyotypically abnormal H7s (“H7*”) with trisomy 12, which is recurrent in hPSC cultures, differentiated H9s (“H9diff”), and multiple passages of H9 and ZAN11. We performed two dimensional hierarchical clustering on this data set (6 controls x 4 subpopulation measurements = 24 data points for each cell line), which enabled visualization of cell-line similarities (Figure 2- 7a). The H7* line expressed all pluripotency markers and had normal hPSC morphology but was the only line that failed to differentiate in response to BMP4, indicative of an acquired abnormality (Figure 2- 7b and Figure 2- 8). The three passages of H9 and ZAN11 cells clustered closely, indicating that the response profile is stable over multiple passages and that the assay is reproducible.
Figure 2-7 Quantitative assessment of cell line-specific endogenous signaling and differentiation
Hierarchical clustering of cell line responses to six conditions: SF, CM, B, BA, TiF, and SF with PD0325901 (Euclidean distance similarity metric, average linkage clustering). Left panel displays sample similarity tree, with samples clustered using a distance threshold of 1.4. Clusters are indicated in red numerals. The plots show the percentage of four indicated subpopulations based on the Oct4 Sox2 code, in response to the indicated stimulus conditions (abbreviations as in Fig 1). Error bars, s.d. (n = 3). The upper plot shows percent primitive streak (d 2) under the indicated conditions in multiple cell lines. ANOVA, p < 0.005. Error bars, s.d. (n = 3). The lower plot shows Troponin T (TnT) expression (d 18) with varying activin A concentrations. Activin A has differential effects on cardiac Troponin T (TnT) expression in ZAN3 and ZAN11 cell lines (p < 0.005, Two-Factor ANOVA with n = 2, independent passages), with Activin A increasing primitive streak in ZAN11s and decreasing primitive streak in ZAN3s (both >95 percent confidence using linear regression). Error bars, s.d. (n = 2). The upper plot shows definitive endoderm induction efficiencies of 12 cell lines. Error bars, s.d. (n = 2). The lower plot shows the correlation of cell line Mesendoderm prediction index values from day 2 to actual definitive endoderm induction efficiencies at day 5 for a panel of cell lines (r = 0.89, p < 0.0001). The left plot shows the response of the H9 line to signaling pathway agonists and antagonists. Error bars, s.d. (n = 3). The right plot shows the estimated endogenous signaling levels of specific pathways for H9 (Eqn. 1, Methods). Estimated endogenous signaling levels for ZAN3 and ZAN11 cell lines. Passage to passage correlation of endogenous signaling profiles. The 60 response outputs (4 readouts x 15 conditions) were obtained for multiple passages for ZAN11 and ZAN3 showing high passage-to-passage correlation (ZAN 11 r = 0.90, ZAN3 r = 0.88). Comparison of ZAN3 and ZAN11 shows low line-to-line correlation (r = 0.52).

We sought to determine if our 2-day factor response profiles could be used to improve differentiation protocols for specific cell lines. Despite the overall similar differentiation profile of ZAN3 and ZAN11 lines (cluster 5, Figure 2- 7a) and similar primitive streak induction frequency in response to BMP and activin A, we observed that when activin A is removed ZAN3s increased primitive streak induction, in contrast to ZAN11 and H9 cells. This trend was confirmed over multiple passages (ANOVA, p = 0.0045) (Figure 2- 7c). To further support the primitive streak commitment of these lineages and to associate this early cell fate with functional commitment, we compared cardiac differentiation efficiency of ZAN11 and ZAN3 in an 18 day protocol (Yang, Soonpaa et al. 2008). We varied activin A concentration during the first four days, as differentiation is known to be sensitive to activin A in a cell-line dependent manner during this period (Kattman, Witty et al. 2011). As predicted by the day 2
results (Figure 2-7c), day 18 cardiac Troponin T (TnT) expression varied with activin A ($p = 0.0045$, Two-Factor ANOVA with $n = 2$, independent passages), with additional activin A increasing cardiomyocyte (TnT$^+$) output in ZAN11 but decreasing cardiomyocyte output in ZAN3 (both $>95\%$ confidence using linear regression) (Figure 2-7c).

To further examine how predictive the day 2 96µCP response profiles are of differentiation over longer periods, we differentiated a panel of 12 hESC and hiPSC lines towards Foxa2$^+$Sox17$^+$ definitive endoderm using a 5 day induction protocol (Figure 2-7d and Figure 2-9a). Using the ratio of %Primitive streak induced in the BA control to %Extraembryonic in the CM condition (“Mesendoderm prediction index”), we found a significant correlation between this day 2 prediction index and cell line definitive endoderm induction efficiency at day 5 (correlation coefficient ($r = 0.89$, $p = 0.0001$) (Figure 2-7d and Figure 2-9a-b). In contrast, pluripotency of the input cell population was not correlated with definitive endoderm induction ($r = 0.29$, $p = 0.42$) (Figure 2-9c).

<table>
<thead>
<tr>
<th>Cell line</th>
<th>Foxa2$^+$Sox12$^+$ (%)</th>
<th>O$^5$' in BA (%)</th>
<th>O$^5$' in CM (%)</th>
<th>Mesendoderm prediction index</th>
<th>Cluster Number (from Figure 4b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA1</td>
<td>76</td>
<td>42.5</td>
<td>0.6</td>
<td>69.26</td>
<td>4</td>
</tr>
<tr>
<td>R306C #15</td>
<td>50</td>
<td>48.5</td>
<td>1.6</td>
<td>30.03</td>
<td>3</td>
</tr>
<tr>
<td>HES2</td>
<td>40</td>
<td>20.5</td>
<td>10.8</td>
<td>1.90</td>
<td>5</td>
</tr>
<tr>
<td>Runx1</td>
<td>38</td>
<td>65.8</td>
<td>9.5</td>
<td>6.95</td>
<td>2</td>
</tr>
<tr>
<td>FOX1</td>
<td>31</td>
<td>47.2</td>
<td>5.1</td>
<td>9.24</td>
<td>2</td>
</tr>
<tr>
<td>B11D</td>
<td>51</td>
<td>26.0</td>
<td>3.1</td>
<td>8.44</td>
<td>4</td>
</tr>
<tr>
<td>110</td>
<td>24</td>
<td>49.6</td>
<td>3.7</td>
<td>13.21</td>
<td>2</td>
</tr>
<tr>
<td>H1</td>
<td>19</td>
<td>24.1</td>
<td>14.2</td>
<td>1.70</td>
<td>2</td>
</tr>
<tr>
<td>B4HA</td>
<td>16</td>
<td>16.3</td>
<td>43.1</td>
<td>0.38</td>
<td>5</td>
</tr>
<tr>
<td>H9p37</td>
<td>16</td>
<td>6.1</td>
<td>11.5</td>
<td>0.53</td>
<td>5</td>
</tr>
<tr>
<td>RTT #3-4 #17</td>
<td>12</td>
<td>18.6</td>
<td>36.5</td>
<td>0.51</td>
<td>5</td>
</tr>
<tr>
<td>T15BM #5</td>
<td>6</td>
<td>0.3</td>
<td>7.3</td>
<td>0.04</td>
<td>6</td>
</tr>
</tbody>
</table>

**Figure 2-9** Definitive endoderm induction across a panel of cell lines
2.3.5 Quantification of cell line endogenous signaling

Our results thus far and several recent studies indicate that variability in endogenous signaling is one major source of cell line differentiation variability. Cell line and passage-specific changes in endogenous activin A (Kattman, Witty et al. 2011) and Wnt (Paige, Osugi et al. 2010) signaling are known to dramatically reduce cardiac induction efficiency. We reasoned that, by measuring responses to saturating agonists and antagonists of specific pathways, endogenous signaling levels could be assessed and quantitatively placed along a spectrum of low activation (endogenous response equivalent to pathway inhibitors) to high activation (endogenous response equivalent to saturating agonists).

We determined saturating concentrations of factors that affect early cell fate decisions from dose curves (Figure 2- 5a). Using H9 cells, we then tested the response to agonists and antagonists of the major pathways regulating hPSC fate (Figure 2- 7e). To obtain an estimate of activation of an endogenous pathway relative to its dynamic range (“%Dynamic Range”), we combined agonist, antagonist, and base line control measurements (Eqn. 1, Methods). The resulting endogenous signaling profile for activin, FGF, EGF, Wnt, and BMP activation in H9 is shown in Figure 2- 7e. We used similar quantitative profiles to compare H9 to ZAN11 and ZAN3 lines, revealing that EGF, FGF, and activin are all differentially endogenously activated (Figure 2- 7f and Figure 2- 10). Comparisons of control and pathway agonist-antagonist response profiles over multiple passages indicate high correlation between passages (ZAN11 \( r = 0.90 \) and, ZAN3 \( r = 0.88 \)), and lower correlation between cell lines (ZAN11 vs. ZAN3 \( r = 0.52 \)) (Figure 2- 7g). In summary, cell lines differ in their response to induction conditions and specific pathway agonists and antagonists, and this response is stable between passages. These response profile measurements offer a quantitative tool for rapidly fingerprinting hPSC signaling profiles.
Figure 2-10 Quantitative comparison of pathway activation across hiPSC lines

As in Figure 2-7e, endogenous signaling profiles were generated for ZAN3 and ZAN11 cell lines.

2.4 Discussion

Heterogeneity in cell response in clonal populations arises from many cell-autonomous and non-cell-autonomous factors. Our previous work and other diverse evidence led us to hypothesize that controlling key sources of microenvironment variance and optimizing colony size and density would allow robust cell response and thus overcome existing limitations in hPSC HTP assays. In comparison to previous assays, our 96µCP platform eliminates the need to seed in CM, results in less hPSC cell fate response variance and faster response kinetics,
and, in conjunction with Oct4 and Sox2 marker analysis, allowed us to screen for cell fate responses to 27 developmental factors, providing the most comprehensive characterization of hPSC cell fate response to cues to date.

Our 2-day HTP stem cell response signature approach is a direct measurement of early transcription factor expression changes to diverse exogenous cues. In the future, combined with cell-line specific epigenetic and gene expression data, this may offer a strategy to decipher the genetic and epigenetic basis of cell-line specific responses to cues. HiPSC cell lines can now be generated at a higher rate than they can be characterized. We propose the 96µCP platform as an effective first pass assay to detect tumorigenic cell lines and quantify neuroectoderm and primitive streak differentiation propensity. Further development may lead to a simple and effective in vitro teratoma surrogate assay.

Finally, population heterogeneity arising from microenvironmental differences applies broadly to adherent cell populations and is increasingly recognized as a major obfuscating factor in drug screening campaigns (Snijder, Sacher et al. 2009). We believe the methods presented herein are widely applicable to adherent cell types. We have developed 96µCP based assays for mouse epiblast stem cells, and hPSC-derived cardiac and endoderm cells with minimal modification to the base media, substrate, and pattern size. Applying these concepts towards rapid characterization of the signaling involved in cell fate decisions of differentiated cell types is a promising strategy to accelerate the drive towards clinical regenerative medicine and drug screening of hPSC disease models.

2.5 Methods

2.5.1 Cell culture

HESC lines H9, H1 and H7 were obtained from the WiCell Research Institute. ZAN3 and ZAN11 were derived from activated CD3+ T cells enriched from umbilical cord blood (see below). HES2 (hESC) was provided by G. Keller (McEwen Centre for Regenerative Medicine/University Health Network). BJ1D (hiPSC) was provided by M. Radisic (University of Toronto). 110 (hiPSC) and CA1 (hESC) were provided by A. Nagy (Samuel Lunenfeld Research Institute). PDX1 (MEL1-derived PDX1-GFP hESC) was provided by D. Melton (Harvard University). Runx1 (HES3-derived Runx1-GFP hESC) was provided by A. Elefanty.
(Monash University). R306C, RTT-Δ3-4 #37, and T158M #5, all Rett syndrome hiPSC disease models, as well as BJ4YA (hiPSC) were provided by J. Ellis (The Hospital for Sick Children). H9, H1, H7, ZAN3, ZAN11, and PDX1 cells were routinely cultured on feeder layers of irradiated MEFs feeders in knockout (KO)-Dulbecco’s modified Eagle’s medium (DMEM) (Invitrogen) with 20% KO-serum replacement (Invitrogen) (KO-DMEM) and supplemented with 4 ng mL⁻¹ FGF-2 (PeproTech). Cells were passaged 1:4-1:6 every 4-5 days, and were disassociated into small clumps using 0.1% collagenase IV (Invitrogen). HES2 and Runx1 were cultured on growth factor reduced Matrigel (MG) in KO-DMEM supplemented with 20 ng mL⁻¹ FGF-2 (PeproTech), and were passaged every 4-5 days using TrypLE Express (Invitrogen). CA1 was cultured on MG using NutriStem (Stemgent) as per manufacturer’s instructions. BJ1D, 110, R306C, RTT-Δ3-4 #37, T158M #5, and BJ4YA were cultured on MG using mTeSR1 (STEMCELL Technologies) as per manufacturer’s instructions. All cell line stocks were confirmed negative for mycoplasma contamination.

2.5.2 Microcontact printing of substrate into 96-well plates

Polydimethylsiloxane (PDMS) stamps were fabricated using standard soft lithography techniques (Peerani, Bauwens et al. 2009), with the exception that liquid PDMS was cast into a Teflon mould before curing, allowing control of the shape of the PDMS stamp (Figure 2-1). PDMS stamps were cast consisting of rectangular base with 24 posts (6 x 4) with micropatterned surfaces configured to enable microcontact printing directly into 96-well plates. The microcontact printing follows a protocol employed previously (Tan, Liu et al. 2004). Substrate solution consisted of either MG diluted 1:30 in phosphate buffered saline (PBS), or a solution of 0.00125% fibronectin (Sigma-Aldrich, F1141) and 0.002% gelatin (Sigma-Aldrich, G9391). Substrate solution was deposited onto the patterned surface of ethanol sterilized PDMS stamps for 4h at room temperature. Stamps were rinsed with ddH2O, dried gently with N2 gas, and placed into tissue-culture treated 96-well plates (Costar). Stamps were incubated in the 96-well plates for 7-10 min in a humidity chamber (Relative humidity 55-70%). The stamps were then removed and substrates were passivated with 5% weight Pluronic F-127 (Sigma-Aldrich) in ddH2O for 1 h.

2.5.3 Seeding hiPSCs onto patterned substrates
hiPSCs were dissociated using TrypLE™ for three min. TrypLE™ was inactivated by adding media containing 20% KO-serum replacement (SR) (Invitrogen). Cells were centrifuged and resuspended in either XVIVO10™ (XV) or SF media both supplemented with 40 ng mL⁻¹ bFGF (R&D), 10 ng mL⁻¹ activin A (R&D), 10 µM ROCK inhibitor Y-27632 (Tocris) and 7 µg mL⁻¹ additional insulin. SF media is a modification of DC-HAIF47 and consists of DMEM/F12, 1x Nonessential amino acids, 50 U mL⁻¹ Penicillin, 50 µg mL⁻¹ Streptomycin, 10 µg mL⁻¹ bovine Transferrin, 0.1 mM β-Mercaptoethanol (all Invitrogen), 2% fatty acid-free Cohn’s fraction V BSA (Serologicals), 1x Trace Elements A, B & C (Mediatech), 50 µg mL⁻¹ Ascorbic Acid (Sigma) and 7 µg mL⁻¹ recombinant human insulin. XV includes XVIVO10™ (Lonza) supplemented with 2 mM L-Glutamine, 0.1 mM β-Mercaptoethanol and 1x Nonessential amino acids. Cells were seeded at 10^5 cells per well (or as described in text) and incubated. After 6 h, cells were washed with PBS two times and incubated a further 42 h with fresh media as indicated (SF supplemented with factors, or MEF-CM).

2.5.4 Immunocytochemistry and image analysis

Plates were fixed for 30 min in 3.7% formaldehyde and permeabilized for 3 min in 100% methanol. Plates were imaged and quantitatively analyzed using the Cellomics Arrayscan VTI platform and Target Activation algorithm (Thermo Scientific). This algorithm generates nuclear masks, provides single cell nuclear intensity values for protein expression (Oct4, Sox2 etc.), DNA content through DAPI staining, as well as spatial x- and y-coordinates of the nuclei centroids. Insufficiently patterned wells were excluded from analysis based on a predetermined number of cells per well. To aid in pattern visualization and quality control, we developed a publicly available MATLAB script to plot these x- and y-coordinates for all wells in a 96-well plate (Nazareth, E. Analysis and visualization of Cellomics data. MATLAB Central File Exchange, http://www.mathworks.com/matlabcentral/fileexchange/43107). Clustering of cells into colonies using Euclidean distance was performed using clusterData (Shoelson, B. clusterData. MATLAB Central File Exchange, http://www.mathworks.com/matlabcentral/fileexchange/35014-clusterdata). Fluorescent images were obtained of Oct4 (1:500; BD), Sox2 (1:500 R&D Systems), Nanog (1:500; Cell Signaling), Tra-1-60 (1:500 R&D Systems), Brachyury (1:200; R&D Systems), Snail (1:200; R&D Systems), and GATA4 (1:200; R&D Systems). Primary antibodies were incubated
overnight in 10 % FBS in PBS at 4°C. AlexaFluor secondary antibodies (1:500; Molecular Probes) were incubated for 1 h in 10% FBS in PBS at room temperature. 16 bit TIFF images were obtained for each channel, contrast adjustment was performed identically across all controls, and channels were combined into pseudo-colored composite images.

2.5.5 hiPSC derivation

Umbilical cord blood samples were collected from consenting donors according to ethically approved procedures at Mt. Sinai Hospital (Toronto, ON, Canada). Activated CD3+ T cells enriched from umbilical cord blood were reprogrammed as described previously (Seki, Yuasa et al. 2012). T cells were enriched from umbilical cord blood using an EasySep human T cell enrichment kit (StemCell Technologies; cat# 19051). They were expanded for four days using Dynabeads human T-activator CD3/CD28 beads (Invitrogen; cat# 111-61D) and 30U per mL recombinant human IL-2 (R&D; cat# 202-IL-010) in OpTmizer T cell expansion serum-free media (Invitrogen; cat#A1048501). Mutant Sendai virus encoding human Oct4, Sox2, KLF4 and c-MYC (kindly provided by DNAVEC Corporation) were added to the cells at MOI 20 on day five. T cell media was replenished on day six and the cells were plated onto irradiated MEF feeders on day seven in hPSC medium with 5 ng mL⁻¹ bFGF. Media was exchanged every two days until hiPSC colonies were picked and characterized three weeks later.

2.5.6 hiPSC cardiomyogenic and endoderm induction

hiPSC cardiomyogenic induction was performed using a serum-free, aggregate-based strategy described elsewhere (Bauwens, Song et al. 2011). A single cell suspension of hiPSC was centrifuged into 400 µm-sized AggreWell™ inserts (Stemcell Technologies) at a density of 500 cells per microwell and cells were allowed to aggregate over night. On day one, mesoderm formation was induced using 5 ng mL⁻¹ bFGF and varying concentrations of Activin A as indicated. On day four, cells were transferred to LowCluster plates (Nunc) and further differentiation towards the cardiac lineage was induced with 10 ng mL⁻¹ VEGF and 150 ng mL⁻¹ DKK1 for four days. Subsequently, cells were maintained in 10 ng mL⁻¹ VEGF and 5 ng mL⁻¹ bFGF. Custom defined media provided in kind by G. Keller was used as the base media. Cells were kept under hypoxic conditions (5 % O₂) from day 0-12 and then transferred to normoxic conditions. Endoderm induction into PDX1⁺ pancreatic progenitor cells was performed as described by Nostro et al. (Nostro, Sarangi et al. 2011). Endoderm induction into
Foa2^+Sox17^+ definitive endoderm was performed as described by Rezania et al. (Rezania, Bruin et al. 2012) adapted to 96-well plates.

2.5.7 Flow cytometry

HiPSC-derived cardiomyocyte aggregates were incubated in collagenase type II (1 mg mL^-1; Worthington, LS004176) in Hank’s Balanced Salt Solution over night at room temperature and pipetted vigorously to obtain a single cell suspension which was then fixed in 4 % paraformaldehyde over night at 4° C. Cells were permeabilized using IntraPrep™ Permeabilization Reagent (Immunotech, A07803). Cardiac Troponin T primary antibody (Thermo Scientific, MS-295-P) was used at 1:200 and AlexaFluor 647 donkey anti-mouse IgG secondary antibody (Molecular Probes, A31571 ) at 1:200. Cells were analyzed using a FACSCanto (BD Biosciences) flow cytometer.

2.5.8 Estimation of endogenous pathway activation

To estimate endogenous pathway activation we obtained the cell fate response to the pathway agonist and antagonist, both at saturating levels as determined by dose-curves. The baseline response (in SF basal media alone) can then be calibrated within this range using the following equation:

\[
%\text{Dynamic Range} = \frac{X_{SF\ ctrl} - X_{Antagonist}}{X_{Agonist} - X_{Antagonist}}
\]

2.5.9 Statistical analysis

Statistics were computed using one way analysis of variance (ANOVA), two-factor ANOVA, or linear regression as indicated. Error bars on plots represent standard deviation (s.d.) of three or more replicate wells except where indicated differently. All statistics were computed in MATLAB using p-values as indicated. Hierarchical clustering was performed with MeV (MultiExperiment Viewer, [http://www.tm4.org/mev/](http://www.tm4.org/mev/)) using euclidian distance as the similarity metric (centered) and centroid linkage as the clustering method.

2.6 Acknowledgements

This work is funded by the Canadian Institutes of Health Research (CIHR) (P.W.Z). E.J.P.N. is supported by a CIHR Frederick Banting and Charles Best Canada Graduate...
Scholarships Doctoral Award. P.B.L. is supported by a Heart & Stroke Richard Lewar Centre of Excellence Studentship award. S.S. is supported by a Vanier Canada Graduate Scholarship. P.W.Z. is supported as the Canada Research Chair in Stem Cell Bioengineering. We thank S. Foster for illustrating Figure 2-1. Ligands provided in kind by the International Stem Cell Initiative. We thank G. Keller (McEwen Centre for Regenerative Medicine/University Health Network), M. Radisic (University of Toronto), A. Nagy (Samuel Lunenfeld Research Institute), D. Melton (Harvard University), A. Elefanty (Monash University), and J. Ellis (The Hospital for Sick Children) for provision of cell lines. We thank Peter Andrews and Martin Pera for helpful discussion.

2.7 Author Contributions

E.J.P.N. designed, performed and analyzed most experiments. J.E.E.O. assisted with immunocytochemistry and software development. P.L. performed cardiac induction experiments. S.S. created hiPSC lines. T.Y. performed endoderm induction experiments. M.A. and T.Y. provided cell culture support. S.K.W.O. provided editorial input on the manuscript. E.J.P.N. and P.W.Z. designed the project and wrote the manuscript.
Chapter 3
A multi-lineage screen reveals mTORC1 inhibition enhances human pluripotent stem cell mesendoderm and blood progenitor induction

Portions of this chapter have been prepared for submission for publication. Co-authors for this chapter include Nafees Rahman, Michael Prakesch, Rima Al-Awar, Ting Yin and Peter W. Zandstra. Authorization to reproduce this work has been obtained from all co-authors.
3 A multi-lineage screen reveals mTORC1 inhibition enhances human pluripotent stem cell mesendoderm and blood progenitor induction

3.1 Abstract

HPSCs self-organize into a range of local microenvironmental configurations, potentially influencing their differentiation propensities. Population-context dependent hPSC responses to exogenous cues represent challenges to the systematic analysis of hPSC response to molecular and chemical perturbations. To address this, we have patterned hPSCs in multi-well plates in configurations designed for rapid and robust cell response to cues, and have applied this platform to screen a collection of 400 small molecule kinase inhibitors for compounds biasing cell fates, simultaneously tracking multiple lineage readouts and measuring resultant yield and purity of these lineages. Phenotype enrichment analysis revealed mTOR inhibitors as strong inducers of mesendoderm cells, and follow-up dose-responses analysis revealed that while mesendoderm cells are induced in a monotonic dose-dependent manner, hPSC self-renewal exhibited a bimodal response. This mesendoderm enhancing effect of rapamycin is blocked by TGFβ or BMP receptor inhibition, and in the presence of both BMP and activin A(BA) and rapamycin the percentage of Brachyury+ mesendoderm cells is 3.2 fold higher than either condition alone, and the number of Brachyury cells per colony had a greater than 6.9 fold expansion than BA alone. Knockdown of Raptor, a mTORC1 component, via siRNA phenocopied the effects of rapamycin. When rapamycin was added to a blood differentiation protocol, CD34+VECAD+ hemogenic endothelium purity and yield were increased 3 fold, with a concomitant enhancement of blood colony forming cells (CFCs). This approach can be applied to leveraging assay optimization and single-cell based analytics to yield biological insights into other micro-environment dependent and multi-population model systems.

3.2 Introduction

HPSCs and their differentiated derivatives offer the exciting opportunity to develop tools to study and treat human diseases. However robust and reproducible control of hPSC fate is required to develop these tools. Small molecules offer one approach to exogenously control hPSC fate. The discovery and characterization of novel small molecules that regulate hPSCs
would be facilitated by effective cell-based phenotypic high-throughput screening (HTS) (Ding and Schultz 2004), which has been applied to hPSCs in a limited manner thus far. Furthermore, emerging data from cell-based hPSC assays has revealed that there are variable and contradictory observations between laboratories, even with matched cell lines and protocols, highlighting the need for standardized assays in order to generate meaningful results (Haibe-Kains, El-Hachem et al. 2013). Although the factors underlying this variability are not completely known, population context has been identified as a major contributor to assay inconsistency (Snijder, Sacher et al. 2012). In hPSC assays, microenvironmental factors such as endogenous ligands, ECMPs, cell-cell contact, and cell subpopulations are strong regulators of cell fate and are spatially heterogeneous (Peerani, Rao et al. 2007). Specifically, local cell density has been identified as a key regulator of population context. Consequently, spatial cell distribution has been shown to impact hPSC self-renewal (Maherali and Hochedlinger 2008) and differentiation trajectory (Cai, Yu et al. 2009, Chambers, Fasano et al. 2009) (Hwang, Chung et al. 2009), as well as the disease phenotype of familial dilated cardiomyopathy induced hPSCs (Sun, Yazawa et al. 2012).

These observations highlight the need for robust assays that combine highly defined cell culture conditions with comprehensive analysis of cell responses to exogenous cues. To that end, we have developed a high-throughput assay which consists of a chemically defined cell-patterning based hPSC culture system and custom analysis software for single cell and colony level analysis for cue-response screening and simultaneous quantification of pluripotent, and early neuroectoderm (NE), mesendoderm, and extra-embryonic/other cells (Nazareth, Ostblom et al. 2013). Colony size, local cell density, medium composition, and substrate have been engineered to allow robust measurement of cell fate responses to exogenous cues. The assay allows Oct4 pluripotency analysis at 48 h compared to 7 d for non-patterned hPSC (Gonzalez, Jennings et al. 2010), can detect both positive and negative regulators of multiple cell fate decisions. Whereas non-patterned based assays failed to detect effects of TGFβ1, BMP4, IGF1, heregulin-β1, and noggin, the patterning based assay accurately classified the positive and negative effects of these compounds on hPSC pluripotency, NE, mesendoderm, and extra-embryonic population induction.

We applied the assay to screen a library of kinase inhibitors for their effect on four early hPSC cell fates. This screen identified mammalian target of rapamycin (mTOR) inhibitors,
such as rapamycin, as having a strong mesendoderm inducing effect on hPSCs in defined conditions. Rapamycin synergizes with BMP4 and Activin A to enhance Brachyury induction more than 3 fold, and concomitantly improves hemogenic endothelium purity and yield 3 fold. Rapamycin also has a separate bimodal effect on hPSCs, enhancing pluripotency at sub-nanomolar concentrations. This study demonstrates the utility of controlling microenvironmental parameters and multiple subpopulation analysis in drug screening assays. The development of such assays should accelerate the fruitful application of HTS in other population context-dependent and multi-cell population cell systems.

3.3 Results

3.3.1 A kinase inhibitor screen of hPSCs reveals lineage-specific regulators

We have previously developed a 48 h hPSC screen with defined medium and substrate that employs control of spatial patterning of cells via micro-contact printing extra-cellular matrix proteins in order to configure the hPSC microenvironment for rapid and robust response to exogenous cues ("96µCP assay") (Nazareth, Ostblom et al. 2013) (Figure 3- 1a). In brief, ECMPs are first micro-contact printed directly into 96-well plates in 200 µm diameter spots arrayed with a 500 µm center-to-center distance (pitch), a single-cell suspension of hPSCs are then seeded into the patterned wells, and after 6 h the wells are washed leaving behind only cells adhered to the patterned ECMPS. Exogenous cues are added to the defined media for 42 h, after which cells are fixed, stained, and analyzed using high-content analysis (HCA). As the screen readout, quantifying Oct4 and Sox2 protein at the single cell level enables simultaneous classification of Oct4+Sox2+ (pluripotent), Oct4-Sox2+ (NE), Oct4+Sox2- (mesendoderm), and Oct4-Sox2- (extra-embryonic/other) cell fates. Additionally, using in-house software we are able to quantify percentages of the subpopulations within each well and also the absolute number of cells of each subpopulation per colony, allowing analysis of factor effects on both yield and purity (Figure 3- 1b).
High-throughput screening of small molecule regulators of hPSC pluripotency, primitive streak, neuroectoderm, and extraembryonic/other cell fate decisions.

(A). Assay design. ECMP is patterned into 96-well plates in 200µm diameter circles, arrayed in a grid pattern with 500µm pitch (center-center distance) using micro-contact printing. A single cell suspension of hPSCs is then generated and seeded into 96-well plates. HPSCs are left to settle and adhere to the ECMP for 6h. The plates are then washed, leaving an arrayed grid of hPSC colonies. Control media or base media with test compounds (400 small molecules kinase inhibitors, at 1 µM or 0.2 µM) is added to each well and incubated for 42h. Plates are then fixed, permeabilized, stained, and imaged. Single-cell Oct4 and Sox2 expression levels are quantified and thresholds are used to classify cells as positive or negative for each transcription factor. x- and y-coordinates of cells are also obtained, allowing cells in colonies to be clustered together and colony-level parameters to be obtained. (B) Schematic of hit rationale. Analysis of subpopulation numbers and purity allows identification of simultaneously high purity and high-yield inducing compounds. (C) Distributions of compounds ranked by efficiency in promoting pluripotency, neuroectoderm, primitive streak, and extraembryonic/other cell fates. Wells with insufficient cell numbers or colonies per well were excluded from analysis. (D) Increase in pluripotent subpopulation purity (%Oct4+Sox2+) vs. increase in yield (#Oct4+Sox2+ cells per colony) obtained from 712 unique treatments. Values from base media alone were subtracted to obtain the relative effect of each treatment, such that the graph origin is defined by treatment with base media alone. Thresholds for purity and yield were used to classify “pluripotency responders”. Arrow indicates a sample responding compound. (E) Increase in early neuroectoderm subpopulation purity (%Oct4-Sox2+) vs. increase in yield (#Oct4-Sox2+ cells per colony) obtained from 712 unique treatments. Thresholds for purity and yield were used to classify “neuroectoderm responders”. (F) Increase in primitive streak purity (%Oct4+Sox2-) vs. increase in yield (#Oct4+Sox2- cells per colony) obtained from 712 unique treatments. Thresholds for purity and yield were used to classify “primitive streak responders”. (G) Calcein staining of cells treated with hit compounds. Calcein staining (a viability marker) is performed on unfixed cells at 48 h. Methanol treated control (non-viable cells) shown as a control. Scale bar, 200 µm. (H) 67 putative hit compounds ranked by increasing viability. N=3. All error bars indicate standard deviation. (I) Sample composite colony images from controls and selected responders. Cells are stained for DNA (DAPI, blue), Oct4 (green), and Sox2 (red). Scale bar, 100 µm.

Response of control conditions, distributed across 10 96-well plates.

We next applied this platform towards screening a collection of 400 small molecule kinase inhibitors. Small molecule control of cell fate is attractive for scale-up purposes, and we also reasoned that such a focused library could be instrumental in elucidating endogenous
regulators of hPSC fate decisions. Compounds were screened at two concentrations (0.2 µM and 1 µM). For further analysis only compounds which met QC criteria were included (> 800 cells per well counted to ensure a significantly large sample size), resulting in 707 unique conditions including 5 controls that were replicated on each plate. Control conditions showed high reproducibility across plates (Figure 3-2) and the overall distributions of the compounds with respect to each subpopulation are shown in Figure 3-1c. In order to visualize factor effect on responding population cell yield and purity, we obtained the percentage cells for each well and also the average number of Oct4+Sox2+ cells per colony in the well. From these values we subtracted the percentage and number obtained with base media alone to get ΔOct4+Sox2+ (%) and ΔOct4+Sox2+ (# cells per colony) respectively. Plotting these two measures allows visualization of effects on purity and yield of each screened factor (Figure 3-1d). Based on desired outcomes (e.g. enhancing percentage and yield of Oct4+Sox2+ cells) we set thresholds for yield and purity to obtain a list of “responders”, which we subsequently performed enrichment analysis on. Similar yield vs. purity visualizations are shown for NE (ΔOct4-Sox2+ (%) and ΔOct4-Sox2+ (# cells per colony), Figure 3-1e) and mesendoderm (ΔOct4+Sox2- (%) and ΔOct4+Sox2- (# cells per colony), Figure 3-1f). Enlarged plots of the threshold regions showing compound names of responders are shown in Figure 3-3. To confirm viability of the endpoint cells, we modified the assay by performing live-cell calcein and Hoechst staining at the 48 h timepoint (Figure 3-1g). Analysis of 66 putative responders revealed high viability across compounds, comparable to control differentiation conditions (Figure 3-1h). Sample images of control conditions and select responders are shown in Figure 3-1i.
Oct4/Sox2 percentages and numbers per colony were obtained for each unique treatment, and treatment with base media alone was subtracted. Select treatments are labeled with the compound name. (A) Responding pluripotency enhancing compounds. Arrow indicates CM control condition. (B) Responding neuroectoderm enhancing compounds. Arrow indicates TiF control condition. (C) Responding primitive streak enhancing compounds. Arrow indicates BA control condition, and BA with FGF2 ("BAF").

3.3.2 Target enrichment analysis identified kinase inhibitor targets for NE and PS responders
Next, we sought to determine if responding compounds were statistically enriched for specific inhibitor targets. To test for enrichment in the responders for each phenotype, we performed the hypergeometric test to obtain $p$-values for the enrichment of each kinase target in the pluripotent, NE, and mesendoderm responder groups (five most enriched targets shown in Figure 3-4a). For the pluripotency enhanced cluster, no target was found to be statistically enriched, indicating that the tested kinase inhibitors are insufficient to rescue hPSCs in the growth factor-free base media which lacks exogenous FGF2 and activin A reported to be necessary for hPSC maintenance (Vallier, Alexander et al. 2005). This is in line with previous studies that indicate that small molecule based maintenance of pluripotency may require combinations of small molecules (Tsutsui, Valamehr et al. 2011), and it is also likely that pathway agonists are required in such a cocktail, and not just inhibitors. For NE responders we obtained the highest enrichment for small molecules targeting ERK/MEK ($p = 2.9 \times 10^{-5}$) (Figure 3-4b). Additionally, we found a high enrichment ($p = 0.001$) of small molecules targeting TGFβ superfamily type I activin receptor-like kinase (ALK) receptors ALK4, ALK5, and ALK7 (Figure 3-4b). Inhibition of TGFβ signaling is long established to result in differentiation of hPSCs towards NE (Vallier, Reynolds et al. 2004, Smith, Vallier et al. 2008). We have previously observed low levels of endogenous TGFβ signaling in the 96µCP assay, and this likely contributes to lower than expected significance of ALK4/5/7 inhibitors. Lastly, in the mesendoderm responders we found the highest enrichment for mTOR inhibitors ($p = 3.0 \times 10^{-10}$) (Figure 3-4c).
Figure 3-4 Target enrichment analysis

(A) Enrichment in pluripotency responders (shown in Figure 1d). \( p \)-values were obtained for each target, using the hypergeometric distribution, and shown in $-\log_{10}$ base 10. Threshold corresponds to \( p = 0.001 \). (B) Similar analysis
for neuroectoderm responders. (C) Similar analysis for mesendoderm responders. (D) Two-dimensional
hierarchical clustering of subpopulation yield and purity data across treatments. Left, sample similarity tree, with
samples clustered. Enhanced primitive streak cluster shown in purple, enhanced pluripotency cluster shown in
blue, and enhanced neuroectoderm cluster shown in red. Right, enhanced primitive streak cluster detailed view,
indicating compound name, concentration, and reported kinase targets. (E) Pluripotent and neuroectoderm
subpopulation percentages from all screened MEK/ERK inhibitors, ranked by pluripotency percentage response.
(F) Pluripotent and neuroectoderm subpopulation cells per colony from all screened MEK/ERK inhibitors, ranked by
pluripotency percentage response. (G) The MEK inhibitor PD-0325901 increases the percentage of early
neuroectoderm, and decreases percentage of pluripotent cells. * indicates significant difference from blank control,
p < 0.001. N ≥ 3. (H) The MEK inhibitor PD-0325901 does not increase the number of early neuroectoderm. *
indicates significant difference from blank control, p < 0.01. N ≥ 3. All error bars indicate standard deviation.

As an alternative to enrichment analysis of responders based on thresholds (in Figure 3-4d-f), we performed unbiased two dimension hierarchical clustering of the screen data, and
performed enrichment analysis on clusters. For each condition, using the four subpopulations (%Oct4+Sox2+, etc.) and four subpopulation per colony values for each condition (average
#Oct4+Sox2+ cells per colony etc.) we performed two dimensional hierarchical clustering (707
conditions, 8 outputs, 5656 data points), which revealed several distinct phenotype clusters
(Figure 3-4d, left). A cluster of compounds that enhanced pluripotency (dendogram
highlighted in blue) also contained the MEF-conditioned media (CM) pluripotency supporting
control (Figure 3-5a). Similarly a cluster of compounds that enhanced NE (dendogram
highlighted in red) also contained the TGFβ inhibitor and FGF2 (TiF) NE inducing control
(Figure 3-5b). The few compounds that enhanced mesendoderm (dendogram highlighted in
purple) also clustered with two mesendoderm induction controls, BMP4 with Activin A ("BA")
and BA with FGF2 ("BAF") (Figure 3-4d, right). In this cluster, 53% of the compounds
targeted mTOR and no other kinase was targeted by more than one compound. Enrichment
analysis of the three clusters revealed identical significantly enriched pathways as our
threshold based method (Figure 3-5c), giving further confidence to the ERK/MEK and mTOR
enrichments found in the screen.
Figure 3-5 Target enrichment in clusters

(A) Enhanced pluripotency cluster detailed view, indicating compound name, concentration. (B) Enhanced early neuroectoderm cluster detailed view, indicating compound name, concentration. (C) Enrichment in pluripotency, neuroectoderm, and in primitive streak clusters obtained through two-dimensional hierarchical clustering. P-values were obtained for each target, using the hypergeometric distribution, and shown in $-\log_10$ base 10. Threshold corresponds to $p = 0.001$. 
3.3.3 ERK/MEK inhibition induces hPSC differentiation towards NE in a BMP-sensitive manner

Broadly looking at all the ERK/MEK inhibitors in our kinase inhibitor library, we observed a general increase in the percent of NE cells at the expense of pluripotent cells, however there was great variance between the different chemical inhibitors (Figure 3-4e). Looking at the numbers of cells per colony induced by the ERK/MEK inhibitors in the library, by ordering the compounds by percentage of NE induced we noted no trend in increasing or decreasing total cell number (Figure 3-4f). We retested PD0325901 and confirmed a significant reduction in percent of pluripotent cells and increased percent of NE cells (both $p<0.001$) (Figure 3-4g). Although PD0325901 reduced the number of pluripotent cells per colony (approx. 13% of the blank control), there was not significantly more NE cells per colony induced relative to base media alone ($p=0.43$) (Figure 3-4h). In the literature, there are mixed reports about the effects of MEK inhibition on neural induction, and it is speculated that variability in endogenous factors, specifically BMP signaling, has hindered interpretation of these previous results (Greber, Coulon et al. 2011).

To investigate this further, we compared the effect of PD0325901 on NE induction with base media alone, with BA, or with 17 other single factors which are agonists or antagonists of major developmental pathways, and the effect of these 19 conditions alone. This mimics adding PD0325901 to hPSCs with different base line endogenous signaling levels. We found that when BMP4 is added with PD0325901 NE induction is reduced from 95% to less than 0.5% ($p<0.002$), confirming that NE is exceptionally sensitive to BMP4, giving further credence to the possibility of endogenous BMP4 abolishing NE induction (Figure 3-6a). Although PD0325901 robustly induced NE across multiple independent passages, we did observe NE induction varied between 74% and 90%, which may possibly be attributed to varying levels of endogenous BMP signaling. Additionally, we have previously assayed the response of a panel of cell lines to a set of conditions including blank base media, PD0325901, and TiF, using the 96µCP assay and staining for Oct4 and Sox2 (Nazareth, Ostblom et al. 2013). Meta-analysis of this data set confirms that PD0325901 and TiF variably induce superior early NE in a cell line dependent manner (Figure 3-6b). We further note that PD0325901 induction of NE was robust to all other factors, including FGF2, Activin A, TGFβ, IGF1, and HRGβ1, congruent with previous proposals that inhibition of MEK, BMP, and
TGFβ signaling (Greber, Coulon et al. 2011) or dual BMP and TGFβ signaling (Kim, Lee et al. 2010) may indeed robustly induce NE regardless of variations in endogenous signaling. Together this data confirm that NE induction is indeed highly variably across hPSC lines, and that this induction is uniquely sensitive to BMP inhibition, highlighting the need for exogenous BMP inhibition to achieve robust NE induction.

Figure 3- 6 MEK inhibitor PD0325901 induced neuroectoderm is sensitive to BMP inhibition but not Activin A inhibition

(A) 19 conditions were screened alone or with added PD0325901, and neuroectoderm induction response was measured to evaluate interactions. PD0325901 induction of neuroectoderm is insensitive to all factors measured except BMP4. * indicate p < 0.002. N=12 for "+blank+blank" condition, all other conditions N=2 if error bars shown, N=1 if no error bars shown. Error bars indicate standard deviation. (B) Cell line comparison of neuroectoderm induction using base media alone ("blank"), TGFβ inhibitor and FGF2 (TIF), and PD0325901. N=3. Meta-analysis of data provided in Nazareth et al. Nature Methods 2013.

3.3.4 mTOR inhibitors induce mesendoderm in hPSCs

Our analysis indicated a strong effect of mTOR inhibition on mesendoderm induction. Rapamycin, the prototypic inhibitor of mTOR, was initially shown to reduce pluripotency in
CM and enhance serum-mediated differentiation towards mesoderm and endoderm (Zhou, Su et al. 2009), however subsequent studies reported no direct effect of rapamycin on hPSCs, and in fact report rapamycin may enhance purity of pluripotent cells (Easley, Ben-Yehudah et al. 2010). We thus sought to elucidate the role of mTOR inhibitors in hPSC cell fate decisions in serum-free defined conditions.
Figure 3- 7 mTOR inhibitors in defined media upregulate mesendoderm via TGFβ and BMP pathways
(A) Nine mTOR inhibitors in the OICR library were tested at 1 and 0.1 μM. Additional inhibitors of kinases implicated in early development but not targeting mTOR ("Other inhibitors") were also tested. Ranking of small molecules by effect magnitude separated mTOR inhibitors from inhibitors of other pathways. * indicates $p < 0.01$, ** indicates $p < 0.001$ compared to blank control. N=3. (B) Ten-point dose curves of mTOR inhibitors rapamycin and AZD-8055 show a dose-dependent increase of %mesendoderm. * indicates $p < 0.05$ compared to blank control. N=3. Blank and BA control response shown as reference lines. (C) Ten-point dose curves of rapamycin show a bimodal increase in the percentage of pluripotent cells. * indicates $p < 0.05$ compared to blank control (0 rapamycin). N=3. (D) Ten-point dose curves of rapamycin and AZD-8055 shows a bimodal increase in the number of pluripotent cells per colony at sub-nanomolar concentrations of rapamycin but not AZD-8055. * indicates $p < 0.05$ compared to blank control (0 rapamycin). N=3. (E) 14 factors (ligands and small molecules) targeting early development pathways were tested alone and with rapamycin to screen for interactions. N=3. (F) TGFβ1 moderately but not statistically significantly enhances the mesendoderm inducing effect of rapamycin, while SB-431542 (an inhibitor of TGFB type I receptor ALK5) and LDN-193189 (an inhibitor of BMP type I receptors ALK2 and ALK3) abolish the mesendoderm-inducing effect of rapamycin. * indicates $p < 0.02$. N=3. (G) TGFβ1, SB-431542, and LDN-193189 do not effect total cell number per colony when added with rapamycin. N=3. All error bars indicate standard deviation.

In order to confirm that mTOR inhibitors enhance mesendoderm induction, we screened 9 mTOR inhibitors at both 1 μM and 0.1 μM. For comparison, we also screened 8 inhibitors of additional pathways implicated in hPSC regulation. When the compounds were ranked according to level of effect on mesendoderm induction %Oct4+Sox2- (%mesendoderm ) the mTOR inhibitors clearly separated from the other inhibitors and the majority of conditions tested significantly enhanced the percent of Oct4+Sox2- cells relative to media alone ($p<0.05$) (Figure 3- 7a). Rapamycin at 1 μM was the strongest responder out of all the conditions tested, and was comparable to the BMP4 and Activin A control. Rapamycin binds to FKBP12 and inhibits the kinase activity of a complex which contains mTOR and raptor (mTORC1) (Sabers, Martin et al. 1995). A separate complex containing mTOR and raptor (mTORC2) is generally thought to be rapamycin insensitive, although this is context dependent (Lamming, Ye et al. 2012). In contrast to rapamycin and its derivatives such as temsirolimus and everolimus which only target mTORC1, next-generation dual mTORC1/2 inhibitors have been developed such as AZD-8055 and KU-63794. It is notable that within the 18 mTOR inhibitor treatments (9 compounds at two concentrations), the top half of ranked responders are enriched for compounds only targeting mTORC1 ($p<0.005$).

To further validate the effect of mTOR inhibitors on hPSCs, we performed 10-point dose curves of two mTOR inhibitors, rapamycin and AZD-8055, and again measured the four early Oct4/Sox2 segregated subpopulations. As expected based on the previous tests, both compounds increased the percentage of mesendoderm cells in a dose-dependent manner (Figure 3- 7b). Surprisingly, when looking at the pluripotency response there was a bimodal response to rapamycin, where at 0.001 and 0.01 nM there was a moderate enhancement of the percentage and number of pluripotent cells, however this enhancement was not seen at
concentrations above 0.1 nM (Figure 3- 7c). The numbers, and not just percentage, of pluripotent cells are enhanced by low concentrations of rapamycin (Figure 3- 7d). This bimodal response was seen with two additional replicates of rapamycin (not shown) but was not seen with the dual mTORC inhibitor AZD-8055 (Figure 3- 7d). Future studies are needed to determine the precise mechanisms of this effect.

![Graph showing percentage of Oct4+ Sox2+ and Oct4- Sox2+ cells with different treatments](image)

**Figure 3- 8 SB-431542 and LDN-193189 antagonize the mesendoderm enhancing effect of rapamycin.**

N=3 . Error bars indicate standard deviation. * indicates p<0.01.

3.3.5 mTOR inhibition induction of mesendoderm is BMP receptor and TGFβ receptor dependent

To gain insight into the mechanism of action of mTOR inhibitors on inducing mesendoderm in hPSCs, we screened 14 agonists and antagonists of early development signaling pathways, both alone and with rapamycin (Figure 3- 7e and Figure 3- 8). TGFβ1 and FGF2 moderately enhanced rapamycin induced mesendoderm, though not significantly. The effects of rapamycin were abolished by the BMP type I receptor (ALK2/3) inhibitor LDN-193189, MEK inhibitor PD-032591, and the TGFβ receptor (ALK4/5/7) inhibitor SB-431542. MEK inhibition has previously been shown to switch differentiation away from mesendoderm (Yu, Pan et al. 2011), our results are in agreement with this and indicate mTOR inhibition is
insufficient to overcome this effect. The LDN-193189 and SB-431542 results imply that rapamycin's effect on mesendoderm is dependent on BMP and TGFβ receptor mediated signaling, which are both required (either exogenously or endogenously) for hPSC differentiation towards mesendoderm (Vallier, Touboul et al. 2009). When either of these inhibitors is added with rapamycin, relative to rapamycin alone there is an abolished Oct4+Sox2- population, no difference in the Oct4+Sox2+ population, and a moderately enhanced Sox2+Oct4- population (Figure 3-7f), with no change in the total number of cells per colony (Figure 3-7g).

Based on the blocking effects of SB-431542 and LDN-193189, we reasoned that rapamycin, which induces mesendoderm alone in serum-free defined media, may act additively or synergistically with TGFβ/BMP signaling. We therefore performed a rapamycin dose curve with and without exogenous BMP4 (10 ng/ml) and activin A (100 ng/ml), and to further confirm mesendoderm induction we stained for Brachyury (Zhang, Li et al. 2008). We observed a dose-dependent increase in the percentage and number of Brachyury+ cells in both BA conditions and without exogenous ligands. The maximum percentage of Brachyury+ cells achieved with rapamycin alone (19%) was comparable to the percent achieved with BA alone (18%), however rapamycin in addition to BA resulted in more than 3 fold higher percentage of Brachyury+ (57%) compared to BA alone ($p<0.005$) (Figure 3-9a). The number of Brachyury+ cells per colony was also enhanced in rapamycin with BA relative to BA alone by 3.8 fold ($p<0.005$) (Figure 3-9b). Sample images are shown in Figure 3-9c. Similar trends were observed with rapamycin or an alternative mTOR inhibitor (temsiriromimus) dose curves performed on non-patterned hPSCs with a BA background (Figure 3-10).
Figure 3-9 mTOR inhibition synergizes with BMP4 and activin in inducing Brachyury+ mesendoderm

(A) Dose curves of rapamycin were performed alone and with BMP4 (10 ng/ml) and activin A (100 ng/ml), and Brachyury expression was quantified. Average percentage of Brachyury cells per well is shown. N=3. * indicates p < 0.05, ** indicates p < 0.005 compared to the equivalent condition without rapamycin. Response of base media alone (“blank”) and BMP4 and activin A alone (“BA”) are shown as reference lines. (B) Same experiment, showing the number of Brachyury cells per colony. * indicates p < 0.05, ** indicates p < 0.005 compared to the equivalent condition without rapamycin. N=3. (C) Images of cells stained for DNA (DAPI, blue) and Brachyury expression (red), with treatments as indicated. Scale bar, 200 µm. (D) The effect of siRNA on Oct4+Sox2- mesendoderm. siRNA was applied in different media conditions as noted to non-patterned hPSCs, and cells were stained for Oct4 and Sox2. N=3. * indicates p < 0.05, ** indicates p < 0.005, *** indicates p < 0.0005. (E) The effect of siRNA on Brachyury expression. siRNA was applied in different media conditions as noted to non-patterned hPSCs, and cells were stained for Brachyury. N=4. * indicates p < 0.05. (F) Same experiment, showing number of Brachyury+ cells per field, and total cells. N=4. ** indicates p < 0.005, *** indicates p < 0.0005. (G) Images of cells stained for DNA (DAPI, blue) and Brachyury expression (red), with treatments as indicated. Scale bar, 500 µm. All error bars indicate standard deviation.

Figure 3-10 mTOR inhibitors enhance Brachyury+ and Oct4+Sox2- induction in a dose dependent manner in non-patterned hPSCs

(A) Dose curve of rapamycin in the presence of BMP4 (10 ng/ml) and activin A (100 ng/ml), applied to non-patterned hPSCs. Resulting Brachyury expression was quantified. Average percentage of Brachyury cells per well is shown. BMP4 and activin A alone shown as reference line. N=3. (B) Same experiment, showing number of Brachyury+ cells per well. N=3. (C) Dose curve of temsirolimus in the presence of BMP4 (10 ng/ml) and activin A (100 ng/ml), applied to non-patterned hPSCs. Resulting Oct4 and Sox2 expression was quantified. Average percentage of Oct4+Sox2- cells per well is shown. N=3. (D) Same experiment, showing number of Oct4+Sox2-+ cells per well. N=3.
3.3.6 Knockdown of Raptor phenocopies the mesendoderm inducing effect of rapamycin

mTOR functions as part of two protein complexes, mTORC1 and mTORC2 (reviewed in (Sabatini 2006)), which are differentially regulated by upstream signals and have different downstream effectors. mTORC1 consists of mTOR, raptor, and mLST8, whereas mTORC2 consists of mTOR, rictor, mLST8, and mSin1. Although rapamycin was long thought to inhibit mTORC1, more recent studies have revealed rapamycin can also inhibit mTORC2, in a cell-type dependent manner (Lamming, Ye et al. 2012). In order to examine the role of each mTOR complex in rapamycin-induced mesendoderm differentiation, we targeted raptor (mTORC1 specific) and rictor (mTORC2 specific) with siRNA. HPSCs were first cultured in pluripotency maintenance media with siRNA added for two days, and then cultured for an additional two days in a media condition (+blank, +rapamycin, or +BA ) with the same siRNA, and Oct4 and Sox2 expression was analysed at day 4 (Figure 3-9d). In base media alone, siRAPTOR enhanced the percentage of Oct4+Sox2- cells relative to control scrambled siRNA (siCTRL) ($p<0.0005$), but siRICTOR had no significant effect. Addition of both siRAPTOR and siRICTOR gave a similar phenotype to siRAPTOR alone, and gave rise to a significantly higher percent of Oct4+Sox2- cells that siRICTOR alone ($p<0.05$). In the presence of rapamycin, siRAPTOR again had a significant effect ($p<0.0005$), and siRAPTOR was not significantly different than siCTRL. This result indicates that rapamycin (at 0.1µM) is not completely inhibiting mTORC1. In the presence of BA, siRAPTOR also enhances the percent of Oct4+Sox2- cells relative to siCTRL ($p<0.0005$). Interestingly, in this media condition siRICTOR also enhanced the percent of Oct4+Sox2- ($p<0.0005$) indicating that the effect of mTORC2 inhibition is dependent on TGFβ/BMP signaling. Focusing on the BA conditions, we repeated the siRNA assay and stained for Brachyury. As expected, siRAPTOR significantly enhanced both percentage (Figure 3- 9e) and number (Figure 3- 9f) of Brachyury+ cells, similarly to rapamycin treatment. Sample images are shown in Figure 3- 9g. siRICTOR had no effect on Brachyury induction, indicating that the effect of rapamycin in Brachyury induction is primarily through mTORC1, and not mTORC2. These results also indicate a divergence
between the Oct4+Sox2- readout and Brachyury expression. As a whole, these results further corroborate that rapamycin is indeed acting via inhibition of mTOR signaling, and not via non-specific effects.

### 3.3.7 mTOR inhibition enhances hPSC differentiation towards mesoderm and not endoderm derived tissues

Based on the ability of mTOR inhibition to enhance the purity and number of mesendoderm cells, we next sought to test the effect of treatment on mesoderm and endoderm downstream progenitors.
Figure 3- 11 Rapamycin enhances blood progenitor induction.

(A) Sample flow cytometry scatter plots of day 8 VECAD and CD34 expression in hPSCs differentiated towards hemogenic endothelium with rapamycin (top) and without rapamycin (bottom). (B) Time course showing VECAD+CD34+ cells throughout day 1-8 of differentiation showing percentage (top) and numbers per well (bottom) with and without rapamycin treatment. N as indicated. Error bars indicate standard deviation. (C) Comparison of the maximum VECAD+CD34+ percentage obtained with or without rapamycin. N=3. Error bars indicate standard deviation. * indicates p<0.05. (D) Comparison of the fold increase in #CD34+VECAD achieved with rapamycin relative to without rapamycin, on day 6 and day 8. N=3. Error bars indicate 99% confidence interval. (E) CFC comparison of definitive blood obtained with and without rapamycin. N=3, pooled independent runs of the HES2 and RUNX1-GFP (two replicates) cell lines. Error bars indicate standard deviation. (F) Four hPSC cell lines were differentiated towards definitive endoderm using a 5 day protocol, and analyzed for Foxa2+Sox17+ expression. Rapamycin added as indicated during the “D0” phase. N=3. Error bars indicate standard deviation.

Figure 3- 12 Differentiation of hPSC cell lines towards hemogenic endothelium is enhanced by rapamycin

(A) Time course showing KDR+ cells throughout day 1-8 of differentiation showing percentage (left) and numbers per well (right) with and without rapamycin treatment. (B) CFC comparison of definitive blood obtained with and without rapamycin. N=3, pooled independent runs of the HES2 and RUNX1-GFP (two replicates) cell lines. * indicates p<0.05 (paired t-test).
To examine if rapamycin could enhance blood induction, hPSCs were differentiated towards hemogenic endothelium (CD34+VECAD+) in chemically defined media (Kennedy, Awong et al. 2012) supplemented with BMP4, SCF, bFGF and VEGF (Pick, Azzola et al. 2007). Rapamycin (0.1 µM) was added from day 0 to day 2 of the 8 day protocol, whereas control treatments were devoid of rapamycin during this time range. Analysis of CD34+VECAD+ cells throughout the time-course revealed that rapamycin addition to the early time points enhances both purity and yield of hemogenic endothelium (Figure 3-11a-b). KDR expression nearly doubled on day 4, though no difference was seen before this time point (Figure 3-12a). The maximum CD34+VECAD+ percentage reached in control conditions was 10%, whereas with rapamycin the purity was enriched to 30% ($p<0.05$) (Figure 3-11c). Rapamycin also enhanced purity of hemogenic endothelium in the HES2 cell line (Figure 3-13). Purity CD34+VECAD+ with rapamycin peaked at day 6, and yield was stable from day 6 to 8. Conversely, purity in control conditions was stable from day 6 to 8, but yield slightly increased from day 6 to 8. The fold increase in CD34+VECAD+ cells with rapamycin was 4 fold on day 6, and 2.5 fold on day 8 (99% confidence intervals shown Figure 3-11d). In order to assess the generation of hematopoietic progenitors in each condition, we determined the frequency of colony-forming cells (CFCs) observed after 14 day differentiation in methylcellulose. Control conditions gave rise to 117 +/- 26 (mean +/- SEM) CFCs per $10^5$ cells, and rapamycin treated conditions 202 +/- 42 CFCs per $10^5$ cells, a 1.7 fold increase ($p<0.05$) (Figure 3-11e, Figure 3-12b). These results demonstrate that rapamycin treatment...
during day 0 to day 2 of hPSC hematopoietic differentiation can enhance the blood progenitor phenotype.

DE differentiation is a multistep process including a 1 day induction phase ("D0") where cells are exposed to activin A (100 ng/ml) and the GSK3β inhibitor CHIR99021 ("CHIR", 2µM). To test the effect of rapamycin on DE induction, we tested various concentrations of rapamycin added to the D0 conditions and measured the resulting Foxa2+Sox17+ DE at day 5. Although DE differentiation protocols can be extremely efficient for particular cell lines (Nostro, Sarangi et al. 2011), cell-line specific differentiation efficacy is known to vary greatly (Nazareth, Ostblom et al. 2013) so we tested four cell lines we have previously shown to have varying DE induction efficiencies. We found that rapamycin dose dependently reduced DE induction in PDX1, HES2, and RUNX1 cell lines, and had no effect on H9 DE induction (Figure 3-11f). These results show that rapamycin does not enhance endoderm differentiation, and thus may be specific to mesoderm induction.

3.4 Discussion

The tremendous potential benefits of hPSCs have merited nearly two decades of intensive research into hPSC cell fate regulation. Despite this, a major bottleneck in scalable and cost effective therapeutic cell production is the lack of efficient protocols to generate differentiated cells with high yield, purity, and function. Additionally, the molecular mechanisms that underpin hPSC fate decisions is poorly understood. To help address this, we have applied an engineered microenvironment based HTP platform to screen a kinase inhibitor library simultaneously for regulators of pluripotency, NE, and mesendoderm, looking at resulting purity and yield of subpopulations.

The screen revealed that mTOR inhibitors added alone to serum free defined media can induce mesendoderm, and our follow-up studies confirmed that addition of rapamycin to existing protocols effectively enhances the formation of mesendoderm, hemogenic endothelium, and blood progenitors. By using a HTP multi-lineage assay for secondary studies, we were able to perform 10-point dose curves that reveal a bimodal effect where rapamycin at low concentrations mildly enhances the purity and number of Oct4+Sox2+ pluripotent cells, whereas higher concentrations reduce the Oct4+Sox2+ population and result in mesendoderm differentiation. These results help reconcile two previous studies. In line with Easley et al. we
observe no negative effect on pluripotency of rapamycin at low concentrations, and in line with Zhou et al. at higher concentrations we observe loss of pluripotency and induction of mesendoderm. However, in contrast to Zhou et al. who observed rapamycin enhances serum mediated EB differentiation towards endoderm and mesoderm, we found rapamycin in fact strongly inhibits definitive endoderm induction. This difference may be serum-mediated or cell-line dependent.

Although initially surprising that sub-nanomolar concentrations of rapamycin have a moderate enhancing effect on pluripotency, this result is in congruence with findings of He et al. that for mouse somatic cell reprogramming, rapamycin at 0.01 and 0.03 nM enhances reprogramming (nearly two fold), but 0.05 nM and higher has no effect or reduces reprogramming (He, Kang et al. 2012). As differentiated cell types such as hPSC-derived extra-embryonic endoderm are known to inhibit pluripotency (Peerani, Rao et al. 2007), rapamycin selectively inhibiting these cell types may lead to increased hPSC cell number, as an alternative to rapamycin supporting the maintenance of hPSCs in a cell autonomous manner.

MEK inhibitors were found to strongly induce NE in our screen. These results were surprising as FGF/MEK signaling has been reported by Vallier et al. to be necessary for hPSC NE specification, as they observed chemical inhibition of FGF receptors inhibits NE induction (Vallier, Mendjan et al. 2009). However, Greber et al. have reported that FGF in fact inhibits NE induction and MEK inhibition via PD0325901 enhances NE induction (Greber, Coulon et al. 2011). Notably, in the mouse system Ying et al. have reported monolayer induction of mouse NE without exogenous inductive stimuli is dependent on autocrine FGF signaling, as FGF receptor inhibition virtually abolishes NE induction (Ying, Stavridis et al. 2003). However, Smukler et al. have observed that in low density minimal conditions mouse embryonic stem cells (mESCs) undergo NE differentiation even under FGF receptor inhibition (Smukler, Runciman et al. 2006). Greber et al. reconcile these observations with the hypothesis that removal of FGF signaling can result in autocrine BMP signaling, in a culture and/or cell-line dependent manner, which may be responsible for the diametrically opposed observations. Our follow-up studies testing NE induction with PD0325901 with a panel of added factors, along with our meta-analysis of TiF, basal media, and PD0325901 NE induction across a panel of cell lines confirm that NE differentiation is highly variable across cell lines and
exceptionally sensitive to BMP inhibition, supporting the necessity for BMP receptor inhibition in "universal" NE protocols as previously observed (Chambers, Fasano et al. 2009).

The methods we apply here are applicable to systems with heterogeneous subpopulations and complex microenvironmental regulation, such as in vitro stem cell and cancer models, and can be applied iteratively. For example, using the hemogenic endothelium produced with rapamycin we are currently using the platform to discover small molecules and biologics that enhance the endothelial to hematopoietic cell transition (manuscript in preparation). Scalable platforms also enable medium scale systematic studies of factor effects at different doses, on large panels of cell lines (as in our endoderm study), and interactions with a panel of agonists and antagonists of signaling pathways ubiquitous in development and highly variable among cell lines. Advancing the application of medium-scale platforms more broadly may aid in increasing the translatability of studies across cell lines and laboratories, and bring coherence to otherwise disparate observations.

3.5 Methods

3.5.1 Cell culture

HESC lines H9 (WA09) H7 (WA07) were obtained from the WiCell Research Institute. ZAN3 and ZAN11 hiPSC lines were derived from activated CD3+ T cells enriched from umbilical cord blood (Nazareth, Ostblom et al. 2013). HES2 (hESC) was provided by G. Keller (McEwen Centre for Regenerative Medicine/University Health Network). BJ1D (hiPSC) was provided by M. Radisic (University of Toronto). CA1 (hESC) were provided by A. Nagy (Samuel Lunenfeld Research Institute). Runx1 (HES3-derived Runx1-GFP hESC) was provided by A. Elefanty (Monash University). R306C, RTT-Δ3-4 #37, and T158M #5 (ref. 44), all Rett syndrome hiPSC disease models, as well as BJ4YA (hiPSC) were provided by J. Ellis (The Hospital for Sick Children). H9, H7, ZAN3, and ZAN11 were routinely cultured on feeder layers of irradiated MEFs feeders in knockout (KO)-Dulbecco’s modified Eagle’s medium (DMEM) (Invitrogen) with 20 % KO-serum replacement (Invitrogen) (KO-DMEM) and supplemented with 4 ng mL⁻¹ FGF-2 (PeproTech). Cells were passaged 1:4-1:6 every 4-5 days, and were disassociated into small clumps using 0.1 % collagenase IV (Invitrogen). HES2 and Runx1 were cultured on growth factor reduced Matrigel (MG ) in KO-
DMEM supplemented with 20 ng mL\(^{-1}\) FGF-2 (PeproTech), and were passaged every 4-5 days using TrypLE Express (Invitrogen). CA1 was cultured on MG using Nutristem® hESC XF (NS) (Biological Industries, cat. no. 05-100-1A) as per manufacturer’s instructions. BJ1D, R306C, RTT-Δ3-4 #37, T158M #5, and BJ4YA were cultured on MG using mTeSR1 (STEMCELL Technologies) as per manufacturer’s instructions. All cell line stocks were confirmed negative for mycoplasma contamination.

3.5.2 Microcontact printing of substrate into 96-well plates

We have previously developed a method for patterning proteins in standard 96-well plates (Nazareth, Ostblom et al. 2013). In brief, PDMS stamps were fabricated using standard soft lithography techniques\(^{45}\), with the exception that liquid PDMS was cast into a Teflon mould before curing, allowing control of the shape of the PDMS stamp. PDMS stamps were cast to fit directly into 96-well plates. The microcontact printing follows a protocol employed previously\(^{46}\). ECM solution consisted of 0.00125 % fibronectin (Sigma-Aldrich, F1141) and 0.002 % gelatin (Sigma-Aldrich, G9391) in phosphate buffered saline (PBS). ECM solution was deposited onto the patterned surface of ethanol sterilized PDMS stamps for 4h at room temperature. Stamps were rinsed with ddH2O, dried gently with N2 gas, placed into tissue-culture treated 96-well plates (Costar) ensuring conformal contact between the plate and PDMS stamp, and incubated in the 96-well plates for 7-10 min in a humidity chamber (Relative humidity 55-70 %). The stamps were then removed and substrates were passivated with 5 % weight Pluronic F-127 (Sigma-Aldrich) in ddH2O for 1 h.

3.5.3 Seeding hPSCs onto 96-well plates and application of small molecules or controls

HPSCs were dissociated using TrypLE\(^{\text{TM}}\) for three min. TrypLE\(^{\text{TM}}\) was inactivated by adding media containing 20 % KO-serum replacement (KOSR) (Invitrogen). Cells were centrifuged and resuspended in NS and 10 µM ROCK inhibitor Y-27632 (Tocris). Cells were seeded at 10\(^5\) cells per well (or as described in text) into 96-well plates, either pre-coated with non-patterned ECM or patterned with ECM, and incubated. After 6 h, cells were washed with PBS two times and incubated a further 42 h with fresh media as indicated (SF supplemented with growth factors or small molecules, or CM). SF media consists of DMEM/F12, 1x Nonessential amino acids, 50 U mL\(^{-1}\) Penicillin, 50 µg mL\(^{-1}\) Streptomycin, 10 µg mL\(^{-1}\) bovine
Transferrin, 0.1 mM β-Mercaptoethanol (all Invitrogen), 2 % fatty acid-free Cohn’s fraction V BSA (Serologicals), 1x Trace Elements A, B & C (Mediatech), 50 µg mL⁻¹ Ascorbic Acid (Sigma) and 7 µg mL⁻¹ recombinant human insulin. The 400 compound library was obtained from the Ontario Institute of Cancer Research (OICR) and is a custom library consisting of small molecule kinase inhibitors, mainly commercially available.

3.5.4 Immunocytochemistry and high-content image analysis

Plates were fixed for 30 min in 3.7 % formaldehyde and permeabilized for 3 min in 100 % methanol. Plates were imaged and analyzed using the Cellomics Arrayscan VTI platform and Target Activation algorithm (Thermo Scientific). This algorithm generates nuclear masks, provides single cell nuclear intensity values for protein expression (Oct4, Sox2 etc.), DNA content through Hoechst staining, as well as spatial x- and y-coordinates of the nuclei centroids. Single cell x-y-coordinate and protein expression data was exported in tab delimited text files and imported into in-house software ContextExplorer (manuscript in preparation) for exploration of colony level details. Fluorescent images were obtained of Oct4 (1:500; BD), Sox2 (1:500; R&D Systems), Brachyury (1:200, R&D Systems) and Hoechst 33342 (Sigma-Aldrich). Primary antibodies were incubated overnight in 10 % FBS in PBS at 4 o C. AlexaFluor secondary antibodies (1:500; Molecular Probes) were incubated for 1 h in 10% FBS in PBS at room temperature. For figures, 16 bit TIFF images were obtained for each channel, contrast adjustment was performed identically across all controls, and channels were combined into pseudo-colored composite images.

3.5.5 96-well siRNA transfection

HPSCs were first seeded into MG treated 96-well plates (non-patterned) at 25,000 cells per well in NS, and incubated for 48 h. For siRNA transfection in 96-well plates, 0.15 µl per well Lipofectamine® RNAiMAX Transfection Reagent (Invitrogen, cat. no. 13778030) and 20 µl per well Opti-MEM® (Invitrogen, cat. no. 31985062) were first mixed in an eppendorf tube and incubated for 5 min at room temperature. Next, 20 µl per well of this solution was transferred to a V-bottom 96-well plate. 0.3 µl per well siRNA (at 20 pm per µl) was then added to each well, mixed, and incubated for an additional 20 min at room temperature ("siRNA mix"). 48 h after hPSCs were seeded into 96-well plates, the media was removed and 20 µl per well of siRNA mix was added along with 40 µl per well fresh NS. After 6 h of
incubation (T=54h) 160 µl of NS was added per well, and the hPSCs were cultured for an additional 42 h. On day 4, media was removed and 20 µl per well of siRNA mix was added (with wells treated with the same siRNA on day 2 and day 4) along with 40 µl per well of treatment media (SF+blank, SF+rapamycin, SF +BA, or NS ). 6 h later 160 µl of the same treatment media was added to each well, and cells were incubated for an additional 42 h before fixing.

3.5.6 Viability analysis

To assess viability, the patterned hPSC assay was modified. Instead of fixing at 48 h, calcein AM (Invitrogen, cat. no. C3099) (1:1000) and Hoechest 33342(1:1000) were added directly to the media, incubated for 30 min, and then imaged.

3.5.7 Differentiation of hPSCs towards blood and definitive endoderm

24-well Aggrewells™ (StemCell Technologies) were manufactured in-house using 400 µm polydimethylsiloxane inserts cast from a silicone master mold. For differentiation, hPSCs on MEFs were dissociated with TrypLEY™ Express treatment and plated onto Geltrex® (diluted 1:50) for 48 hours. After 48 hours, TrypLEY™ was added for 5 minutes and thereafter quenched with 50%FBS in DMEM/F12. Single cells suspensions were seeded and centrifuged into Aggrewells™ supplemented with ROCK inhibitor Y-27632 (for one day) at 1500 rpm for 5 minutes. Cells were cultured for 8 days in hypoxia supplemented with cytokines with base media comprising of StemPro34 (Invitrogen), ascorbic acid (50 µg/ml; Sigma), L-glutamine (1% v/v, Invitrogen), penicillin/streptomycin (1% v/v, Invitrogen), 1-monothioglycerol (4x10⁻⁴ M; Sigma), transferrin (150 µg/ml; Roche). The cytokines comprised of BMP4 (40 ng/ml, R&D), VEGF (50 ng/ml, R&D), SCF (40 ng/ml, R&D), bFGF (5 ng/ml, Peprotech), rapamycin (0.1 uM, SIGMA). On day 2, hPSCs-derived cells were transferred to low-cluster six well plates wherein media was exchanged without rapamycin and rock inhibitor and cells were cultured for an additional 4 days. Endoderm induction into Foxa2⁺Sox17⁺ definitive endoderm was performed as described by Rezania et al.⁵¹ in 96-well plates.

3.5.8 Colony forming cell (CFC) assay

A minimum of 150,000 cells were seeded on day 8 in 35 mm Greiner dishes in methylcellulose based MethoCult™ H4435 Enriched media (Stem Cell Technologies).
Samples were scored based on morphology 14 days after plating as described by technical manuals provided by StemCell Technologies.

### 3.5.9 Statistical analysis

Statistics were computed using one way analysis of variance (ANOVA), two-factor ANOVA, or linear regression as indicated. Error bars on plots represent standard deviation (s.d.) of three or more replicate wells except where indicated differently. All statistics were computed in MATLAB using p-values as indicated. Enrichment analysis was performed using the hypergeometric distribution in Microsoft Excel. Hierarchical clustering was performed with MeV (MultiExperiment Viewer, http://www.tm4.org/mev/) using euclidian distance as the similarity metric (centered) and centroid linkage as the clustering method. Exploratory analysis was performed with Microsoft Excel and Tableau (Tableau Software, www.tableau.com).

### 3.6 Acknowledgements

This work is funded by the Canadian Institutes of Health Research (CIHR) (P.W.Z). E.J.P.N. is supported by a CIHR Frederick Banting and Charles Best Canada Graduate Scholarships Doctoral Award. P.W.Z. is supported as the Canada Research Chair in Stem Cell Bioengineering. We thank G. Keller (McEwen Centre for Regenerative Medicine/University Health Network), M. Radisic (University of Toronto), A. Nagy (Samuel Lunenfeld Research Institute), D. Melton (Harvard University), A. Elefanty (Monash University), and J. Ellis (The Hospital for Sick Children) for provision of cell lines.

### 3.7 Author Contributions

E.N. designed, performed and analyzed most experiments. N.R. performed all blood induction experiments. T.Y. provided cell culture support. M.P. and R.A. provided the kinase inhibitor library and related support. E.N. and P.W.Z. designed the project and wrote the manuscript.
Chapter 4
Tools for high-throughput screening human pluripotent stem cells in configurable microenvironments

Portions of this chapter have been prepared for submission for publication. Co-authors for the protocol include Mukul Tewary, Orla O’Shea, Shalinee Khadun, Lyn Healy, Glyn Stacey, Dominika Dziedzicka, Mieke Geens and Peter W. Zandstra. Co-authors for the ContextExplorer work include Joel E. E. Ostblom and Peter W. Zandstra. Authorization to reproduce this work has been obtained from all co-authors.
4 Tools for high-throughput screening human pluripotent stem cells in configurable microenvironments

4.1 A multi-lab validated protocol for simultaneous and reproducible quantification of human pluripotent stem cell pluripotent, neuroectoderm, mesendoderm, and extra-embryonic cell fate responses to cues

4.1.1 Abstract

Human pluripotent stem cells (hPSCs) can undergo differentiation towards multiple germ layers and give rise to diverse cell types and tissues with immense medical potential. There is a need for simple, quick, cost-effective, and reproducible assays of hPSC multi-lineage differentiation response to cues. Through the International Stem Cell Initiative (ISCI) consortium we have developed a multi-laboratory validated protocol for a 48 h assay to quantify hPSC cell fate response to cues that simultaneously measures resultant pluripotency, neuroectoderm, mesendoderm, and extra-embryonic/other cell populations. The assay is performed in 96-well plates, and makes use of patterning hPSCs into defined microenvironments, which aids in reproducibility and interpretability of results. Although compatible with high-throughput (HTP) screening approaches, this protocol is designed to aid transition of HTP, high-content (HCA), and micro-technology approaches to broader scientific communities in various medium-scale applications. The protocol can be applied to diverse cell types, including hPSC derivatives. We detail several applications of this assay, such as quantitatively "fingerprinting" hPSC phenotype response to standardized control conditions. Such standardized and quantitative fingerprints are stable between passages, variable between cell lines, predict endoderm induction efficiency, can be used to rank cell lines by lineage induction efficiency, and can rapidly flag neoplastic/karyotypically abnormal cell lines.

4.1.2 Introduction

HPSCs offer limitless potential to improve human health through enhancing regenerative medicine, developmental biology, disease modeling, and drug discovery. A major bottleneck to
these efforts is efficient differentiation of hPSCs towards relevant cell types. It has long been observed that hPSCs have varying differentiation propensities, often necessitating tedious cell-line specific optimization of differentiation protocols. Additionally, there are serious concerns about the reproducibility of current hPSC protocols between laboratories. In the first multi-laboratory comparative study of published defined culture systems for hPSC expansion, the International Stem Cell Initiative (ISCI) recently reported remarkable variability across labs, with two out of eight tested published media formulations reproducibly maintained hPSCs across laboratories. Similar results have been reported in other biomedical fields, including an analysis of large cancer cell line screens which found that drug response data between laboratories are so highly discordant that meaningful gene-drug associations is precluded (Haibe-Kains, El-Hachem et al. 2013). One possible source of this observed assay variance is variability in population context (such as how the cells are distributed in a culture vessel), which is well established to confound assay results in large scale screens (Snijder, Sacher et al. 2009, Snijder, Sacher et al. 2012). Population context has proven to be a particularly strong regulator of hPSCs and derivatives (reviewed in Chapter 1), with reprogramming to pluripotency, differentiation towards neural, pancreatic, and cardiac cell types all being reported to have strong population context effects. As the number of hPSC cell lines continues to grow rapidly, there is a need for simple, quick, cost-effective, and reproducible assays of hPSC response to diverse factors.

Towards addressing this, herein we detail a 48 h assay for quantitatively assessing hPSC response to cues by simultaneously measuring resulting pluripotency, neuroectoderm, mesendoderm, and extra-embryonic/other cell populations. This protocol entails expanding hPSCs, harvesting hPSCs as a single cell suspension, seeding these hPSCs into 96-well plates with arrayed extra-cellular matrix (ECM), subsequently applying various control conditions, and then fixing the cells and performing fluorescent image based analysis to obtain single-cell measurements ("high-content analysis" (Singh, Carpenter et al. 2014), HCA) (Figure 4- 1a). Simultaneous tracking of four major developmental subpopulations is enabled by costaining for Oct4 and Sox2 protein, which we have previously demonstrated forms a binary code that discriminates between pluripotent (Oct4+Sox2+), early neuroectoderm (Oct4-Sox2+), mesendoderm (Oct4+Sox2-), and extra-embryonic/other populations (Oct4-Sox2-) (Nazareth, Ostblom et al. 2013) (Figure 4- 1b). We assume no access to traditional high-throughput
screening facilities (robotics) or proprietary software. For HCA we provide several open-source resources, including our in-house developed open-source ContextExplorer software.

**Figure 4-1 Protocol overview**

(A) HPSCs are first expanded, until ready for passaging. In parallel, 96-well plates with ECM patterned in 200 µm diameter spots, arrayed with 500 µm pitch is manufactured, either using microcontact printing or UV-lithography based approaches. A single cell suspension of hPSCs is generated and seeded onto the plates, and left to incubate for 6 h. Plates are then washed with PBS, and test conditions can be applied to wells, and incubated for an additional 42 h. Plates are then fixed, permeabilized, and high-content analysis is performed to determine Oct4 and Sox2 single-cell protein expression. (B) Oct4 and Sox2 form a binary code that discriminates 4 early hPSC cell fates.
Following this protocol, hPSC cell lines were assayed on patterned 96-well plates in two independent laboratories. (A) VUB14 cell line (VUB). (B) NIBSC5 cell line (UKSCB).

Figure 4- 2 Sample x-y- scatters of all cells in 96-well plates
A unique aspect of this work is that the protocol has been developed and validated though close collaboration with several laboratories in the ISCI consortium. Patterned 96-well plates were manufactured at the University of Toronto (UT), shipped to the UK Stem Cell Bank (UKSCB) and the Free University Brussels (VUB) where assays were performed, and fixed plates were shipped back to UT for analysis. Iterations of this process proved critical for refinement of the protocol. Following the protocol, at the VUB the VUB01, VUB02, VUB07, and VUB14 cell lines were assayed for response to control conditions, with cells seeded at 4 concentrations. For the VUB14 assay, the x-y- scatters of every cell in a 96-well plate are shown in Figure 4- 2a. Similarly, at the UKSCB the NIBSC5 cell line was assayed, and the resulting x-y- scatters are shown in Figure 4- 2b. Based on patterning fidelity, optimal cell seeding density was selected for further analysis. Sample colony images from NIBSC5 and VUB lines are shown in Figure 4- 3. Diverse response across control conditions were observed, in agreement with expected results. Note that cell line variation to controls is expected, and can be quantified. For the four VUB cell lines, we show the response data, which matches expected results (Figure 4- 4a). Pluripotency maintenance control media (NS) maintains high Oct4+Sox2+ expression and trophectoderm/extra-embryonic endoderm inducing control media (B) dramatically reduces the frequency of Oct4+Sox2+ cells. Neuroectoderm inducing control media (FTi) upregulates %Oct4-Sox2+, although this varies across cell lines and also in a seeding density-dependent manner (Figure 4- 4b). For example, the VUB14 line, neuroectoderm induction is negatively correlated with higher seeding density (Pearson’s $r = -0.92$, $p = 5.8 \times 10^{-7}$). The %Oct4+Sox2- mesendoderm population is only enhanced in the B and mesendoderm inducing control media (BA) conditions as expected. B also upregulates the %Oct4-Sox2- population as expected. Lastly, we show that the results from the screen can be transformed to a Mesendoderm Prediction Index (see Equation 1), which we have shown can predict day 5 endoderm induction efficiency (Nazareth, Ostblom et al. 2013) (Pearson's $r = 0.89$, $p = 0.0001$) (Figure 4- 5). This work provides proof of principle that high-throughput micro-technology based HCA assays can be transitioned to a larger non-specialized scientific community.
Figure 4: Subpopulation frequency in response to control conditions.

(A) The 4 Oct4/Sox2 subpopulations are shown in response to 5 control conditions. Data for 64,000 cells per well initial seeding density is shown. Error bars indicate standard deviation. N=4. (B) Cell lines vary in neuroectoderm induction efficiency. In the VUB14 line, neuroectoderm induction is negatively correlated with higher seeding density (Pearson’s $r = -0.92$, $p = 5.8 \times 10^{-7}$). N=4.
The screen readouts can be used to calculate the Mesendoderm prediction index, which is correlated with definitive endoderm (%Foxa2+Sox17+) frequency. Previously measured results are shown in blue, and %Foxa2+Sox17 predictions are shown in red.

In our previous work (Nazareth, Ostblom et al. 2013) we have demonstrated several applications of this assay that can be of utility to the larger hPSC scientific community. By assessing the response of a panel of hPSC cell lines to a defined set of controls, we have shown that the phenotypic response fingerprint of each line is stable between passages, variable between cell lines, predicts endoderm induction efficiency, can be used to improve differentiation of cell-lines reticent to cardiac induction, can be used to group similar cell lines (as defined by response to exogenous stimuli), can rank cell lines by lineage induction efficiency. HPSCs that maintain pluripotent markers in the presence of strong differentiation signals can be efficiently flagged as being potentially neoplastic/karyotypically abnormal. This assay can also be used for dose curves, and we have shown by performing dose curves for 27 developmentally related signaling factors. By staining the populations induced in the control conditions with a library of antibodies the assay can also assess single-cell expression of a...
large panel of proteins. By assessing responses to pathway agonists and antagonists and comparing to baseline conditions, this assay can be used to quantify baseline Activin, FGF, EGF, Wnt, and BMP levels. We have also used this assay to compare ECM, media, and media components in optimization studies. Additionally (in Chapter 3), we have shown this assay can be used to screen kinase inhibitor libraries (400 compounds at 2 concentrations), and for mechanism of action studies (by screening leads with various pathway agonists and antagonists). The assay can be adopted minimally to also screen for the effects of siRNA knockdown, with and without additional exogenous cues. This assay is also compatible with live-cell imaging, for example by using live-cell calcein staining instead of Oct4 and Sox2 staining this assay has been modified to readout cell viability. In the future, linking gene expression and other "omics" data to reproducible response data from diverse hPSC lines may also enable a better mechanistic understanding of cell-line differences (drug-gene relationships). Lastly, we note that the protocol detailed here can be modified for diverse applications on many cell types as in addition to hPSCs we have performed similar assays by patterning various cell types (Nazareth, Ostblom et al. 2013) including hPSC derived endoderm, hPSC derived cardiomyocytes, and mouse epiblast stem cells (mEpiSCs).

4.1.3 Materials

4.1.3.1 Reagents

WA-09 (H9) hESC line (passage 35–45) or other hPSC lines

Note: All relevant institutional and governmental regulations for the use of hPSCs must be followed.

Mitomycin C–treated murine embryonic fibroblasts (MEFs; Chemicon, cat. no. PMEF-N)

Gelatin solution, 0.1% (wt/vol) (Millipore, cat. no. ES-006-B)

Growth Factor Reduced Matrigel® (Matrigel; BD Biosciences, cat. no. 354230) Note: This should be separated into aliquots of 1 mL and stored at −20 °C.

DMEM/F12 medium (Invitrogen, cat. no. 11330-032)

KnockOut Serum Replacement (KSR; Invitrogen, cat. no. 10828-028)
Nutristem® hESC XF (Biological Industries, cat. no. 05-100-1A)

STEMdiff™ APEL™ Medium (APEL; Stemcell Technologies, catalogue #05210)

l-Glutamine (Invitrogen, cat. no. 21051-016)

Penicillin-streptomycin liquid (Invitrogen, cat. no. 15140122)

MEM minimum non-essential amino acids solution (NEAA; Invitrogen, cat. no. 11140-050)

2-Mercaptoethanol (Invitrogen, cat. no. 21985-023) Note: 2-Mercaptoethanol is toxic if inhaled, ingested or when in contact with skin.

L-Ascorbic acid (Sigma, cat. no. A4403)

Trace Elements A (Corning, cat. no. 25-021-CI)

Trace Elements B (Corning, cat. no. 25-022-CI)

Trace Elements C (Corning, cat. no. 25-023-CI)

TrypLE™ Express Enzyme (1X) (TrypLE, Invitrogen, cat. no. 12605036)

FGF2 (R&D Systems, cat. no. 233-FB-001MG/CF)

BMP4 (R&D Systems, cat. no. 314-bp)

Activin A (R&D Systems, cat. no. 338-AC-01M)

Insulin (Sigma-Aldrich, cat. no. I9278)

Transferrin (Invitrogen, cat. no. 11107-018)

Rho kinase inhibitor Y-27632 (Tocris Bioscience, cat. no. 1254)

TGFβ receptor inhibitor SB-431542 (Tocris Bioscience, cat. no. 1614)

Trypan blue solution (Sigma-Aldrich, cat. no. T8154)
PBS (Wisent, cat. no. 311-010-CL)

Bovine serum Albumin, 20% (Multicell, cat. no. 809-098-Q1)

DMSO (Sigma-Aldrich, cat. no. D2650)

FBS (Invitrogen, cat. no. 16140-071)

Primary Oct4 antibody (BD Biosciences, cat. no. BD611203)

Primary Sox2 antibody (R&D Systems, cat. no. MAB2018)

Oct4 secondary antibody (Alexa Fluor 555 Goat anti-Mouse IgG1, Molecular Probes, cat. no. A-21127)

Sox2 secondary antibody (Alexa Fluor 647 Goat anti-Mouse IgG2a, Molecular Probes, cat. no. A-21241)

Hoechst 33342 (Sigma-Aldrich, cat. no. H1399)

Paraformaldehyde solution (EMD Chemicals, cat. no. FX0410-5)

100% Methanol, ACS grade (EMD Chemicals, cat. no. MX0475)

Sterile double distilled water (ddH2O, filtered using Millipore Q-Gard11, cat. no. QGARDOOR1)

4.1.3.2 Equipment

Sterile biosafety tissue culture hood

Chemical fume hood

Sterile tissue culture incubator (37 °C, 5% CO2 and 100% humidity)

Cell culture centrifuge

Plate-adapted inverted microscope

Glass hemocytometer
Cell culture dishes (10 cm)
Falcon tubes (15 ml)
Falcon tubes (50 ml)
A centrifuge to spin the cells
Bottle top filters (0.22 μm) for all media (Millipore, cat. no. SCGPT02RE)
Tissue culture cell scraper (Sarstedt, cat. no. 831830)
Cell strainer (40 μm)
12-channel multichannel pipette
sterile 50 ml reagent reservoirs (VWR International, cat. no. 89094-680)
96-well Deep Well plates, 1.2 ml volume per well, sterile/autoclaved (Deep Well plates; VWR International, cat. no. 82006-448)
Automated microscope (e.g. Cellomics™ Arrayscan VTI)

4.1.3.3 Reagent setup

Note: All media should be made in a sterile biosafety tissue culture hood and then sterile filtered (0.22 μm), unless otherwise noted.

20% KOSR medium (100 ml) Mix DMEM/F12 (93 ml) and 20% (vol/vol) KSR (20 ml). This medium should be kept at 4 °C and used within 1 month.

Seeding media (100 ml) Mix Nutristem® hESC XF, 1x Penicillin-streptomycin (1 ml), and 10 μM Y-27632. This medium should be kept at 4 °C and used within 1 day.

Serum-free base media (SF) (1000 ml) Mix DMEM/F12 (930 ml), 0.5% (vol/vol) BSA (25 ml), 1x NEAA (10 ml), 1x Penicillin-streptomycin (10 ml), 0.1 mM 2-Mercaptoethanol (10 ml), 50 μg mL⁻¹ ascorbic acid, 1x Trace Elements A (1 ml), 1x Trace Elements B (1
ml), 1x Trace Elements C (1 ml), and 10 µg mL⁻¹ transferrin, and 7 µg mL⁻¹ insulin. This medium should be kept at 4 °C and used within 1 week.

Note: As an alternative, APEL media can be used with 1x Penicillin-streptomycin.

**Pluripotency maintenance control media (NS) (10 ml)** Mix Nutristem® hESC XF and 1x Penicillin-streptomycin (1 ml). This medium should be kept at 4 °C and used within 1 day.

Note: Alternatively, mTeSR1™ or MEF-conditioned media can be used.

**SF with BMP4 - trophectoderm/extra-embryonic endoderm inducing control media (B) (10 ml)** Mix SF with 100 µg mL⁻¹ BMP4. This medium should be kept at 4 °C and used within 1 day.

**SF with BMP4 and Activin A - mesendoderm inducing control media (BA) (10 ml)** Mix SF with 10 µg mL⁻¹ BMP4 and 100 µg mL⁻¹ Activin A. This medium should be kept at 4 °C and used within 1 day.

**SF with FGF2 and SB-431542 - neuroectoderm inducing control media (FTi) (10 ml)** Mix SF with 40 µg mL⁻¹ FGF2 and 10 µM SB-431542. This medium should be kept at 4 °C and used within 1 day.

**1:40 GFR Matrigel** Thaw aliquot of Matrigel (1 mL) overnight at 4 °C and add DMEM/F12 (39 ml). This medium should be kept at 4 °C and used within 1 day. Note: At all times outside the fridge this medium should be kept on ice.

**3.7% Formaldehyde** In a fume hood, prepare 3.7% (vol/vol) formaldehyde in PBS. Note: Formaldehyde is toxic.

**10% FBS blocking solution** Prepare 10% FBS in PBS. Note: no filtering required.

4.1.4 Procedure

**Expansion of hPSCs**
1. Culture 6 10-cm dishes of hPSCs using standard culturing methods, either on feeders or feeder free.

Note: It is critical to maintain the hPSCs as undifferentiated, and low-passage numbers are preferable. Cells should be used when ready for passaging, and when not too over-confluent. Seeding one 96-well plate will require 5.4 million cells, approximately 3 10-cm dishes.

Note: Monolayer cultures, especially single-cell adapted hPSCs, may require longer incubation in TrypLE to achieve a single cell suspension. The TrypLE incubation time can be reduced by using these cells earlier before confluence.

**Preparation of 96-well plates with patterned GFR Matrigel**

2. Pattern 1:40 GFR Matrigel into 96-well plates arrayed in 200 µm diameter spots, with 500 µm pitch (center-center distance) using micro-contact printing (Azioune, Storch et al. 2009). Store patterned plate with PBS in wells to ensure Matrigel does not dry out, and store at 4 °C.

Note: A detailed protocol for micro-contact printing Matrigel onto slides is available in ref. (Peerani, Bauwens et al. 2009), which can be simply modified to pattern Matrigel directly into standard tissue culture polystyrene 96-well plates as described in ref. (Nazareth, Ostblom et al. 2013). Alternatively, a UV-lithography based approach for manufacturing 96-well plates with patterned extracellular matrix proteins is available in ref. (Azioune, Storch et al. 2009). This hPSC assay protocol has been validated in-house using micro-contact printing, and validated across multiple laboratories UV-lithography (modified version of ref. (Azioune, Storch et al. 2009), manuscript in preparation). For using UV-lithography plates instead of micro-contact printed, the only modification needed is an alternative washing step (see note in step 21).

**Generating a single cell suspension of hPSCs**
3. Warm TrypLE in a 37 °C water bath. Warm 20% KOSR medium, seeding media, and 96-well plates with patterned Matrigel at room temperature in a tissue culture hood.

4. Remove media from hPSCs, and add 0.5 ml TrypLE per 10 cm dish. Ensure TrypLE coats all cells by rocking the dish back and forth, and then immediately transfer to a 37 °C incubator.

   Note: Steps 4-13 should be completed quickly, as hPSCs are extremely sensitive to cell death and differentiation while in a single cell suspension.

5. Incubate hPSCs in TrypLE in the incubator for 4 minutes. Ensure cells are detached by rocking the plate and visually inspecting that cell sheets do indeed lift off. Cell sheets should detach when the vessel is shaken.

   Note: Cell adhesion after seeding is very sensitive to TrypLE, cells will not adhere properly to patterned surfaces if TrypLE is left for too long. Even 5 minutes can result in cells not adhering at the 6 h timepoint.

   Note: Visual inspection that TrypLE has been applied for sufficient time is especially critical for over-confluent and/or monolayer cultures.

6. Transfer hPSCs back to the tissue culture hood and quench the TrypLE by adding 2.5 ml 20% KOSR medium per well.

   Note: KOSR containing media must be used for this step, SF media is insufficient for quenching the TrypLE.

7. For each dish, use a cell scraper to loosen all cells off the vessel surface, and then transfer all media and cells into a 50 ml Falcon tube. Pipette several times to mechanically disassociate all clumps and create a single cell suspension.

8. Spin down cells in the 50 ml Falcon tube, for 5 minutes at 200 g (approx. 1000 rpm), to make a cell pellet. Transfer hPSCs back to the tissue culture hood and carefully aspirate supernatant.
Note: As the supernatant contains TrypLE it is imperative to remove as much as possible. Tilt the 50 ml Falcon tube to ensure all aspiration occurs far away from the cell pellet.

9. Resuspend cells in seeding media, ensuring that the cell density is at least 1.2 million cells per ml (e.g. resuspend in 1 ml per 10 cm input hPSC dish). Optionally, if cell clumps are observed the suspension can be filtered with a 40 µm strainer.

10. Transfer hPSCs from the 50 ml Falcon tube into a sterile 50 ml reagent reservoir ("cell reservoir").

11. Agitate cell suspension, ensuring cells are homogeneously mixed, and immediately take a 20 µl sample for a cell count. Repeat.

Note: HPSCs settle quickly, and an accurate cell count is imperative, so cell samples for counts must be taken with great attention.

12. Count cells, and then dilute hPSC suspension with additional seeding media so a final concentration of 100,000 cells per 100 µl is obtained.

Seeding hPSCs into patterned 96-well plates (T=0)

13. Using a multichannel pipette, transfer the PBS from the patterned 96-well plates into a 50 ml reagent reservoir ("waste reservoir").

Note: A multichannel pipette should be used for all steps involving transferring to and from the patterned 96-well plates.

14. Different volumes of the cell suspension will be seeded into the rows of the patterned 96-well plate, but first cell-free seeding media must be transferred into the rows so the total volume after adding the cell suspension is 100 µl. Transfer 20 µl per well of cell-free seeding media into row C of the 96-well plate. Similarly transfer 36 µl per well into row D, 52 µl per well into row E, and 68 µl per well into row F of cell-free seeding media into the 96-well plate.
Note: Alternatively if a full plate is to be used, transfer 20 µl per well into rows A and E, 36 µl per well into rows B and F, 52 µl per well into rows C and G, and 68 µl per well into rows D and H.

15. Transfer 80 µl of hPSC cell suspension into each well of row C of the 96-well plate, to obtain 80,000 cells seeded per well. Similarly transfer 64 µl per well (64,000 cells per well) into row D, 48 µl per well (48,000 cells per well) into row E, and 32 µl per well (32,000 cells per well) into row F of the 96-well plate.

Note: Alternatively if a full plate is to be used, seed 80 µl per well into rows A and E, 64 µl per well into rows B and F, 48 µl per well into rows C and G, and 32 µl per well into rows D and H.

Note: Seeding at a range of densities will mitigate issues arising from too low or too high cell densities. Typically only one row will be at a suitable density for appropriate cell patterning and phenotypical response to control conditions, the rest of the rows will not be suitable for extracting usable response data.

16. Ensure all wells with hPSCs have a 100 µl volume. Fill all unused wells with 300 µl PBS.

17. Leave the plate out at room temperature for exactly 20 minutes, allowing most of the cells to settle evenly, and then gently transfer the plates in the tissue culture incubator. Avoid keeping the plate out for much longer than 20 minutes, this affects the cells negatively. Incubate the plates for 6 h.

Note: If plates are immediately moved to the incubator, then convection currents move suspended cells towards the edges of the well, resulting in uneven patterning. Room temperature incubation allows cells to settle before being transferred to the incubator. Minimize disturbing the plate, as this may shift the cell distribution. Chose an incubator with minimal traffic (door opening and closing) and place plates near the back of the incubator. Do not stack plates.
Addition of control media to wells (T=6h)

18. Warm SF, NS, B, BA, FTi at room temperature in a tissue culture hood.
   Note: B, BA, and FTi can be prepared slightly before this timepoint.
19. Transfer 500 µl per well of control media into wells of a deep-well plate, as outlined in Table 4-1.
   Note: Alternatively if a full patterned 96-well plate is to be used, transfer 1000 µl per well.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
<td>FTi</td>
<td>SF</td>
<td>SF</td>
<td>NS</td>
<td>BA</td>
<td>B</td>
<td>FTi</td>
<td>SF</td>
<td>SF</td>
<td>NS</td>
<td>BA</td>
</tr>
</tbody>
</table>

Table 4-1 Deep-well plate layout.

20. Transfer the 96-well plates with patterned hPSCs from the tissue culture incubator to the tissue culture hood. Cells should be evenly distributed across the bottom of each well. Optionally, take bright field images of wells seeded with different densities. Transfer PBS into a 50 ml reagent reservoir ("PBS reservoir").
   Note: Steps 20-24 should be performed quickly so that hPSCs are not left out of the incubator for too long. For multiple plates one plate at a time should be washed and fed.
21. To wash the patterned wells of non-adherent hPSCs, transfer all media from the wells containing hPSCs in the patterned 96-well plates into a new 50 ml waste reservoir ("waste reservoir"), and transfer 100 µl of PBS per well into the now empty wells. Repeat this wash step once more.
   Note: All wash steps on 96-well plates should be performed so that the multichannel tips do not scrape the patterned surface of the well. Media should be removed from the bottom edge of the well, and added gently.
   Note: For UV-lithography plates, this wash step is modified as follows. Without removing any media, add 100 µl of PBS and remove 100µl from the solution. This dilutes out the media removes non-adherent cells. Repeat 4-5 times.
22. Remove all PBS from wells containing hPSCs. Arrays of patterned hPSCs should be visible.

23. Using the layout of control media shown in Table 4-2, use a multichannel pipette to transfer 100 µl of control media one row at a time from the deep-well plate onto rows C and E. To reverse the order of the controls, rotate the patterned hPSC plate and transfer control media into row D and F.

Note: If using a full patterned 96-well plate, similarly transfer the controls from the deep-well plate in an alternating fashion. In this way a full 96-well plate can be fed in 8 pipette steps.

Table 4-2 Layout of control media applied to patterned hPSCs.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>B</td>
<td>FTi</td>
<td>SF</td>
<td>SF</td>
<td>NS</td>
<td>BA</td>
<td>B</td>
<td>FTi</td>
<td>SF</td>
<td>SF</td>
<td>NS</td>
<td>BA</td>
</tr>
<tr>
<td>D</td>
<td>BA</td>
<td>NS</td>
<td>SF</td>
<td>SF</td>
<td>FTi</td>
<td>B</td>
<td>BA</td>
<td>NS</td>
<td>SF</td>
<td>SF</td>
<td>FTi</td>
<td>B</td>
</tr>
<tr>
<td>E</td>
<td>B</td>
<td>FTi</td>
<td>SF</td>
<td>SF</td>
<td>NS</td>
<td>BA</td>
<td>B</td>
<td>FTi</td>
<td>SF</td>
<td>SF</td>
<td>NS</td>
<td>BA</td>
</tr>
<tr>
<td>F</td>
<td>BA</td>
<td>NS</td>
<td>SF</td>
<td>SF</td>
<td>FTi</td>
<td>B</td>
<td>BA</td>
<td>NS</td>
<td>SF</td>
<td>SF</td>
<td>FTi</td>
<td>B</td>
</tr>
<tr>
<td>G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

24. Transfer plates back into the tissue culture incubator. Optionally, take bright field images of wells seeded with different densities, hPSCs should be visibly patterned after the wash steps, with little or no cells between the spots.

Fixing and staining for immunocytochemistry (ICC)

25. Transfer the 96-well plates from the tissue culture incubator into a chemical fume hood.

26. Transfer PBS, 3.7% formaldehyde, 100% methanol, and 10% FBS blocking solution into separate 50 ml reagent reservoirs in the fume hood.

27. Remove all media from the plates into a 50 ml reagent reservoir ("cell waste reservoir").
28. To fix cells, add 50 µl 3.7% formaldehyde per well. Incubate the plates at room temperature for 30 min.

29. Remove all formaldehyde from the plates and transfer the formaldehyde into a new 50 ml reagent reservoir ("formaldehyde and methanol waste").

30. To permeabilize cells, add 50 µl 100% methanol per well. Incubate the plates at room temperature for 3 min.

31. Remove all methanol from the plates and transfer the methanol into the formaldehyde and methanol waste reservoir.

32. Wash by adding 100 µl PBS, then transferring the PBS to the formaldehyde and methanol waste reservoir. Perform two additional PBS wash steps and transfer the PBS waste into a new 50 ml reagent reservoir ("PBS waste")

   Note: PBS from the first wash is treated as formaldehyde and methanol waste.

   Ensure to follow your institutions waste guidelines for all formaldehyde, methanol, and cell waste.

   Note: At this point 100 µl PBS can be added to all wells and the plates can be stored at 4°C. Ideally ICC staining and analysis should be completed within a week, however fixed plates can be stained up to months later, with some reduction in resulting image quality.

33. To block non-specific adhesion of primary antibodies to cells, add 50 µl 10% blocking solution per well. Incubate for 1h at room temperature, or overnight at 4°C.

34. Dilute primary Oct4 and Sox2 antibodies 1:500 in 10% FBS blocking solution. Add 35 µl of diluted antibody solution per well. Incubate for 1h at room temperature, or overnight at 4°C.

35. Wash plates 3 times with PBS.

36. Dilute Hoechst 33342 (1:1000), 555 Goat anti-Mouse IgG1 secondary antibody (1:500), and 647 Goat anti-Mouse IgG2a secondary antibody (1:500) in 10% FBS blocking solution. Add 35 µl of diluted Hoechst/secondary antibody solution per well. Incubate for 1h at room temperature in the dark.

   Note: Leaving secondary solutions for longer will increase non-specific staining.
37. Wash plates 3 times with PBS. Leave PBS in wells and store plates at 4°C and in the dark until ready for fluorescent image acquisition.

Note: Image acquisition should ideally be performed immediately after secondary staining, but plates can be left for up to 1 week.

**Image acquisition and high-content analysis**

38. Image DNA, Oct4, and Sox2 using DAPI (350 nm excitation max, 470 nm emission max), TRITC (555 excitation max, 580 emission max), and Cy5 (650 excitation max, 665 emission max) common filter sets respectively. Images can be saved manually for small scale experiments, but is ideally performed on an automated microscopy system, such as the Thermo Scientific™ Arrayscan™ VTI high content platform. Four fields per well using a 10x objective is sufficient, or ideally more fields if scan times are not an issue.

Note: Visually inspect wells to ensure staining, general patterning fidelity, and cell viability is suitable. Controls can be rapidly visually compared to obtain qualitative phenotypes (e.g. B abolishes Oct4 and Sox2 expression, NS maintains Oct4 and Sox2 expression).

39. Perform single cell analysis using suitable software identify cells and quantify single cell Oct4 and Sox2 average intensities. Several open-source solutions are available for image analysis including ImageJ (Collins 2007) (http://imagej.nih.gov/ij/) for routine image processing, or CellProfiler (Carpenter, Jones et al. 2006) (http://www.cellprofiler.org/) specifically designed to analyse images to quantifying cell phenotypes.

Note: If using the Arrayscan™ platform, the Target Activation algorithm can automatically outline single cell nuclei and quantify Oct4 and Sox2 average nuclear intensities.

Note: For downstream analysis it is recommended to export the single cell Oct4 and Sox2 intensity data in a plain text file (e.g. a comma-separated values (CSV) file).

40. Visually inspect images to determine suitable Oct4 and Sox2 intensity thresholds.
41. Analyse single cell data to verify Oct4 and Sox2 intensity thresholds, and determine %\textit{Oct}4+Sox2+, %\textit{Oct}4+Sox2-, %\textit{Oct}4-Sox2+, and %\textit{Oct}4-Sox2- per well.

Note: CellProfiler Analyst (Jones, Kang et al. 2008) (http://www.cellprofiler.org/) is open-source software for data analysis and exploration of large screens image sets, and is a companion tool to CellProfiler (Carpenter, Jones et al. 2006). We have also developed open-source software, ContextExplorer (https://github.com/joelostblom/context_explorer, manuscript in preparation) which is particularly streamlined for automatic detection of colonies from single cell data, and analysis of % Oct4+Sox2+, #Oct4+Sox2+ etc. on a well-level and per-colony level. Either software accepts CSV files of single cell data as input.

42. Average replicates and analyse results.

Note: Regardless of the software used for high-content analysis, to explore the generated data we also recommend using drag and drop visualization tools, such as free Tableau Public software (https://public.tableau.com/s/).

43. The following equation can be used to obtain the Mesendoderm Prediction Index, which can be compared to other cell lines in ref. (Nazareth, Ostblom et al. 2013) and is shown to be highly correlated to day 5 endoderm induction efficiency (Nazareth, Ostblom et al. 2013):

\[
\text{Mesendoderm Prediction Index} = \frac{BA \text{ condition } \%Oct4 + Sox2 -}{NS \text{ condition } \%Oct4 - Sox2 -}
\]

4.1.5 Author Contribution

E.J.P.N. designed the protocol, coordinated experiments, and analyzed all experiments. M.T. developed and manufactured UV-lithography patterned 96-well plates and adapted the protocol for such plates. O.O. and S.K. performed all experiments at the UKSCB. L.H. and G.S. supported all work at the UKSCB. D. D. and M. G. performed all experiments at the VUB. E.J.P.N. and P.W.Z. designed the project and wrote the manuscript.

4.1.6 Acknowledgements
This work is funded by the Canadian Institutes of Health Research (CIHR) (P.W.Z). E.J.P.N. is supported by a CIHR Frederick Banting and Charles Best Canada Graduate Scholarships Doctoral Award. P.W.Z. is supported as the Canada Research Chair in Stem Cell Bioengineering. Collaborative efforts were supported through funding from the International Stem Cell Initiative. We thank Peter Andrews, Steve Oh, and Martin Pera for helpful discussion.

4.2 ContextExplorer: Open-source software for spatial analysis in high-throughput screens

4.2.1 Abstract

A growing body of evidence highlights the importance of the cellular microenvironment as a regulator of phenotypic and functional cellular responses to perturbations. ContextExplorer is a software tool to facilitate investigation of microenvironmental parameters, such as local cell density and position within a colony, and their influence on protein expression and cellular phenotype in high content image analyses. The software is graphical user interface (GUI) based for ease of use by a broad scientific community, and allows for batch processing of high-content analysis (HCA) multi-well plate data making it amenable for analysis of high-throughput (HTP) screen data. We demonstrate the utility of ContextExplorer in the analysis of human pluripotent stem cells (hPSCs) patterned in colonies of defined size and shape within several multi-well plates. Automatic colony identification correlates well with manual colony count (Pearson $r = 0.99$). The software can also automatically count hPSC colonies grown on MEFs (Pearson $r = 0.91$). Automatic detection and virtual superimposing of such colonies enables statistical analysis of spatial trends, revealing that Oct4 and Sox2 intensity levels vary dependent on their position within the colony. Cells expressing high levels of these proteins are most frequently located in the interior of colonies, whereas cells with lower expression tend to reside along the border of the colony. By classifying cells based on Oct4 and Sox2 expression, we find that pluripotent cells are more prevalent in the center of the colonies and extraembryonic fated cells increase in frequency towards the colony perimeter. Lastly, we apply the software to automatically extract data for quality control (QC) of HTP screens by extracting cells per colony, cells per well, and colonies per well in a multi-plate screen, which revealed patterning fidelity differences between several hPSC cell lines. ContextExplorer can
readily be used in conjunction with various existing image analysis programs such as CellProfiler. It is distributed as an executable for Windows, written in Python and Qt under an open-source license. ContextExplorer is a powerful, simple to use, and HTP-compatible tool to seamlessly interrogate spatially variable response data from drug development and fundamental research programs.

4.2.2 Introduction

Emerging evidence stresses the importance of a cell's local microenvironment as a regulator of cellular phenotype and heterogeneity within cell populations. Microenvironmental parameters such as mechanical forces, cell to cell contact and endogenous signaling all vary between cells at different positions in a well (McBeath, Pirone et al. 2004, Peerani, Rao et al. 2007). Such spatially heterogeneous factors have been shown to explain the majority of variability in the efficiency of endocytosis and the vulnerability to viral infection (Snijder, Sacher et al. 2009), influence epithelial tissue growth (Kim, Kushiro et al. 2009), impact the expression of angiogenic factors in tumor cells (Kumar, Kuniyasu et al. 1998), and mouse (Davey and Zandstra 2006) and human PSCs (Peerani, Rao et al. 2007). Microenvironmental heterogeneity is also a potential confounding factor contributing to contradictory findings in the response to key signaling pathway activity in different cell types (Jong, Koster et al. 2009, Akopian, Andrews et al. 2010, Snijder, Sacher et al. 2012) and could limit the interpretability and reproducibility of experiments. A comparative analysis (Haibe-Kains, El-Hachem et al. 2013) of two large scale pharmacogenomic studies, the Cancer Genome Project (CGP) (Garnett, Edelman et al. 2012) and Cancer Cell line Encyclopedia (CCLE) (Barretina, Caponigro et al. 2012), revealed a surprisingly poor correlation between cell line drug response phenotypes between laboratories, which precludes meaningful extraction of drug-gene relationships. Even when matched protocols and cell lines with highly correlated gene expression profiles were used there was still a poor correlation, highlighting the need to standardize drug response assays and analysis. Although in this study the exact source of variation is unknown, a separate analysis of single cell data from 45 high-throughput (HTP) screens revealed that population context is indeed a ubiquitous source of variation in screens, and accounting for population context can improve variation between cell lines and laboratories (Snijder, Sacher et al. 2012). Understanding population heterogeneity is acknowledged as being critical to biomedical research (Altschuler and Wu 2010, Pelkmans 2012).
2012), but the adoption of controlling for microenvironmental variables as a source of such heterogeneity has been slow within the scientific community. This is partially a consequence of limited technical ability to measure and control a cell’s population context (including variability in cell distribution on surfaces) in combination with insufficient appreciation of the significance of variation in the local microenvironment.

The increasing affordability of high content screening (HCS) instruments, emergence of core screening facilities and technological advancements such as micropatterning in multi-well plates (Azioune, Storch et al. 2009), allow us to investigate population context dependent variables with unprecedented throughput and veracity (Xia and Wong 2012, Nazareth, Ostblom et al. 2013). By patterning of extracellular matrix (ECM) proteins, cells can be restricted to adhere to an array of spots of predefined shapes and sizes (Folch, Jo et al. 2000, Fink, Thery et al. 2007). An advantage of such patterning is enhanced control over microenvironmental variation within each well. HPSCs growing in patterned colonies express varying levels of pluripotency markers Oct4 and Sox2 in a colony size dependent manner (Peerani, Rao et al. 2007, Nazareth, Ostblom et al. 2013). Cells growing in colonies of defined size and shape lend themselves well to analyses of intra-colony spatial variation in cell phenotype. Elucidating the impact of the population context dependent variables on cellular phenotype will not only add to our understanding of fundamental cell biology, but will also allow us to optimize culture conditions and cell assays, provide possible explanations for current seemingly conflicting research and may provide powerful information to in silico models. These aspects are critical to next-generation drug development strategies and systems biology approaches.

We have previously developed a high-throughput (HTP) platform to homogenize microenvironmental variation between wells in a high throughput setting by having cells adhere to micropatterned ECM spots of defined shapes in multi-well plates (Nazareth, Ostblom et al. 2013). As an addition to this platform we have developed a computational tool, ContextExplorer, which facilitates colony level analyses and cell patterning quality control. There are many software solutions for analyzing features of imaged cells, both open source (Carpenter, Jones et al. 2006, Jones, Kang et al. 2008, Misselwitz, Strittmatter et al. 2010) and commercial, and some implementations are capable of identifying arrays of cells on glass slides (Bauer, Kim et al. 2012). Our software specifically aims to improve the HTP workflow of analyzing patterned multi-well plates, control for consistent patterning of cells and facilitate
colony identification and colony level analyses of patterned plates, all in a GUI-based and open-source platform. Here, we demonstrate the utility of ContextExplorer by examining the impact of population context variables on hPSC fate choices.

4.2.3 Methods

4.2.3.1 Cell culture

HESC lines H9 (WA09) were obtained from the WiCell Research Institute. H9 were routinely cultured on feeder layers of irradiated murine embryonic fibroblast (MEF) feeders in knockout (KO)-Dulbecco’s modified Eagle’s medium (DMEM) (Invitrogen) with 20 % KO-serum replacement (Invitrogen) (KO-DMEM) and supplemented with 4 ng mL⁻¹ FGF-2 (PeproTech). Cells were passaged 1:4-1:6 every 4-5 days, and were disassociated into small clumps using 0.1 % collagenase IV (Invitrogen). All cell line stocks were confirmed negative for mycoplasma contamination.

4.2.3.2 Microcontact printing of substrate into 96-well plates

We have previously developed a method for patterning proteins in standard 96-well plates (Nazareth, Ostblom et al. 2013). In brief, PDMS stamps were fabricated using standard soft lithography techniques (Kane, Takayama et al. 1999), with the exception that liquid PDMS was cast into a Teflon mould before curing, allowing control of the shape of the PDMS stamp. PDMS stamps were cast to fit directly into 96-well plates. ECM solution consisted of Matrigel diluted 1:30 in phosphate buffered saline (PBS). ECM solution was deposited onto the patterned surface of ethanol sterilized PDMS stamps for 4h at room temperature. Stamps were rinsed with ddH2O, dried gently with N2 gas, placed into tissue-culture treated 96-well plates (Costar) ensuring conformal contact between the plate and PDMS stamp, and incubated in the 96-well plates for 7-10 min in a humidity chamber (Relative humidity 55-70 %). The stamps were then removed and substrates were passivated with 5 % weight Pluronic F-127 (Sigma-Aldrich) in ddH20 for 1 h.

4.2.3.3 Seeding hPSCs onto patterned substrates

HPSCs were dissociated using TrypLE™ for three min. TrypLE™ was inactivated by adding media containing 20 % KO-serum replacement (SR) (Invitrogen). Cells were centrifuged and resuspended in Nutristem® hESC XF (Biological Industries, cat. no. 05-100-
1A) and 10 µM ROCK inhibitor Y-27632 (Tocris). SF media consists of DMEM/F12, 1x Nonessential amino acids, 50 U mL⁻¹ Penicillin, 50 µg mL⁻¹ Streptomycin, 10 µg mL⁻¹ bovine Transferrin, 0.1 mM β-Mercaptoethanol (all Invitrogen), 2 % fatty acid-free Cohn’s fraction V BSA (Serologicals), 1x Trace Elements A, B & C (Mediatech), 50 µg mL⁻¹ Ascorbic Acid (Sigma) and 7 µg mL⁻¹ recombinant human insulin. Cells were seeded at 10⁵ cells per well (or as described in text) and incubated. After 6 h, cells were washed with PBS two times and incubated a further 42 h with fresh media as indicated (SF supplemented with factors, or CM).

4.2.3.4 Immunocytochemistry and high-content image analysis

Plates were fixed for 30 min in 3.7 % formaldehyde and permeabilized for 3 min in 100 % methanol. Plates were imaged and analyzed using the Cellomics Arrayscan VTI platform and Target Activation algorithm (Thermo Scientific). This algorithm generates nuclear masks, provides single cell nuclear intensity values for protein expression (Oct4, Sox2 etc.), DNA content through Hoechst staining, as well as spatial x- and y-coordinates of the nuclei centroids enabling high-content analysis (HCA). Fluorescent images were obtained of Oct4 (1:500; BD), Sox2 (1:500 R&D Systems), and Hoechst 33342 (Sigma-Aldrich). Primary antibodies were incubated overnight in 10 % FBS in PBS at 4° C. AlexaFluor secondary antibodies (1:500; Molecular Probes) were incubated for 1 h in 10% FBS in PBS at room temperature. For figures, 16 bit TIFF images were obtained for each channel, contrast adjustment was performed identically across all controls, and channels were combined into pseudo-colored composite images. Following HCA, single cell x-y- coordinate and Oct4 and Sox2 intensity data was exported in tab delimited text files and imported into ContextExplorer for exploration of colony level details.

4.2.3.5 Statistical analysis

Error bars on plots represent 95 % confidence intervals (CI) of well replicates and are calculated through bootstrapping the population 1000 times. Error bars and Pearson correlations are computed as implemented in the Python package seaborn (https://github.com/mwaskom/seaborn/tree/v0.5.0). T-tests were performed using the Student's t-Test, with a two-tailed distribution and unequal variance.

4.2.3.6 Radial analysis
ContextExplorer was used to cluster cells in colonies, using the DBSCAN algorithm (Ester, Kriegel et al. 1996). To superimpose colonies, relative coordinates to the colony mean were calculated by subtracting the mean coordinates. The single cell relative coordinates are exported as a CSV file, along with distance from the mean, and this distance in terms of percent of the maximum colony radius. To plot the spatial trend in protein expression values, the relative coordinates were binned in ten groups from the colony center to the edge and the mean of each group was plotted.

4.2.4 Results

4.2.4.1 ContextExplorer enables automated colony Identification

To apply ContextExplorer towards analysis of intra-colony transcription factor trends in hPSCs, we first patterned hPSCs in 200 µm diameter colonies in 96-well plates and assessed their response to 42 h of SF media alone which induces a mix of pluripotent and neuroectoderm cells (Nazareth, Ostblom et al. 2013), SF with added BMP4 (SF+B) which induces trophectoderm and primitive endoderm (Xu, Chen et al. 2002, Vallier, Touboul et al. 2009) or MEF conditioned media (CM) which maintains pluripotency (Xu, Inokuma et al. 2001). At the end of the assay, single cell Oct4, Sox2, and x-y-coordinates were obtained using HCA, and this data was imported into ContextExplorer (Figure 4- 6a). Oct4 and Sox2 can be thresholded to classify cells at early timepoints (48 h) into lineage subpopulations: Oct4+Sox2+ for pluripotent, Oct4+Sox2- for primitive streak, Oct4-Sox2+ for neuroectoderm, Oct4-Sox2- for extraembryonic and other tissue. Given thresholds, ContextExplorer can perform this classification and obtain, for example, number and percentage of Oct4+Sox2+ cells in each colony in a set of multi-well plates.
**a**

0 h: Seed hPSC suspension into patterned wells
6 h: Wash and apply test conditions
48 h: Fix, permeabilize, stain and image
Automated colony analysis
Post analysis

**b**

[Image of a software interface with settings and options for colony detection and analysis]

**c**

Images showing different stages of colony growth and development.

**d**

Images with annotations indicating different categories of colonies.

**e**

Graph showing correlation with Pearson r = 0.99; p = 3.2e-95

**f**

Graph showing correlation with Pearson r = 0.91; p = 1.2e-95
A) Experimental timeline. Following standard immunocytochemistry procedures, samples are imaged using high-content analysis (HCA) approaches to obtain single-cell x-y coordinates and marker expression intensity values. These values are exported and processed by ContextExplorer to threshold individual cells as responders (e.g. Oct4+ or Oct4-) and cells are clustered into colonies using user-defined parameters. Colonies can be excluded from analysis using user-defined thresholds. Well-, colony-, and single cell-level analysis is exported for further analysis (“post analysis”). B) GUI of the ContextExplorer. Plates can be processed in batches. C) Patterned hPSCs colonies stained with Oct4 (green), Sox2 (red) and DAPI (Blue). Scale bar indicates 200 µm. 9 fields imaged using a 10x objective lens were stitched together. D) x-y scatter plots of single cell nuclei locations, and clustering of these cells into colonies. Color can be used to indicate assigned phenotype as indicated. E) Comparison of manual colony count and ContextExplorer automatic count for patterned hPSCs. N=120 wells in two 96-well plates. F) Comparison of manual colony count and ContextExplorer automatic count for hPSC colonies (Oct4+) grown in MEFs (Oct4-). N=54 wells of a 96-well plate with H9 hPSCs growing in MEFs.

A cell colony can be defined as a group of cells clustered together at a higher density than the surrounding cells. Especially in ECM patterned wells, this is suitable for separating colonies from each other. There are many existing algorithms for identifying clusters of spatial data points (Ester, Kriegel et al. 1996, Ankerst, Breunig et al. 1999, Xu and Wunsch 2010, Pedregosa, Varoquaux et al. 2011, Karami and Johansson 2014). As an alternative to automatic parameter estimation, ContextExplorer allows for Eps and MinPts to be adjusted interactively and visual feedback is immediately given through the graphical user interface (GUI) (Figure 4-6b). ContextExplorer also includes filters for colony size, density and roundness to refine the colony identification procedure. Upon rapid optimization of the clustering parameters, one can obtain nearly identical results to visually inspecting the images from each well in a fraction of the time (Figure 4-6c-e). The colony identification is also accurate for hPSCs grown on a feeder layer of MEFs using standard culture techniques (e.g. not patterned) (Figure 4-6f). For hPSCs on feeders, an Oct4 threshold is simply applied to exclude Oct4- cells (such as MEFs) before clustering the remaining Oct4+ cells into colonies.

4.2.4.2 Comparison of cells within colonies and not in colonies

The SF, SF+B, and CM conditions induced differences in Oct4 and Sox2 expression, but within wells there are also variations in these transcription factors (Figure 4-7a-b). To analyse the spatial organization of this variation, we first compared the absolute expression levels of Oct4 and Sox2 between cells inside and outside colonies in three conditions (Figure 4-7c-d). In the presence of BMP4 in the growth media, neither Oct4 nor Sox2 expression is induced and the expression values are similar for cells outside and inside colonies. The CM condition induces the expression of both Oct4 and Sox2 in cells within colonies, while cells outside colonies express the markers to a lesser extent. This is visible both in a change of the
distribution shape and a shift in means of the Sox2 and Oct4 intensities. In SF media, Sox2 expression is elevated in cells inside colonies compared to cells that adhered outside patterns, while this distinction is less pronounced for Oct4. These differences in absolute fluorescent intensities suggest that cells inside and outside colonies do not respond similarly to added factors in the media, or it could be a result of differential cell migration.
Figure 4-7 Differences in fluorescent protein intensities between cells inside and outside colonies

A) Density distributions of the Sox2 intensity values of the CM, SF, and SF-B conditions. B) Density distributions of the Oct4 intensity values of the CM, SF, and SF-B conditions. C) Density distributions and means of Sox2 intensity values comparing cells inside and outside colonies within the same condition. D) Density distributions and means of Oct4 intensity values comparing cells inside and outside colonies within the same condition. E) Frequencies of
each cell type compared between cells inside and outside colonies for each condition. Error bars represent 95% CI between well replicates.

To link these changes in transcription factor intensities to specific cell fates, we next compared the frequencies of the four early lineage subpopulations, which further highlights the difference between cells inside and outside colonies (Figure 4-7e). In CM, over 90% of cells inside colonies maintain pluripotency while only 60% of cells outside colonies show the same phenotype. In the SF condition, pluripotency is maintained in 40% of cells inside colonies versus 20% of cells outside colonies. There are also differences in the proportion of Oct4-Sox2+ cells, which are more common inside colonies and there are five times the proportion of double negative cells outside colonies than inside colonies. No differences are seen in the SF+B condition, which is expected since this condition induces a double negative phenotype in all cells, the same as what appears to be the dominating phenotype for cells outside colonies. These data suggest that cells outside colonies are more likely to lose expression of both Oct4 and Sox2 than cells inside colonies in the same condition. However, the phenotype frequencies of the cells outside colonies still vary significantly between conditions, indicating that the population context is a contributing, but not decisive factor of hPSC fate choice in our assay.

Excluding cells adhering outside patterned ECM spots in the analysis, eliminates variation due to varying amounts of non-specifically adhered cells between replicate wells. To test if the elimination of this variable amount contributes to a higher quality assay, we determine the platform Z' score, which is a metric to assess the viability of a high throughput assay and tests the ability of the assay to distinguish between a positive and a negative response (Zhang, Chung et al. 1999).

4.2.4.3 Analysis of transcription factor intra-colony variation

Radial gradients of protein expression have been reported in non-patterned and patterned colonies of hPSCs (Davey and Zandstra 2006, Peerani, Rao et al. 2007, Warmflash, Sorre et al. 2014). Assessing spatial trends manually by visual inspection or manual data analyses is feasible in a platform with lower throughput, but becomes error-prone and time consuming in a HTP system, where hundreds of colonies need to be summarized within replicate wells in the same condition. To facilitate the assessment of whether there is spatial bias in the expression of the proteins, ContextExplorer can superimpose colonies within replicate wells and display an averaged representative colony per condition (Figure 4-8a). This analysis can be applied to
patterns of arbitrary size and shape. For circular patterns, radial analysis can be performed by binning cells within specific distances of the center. In our data set, we observe a clear radial gradient of both Sox2 and Oct4 expression in colonies in CM and SF (Figure 4-8b). To aid in visualization of spatial differences, and not differences attributed due to exogenous factors, all intensities are normalized relative to expression at the center of the colony. In CM and SF, Sox2 expression is linearly decreasing towards the edge and cells along the perimeter have only half the fluorescent intensity of cells in the center of the colony. In CM, Oct4 expression is more or less consistent up to 60% of the radius, where it starts dropping off linearly until it reaches the perimeter. However, in SF, the expression of Oct4 displays a bimodal spatial organization where cells at around 60–70% from the center on average expresses 10–20% more protein than cells in the middle. These differences are statistically significant as can be seen by the 95% confidence intervals (CI). With the addition of BMP4 in the media, the expression of Sox2 and Oct4 is stabilized throughout the colony, consistent with BMP4 inducing differentiation and overriding any local pluripotency supporting signals. Lastly, to examine how these intensity levels translate to cell fate, we applied thresholds to Oct4 and Sox2 and observed the frequency of each subpopulation as distance from colony center varied (Figure 4-8c). For SF+B, the Oct4-Sox2- extraembryonic/other population is equivocally high throughout the colony. In CM conditions, Oct4+Sox2+ pluripotent population is high at the colony center but dips near the colony edges, with a concomitant increase in Oct4-Sox2- cells. The SF condition has 70% Oct4-Sox2+ neuroectoderm fated cells at the colony center, which linearly decreases towards the periphery of the colony. The pluripotent Oct4+Sox2+ population is at 20% near the colony center, peaks at 50%, and then declines to 30% at the colony edge. The Oct4-Sox2- are low in the center (<10%), but after 80% of the way to the colony edge this population increases, and quadruples to 40% of the cells at the colony edge. This observation has also previously been reported in early feeder-free hPSC cultures, with hPSCs undergoing epithelial-mesenchymal transition (EMT) and attaining a mesenchymal-like phenotype at the colony periphery (Ullmann, In’t Veld et al. 2007). The inverse relationship between neuroectoderm cells at the edge and extraembryonic cells at the periphery may be explained by gradients in GDF3/BMP (Peerani, Rao et al. 2007), with GDF3 being secreted by hPSCs and achieving maximum concentration near the colony center, and diminishing towards the edges, as predicted by modeling.
Figure 4-8 Superimposition of colonies enables statistical analysis of radial trends

A) For each well, all cell-coordinates can be obtained, and individual colonies can be identified. Colonies can then be superimposed with respect to the center of each colony. Cell intensities can be averaged in hexagonal bins to visualize the overall expression pattern of colonies grouped by well or condition, enabling visualization of spatial trends in arbitrarily shaped colonies. B) Radial patterns in Sox2 and Oct4 intensities, relative to colony centroid. Automatically generated overlays of 200 µm diameter patterns and line plots of fluorescent intensity vs normalized
distance from colony center for Sox2 intensity (left panel) and Oct4 intensity (right panel). Colony overlays represent intensity values normalized to the center of the colony, where cells are binned spatially in a hexagonal grid. Line graphs display the mean of cells binned in rings from the colony center to the edge. Error bars represent the 95% CI between wells for cells in the same bin. C) The spatial distribution of the Oct4/Sox2 defined cell types.

4.2.4.4 Quality control of patterns

An additional application of ContextExplorer is in performing spatial analysis of batches of plates for quality control (QC) purposes. We have previously developed a multi-laboratory validated protocol for hPSC 96-well patterned assays (Chapter 4), which involved multiple laboratories performing HTP hPSC assays using multiple cell lines. To assess cell line patterning fidelity, and demonstrate the batch QC capabilities of ContextExplorer, we analysed data from 4 cell lines (VUB01, VUB02, VUB07, VUB14), patterned in 4 full 96-well plates (1 cell line per plate), seeded at 4 seeding densities (32000, 48000, 64000, 80000 cells per well), and treated with SF, SF+B, and NS. For each well, ContextExplorer was used to calculate cells per colony (Figure 4-9a), cells per well (Figure 4-9b), and colonies per well (Figure 4-9c). In general, as more cells are seeded there was an increase in cells per well and colonies per well, but interestingly the cells per colony did not increase as much as expected. This indicates that seeding additional cells results in robust patterning of additional colonies, but not an increase in cells per colony. This may be the result of cells clustering together at a saturating density on select ECM spots during the seeding phase, rather than spreading out evenly. In addition, we sought to determine if patterning fidelity varied between cell lines. We found one cell line, VUB07, had markedly reducing patterning fidelity, and a significantly lower number of colonies per well (when averaged across the entire plate) than the other cell lines ($p < 0.001$ compared to VUB14) (Figure 4-9d).
Figure 4-9 Automatic colony detection facilitates patterning fidelity analyses

Three criteria are used to investigate patterning fidelity between four hPSC lines – **A**) cells per colony, **B**) cells per well and **C**) colonies per well. Colored dots are individual well measurements and black bars represents the mean values. Four hPSC lines (VUB01, VUB02, VUB07, and VUB14) were assayed in 200 µm diameter patterned plates. Cells were seeded at varying densities as noted, and response to SF+B, NS, and SF was assayed. 16 fields were imaged for each well and DAPI images were used to extrapolate cell x-y coordinates, and ContextExplorer extrapolates from this cells per colony, cells per well, and colonies per well. **D**) Average number of colonies per well varies across cell lines. * indicates $p < 0.001$, Students t-Test with two-tailed distribution and unequal variance.
4.2.5 Conclusions

There is overwhelming evidence that increased control and monitoring of population context parameters is needed to improve assay reproducibility and utility, however addressing this challenge has proven difficult for the broad biomedical community. We have separately developed cell patterning tools for control of assay population context parameters, here we demonstrate a software tool for improved monitoring of microenvironment variables. ContextExplorer is GUI-based, compatible with HTP screening and standard fluorescent microscopy based assays, open-source, and easy to use, which will enable adoption by a broad (non-specialist) scientific community.

HTP analysis of population context parameters can be a powerful tool to investigate the cellular microenvironment. Using hPSCs as a model system, ContextExplorer was utilized to explore and statistically characterize marked radial trends in transcription factor expression. As such trends are ubiquitous in biological systems, ContextExplorer can be a tool to quantify and interrogate the mechanisms of such trends in future and past HCA data sets. Once single-cell population context data is obtained (local cell density, being at the edge of a colony etc.) modeling tools such as Bayesian analysis can be used to discover causal links between such parameters and cell fate decisions (Snijder, Sacher et al. 2009), thus partially explaining the mechanistic causes of observed "noise". ContextExplorer may be particularly powerful with dynamic live-cell data. As an example, Hoechst and TRA-1-60 can be used to live-cell image pluripotent colonies (changes in colony number and size) through time, during routine hPSC culture or as they emerge during induced pluripotent stem cell (iPSC) induction. Additionally, to date the practical problems related to analyzing HCA data sets has led to the simplification of assays carrying valuable multivariate phenotypic information, with assays routinely being reduced to one simple output (Singh, Carpenter et al. 2014). We predict that advanced data analysis methods that enable full multiparametric data to be harvested for entire cell populations will enable HCA to finally reach its potential.

By importing and exporting data as CSV files, ContextExplorer is easily integrated with existing HCA tools such as CellProfiler (Carpenter, Jones et al. 2006) and CellProfiler Analyst (Jones, Kang et al. 2008). Exported data can also easily be explored programatically or in drag-and-drop exploratory data analyses software such as Tableau Public (Tableau Software, www.tableau.com). To further expand the utility of our software, we have also included
automatic thresholding, batch analysis and a graphical summary of the data via exported PDF files. Thresholding has earlier been done by visual inspection, but ContextExplorer allows for this to be set either as a percentage positive cells in a user-defined population or by displaying the histogram of intensities and scatter plots for the user to identify thresholds. ContextExplorer is built in Python and Qt, distributed as an executable for Windows, the source code runs under Linux, OS X and Windows and is distributed under the MIT open source license and available online at https://github.com/joelostblom/context_explorer.

4.2.6 Author contributions

J.E.E.O. developed software and performed all analysis. E.J.P.N. designed and performed experiments. J.E.E.O., E.J.P.N. and P.W.Z. designed the project and wrote the manuscript.

4.2.7 Acknowledgements

We thank Dominika Dziedzicka and Mieke Geens (Free University Brussels, Brussels, Belgium) for data from experiments performed on VUB01, VUB02, VUB07, and VUB14 hPSC lines.
Chapter 5
Discussion and Future Work
5 Discussion and Future Work

5.1 Contributions

5.1.1 Understanding and controlling cell-to-cell variability in a HTP platform

Using hPSCs, it was demonstrated that the 96-well assay developed by a population-context driven approach has many powerful applications in cell biology, and is translatable to the broad cell biology research community (i.e. not just specialists in micro-technologies). The platform consists of a method to pattern cells in 96-well plates, an optimized cell fate factor screening assay protocol, defined media and substrate, and a single-cell imaging and data analysis pipeline. Colony size, cell density, media composition and substrate were optimized to allow robust cell fate responses to exogenous cues to be measured in 48 h. Single cell protein expression profiling across diverse control conditions revealed Oct4 and Sox2 costaining as a screen output can discriminate pluripotent, neuroectoderm, primitive-streak, and extraembryonic subpopulations. This platform was applied to quantify the phenotypic response to a panel of factors (“fingerprint”), and it was demonstrated that this cell-line specific fingerprint is stable between passages and is very informative. This platform can also be used for dose curves, which was demonstrated by characterizing hPSC response to dose curves of 27 factors. By staining the populations induced in the control conditions with a library of antibodies the assay can also assess single-cell expression of a large panel of proteins. By assessing responses to pathway agonists and antagonists and comparing to baseline conditions, this platform can be used to quantify baseline Activin, FGF, EGF, Wnt, and BMP levels. This assay has also been applied to aid in optimization of ECM and media components. Additionally, in Chapter 3 it was shown that this platform can be used in lead discovery applications, demonstrated by screening a kinase inhibitor library (400 compounds at 2 concentrations). The platform can also enhance secondary studies and mechanism of action studies, which was shown by screening rapamycin with various pathway agonists and antagonists to discover dependencies on endogenous BMP and TGFβ signaling, and which is also shown by targeted siRNA knockdown, with and without additional exogenous cues, in further mechanism of action studies. This platform is also compatible with live-cell imaging.
for example by using live-cell calcein staining instead of Oct4 and Sox2 staining the assay has been modified to readout cell viability. Chapter 4 provides proof of principle that HTP microtechnology based HCA assays, and thus these applications outlined, can be transitioned to the broader scientific community.

This approach and platform which has been demonstrated using hPSCs as a model system can be applied to other stem cell model systems with heterogeneous subpopulations and complex microenvironmental regulation, such as in vitro stem cell and cancer models, and can be applied iteratively. Various cell types in addition to hPSCs can be patterned (such as hPSC derived endoderm, hPSC derived cardiomyocytes, and mEpiSCs). Furthermore, this enables a step-wise screening strategy to systematically improve production of hPSC derivatives. As a demonstration of this, others in our group (Nafees Rahman, manuscript in preparation) have applied this approach to hPSC derived hemogenic endothelium. Hemogenic endothelium was patterned in 96-well plates using microcontact printing, and causal links between colony size and blood formation were discovered. These findings have subsequently been used to screen hemogenic endothelium in patterned 96-well plates, in microenvironments engineered for enhancing the screen readout, in lead discovery studies for blood progenitor enhancement. The enhanced hemogenic endothelium production (3 fold yield and purity increase) developed in Chapter 3 feeds forward into these studies. In addition, PDX1+ pancreatic progenitor differentiation from progenitors has also been anecdotally shown to be spatial context dependent (Cai, Yu et al. 2009), so screening definitive endoderm or derivatives in controlled spatial context may be a powerful approach to improve pancreatic cell production. As heterogeneity arising from microenvironmental differences applies broadly to cell populations, it is anticipated this approach can be applied to many other adherent pluripotent and multipotent cell types at various differentiation bifurcation points.

5.1.2 A novel method for patterning proteins into standard 96-well plates

In Chapter 2 a novel method of microcontact printing of proteins in user-defined configurations in 96-well plates is developed, this invention can be applied to study many different cell types and model systems. Microcontact printing approaches to controlling cell distribution were developed more than two decades ago (Singhvi, Kumar et al. 1994) and have
diverse applications, our adaption of this approach to multi-well plates will aid in the transition of these applications to HTP platforms. Importantly, patterning technologies enables the study of causal links between cell distribution and phenotype. Our advancement enables causal study of these links to be performed in a HTP paradigm, which is critical in order to efficiently determine mechanisms of such links, and place these links in a broader framework of additional contributing factors. As a demonstration, in Chapter 2 by patterning hPSCs in different colony sizes it is shown that different colony sizes induce different hPSC fates, with larger colonies maintaining pluripotency, moderate size colonies resulting in a mix of pluripotent and differentiated cells, and smallest colonies resulting in complete loss of pluripotency markers. However, by screening these different microenvironments (4 diameters of spot sizes) with different soluble factor environments (5 media conditions), and in triplicate, it was shown that soluble factors such as BMP4 can override the supportive factors in large diameter colonies. As an additional example, in Chapter 4 data is presented from external laboratories screening 5 hPSC lines (with different genetic backgrounds), at 4 seeding densities, in 5 soluble microenvironments, and the spatial distributions of 4 resulting subpopulations was measured. This data set was mined to find that in the VUB14 line, neuroectoderm induction is negatively correlated with higher seeding density (Pearson’s r = -0.92, p = 5.8 x 10^-7). In future studies, additional single cell parameters can be obtained using ContextExplorer such as local cell density for each cells (as measured by cells per colony), radial distance from the colony edge for each cells, and well population size. Bayesian network learning is a powerful approach to find causal links between such population context variables (genetic differences, soluble factors, local cell density, radial distance to colony edge, well population size) to resultant cell phenotype (Oct4 and Sox2 expression), as shown by Snijder et al. who used this approach to explain 74% of the variance seen in simian virus 40 infection (Snijder, Sacher et al. 2009).

5.1.3 Understanding and controlling cell-to-cell variability at the bench scale

Reproducible assays are a prerequisite for truly elucidating the underlying regulatory networks that link cell-extrinsic factors, signal transduction, transcription factor regulation, and cell fate decisions. In Chapter 1 the need to control and quantitatively measure the stem cell microenvironment in stem cell assays was highlighted, however there are key barriers towards
this aim. The lowering cost of automated microscopy systems, and the development of open-source software tools have made HCA more accessible, but controlling cell patterning is still not practical beyond a small specialist field. To address this, extensive efforts were made to ensure the multi-well patterning approaches developed are not just used in the HTP screening paradigm (i.e. in screening labs), but also in bench-scale research paradigm for various routine stem cell assays. The protocol validated in multiple labs in Chapter 4, as well as the user-friendly ContextExplorer software for spatial analysis and QC is the major contribution towards this.

A remaining barrier is the access to 96-well plates with patterned ECM. A method for patterning 96-well plates is provided in Chapter 2 that employs PDMS stamps made to pattern directly into one quarter of a 96-well plate. This stamp is made by simply curing liquid PDMS in a Teflon mould on top of a silicone master. PDMS molding is considered inexpensive, procedurally simple, and does not require strict regulations of the laboratory environment (Kane, Takayama et al. 1999). Manufacturing of the Teflon mould was done via machining tools, and machine shops are widely accessible to biological labs. As an alternative, 3D printing could also be used to manufacture the Teflon mold. The remaining barrier is the silicone master (which has 200 µm diameter pits, the inverse of the pattern on the PDMS stamp) which was manufactured in a clean room using photolithography (Peerani, Bauwens et al. 2009). The 200 µm feature size allows for more accessible alternatives such as laser printing toner onto printed circuit board (PCB) substrates (Abdelgawad, Watson et al. 2008), or micromilling polystyrene via a computer numerical control (CNC) milling machine (Guckenberger, de Groot et al. 2015). As an alternative to individual laboratories manufacturing patterned ECM plates, manufacturing 96-well plates patterned using UV-lithography was also pursued in-house (with Mukul Tewary) that is a modification of a previously published protocol (Azioune, Storch et al. 2009). This approach is scalable, and as such is more amenable to commercialization, or provision of patterned plates to laboratories via a central facility as was done for the multi-lab studies in Chapter 4.

5.1.4 Oct4 and Sox2 based discrimination of early development subpopulations
A key enabler of our subpopulation tracking approach is the novel use of Oct4 and Sox2 as a binary code to discriminate four major early cell fates in human development: Oct4⁺Sox2⁺ for pluripotency, Oct4⁻Sox2⁺ for early neuroectoderm, Oct4⁺Sox2⁻ for early primitive streak, and Oct4⁻Sox2⁻ for early extraembryonic committed and other tissues. This classification is congruent with previous work showing that Oct4 and Sox2 along with Nanog are the major regulators of the core pluripotency transcription network (Boyer, Lee et al. 2005), Oct4 and not Sox2 is transiently expressed in the primitive streak (Mossman, Sourris et al. 2005), Sox2 is maintained during neuroectoderm development while Oct4 is rapidly downregulated (Chng, Teo et al. 2010), and that upon extraembryonic differentiation Oct4 and Sox2 are both rapidly downregulated (Xu, Chen et al. 2002).

The Oct4/Sox2 binary code was discovered via single cell protein expression profiling performed on the 96-well patterned hPSC platform, this approach can be repeated with other cell types to find similarly useful transcription factor combinations to discriminate subpopulations. Characterizing such transcription factor codes would not only be useful in HTP screens and general assays, but also could point towards important mechanistic transcription factor regulation during differentiation. Interestingly, our Oct4/Sox2 code is supported by studies of mPSC transcription factor dynamics and binding studies where it was found that Oct4 and Sox2 together maintain pluripotency, Oct4 alone suppresses neuroectoderm differentiation and promotes mesendoderm, and Sox2 alone reversely promotes neuroectoderm and inhibits mesendoderm (Thomson, Liu et al. 2011). Our findings indicate that similar transcription factor regulation occurs during human development, further exploration of transcription factor regulation and binding dynamics is needed to confirm this.

5.1.5 Quantitative measurement of cell line fingerprints and base line signaling levels

In Chapter 2 the 96µCP platform was applied to fingerprint hPSC response to a panel of controls, and it was demonstrated that this cell-line specific fingerprint is stable between passages, can vary between cell lines, predicts endoderm induction efficiency, can be used to improve differentiation of cell-lines reticent to cardiac induction, can be used to group cell lines that respond similarly to exogenous cues, and can rank cell lines by lineage induction efficiency. Additionally, hPSC lines that maintain pluripotent markers in the presence of strong
differentiation signals can be efficiently screened and flagged as being potentially neoplastic/karyotypically abnormal. As hiPSC lines are generated at a higher rate than they can be appropriately characterized, I propose that with further development this fingerprint assay could be an effective way for stem cell banks and laboratories generating hPSC lines to quickly and quantitatively assess new lines. In particular, the 96µCP platform could be an effective first pass assay to detect tumorigenic cell lines and quantify neuroectoderm and primitive streak differentiation propensity.

Related to this, with modification, the 96µCP may also lead to a simple and effective in vitro teratoma surrogate assay. Despite being the gold standard for testing hPSC lines, there is great inconsistency in the methods employed for teratoma assays and reporting results and the assays are timely, costly, and they use many mice which raises ethical concerns, leading some to question its value in proving pluripotency (Muller, Goldmann et al. 2010). The rapid expansion in derivation of new hPSC lines has further exacerbated these issues and has motivated the development of simple, cost-effective in vitro assays (Buta, David et al. 2013). The ISCI consortium is working towards testing alternatives, in fact the multi-lab study in Chapter 4 was performed in this context in parallel to several proposed alternatives. To meet these needs, the 96µCP assay could be modified to fix at a later timepoint, stain for additional markers, and use additional control conditions, the key goal being to definitively characterize neuroectoderm, mesoderm, and endoderm lineages ("tri-lineage assay").

In Chapter 2 an assay to quantify activation of an endogenous pathway relative to its dynamic range ("%Dynamic Range") is demonstrated, which combines measurements of hPSC response to pathway agonists, antagonists, and base media alone. Comparisons of control and pathway agonist-antagonist response profiles over multiple passages indicate high correlation between passages, and lower correlation between different cell lines. Differences in endogenous pathway activation in hPSC lines has long been observed anecdotally, our method provides a quantitative way to measure these differences. Additionally, it has been observed that for BMP activation, pSMAD1 activation diverges from ID1 activity, and hence does not accurately reflect BMP transduction (Xu, Peck et al. 2005). Thus, using a functional response (i.e. cell fate response) to quantify pathway activation and inhibition is a simple and attractive strategy to measure endogenous pathway activation, as an alternative to phosphoprotein
measurements which are highly time sensitive, and require cell starvation periods that in themselves could dramatically alter stem cell fate.

5.1.6 Insights into hPSC signaling

Application of the 96µCP platform to hPSC biology has revealed several insights into hPSC signaling. In Chapter 2, 27 factors prominent in hPSC studies, as selected by the ISCI consortium, were characterized across a wide range of doses, and their dose-dependent effects on multi-lineage differentiation were reported, revealing new dose- and lineage-specific optimal concentrations. Additionally, the optimized concentration of LDN-193189, a small molecule BMP inhibitor, was applied to significantly enhance induction of later-stage endodermal PDX1⁺ pancreatic progenitor cells. This published data set is the most comprehensive characterization of the effect of these factors on hPSCs to date, and importantly the study was performed in highly-defined conditions, which aids in interpretability.

In Chapter 2 it is also demonstrated that DKK1 inhibits pluripotency and enhances neuroectoderm induction in hPSCs, in contrast to previous studies by Dravid et al. who report Dkk1 has no effect on hPSC pluripotency (Dravid, Ye et al. 2005). In the presence of pluripotency supporting factors (FGF or activin A) this effect of DKK1 is masked, and this reconciles our observation with Dravid et al. who performed their DKK1 assay on hPSCs cultured on MEFs (which secrete FGF and TGFβ activating ligands) and measured alkaline-phosphatase colony formation after 26 days. This study serves as an illustration of how endogenous ligands obfuscate assay results and interpretation, and highlights the need to assess response to factors in diverse exogenous signaling environments in order to allow for more meaningful interpretation and account for the highly variable background endogenous signaling in hPSC cultures.

5.1.7 mTOR inhibition enhances generation of hemogenic endothelium

In Chapter 3 the 96µCP platform is applied in a HTP lead discovery program, and found that rapamycin alone added to serum-free defined media induces Brachyury+ mesendoderm cells, and furthermore can be used to improve hemogenic endothelium purity and yield 3 fold, with a 1.7 fold increase in CFC yield. This study not only provides a powerful small molecule
to enhance hPSC derived blood induction, but also serves as proof of principle of the utility of controlling microenvironmental parameters and measuring multiple subpopulations in drug screening assays. Additionally this study demonstrates application of the 96µCP assay in many secondary and mechanistic studies, for example to identify the interactions of mTOR inhibition and TGFβ and BMP signaling and to confirm the target of rapamycin via siRNA targeting of the mTORC1 component Raptor. It was found that rapamycin at 0.001 nM and 0.01 nM moderately enhances pluripotency, further studies are needed to identify the mechanism of this effect. This work reconciles previous studies as in agreement with Easley et al. (Easley, Ben-Yehudah et al. 2010) no negative effect on pluripotency is observed with low concentrations of rapamycin, and in agreement with Zhou et al. (Zhou, Li et al. 2012) and Sachlos et al. (Sachlos, Risueno et al. 2012) at higher concentrations a loss of pluripotency is observed. This underscores the utility of having a HTP hPSC assay to perform dose curves on rapamycin, which neither of the previous studies performed. In contrast to Zhou et al. who observed rapamycin enhances serum mediated EB differentiation towards endoderm and mesoderm, it was found rapamycin in fact strongly inhibits definitive endoderm induction in a cell-line dependent manner. This difference may arise from the use of serum, or may arise from the cell lines used for the Zhou study. Note that for this endoderm assay was adapted for 96-well plates from a previous protocol (Rezania, Bruin et al. 2012), which aids in multi-cell line comparisons.

Determining if mTOR inhibition can enhance derivation of other mesoderm derivatives (cardiac, muscle) is also a priority. Recent work from Mendjan et al. indicates that instead of multipotent mesoderm that can pattern into different mesoderm subtypes, different methods of inducing Brachyury+ mesoderm each give rise to restricted mesoderm that patterns into restricted subtypes ("restricted mesoderm model") (Mendjan, Mascetti et al. 2014). Furthermore Nanog and CDX2 repress each other, and Nanog is required for anterior subtypes (including cardiac and lateral plate) but blocks somitic subtypes, whereas CDX2 is required for somitic subtypes and blocks anterior subtypes. Hemogenic endothelium arises from lateral plate mesoderm, so one possibility is that mTOR inhibitors enhance Nanog+Brachyury+ anterior restricted mesoderm. Our preliminary studies indicate that rapamycin in fact hinders cardiac differentiation, so may specifically enhance non-cardiac mesoderm. As the 96µCP platform enables HTP primitive streak assays, one interesting future screen to rapidly
interrogate this specification would be to screen BMP+Activin, with and without various pathway agonists/antagonists, and with and without rapamycin, and co-stain for Brachyury and HAND1 (posterior primitive streak), EOMES (anterior primitive streak), and CDX1 (late primitive streak). Costaining could be performed using multiplexing approaches (see below), or replicate plates could be stained with different marker combinations.

5.1.8 Software for analysis of spatial context in HTP data

For Chapter 2, MATLAB code was developed to assist in analysis of HCA data, and although the code was made available it was not user friendly. In Chapter 4 ContextExplorer, GUI-based population context analysis software designed for ease of use by a broad scientific community, is described which allows for batch processing HCA multi-well plate data making it amenable for analysis of HTP screen data. This software is applied to analyze patterned and non-patterned hPSC data. The software can automatically detect colonies, either patterned or hPSC colonies on MEFs (by using differential marker expression). This colony detection feature enables virtual superpositioning of colonies to extract spatial trends, which revealing that Oct4 and Sox2 intensity levels vary dependent on their radial position within the colony. The utility of ContextExplorer in QC applications, analysis of local cell density (cells per colony) and other spatial parameters is also demonstrated. ContextExplorer can readily be used in conjunction with various existing image analysis programs such as CellProfiler (Carpenter, Jones et al. 2006). To date the practical problems related to analyzing HCA data sets has led to oversimplification of assays to one parameter, discarding valuable phenotypic information such as spatial context (Singh, Carpenter et al. 2014), ContextExplorer was designed to aid in addressing this. Powerful future extensions of ContextExplorer include addition of spatial context parameters to measure such location on a colony edge (binary output for each cell) and compaction (i.e. local cell density at ~20 µm scale). As mentioned above, linking such parameters with additional experimental perturbations using Bayesian networks would be a powerful way to identify causal effects in the complex stem cell microenvironment.

5.2 Future directions

5.2.1 Strategies towards growth-factor free culture systems for scalable expansion of hPSCs
A key factor for commercial success of hPSC based cell therapies and drug screening applications is the ability to expand hPSCs robustly and cost-effectively. Towards this goal several avenues are being pursued, including advancement of bioprocess techniques in automated closed-process systems for culturing adherent cells (such as CompaT SelecT, TAP Biosystems), and bioreactors (Baptista, Fluri et al. 2013); however, regardless of the bioprocess cost-effective media for robust hPSC expansion will be critical and may be capable of reducing the material cost of expanding hPSCs 30-60% (Jenkins and Farid 2015).

Removing the requirement of growth factors from media, which may be feasible by finding small molecule cocktails, would improve the cost-effectiveness of media and also eliminate a major source of variability. This may be particularly feasible as hPSCs secrete autocrine pluripotency supportive factors such as FGF2 (Dvorak, Dvorakova et al. 2005) and GDF3 (Peerani, Rao et al. 2007). As one strategy to accomplish this, there is an indication that screening a select set of factors, or combinations of such factors, may be a powerful approach. This approach is exemplified by Watanabe et al. (Watanabe, Ueno et al. 2007) where screening a select set of caspase inhibitors, growth factors, trophic factors and kinase inhibitors resulted in characterization of Y-27632, which is in widespread use today.

A major challenge to this approach is identifying the select core small molecules to include in such an approach. The BMP receptor inhibitor LDN-193189, characterized in Chapter 2, would be important to include, surprisingly small molecule BMP inhibitors have not been employed in hPSC expansion media to date. Low doses of rapamycin also would be of interest. IDE1 and IDE2, which activate TGFβ signaling (through unknown mechanisms) and improve definitive endoderm induction (Borowiak, Maehr et al. 2009), are also attractive compounds in a small molecule cocktail to expand hPSCs, and remarkably their effect on hPSC expansion has not yet been characterized. Similarly, Tzv and Ptn (Xu, Zhu et al. 2010), which are reported to enhance phosphorylation of FGFRs, IGFRs, EGFRs, PI3K, AKT and MAPK would also be intriguing to screen. Screening extracellular matrix components in combination with soluble factors may also be of utility, as these culture components are well known to interact (Flaim, Teng et al. 2008). Note that both microcontact printing and UV lithography approaches of ECM patterning are amenable to patterning different ECM compositions in each well of a 96-well plate. Similarly, polymer surface chemistry is known to interact with adsorbed proteins, making this another intriguing variable in such a screen (Mei, Saha et al. 2010).
concentrations of insulin would be another important factor to screen, as Wang et al. have reported that IGF1R inhibition results in increased differentiation (Wang, Schulz et al. 2007), and at the high concentrations of insulin present in hPSC formulations, insulin is known to activate IGF1R signaling (Freund, Ward-van Oostwaard et al. 2008).

To optimize the combinations and dosing, the approach of Tsutsui et al. could be followed, where several rounds of combinatorial screening are performed with the results of each round input into a differential evolution algorithm to determine the next set of conditions to test (Tsutsui, Valamehr et al. 2011). As discussed in Chapter 1, major limitations of the approach of Tsutsui et al. were the selection of compounds and this library size (5 compounds), the size of each round (maximum 32 conditions were tested per round), the number of rounds, and the output used in the screen (initially alkaline phosphatase was used). The 96µCP HTP assay with proven predictive readout in conjunction with improved library design would address these concerns. An alternative approach to such a differential algorithm strategy is application of Design of Experiments, which was recently used in an attempt to optimize hPSC culture media (Marinho, Chailangkarn et al. 2015), but has similar limitations (only two factors were varied, FGF2 and heregulin-β-1, and only 9 conditions were assayed).

To date similar screens have often failed when validation is attempted past a few passages. There are also strong indications that maintaining pluripotency during passaging is a key obstacle for a culture system sustaining long-term pluripotency. As detailed in Chapter 1, Li et al. have observed that blocking cell-cell adhesion with EGTA abolishes the survival effect of Y-27632, indicating that ROCK inhibition decrease of apoptosis is indirect and mediated by increasing cell-cell contact (Li, Krawetz et al. 2009). Additionally, Eiselleova and colleagues have revealed that hPSCs have intracrine FGF2, and propose a model where this intrinsic FGF2 maintains hPSCs, and exogenous FGF2 is mainly required to enhance adhesion and survival during passaging (Eiselleova, Matulka et al. 2009). I propose screening separately for adhesion (i.e. what factors increase hPSC adhesion from suspension after a few hours) and maintenance of pluripotency (of cells already adherent), with the resulting culture system having separate media for when cells are passaged and for routine post-passaging media exchanges. An adhesion screen would be simple, however none has been reported to date. Tzv and Ptn (Xu, Zhu et al. 2010) are known to enhance cell adhesion, and act similarly to ROCK inhibitors. Additional strategies to improve cell adhesion include culturing cells at high surface
density, and chemical and geometrical optimization of the surfaces via UV based patterning as performed by Saha et al. to improve early cell aggregation (Saha, Mei et al. 2011).

Instead of improving culture conditions for hPSCs, using alternate pluripotent populations that are more amenable to expansion has emerged as a fascinating alternative option for scalable expansion of pluripotent cells. Naïve hPSCs have properties similar to mouse naïve PSCs, except for a notably higher cloning efficiency (up to 88% with ROCK inhibitor, compared to 22% for primed PSCs) and shorter doubling time making them more amenable to scale up (Gafni, Weinberger et al. 2013, Chen, Aksoy et al. 2015). Recently human region-selective pluripotent stem cells (rsPSCs) have been derived, these cells are cultured in FGF2 and Wnt inhibitor IWR1 and also have a high cloning efficiency (Wu, Okamura et al. 2015). Indeed, mEpiSCs have been derived from various developmental stages with distinct transcription profiles and predispositions to differentiate towards specific germ-layers (Kojima, Kaufman-Francis et al. 2014). A crucial challenge will be to define growth conditions that allow reproducible, robust, long-term maintenance of these alternative pluripotent cells in genetically unmodified human cells. For novel culture conditions these studies highlight the need to investigate the specific PSC state, and not just verify maintenance of pluripotency. Additionally, reports indicate that modifying differentiation conditions for specific PSC states is also likely required (Zhou, Li et al. 2010). To meet these challenges effectively, the approaches performed to characterise hPSCs in Chapter 2 and Chapter 3 can readily be translated to characterize signaling-fate relationships in these quickly emerging alternative PSCs states.

5.2.2 Multiplexing in HCA

A major bottleneck in using fluorescent imaging approaches to study cellular systems has been the limited multiplexing available to date. For example, the HCA approaches employed in this work use a maximum of 3 fluorescent channels (e.g. to image Hoechst, Oct4, Sox2). Although 4-5 total channel are available to HCA platforms, channels tend to bleed through and require tedious correction. This is severely limiting, particularly in heterogeneous culture systems where multiple markers are needed just to identify the cell type. In an exciting development, several studies have begun using cyclic imaging and bleaching approaches to dramatically enhance multiplexing abilities. Photo-bleaching approaches have been used for enhanced multiplexing, Schubert et al. have used this approach with robotics-based cyclic
imaging and bleaching to image 2100 antigens simultaneously in a tissue section (Schubert, Gieseler et al. 2009), however these techniques are not widely employed in the HCA field due to implementation hurdles and the inefficient and variable quenching. Gerdes et al. have developed a multiplexed fluorescence microscopy method (MxIF) involving chemical inactivation of fluorescent dyes after each image acquisition round, using a proprietary formulation (Gerdes, Sevinsky et al. 2013). The quenching takes approx. 15 min and can reduce fluorescent to less than 2% of original intensity, enabling the reuse of dyes for more than 50 rounds (no saturating limit was reached), and enabling co-expression of 61 protein antigens to be analysed at the single cell level in 747 colorectal cancer subjects (Gerdes, Sevinsky et al. 2013). As one remarkable demonstration of the power of multiplexing in HCA, Chen et al. have recently developed multiplexed error robust FISH (MERFISH), which combines combinatorial FISH labeling, Hamming encoding schemes, and iterations of hybridization and imaging rounds to enable the simultaneous imaging of ~1000 RNA species in a single cell, with subcellular resolution (Chen, Boettiger et al. 2015). Broader implementation of these proof of principle studies will revolutionize the HCA field and stem cell biology, as this increase in multiplexing offers an exponential increase in information.

5.2.3 Systems models of hPSC fate decisions

Effective control of stem cell fate will require tight regulation of nonlinear complex networks. However, identification of key control nodes has been challenging due to the lack of understanding of the cell fate regulatory networks in cells. Due to the number of possible combinations, brute force HTP screening alone will not be sufficient to effectively control cells, but likely will need to be used in conjunction with system biology approaches. For example, choosing 6 out of 400 possible compounds, there are $5.5 \times 10^{12}$ combinations to test. Statistical methods, such as partial least squares regression (PLSR), can also be used to model the relationship between perturbations, signaling events and cell fate outcomes. Janes et al. used PLSR to link 7980 intracellular signaling measurements to 1440 apoptosis outputs in response to combinations of cytokines (Janes, Albeck et al. 2005). Remarkably, the complex input/output relationship could be reduced using principal component analysis to determine two groups of cytokines ("canonical basis set") that linearly controlled apoptotic response, demonstrating the power of such statistical approaches, especially when little mechanistic information is known about the underlying cell fate control network. When more information
is available, more mechanistic modeling such as ordinary differential equation (ODE) modeling can be used. Lehar et al. have also demonstrated that the response to combinations of perturbations can also be used to infer the underlying network motifs (Lehar, Zimmermann et al. 2007). Specifically, the response surface topology can be compared to computationally predicted surfaces arising from alternative network topologies, such as parallel and serial conformations, feedback loops, and branched conformations. With or without explicit cell fate models (such as ODEs or signaling networks), biological search algorithms (reviewed in (Feala, Cortes et al. 2010)) can also be employed to optimize drug combinations. These search algorithms iteratively obtain HTP assay measurements and feed these into an algorithm which generate new combinations to test, systematically characterizing the control landscape of the cell fate network. For optimal control of stem cell fate, it is clear that approaches beyond brute force are needed, it is hoped that increased use of HTP assays will enable such approaches and enable the application of PLSR, ODE, network motif modeling, and biological search algorithms to gain a mechanistic understanding of cell fate decisions.

Saez-Rodriguez et al. have developed a computational approach to generate predictive Boolean models of mammalian signaling, which is implemented in free software (CellNetOptimizer; CNO) (Saez-Rodriguez, Alexopoulos et al. 2009). The approach entails first obtaining a large scale generic protein signaling network from the literature (e.g. ~85 node Boolean model), obtaining experimental data on the cell type of interest (e.g. ~1000 cue-response measurements of signaling molecules such as phosphoproteins), and then using CNO to fine-tune the signaling network to the data using a genetic algorithm approach to add and remove links between nodes to better fit the data. It was found that trained models were markedly more predictive, and the approach was used to identify new interactions between nodes. This results in a cell type specific model. The authors have also applied this approach to derive networks of primary and transformed hepatocytes, and the approach enabled a mechanistic understanding of response differences between these cell types (Alexopoulos, Saez-Rodriguez et al. 2010). Comparing models of hPSC cell lines, or different PSC states could be a powerful way to gain insight into cell fate regulation differences. The key barrier to applying this approach to hPSCs has been the lack of sufficient HTP data. A combinatorial screen on the H9 cell line has been performed, looking at over 300 conditions consisting of all 1 and 2 factor combinations of 24 factors, and measuring the resultant Oct4 and Sox2
phenotype. Using the CNO approach a preliminary model was developed, and fit to the existing data set (Figure 5-1). Additional signal transduction response data, and literature based refinement of the initial signaling network may yield a Boolean model of signal transduction in H9s. Repeating this process with another cell line and comparisons of the derived signal transduction models could be an intriguing strategy to obtain predictive mechanistic models of hPSC fate decisions, and understand the network differences that give rise to cell line differences. As an alternative modeling approach that doesn't require an initial network, PLSR modeling could be used to link signal transduction and cell fate data to applied cues (Miller-Jensen, Janes et al. 2007).

**Figure 5-1 Signal transduction model H9s based on combinatorial screen response data**

(A) A 300+ condition screen of all 1- and 2- factor combinations of 24 developmental pathway agonists and antagonists. % and # of 4 Oct4/Sox2 subpopulations are measured. (B) Preliminary fit of an input signaling network. Using CNO, specific edges are iteratively added or removed using a genetic algorithm to fit the combinatorial Oct4/Sox2 data, resulting in a “H9 specific” signalling network, a predictive Boolean model of how H9s make cell fate decisions. Edges can then be experimentally tested. Work with Calvin Lao, Camille Terfve and Julio Saez-Rodriguez.

5.3 Study Insights
A general method of microcontact printing directly into 96-well plates was invented, and this was applied to enable HTP applications of patterned hPSCs. Colony size, cell density, media composition and substrate were developed to allow robust cell fate responses to exogenous cues to be measured in 48 h. Single cell protein expression profiling across diverse control conditions revealed Oct4 and Sox2 costaining as a screen output can discriminate pluripotent, neuroectoderm, primitive-streak, and extraembryonic/other fated subpopulations. Using the optimized assay and Oct4/Sox2 subpopulations as a readout (96µCP assay) 27 factors agonizing and antagonizing key developmental pathways were characterized across a wide range of doses, and their dose-dependent effects on multi-lineage differentiation were reported, the most comprehensive characterization of the effect of these factors on hPSCs to date. Additionally, hPSC cell-line specific response to a panel of controls ("fingerprint") was shown to be stable between passages, discriminate between cell lines, and be predictive endoderm induction efficiency. A method to quantify endogenous pathway activation relative to dynamic range (“%Dynamic Range”) was demonstrated, which combines measurements of hPSC response to pathway agonists, antagonists, and base media alone. The 96µCP assay was next applied in a HTP lead discovery program, and found that small molecule mTOR inhibitors alone added to serum-free defined media induces Brachyury+ mesendoderm cells, and furthermore can be used to improve hemogenic endothelium purity and yield 3 fold, with a 1.7 fold increase in CFC yield. This study serves as proof of principle of the utility of controlling microenvironmental parameters and measuring multiple subpopulations in drug screening assays. To aid in transition of these approaches to the broader hPSC scientific community, a protocol for the 96µCP assay was developed and validated with multiple external laboratories. In addition, ContextExplorer, GUI-based population context analysis software, is described which allows for batch processing HCA multi-well plate data making it amenable for analysis of HTP screen data. This approach and platform which has been demonstrated using hPSCs as a model system can be applied to other stem cell model systems with heterogeneous subpopulations and complex microenvironmental regulation, such as in vitro stem cell and cancer models, and can be applied iteratively in a stepwise approach. With further development the 96µCP assay may also lead to a simple and effective in vitro teratoma surrogate assay. Ultimately, these technologies and applications developed enable a more comprehensive elucidation of the underlying regulatory networks that link cell-extrinsic factors, signal transduction, transcription factor regulation, and cell fate decisions.
6 References


pluripotent stem cells identifies compounds that rescue IKBKAP expression. "Nat Biotechnol 30(12): 1244-1248.


