Nanotechnology & human stem cells: applications in cardiogenesis and neurogenesis

Martin Lyubomirov Tomov
University at Albany, State University of New York, mtomov@albany.edu

The University at Albany community has made this article openly available. Please share how this access benefits you.

Follow this and additional works at: https://scholarsarchive.library.albany.edu/legacy-etd

Part of the Biomedical Engineering and Bioengineering Commons, and the Nanoscience and Nanotechnology Commons

Recommended Citation
https://scholarsarchive.library.albany.edu/legacy-etd/1742

This Dissertation is brought to you for free and open access by the The Graduate School at Scholars Archive. It has been accepted for inclusion in Legacy Theses & Dissertations (2009 - 2024) by an authorized administrator of Scholars Archive. Please see Terms of Use. For more information, please contact scholarsarchive@albany.edu.
NANOTECHNOLOGY & HUMAN STEM CELLS:
APPLICATIONS IN CARDIOGENESIS AND NEUROGENESIS

By

Martin L. Tomov

A Dissertation Submitted to the University at Albany, State University of New York
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

College of Nanoscale Science and Engineering

2016
Thesis Committee:

Janet L. Paluh Ph.D.
Thesis Research Advisor
State University of New York Polytechnic Institute
Colleges of Nanoscale Science and Engineering

Jose Cibelli, D.V.M., Ph.D.
External Thesis Research Advisor
Michigan State University
LARCEL, Laboratorio Andaluz de Reprogramación Celular, Spain.

Scott Tenenbaum, Ph.D.
State University of New York Polytechnic Institute
Colleges of Nanoscale Science and Engineering

Xinxin Ding, Ph.D.
State University of New York Polytechnic Institute
Colleges of Nanoscale Science and Engineering

Yubing Xie, Ph.D.
State University of New York Polytechnic Institute
Colleges of Nanoscale Science and Engineering
Acknowledgements

I would like to thank my family and friends for their support and discussion. My friends in the lab and the constellation, both students and faculty, who provided so much in terms of help, collaborations, discussions, and encouragement. Our collaborators, without whom this work would have been impossible. My committee members and my adviser for helping when I got stuck, or needed another opinion on the project.

Dedications:

I would like to dedicate my work to my parents Ivanka and Lyubomir, my grandparents Yanka and Velin, my other grandmother Magdalena, and especially my fiancée, Jessica. Without their help and encouragement, I would not have been able to succeed in this endeavor.
Table of Contents

Chapter 1: Thesis Overview

Specific Aim 1 and Experimental Procedures ...................................................... 1
Specific Aim 2 and Experimental Procedures ...................................................... 3
Specific Aim 3 and Experimental Procedures ...................................................... 5
Clinical Relevance ............................................................................................... 7
Thesis Introduction ................................................................................................ 9

Techniques Used .................................................................................................. 14

Cell Reprogramming............................................................... 14
qRT-PCR Analysis ........................................................................ 15
\textit{In vivo} teratoma formation assay ........................................... 15
hESC and ED-iPSC Line Maintenance and EB Formation ............... 15
Directed Differentiation of hESC and ED-iPSC ........................................ 16
Immunocytochemical Analysis (ICC) ................................................... 20
SU-8 Processing ........................................................................... 21
Polydimethylsiloxane (PDMS) Casting .............................................. 22
Micro-Contact Printing .................................................................. 22
3D Model Design and 3D Printer Rapid Prototyping ....................... 22
Nucleic Acid Preparation for RNA-Seq and ChIP-seq .................... 23
RNA-Seq Data Analysis .................................................................. 24
ChIP-seq Data Analysis .................................................................. 25

References .............................................................................................. 26

Chapter 2: The Human Embryoid Body Internal Architecture Complexity Has Significant Implications for the Differentiation Potential of Stem Cells

State of the Field .......................................................................................... 33
Hypothesis ................................................................................................. 34
Strategy ......................................................................................................... 35
Dissertation Abstract

Human stem cell research holds an unprecedented promise to revolutionize the way we approach medicine and healthcare in general, moving us from a position of mostly addressing the symptoms to a state where treatments can focus on removing the underlying causes of a condition. Stem cell research can shed light into normal developmental pathways, as we are beginning to replicate them in a petri dish and can also be used to model diseases and abnormal conditions. Direct applications can range from finding cures for single or multigene diseases to demonstrating that we can replace these genes with a normal copy. We can even begin to model lifelong conditions such as aging by iPSC technology by relying on fetal, young, adult, and centenarian populations to provide insights into the process. We have also begun to understand the microenvironment in which specific cell populations reside. Being able to replicate the chemical, physical mechanical, and spatial needs of those cells, research groups are successfully generating full organs using cadaver scaffolds of heart and kidney, and there is promising research to reach the same success with other organs, such as the liver, and pancreas. Advances in those areas open an enormous potential to study organs, organoids, organ valves, tubes or other functional elements such as beating cardiomyocytes in vitro.

There is also the need to evaluate the whole genome of induced and differentiated cells, with its myriad of interacting pathways. Bioinformatics can help our understanding of embryogenesis, organ differentiation and function. It can also help optimize our stem cell and bio-scaffold tools to advance closer to functional organs and tissues. Such a combination approach will also include pluripotency evaluation and multi-lineage
differentiation, as well as platforms that may assist in cell therapies: 3D structures, micro-ribbons, directed patterning to name a few. There is now a clearer path forward with stem cell research than was ever before possible. My research has made fundamental contributions to the stem cell field by detailed analysis of uniformly generated 3D stem cell intermediates that are embryoid bodies. I have also contributed to the derivation of the first fully characterized ethnically diverse induced pluripotent stem cells from minority populations (ED-iPSCs), and advances in generating functional beating cardiomyocytes in vitro to aid cardiomyoplasty therapies. My work has also explored scaffolds for directing neural cell assembly or encouraging self-assembly for applications in CNS neurodegeneration, addiction, and spinal cord injury. These contributions to the field are outlined in my Specific Aims below and detailed in the chapters of my thesis.
Chapter 1:

Thesis Overview

Specific Aim 1: Advance stem cell technology by expanding stem cell resources to include population diversity by derivation and comprehensive multi-lineage and bioinformatic evaluation of ethnically diverse iPSCs.

Hypothesis: The lack of diversity in resources for stem cell biology will limit its full application in the medical field by preventing analysis of contributing SNPs and known ethnic contributions to therapy resistance.

Overview: Through a collaborative effort, induced pluripotent stem cell lines were derived by reprogramming human fibroblasts from ethnically diverse donors (ED-iPSCs). My analysis of these lines for pluripotency included in vitro multi-lineage commitment and terminal differentiation assays including both functional and cell biological analysis. This work required optimization of established protocols, bioinformatics analysis, and both 2D cultures and 3D microenvironments using lithography-templated devices. The human embryonic stem cell (hESC) line WA09, served as a differentiation control. This work is addressed in Chapters 1-4, and Chapter 6.

Aim 1.1 – Determine the importance of parameters such as size and shape in discrete embryoid body (EB) on the comparative differentiation potential of pluripotent stem cells using uniform EBs generated in high throughput custom microarrays.
Aim 1.2 – Comparative evaluation of newly generated ED-iPSC replicate lines within and across ethnicities by whole genome transcriptome and epigenetic bioinformatics analysis of pluripotent 2D and 3D cultures and multi-lineage early commitment to develop a bank of high quality ethnically diverse stem cell lines made available to researchers.

Aim 1.3 – Demonstrate targeted terminal differentiation towards endodermal pancreatic cells, mesodermal smooth muscle and cardiomyocyte cells, and ectodermal retinal pigment epithelial (RPE) cells, astrocytes and cortical pyramidal neurons, which can be used as a springboard in future projects, and as verification of pluripotency capability of the ED-iPSC lines.

**Experimental Procedures:** Aim 1.1. The custom microarray design was empirically determined and generated in a 10x10 grid formed out of polydimethylsiloxane (PDMS) with wells of defined depth and diameter as follows: I tested 200, 300, 400, and 500 µm diameter wells, with depths of 100, 200, and 400 µm to identify the best parameters for loading, uniformity, and reproducibility to culture stem cells for generating 3D embryoid bodies (EBs). Initial analysis indicated differences in the internal architecture of the EB that necessitated comparing the state of the stem cells used to seed the wells and generate the EBs, namely single cells suspension or mechanical passaging to generate 2D clusters, or pre-formed 3D early aggregates.

Uniform 200 um EBs were determined to be most effective in differentiation and all EB formation was independent of state of the cells used in seeding. Aim 1.2. Total RNA for whole genome transcriptome and epigenetic was prepared for 2D and 3D pluripotent cultures and processed at SUNY Buffalo to generate high quality bioinformatics database that would be explored in the following Aims. Aim 1.3. The most recent protocols available
in literature from expert labs were applied to successfully differentiate the control WA09 hESC line and then applied to multiple ED-iPSC lines to generate endodermal pancreatic lineage cells, mesodermal smooth muscle and cardiomyocytes, and ectodermal cortical neurons, astrocytes, and retinal pigment epithelium (RPE) cells.

**Specific Aim 2:** Advance potential of iPSC technology for cardiomyoplasty by evaluating replicate ED-iPSC lines for functional cardiomyocyte formation through improved analysis platforms and evaluation of biomarkers, contractility, and epigenetic and transcriptome differences in cardiogenesis.

**Hypothesis:** The ability to generate cardiomyocytes alone is insufficient to advance therapies in cardiomyoplasty, thus a better understanding of cardiomyocyte developmental pathways and means to compare generated cardiomyocytes will benefit clinical applications and *in vitro* projects.

**Overview:** Ethnic genetic background is known to contribute to cardiac disease and treatment efficacy. Cardiomyocyte protocols in Aim 1 will be used or optimized to understand underlying biological, bioinformatics and microenvironment contributions to cardiomyocyte formation in 2D colonies, or 3D microarchitectures, which includes acquisition of functional features of tropomyosins and beating, as well as epigenetic parameters that may control differentiation efficacy. This work is addressed in Chapter 5 and Chapter 6.

Aim 2.1 – Optimization of cardiomyocyte differentiation protocols and functional analysis methods including design and fabrication of lithography-based platforms for analysis of
3D cultures to aid in the evaluation of newly generated human ED-iPSC lines for terminal differentiation into functional cardiomyocytes as determined by Troponin-T and beating activity.

**Aim 2.2** – Determination of critical parameters of pluripotency, ethnicity, transcriptome, and application of bioinformatics analysis to understand ED-iPSC differentiation into functional cardiomyocytes in order to generate a more detailed database as the lines are made available to other researchers.

**Experimental Procedures:** **Aim 2.1** Optimization of a narrow range of GSK3/Wnt pathway inhibition was needed along with 3D self-assembly (versus 2D culture) are major factors in generating the beating phenotype and polymerized Troponin-T expression. However that optimal concentration varied between the newly generated iPSC lines. In traditional 2D culture, 3D beating clusters are formed in distributed mounds within a confluent cell sheet of non-beating cells. Therefore, I adapted the high throughput PDMS grids developed in Aim 1 and confirmed that efficiency of 3D cardiac differentiation *in vitro* is enhanced by seeding stem cells into grids and monitoring beating and troponin-T by ICC. The initial experiment used a mixed WA09 hESC beating and non-beating cardiomyocyte culture seeded into a 200 µm diameter 400 µm deep grid and self-assembled over a 14 day period. These were sufficient based on other tested differentiation pathways, including smooth muscle, which is another mesoderm-derived tissue type. **Aim 2.2.** Bioinformatics included evaluation within and across ethnicities comparing replicate lines, focused on gene ontology and differential expression of genes and gene pathways.
Specific Aim 3: Apply nanotechnology to develop and optimize platforms for biomimicry of the human central nervous system cerebral cortex microarchitecture for applications in neurodegenerative diseases.

Hypothesis: The complexity of CNS microarchitectures and inability to study them in *in vitro* models hinders advancing biomedical therapies. There is an acute need to develop both directed and self-assembly models to advance CNS therapeutic outcomes.

Overview: Minimal regional models of neural tissue, based on relevant stem cell derived neurons and support glia cells, will be generated and used to analyze neuron-glia interactions in microenvironments – cortical brain, prefrontal cortex, and spinal cord – including reactions to perturbed conditions. Both 2D co-cultures and 3D microarchitectures were generated and various biocompatible scaffolds such as microstrands, lithography-templated microarrays, and patterned surfaces for self-assembly, assisted self-assembly, and directed patterning were tested to study neural cortical development and plasticity. This work is elaborated and expanded on in Chapter 5 and Chapter 6.

Aim 3.1 – Optimization of neural cell differentiation protocols with the ED-iPSC lines for cortical pyramidal neurons and astrocytes in order to evaluate their morphology, immunocytochemistry and functional activity for use in models of neurodegeneration.

Aim 3.2 - Development of 2D and 3D microarchitectures of co-cultured human neural cells by assisted assembly to evaluate patterned versus normal self-assembly and apply this to understanding homeostasis and plasticity changes during normal and pathological CNS changes in normal, disease, and aging models.
**Experimental Procedures:** **Aim 3.1.** Pyramidal neurons take on average 45-60 days to differentiate from pluripotent stem cells based on reported results in literature, with cell death during the initial steps, leading to reduced numbers of cells in the final culture. Astrocytes typically require around 30 days to terminally differentiate. This time gap introduces a need for careful planning in designing experiments, and maintaining the cultures leading up to experiments for Aim 3.2 with 3D microstructures and co-culture studies. The range and number of experiments that can be conducted within each differentiation set must be considered.

To partially address this I have modified the protocol by substituting a newly developed commercial media that generates stable, self-renewing neural progenitors reliably, thus reducing the time to terminally differentiated pyramidal neurons and astrocytes to around 30 days for each. Importantly, the generated “neural stem cells” can be maintained for extended periods, at least 20 days, without loss of multipotency, and can be cryo-stored, allowing for high throughput generation of neural cell cultures at significantly reduced time. All ED-iPSC lines formed cortical neurons and type 1 and type 2 astrocytes with similar efficiency. The characteristic neural rosettes that are positive for Nestin formed for all lines, but with slightly differing efficiency. A more detailed analysis that includes cytoskeletal parameters and functional transport along axons is also underway. **Aim 3.2.** The neural precursors readily attach to Matrigel and PLL + Laminin for adherent differentiation protocols to pyramidal neurons or astrocytes. The differentiating neurons are more sensitive to re-attachment to coated surfaces. This is in contrast to astrocyte differentiation, in which differentiating cells were capable of lifting off and readily re-attached to form a monolayer for over 12 passages. *In vitro, ECM*
complexity is often minimized, but *in vivo* and in 3D tissue models greater complexity must be considered.

**Clinical Relevance:** The critical pipeline from the bench to the clinic for cell-based *in vivo* therapies requires identifying optimal cell types needed and how to effectively integrate into the target site. This includes controlling parameters such as transplanted cell survivability and retention, as well as restoring normal functional integration between the organ and the introduced cells. To more rapidly address these challenges, there is a critical need to develop *in vitro* functional tissue models that reflect normal function, or disease in a dish. Such animal-free platforms also accelerate translation to clinic for initial drug testing and optimizing novel medications from the lab to the patient.

My research on stem cell derived cardiomyocytes is aimed towards optimizing the key cell type needed for repair of cardiac infarcts that are a prominent subset of cardiac disease. It is important to note that the optimal cell type, the cardiomyocyte, has both functional and biomarker-based parameters that are anticipated to affect the efficiency of cardiomyocyte replacement therapies in cardiac disease. For example, in clinical trials there are currently several sources (iCs, iPSCs, ESCs) of cardiomyocytes being tested *in vivo*. Not all result in the same outcome. Thus, while these sources all generate functional cardiac tissue, the differences observed in their effectiveness may relate to different electrophysiology and epigenetic profiles, which differ between fetal or adult. That is, ESC derived cardiomyocytes may be more fetal in character, whereas iPSCs used from the patient may be more age and developmentally matched. This suggests that there are important functional transitions as an organ matures during development, which may impact how tissues mature. Currently, it is unclear how to fully define fetal versus adult
cardiomyocytes beyond perhaps electrophysiological profiles. It may be possible to transition derived cardiomyocytes from the fetal state to the fully mature adult state, and \textit{in vitro} models will be most useful in such studies because of the ability to readily evaluate function and transcriptomes.

The clinical needs in regard to CNS-based injury or disease are expanding rapidly. Ectoderm-derived support glia cells and neurons have the ability to replicate the interplay between support cells and neurons. We can use this information to establish \textit{in vitro} models to investigate the proper queues to understand the complexity of normal developmental function versus disease states in a 3D co-culture that recapitulates the cortical complexity \textit{in vitro}. This is proving true even for diseased states that occur over decades, which were once thought to be impossible to simulate \textit{in vitro}. As with cardiac therapies, it is expected that important functional transitions during development may impact maturation and functionality in the CNS in regard to stem cell derived therapies. For example, during early development, neural networks are generated but must be pruned for robust signaling. Whether it is important to include pruning in stem cell derived neural cultures is not yet known. These and other potential challenges will need to be addressed if stem cell-based therapies directed at cardiac and neural engineered tissues for \textit{in vivo} transplantation are to be successful, as well as when applied to \textit{in vitro} functional models for disease, normal development, and patient-specific treatments. Thus the transition between discoveries at the bench and their clinical translation with stem cells is much closer than possible compared to many other fields.
Thesis Introduction

Stem cell research coupled with nanotechnology and bioengineering [1-5] offer unique capabilities to biomimic in vitro multicellular formation of embryoids to tissues to organoids and organs. Most recently, human stem cell biology is being applied towards a range of prominent debilitating and untreated diseases such as Parkinson’s disease [6-9], cardiomyopathies [10-14], spinal injuries [15-19] and age-related macular degeneration [20, 21, 22]. The capacity of human embryonic stem cells (hESCs) [23] and induced pluripotent stem cells (iPSCs) [24, 25, 26] to differentiate along multiple lineages makes these cells highly desirable for use in biomedical advances. Pluripotency is evaluated in vivo by teratoma formation [27], as well as by multi-lineage differentiation in vitro, both of which together provide the most thorough analysis. For in vitro analysis, the generation and differentiation of 3D cell aggregates called embryoid bodies (EBs) [28] are often used. The use of EBs has also proven to be a successful strategy to generate organoids for modeling cardiac organ development [13, 14], brain organoids [30, 31] and intestinal 3D tissue constructs [32]. The growing reliance on EBs and other 3D cultured tissues for bioengineering applications and in disease studies includes applications in drug screening and analysis of tumor formation [33, 34], and for regenerative medicine [35].

One of the important hurdles facing 3D cell-cell assembly in vitro is the ability to control the temporal formation, shape and size of these multicellular assemblies to limit variability that is inherent in non-templated cell aggregates and which can affect statistical evaluation. Lithography-based bioengineering devices offer a solution to this non-
uniformity and allow high throughput templating. Additionally, lithography-based devices are an attractive approach, because they can be manufactured in large numbers with virtually identical output from batch to batch, coupled with the ability to control topological features in a range from micro- down to nano-scale. It thus becomes feasible to generate fully defined niches, specifically tailored to a particular cell-tissue experiment, or medical device need.

In addition to developing effective platforms to reproducibly analyze cells and tissues in synthetic environments my research has been instrumental to fully evaluate newly derived ethnically diverse-induced pluripotent stem cells (ED-iPSC) lines of low passage number from African-American, Hispanic-Latino, Caucasian, and Asian ethnicities by characterization of their “stemness”, in vitro. Stemness is defined as the ability to self-renew and multilineage differentiation potential. The ED-iPSC lines have undergone core quality control in the Cibelli lab and are free of mycoplasma, endotoxins, bacteria, yeast, mold and viruses. They exhibit normal karyotype after low passage maintenance in vitro on human foreskin fibroblasts (HFF) and in vivo exhibit pluripotency capability by teratoma formation. My research moved these cells to feeder-free extracellular xenofree matrices for subsequent RNA isolation and bioinformatics analysis at SUNY Buffalo (Dr. Michael Buck). To comparatively evaluate differentiation, I first optimized and evaluated EB formation. By design and fabrication of a custom lithography-templated microarray I was able to generate uniform EBs by size and shape and track formation. My subsequent differentiation analysis used these size-optimized uniform EBs to examine tri-lineage commitment in vitro into ectoderm, mesoderm, and endoderm with further terminal differentiation into neurons, glia, smooth muscle, cardiomyocytes, and endoderm pancreatic progenitors. My research focused on understanding transcriptome,
epigenetic and ethnic genetic contributions in cardiomyocyte functional differences I am observing, as well as to biomimic neuron-glial microarchitectures in culture in efforts to understand developmental plasticity and aging of the human cortex. The new human pluripotent minority lines, along with their extensive characterization under uniform platforms, generates one-of-a-kind new resource for New York, national, and international biomedical research that is expected to be important for stem cell-based translational clinical efforts.

The promise of stem cells to recapitulate aspects of normal or impaired development, mature cellular function and disease states provides hope for new insights applicable to regenerative medicine, cell and tissue therapies, and drug studies in human health care. Stem cell research is already providing important information on monogenetic and metabolic diseases by utilizing patient iPSCs for disease mechanism studies surrounding the affected cell types [36-38]. Analysis of early or late onset monogenetic diseases includes neurodegeneration, such as the role of SMN1 in early motor neuron death in a spinal muscular atrophy patient [39] or late onset familial Parkinson’s disease that is exacerbated by mutation of LRRK2 in dopaminergic neurons of disease patients [40, 41]. Disorder of carbohydrate metabolism in type I diabetes with insufficient production of insulin hormone may also be approachable through stem cell therapies [42].

By expanding the availability of iPSC lines that reflect new ethnic groups, and disease protection or predisposition, we create an opportunity to further refine our understanding of how gene expression can elicit phenotypically different responses in disease onset or drug treatments. The discoveries made through my thesis work incorporating the ethnically diverse iPSC lines I am characterizing will assist in identifying ethnic genetic contributions for improved biomedical care. As well my analysis of multifunctional tissue
platforms for cardiomyocytes and neuron-glia microarchitectures, described in the next paragraphs, will advance biomedical understanding of human disease and assist in improving therapies.

Stem cells allow us to potentially recapitulate aspects of organ development in a dish and therefore damage to those tissues and potential for recovery. The cardiomyocytes I generated will inform on early stages of heart development and may be particularly useful for cardiomyopathies, such as myocardial scarring. Being able to generate cardiac cells with high efficacy is therefore one of my goals. My thesis work is already providing insights into the critical need for a 3D microenvironment to promote development of Troponin-T banding that indicates polymerization and is required to generate functional beating heart tissue \textit{in vitro}. It known that certain drugs for treating cardiomyopathies have an ethnic contribution that affects efficacy of the drug. For example the drug BiDil is known to be more effective on African Americans compared to Caucasians, and that has been potentially linked to a small nucleotide polymorphism (SNP) on chromosome 8 [55]. The cardiomyocytes I have developed could provide an opportunity to analyze this locus and also possibly discover other contributing factors. Such studies would be useful in developing novel heart drug candidates tailored to the patient.

The brain is highly complex organ that has to balance biological functions and high order cognitive processes, such as memory, learning and thought. Its analysis on a whole is daunting; there are approximately 86 billion neurons, including perhaps 1000 types, and approximately 100 billion glia in the human brain. These neural cells occupy different regions of the brain that exhibit unique microarchitectures and abilities to cross-
communicate and exhibit plasticity. Despite this complexity, much of brain tissue function can be broken down in the interactions between neurons and their support glia.

The interest in neuron and glia communication is not a new one, however only recently have we been able to use human stem cell derived neural cells to ask precise questions. Only as recently as 2012, were human pyramidal neurons derived and shown to function in the neural circuitry when engrafted into a mouse brain. In my research, I have derived human pyramidal neurons using a modified version of this protocol and validated these cells initially by unique morphology and immunocytochemistry (ICC). To study neuron-glia communications, I am establishing a range of reduced to more complex systems that allow me to assemble and study microarchitectures with these cells. My goal is to be able to biomimic in vitro the interactions between neurons and glia to assist the study of normal brain plasticity versus neurodegeneration, such as the observed age-related loss of neurons and plasticity in Alzheimer’s disease, or loss of neuronal functions due to physical trauma.

Applying techniques that have been pioneered by the semiconductor industry to regenerative medicine is an attractive approach to generate the highly complex platforms that would be necessary to translate advances in the lab to the clinic in the future. The work done in fulfilment of this dissertation could facilitate future projects that aim to develop patient-specific pharmaceuticals and regenerative treatments for diseased organs. Generating a previously unavailable stem cell resource that is the panel of ethnically diverse induced pluripotent stem cells would allow for researchers to have access to high quality stem cell lines of underrepresented ethnicities, which can be used to study how to optimize and tailor treatments to the epigenetic background of a population and extend that knowledge to helping individual patients with treatments that are
optimized for their unique condition. In the long run, this will improve healthcare, and reduce medical costs due to the current imperfect one-size-fits-all approach.

Techniques Used

Reprogramming fibroblasts into induced pluripotent stem cell (ED-iPSC) lines.

Human fibroblasts for iPS cell derivation were obtained from Coriell Institute (Camden, New Jersey) and reprogrammed using a single polycistronic vector using four-factor 2A (4F2A) doxycycline (DOX)-inducible lentivirus encoding mouse cDNAs for Oct4, Sox2, Klf4, and c-Myc. The 5 human fibroblast lines were transduced by viral particles in xeno-free human fibroblast culture medium in the presence of polybrene (8 μg/mL). Forty-eight hours after infection, less than 15% of fibroblasts tested immunopositive for viral-derived OCT4. The medium was replaced two days after infection, and then daily, with xeno-free hES medium plus doxycycline (1 μg/ml) formulated to maintain stem cell pluripotency. After 35 days of culture, small cell clumps distinguishable from the fibroblast morphology appeared. Those that formed cell colonies with hESC-like morphology were mechanically isolated and passed on to mitotically inactivated xeno-free human foreskin feeder cells (ATCC PCS-201-010). Overall reprogramming efficiency by this method was calculated to be 0.002 ~ 0.004 %. The iPSC colonies were expanded for several passages under xeno-free conditions without doxycycline and evaluated for expression of markers of pluripotency.
qRT-PCR analysis of pluripotency of newly generated ED-iPSC lines.

Quantitative PCR analysis was done by isolation of total RNA from the hESC or iPSC lines and parental fibroblast lines. Primers for use in qPCR were first validated by maximally amplifying cDNA from a range of samples to confirm that a single PCR reaction product was produced and that the amplicon was of the predicted length. Each sample was analyzed by triplicate by an ABI PRISM 7000 sequence detection system, using the system’s software. The expression of gene of interest was normalized to GAPDH in all cases and compared with human ES cells (WA09 [H9], Wicell).

In vivo teratoma formation assay

iPSCs were injected subcutaneously in the flank region of NOD SCID gamma (NSG) mice (The Jackson Lab). After 12-24 weeks, teratomas formed from 10 iPSC lines, and tumors were excised & fixed in 10% normal buffered formalin (NBF) overnight. The samples were processed for histology by the Division of Human Pathology at MSU. Hematoxylin- and eosin (H&E)-stained sections were examined under a microscope.

hESC and ED-iPSC line maintenance, 2D and 3D passaging, and microarray formation of uniform embryoid bodies.

Induced pluripotent stem cell lines (ED- iPSCs) maintained on human foreskin fibroblast feeders were transferred to feeder-free conditions in non-tissue culture treated dishes coated with xenofree Vitronectin-XF (StemCell Technologies, Vancouver, Canada) or 1:100 Matrigel (10 mg/ml; BD Biosciences, San Jose, CA). Cells were maintained in mTeSR2 or mTeSR-E8 complete media (StemCell Technologies, Vancouver, Canada). Media was replaced on day 1 after the first passage of the series and cells grown
overnight. On day 2, slow release bFGF2 beads (20 microliters of PLGA beads loaded with bFGF2; StemBeads; Stem Culture Incorporated, Rensselaer, NY) were added with fresh media. Media changes were done every 2-2-3 days with StemBeads FGF2.

Preparation of uniform sized EBs from iPSCs colonies was done in custom lithography template microarrays (LTA) generated in-house. Chemical dissociation of the stem cell colonies into single cell suspension was done before and loading of the cells into LTA- polydimethylsiloxane (PDMS) grids in mTeSR2 or mTeSR-E8 media in the presence of 10 μM Rock inhibitor (Sigma-Aldrich, St. Louis, MO) at day 0. Stem cells were maintained in grids for five days with media changes every two days.

Generation of 200 μm uniform templated EBs (t-EBs) was done in custom microarrays of polydimethylsiloxane (PDMS) as previously described. To load microarrays we used chemically dissociated stem cell colonies into single cell suspension using the Gentle Cell Dissociation Agent, followed by suspension into mTeSR2 media containing 10 μM Rock inhibitor (Sigma-Aldrich, St. Louis, MO). t-EBs formed within five days in the grids with media changes every two days and removed by liquid expulsion with a p1000 micropipette.

**In vitro directed differentiation of hESC and ED-iPSC into astrocytes, pyramidal neurons, retinal pigment epithelial cells and pancreatic progenitors.**

Templated EBs (t-EBs), or single cells suspension of hESC WA09 or ED-iPSCs were suspended in mTesR2 or mTeSR-E8 media with ROCK inhibitor (StemCell Technologies; Vancouver, Canada) and incubated overnight on Matrigel or Vitronectin-XF coated TC-treated dishes. On day 1 of differentiation, the maintenance media was replaced with
Neural Induction Media (StemCell Technologies, Vancouver, Canada) and neural stem cells were generated per the protocol over 7-10 days. The resultant rosettes were then chemically selected and replated onto Matrigel-coated dishes in Neural Induction Media (NIM), supplemented with 10 µM ROCK inhibitor overnight.

For astrocyte generation, adhered neural stem cells were incubated in Astrocyte Media (Sciencell Research Laboratories, Carlsbad, CA). Complete media changes were done every 3-4 days. Successful differentiation was confirmed by immunocytology to identify multiple astrocyte specific biomarkers that are A2B5, CD44, Vimentin, and GFAP at day 17, day 20, day 24, and day 31 counting from the pluripotent cell state. Negative controls used here were Tuj1 and Oct4a. By day 7 (total day 17 in the protocol) in Astrocyte Media, the bulk of cells were GFAP and CD44 positive.

To generate neurons, t-EBs were plated onto Matrigel-coated dishes in mTeSR2 complete medium, allowed to adhere overnight, and then media changed to default defined medium supplemented with B27 and Noggin. Starting at day 2, 10 µM Cyclopamine was added to the media every two days for eight days. After that, the media was then replaced with DDM+B27 medium at two day intervals for 10 more days. Neural progenitors were then manually re-plated to Matrigel-coated dishes in DDM+B27 supplemented with 10 µM ROCKi. After five days, 50% of the medium was replaced with Neurobasal-A medium supplemented with B27 and 2 mM GlutaMAX, and 50% of the media was replaced with fresh media every 5 days until day 55+ of differentiation. Immunocytology of multiple neuronal biomarkers A2B5, Tuj1, MAP2, and VGlut1 was performed at critical media transition steps during the differentiation protocol.
For retinal pigment epithelial (RPE) cell generation, an adapted protocol was used. Starting with pluripotent stem cells at 80-95% confluency, the media was changed to mRPE media (RPMI + B27 + l-Glut + p/s + NEAA) and supplemented with 10mM Nicotinamide (NIC). Media was changed every other day for 14 days. At day 14, the mRPE media was supplemented with 10mM NIC and 140ng/ml Activin-A. At day 21, 250nM Retinoic Acid was supplemented to the media. Media was changed every other day until characteristic RPE cell morphology was achieved, usually between days 28 and 55. Once a confluent RPE sheet was formed, media was switched back to mRPE + NIC (10mM) and changed every 3-4 days.

To generate pancreatic progenitors, t-EBs, or single cells suspension were suspended in mTeSR2 media supplemented with 10µm ROCK inhibitor (StemCell Technologies, Vancouver, Canada) and seeded into Matrigel-coated dishes overnight. On day 1 of differentiation, the mTeSR2 media was replaced with media from the StemDiff endoderm differentiation kit (StemCell Technologies) and the protocol from the kit was followed. At day 5 the media was replaced with DMEM media (Life Technologies, Grand Island, NY) supplemented with B27+Insulin, GlutaMAX, pen/strep, and 50 ng/ml of KGF. The cells are incubated overnight, followed complete media refresh and an additional 2 more days incubation before the media was replaced and cells were grown in DMEM media supplemented with B27+Insulin, 2 uM Retinoic Acid, 250 nM Cyclopamine and 50 ng/ml Noggin. The cells were grown in this media for 4 days, with a complete media change at day 2. The differentiating pancreatic cells were maintained in DMEM media supplemented with B27+Insulin with media changes every 3 days. ICC was done to assess for the pancreatic biomarkers Pax6 and PDX1.
In vitro directed differentiation of ED-iPSC to cardiomyocytes and smooth muscle cells.

Templated embryoid bodies (t-EBs) were generated as previously described above from single cells suspensions of hESC WA09 or ED-iPSCs. The t-EBs were then removed from the microarrays, seeded onto Matrigel coated plates and incubated overnight before initiating the differentiation protocol. Single cell suspensions of either hESC WA09 or ED-iPSCs were also used as an alternate starting material for differentiation. To initiate differentiation, on day 1 mTeSR2 was replaced with RPMI media containing B27 supplement without Insulin, Glutamine, pen/strep, Non-Essential Amino Acids, 50 µg/ml of Ascorbic Acid, and 0.01mg/ml Matrigel. For Wnt pathway activation, the GSK3 inhibitor CHIR99021 was added to the media and cells are incubated for 24 hours. The media was then replaced with RPMI + B27-Insulin media and the cells incubated for an additional 48 hours. On day 3 of differentiation, 50% of the media was replaced with fresh RPMI + B27-Insulin media containing a second Wnt inhibitor, IWP2 and incubated for 48 hours. On day 5 of differentiation, the media was completely replaced with fresh RPMI + B27-Insulin media and incubated for 2 days. From day 7 onward, the media was replaced with RPMI + B27+Insulin and complete media changes were done every two or three days until robust contractions were observed. ICC was performed to assess expression of the cardiac-specific biomarkers Troponin-T, Nkx2.5 and GATA4.

Generation of smooth muscle cells was done by protocol optimized from literature. Single cell suspension of pluripotent stem cells were plated onto Matrigel, or Vitronectin-XF coated TC-treated dishes at around 60-80% confluency in mTeSR2 supplemented with 10µM ROCK inhibitor overnight. At day 1 of differentiation, cells were incubated in
Smooth Muscle Media I (DMEM + 15% FBS + GlutaMAX + NEAA + pen/strep) supplemented with 10 µM Retinoic Acid (RA). Media was completely changed every day for 10 days. At day 11, putative smooth muscle cells were washed with HBSS was incubated in Smooth Muscle Media II (DMEM + 15% KOSR + GlutaMAX + NEAA + pen/strep). Media was completely changed every day for 10 days. After day 21, the smooth muscle cells were maintained in Media II, changing it every 3 to 4 days. Cells were stained for alpha smooth muscle actin (α-SMA) to assess differentiation.

**Optimized cardiomyocyte differentiation protocol for ED-iPSCs lines.**

We hypothesized that a key step in our cardiac differentiation is the initial Gsk3 inhibition (CHIR99021), as has been reported previously. Experimentally, we observe that successful versus failed differentiation conditions are dependent on the number of cells that die during the initial step of the process and how they change morphology during this critical step. This generated differentiated cells were able to develop banded Troponin-T and spontaneously beating areas, as well as induce beating once the cells were forced into 3D cultures in the LTA-PDMS grids.

**Immunocytochemical analysis of pluripotent and differentiated cell cultures.**

For immunocytochemistry (ICC) of biomarkers in iPSC colonies, cells were fixed using 4% paraformaldehyde and permeabilized by 0.1% Triton-X 100 for 10min before antibody addition. Phase imaging for in vitro differentiated samples was done on a Nikon 80i epifluorescence microscope using a cooled QICam CCD camera. Fluorescent images were obtained on a Leica SP5 Laser Scanning Confocal Microscope and a Zeiss AxioObserver Z1 Inverted Microscope with Colibri LED illumination. Images were
captured with a Hamamatsu ORCA ER CCD camera and Zeiss Axiovision Rel 4.8 acquisition software. Figures were compiled using Adobe Photoshop (Adobe Systems Inc., San Jose, CA) and Microsoft PowerPoint (Microsoft Corp., Redmond, WA) software. The immunocytology of 2D cell cultures or three dimensional EBs was done by fixing the samples in 4% paraformaldehyde as detailed above. Immediately before incubation with antibodies, the cells were permeabilized with PBS + 0.5% Triton X-100 and nonspecific binding was blocked in 1% BSA in HBSS. After staining, the prepared samples were mounted in ProLong Gold antifade reagent at 20°C overnight in the dark before imaging immediately or storing at 4°C for up to 3 months.

**SU-8 processing to generate device moulds.**

SU-8 (Microchem) was processed using standard lithography techniques, expanded to include a post-exposure bake at 150°C for 3 hours, which relieves stress within the cross-linked SU-8 and eliminates internal and external fractures formed during the fabrication process. To generate 3D moulds to use for casting of our engineered templated devices. The master moulds were further treated with Sigmacote (Sigma) to passivate the surface and facilitate peeling off of the PDMS-based devices.

Components of the templated devices that were made from SU-8 were manufactured via the above protocol. An additional plasma treatment for 30-45sec step was introduced after processing, immediately before device assembly to turn the surface temporarily hydrophilic so that liquid and cell suspensions could flow through the fabricated channels and patterning openings.
Polydimethylsiloxane casting for devices and templated surfaces.

Polydimethylsiloxane (PDMS) was prepared in 10:1 (general), 15:1 (cardiac), and 20:1 (neural) ratios of elastomer to cross-linker, depending on the cell type that would be integrated within the device. After careful peeling off of the castings, the devices were sterilized under UV for 30min and stored until ready to use. Just prior to assembly and loading, each device was carefully cut out from the casting and assembled per design. Assembled devices were then again sterilized for 30min under UV, and plasma-treated for 30-45sec to facilitate cell and 3D culture loading.

Micro-contact printing to generate functionalized patterned surfaces.

Micro-contact printing was done on plasma-treated glass surfaces per established protocols, using PDMS stamps manufactured at CNSE’s cleanroom suite. Assuring specific surface functionality and pattern durability was done by silanizing the glass surface after treatment with an appropriately terminated silane, either –CHO, for visualization, or –HN₂ for specific pattern functionality. Silanization of glass and PDMS surfaces was done after plasma-cleaning them, using in-house protocols to produce CVD silane layer on the surface.

3D model design and rapid prototype optimization of engineered devices.

Common 3D modeling and design software, such as Blender, AutoCAD, and 123D were used to generate concepts for novel devices and to produce 3D printable versions of those designs, either as actual product, or as a mould to cast the final device out of the appropriate material. The 3D printed devices were produced and optimized in the CNSE
nanobioengineering constellation’s rapid-prototyping suite utilizing the Form1+ stereolithographic printer (FormLabs).

Nucleic acid preparation for RNA-Seq and ChIP-seq.

Total RNA isolation for transcriptome analysis of the eight ED-iPSC lines and the four initial fibroblast cell lines were done using the Ambion PureLink Mini RNA isolation kit (Life Technologies, Grand Island, NY). Total RNA integrity value following isolation was measured using the Agilent Technologies 2100 Bioanalyzer and was equal to 8-10 for all tested samples, indicating high quality pure RNA preparation. mRNA preparation, library construction, and purification were done according to the TruSeq RNA Sample Preparation v2 low sample (LS) protocol (Illumina, San Diego, CA). Libraries were normalized based on qPCR values and pooled using the TruSeq SR Cluster Kit v3-cBot-HS (Illumina, San Diego, CA). Pooled samples were four-plexed and sequenced with the HiSeq 2500 v3 sequencer (Illumina, San Diego, CA) producing 30 to 50 million single-end 50 bp reads.

Chromatin extractions for ChIP-seq analysis of the ED-iPSC cell lines were done using the Chromatin Shearing Optimization Kit - Low SDS (Diagenode, Denville, NJ). Cells were fixed with 1% formaldehyde for 10 min before sonication. Shearing efficiency was evaluated and the whole amount of each purified sheared chromatin sample was run on a 1.5% agarose gel to observe a required smear of DNA fragments between 100-350 bp. Following, ChIP DNA was prepared with the Auto Histone ChIP-seq kit protein G (Diagenode, Denville, NJ) using 1 million cells, 2.4 μg H3K27ac and 3 μg H3K4me1 premium polyclonal antibody (Diagenode, Denville, NJ) for each 100 μl immunoprecipitation reaction. ChIP DNA was quantified using the Quant-iT PicoGreen ds
DNA Assay Kit (Invitrogen, Carlsbad, CA) and used for library construction with the ThruPLEX™-FD Prep Kit (Rubicon Genomics, Ann Arbor, MI). Constructed ChIP-seq libraries were purified and library quality control was performed with the Agilent Technologies 2100 Bioanalyzer. Pooled samples and reference inputs were four-plexed and sequenced with the HiSeq 2500 v3 sequencer (Illumina, San Diego, CA) producing 20 to 56 million single-end 50 bp reads.

**RNA-Seq data generation and analysis.**

Raw reads were analyzed with FASTQC (v. 0.11.3) for quality control. Overrepresented (e.g. adapter and similar technical) sequences remaining in the raw reads were assessed and subsequently removed using the fastq-mcf module (v 1.04). Low quality base threshold was set at 25. Following technical sequence and low quality base removal using reads that were shorter than 20bp were filtered out. Transcript quantification was done using a reference genome by RSEM (v. 1.2.22) with default parameters, which utilizes Bowtie2 for read alignment. Differential expression analysis was done using EBSeq (v. 1.10), which uses an empirical Bayes model. Expression levels were subject to median normalization and genes with a fold change greater than 2 and a posterior probability of differential expression >= 0.95 were chosen (a target false discovery rate of 5%) were considered as significantly differentially expressed.

For each comparison, EBSeq model convergence was manually verified. Clustering of samples and/or genes was done using the Unweighted Pair Group Method with Arithmetic-mean method and Pearson's correlation as the distance measure. The Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.7 was used for functional analysis of the gene lists interrogating Biological Process (BP), Molecular
Function (MF), and Cellular Component (CC) Gene Ontology (GO) categories and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. Biologically relevant categories that are over-represented in the gene set and therefore may be of further interest were assessed using the Expression Analysis Systematic Explorer (EASE) score in the DAVID tool. We picked GO categories that have EASE scores of 0.05 or lower as significantly over-represented.

**ChIP-seq data generation and analysis.**

Genomic sites enriched for the H3K27ac and H3K4me1 post-translational histone tail modifications common to one or more of the eight investigated ED-iPSC cell lines, were isolated and merged with a 200 bp window using Bedtools v2.17.0. Coverage was estimated for each of these intervals using Bedtools v2.17.0 and normalized to the total number of reads per one million (rpm) for all samples and reference inputs using SAMtools v1.1. Final standardized coverage values for all samples and enriched genomic sites were plotted as a heatmap with the R package Pheatmap version 0.7.7. GO biological processes related to differentially marked genomic sites were identified using GREAT.
Chapter References:


Chapter 2:

The Human Embryoid Body Internal Architecture Complexity Has Significant Implications for the Differentiation Potential of Stem Cells

State of the Field

Three dimensional culture is an exciting development in the stem cell field, as they have been shown to recapitulate differentiation conditions in vitro more closely to in vivo than traditional 2D culture. In pluripotent stem cell differentiation, embryoid bodies (EBs) provide a three-dimensional multicellular precursor in lineage specification in a large number of differentiation protocols. In addition, the internal structure of EBs is not well characterized yet is predicted to be an important parameter to differentiation. Unfortunately, the nature of the most common EB generation techniques come with inherent limitations that makes adoption of these 3D cultures in the biomedical sector, and as high throughput alternative to traditional culture methods difficult. One common method of EB generation, the free suspension, generates many EBs relatively easily, but the generated 3D cultures vary tremendously in size and shape and are prone to fusing while in culture. Another common method of EB generation, the hanging drop, eliminates size and shape variability, but it is very time intensive, requires high skill level, and is relatively low yield.

These methods are appropriate for early research strategies to develop protocols and proof-of-principle treatments, but in order to move the field from the bench to the
Recent advances in bioreactors have allowed for size and shape control, while maintaining high throughput, but fusing is still an issue and this method requires high upfront investment, in addition to specialized tools and skills. While bioreactors offer a lot of promise to generate meaningful numbers of relatively uniform spheroids for tissue engineering applications, the high upfront investment that the system requires makes it an unsuitable option for smaller scale projects, such as benchtop research and process optimizations. Also, in cases where multiple different 3D cultures need to be generated at the same time in a limited space. While bioreactors would certainly find their place in manufacturing and scale-up of processes, lithography templated devices would serve an important role to bridge the current gap between the very high throughput required in biotech production or the clinical setting and the initial stages of research and optimization needed to drive new technologies and discoveries to the patient bedside.

The outlined hurdles make it hard to reproducibly use EBs as a reliable platform to initiate differentiation or in other downstream applications that require stem cells. Therefore, in order to utilize the benefits of EB 3D culture for stem cell development, the major drawbacks of the method need to be minimized, while 3D culture specific aspects, such as size and shape of spheroid microarchitecture need to be characterized more fully.

**Hypothesis**

The embryoid body was initially developed as a 3D analog of early-stage gastrulation, which was predicted to aid in differentiation efficacy and as a tool for organoid development. Surprisingly, even though the structure of EBs was known, little else was known of the functional parameters of EBs. This included importance of size, shape and
internal structure architectures, as well as changes that occur in their internal architecture during transition from 3D spheroid to 2D cell layer as commitment progresses during differentiation. Therefore, we hypothesize that in order to use pluripotent templated EBs of defined sizes as an efficient and reproducible initial step in multiple differentiation and organoid development strategies, we need to understand what are their internal morphologies and how do they affect the 3D transition to 2D during endoderm, ectoderm, and mesoderm commitment.

**Strategy**

We propose to develop a platform that can produce high numbers of uniform 3D cultures by using lithography inspired techniques. Major advantages of lithography templated platform over bioreactors would be the price, ease of use, and relatively low upfront cost to generate templated EBs. The produced platform will be used to generate templated stem cell embryoid bodies (EBs), which are a common precursor in multiple differentiation strategies. We plan to address the issues of size and shape variation, as well as minimize formed EB fusion while in culture prior to differentiation, identified previously as the main issues in moving 3D pluripotent culture into high throughput applications. In order to generate uniform EB cultures, we will develop a lithography-templated array made out of polydimethyl siloxane (LTA-PDMS) which is able to generate hundred templated spheroids made of pluripotent stem cells of defined diameter, maintain their shape, and prevent EB fusion during culturing.
We will use custom SU-8 moulds to generate transparent lithography-templated arrays of polydimethylsiloxane (LTA-PDMS) for high throughput analysis of human embryonic stem cell (hESC) EB formation and internal architecture. EBs will be formed in 200 and 500 µm diameter microarray wells by use of single cells, 2D clusters, or 3D early aggregates and compared for internal structure and multi-lineage pluripotency potential. Initial work in this area has shown that templated arrays can generate similar spheroids of pluripotent stem cells. We propose to generate a platform that in addition to templating high throughput EBs, could also be used to track spheroid formation, and be used to evaluate microarchitecture in 3D cultures.

Acknowledgements and Contributions:

I conducted the experiments, figure assembly, data collection and analysis, as well as manuscript writing. Zach Olmsted contributed to data collection, figure assembly and manuscript writing. Janet Paluh designed and supervised experiments, analysis and manuscript and figure preparations. Copyright permission for the purposes of this dissertation can be found at: [http://media.wiley.com/assets/1540/98/ctavchglobal.pdf](http://media.wiley.com/assets/1540/98/ctavchglobal.pdf)

Parts of this work have been published in:

Abstract

In pluripotent stem cell differentiation, embryoid bodies (EBs) provide a three-dimensional multicellular precursor in lineage specification. The internal structure of EBs is not well characterized yet is predicted to be an important parameter to differentiation. Here we use custom SU-8 moulds to generate transparent lithography-templated arrays of polydimethylsiloxane (LTA-PDMS) for high throughput analysis of human embryonic stem cell (hESC) EB formation and internal architecture. EBs formed in 200 and 500 µm diameter microarray wells by use of single cells, 2D clusters, or 3D early aggregates were compared. We observe that 200 µm EBs are monocystic versus 500 µm multicystic EBs that contain macro-, meso- and micro-sized cysts. In adherent differentiation of 500 µm EBs, the multicystic character impairs the 3D to 2D transition creating non-uniform monolayers. Our findings reveal that EB core structure has a size dependent character that influences its architecture and cell population uniformity during early differentiation.
Introduction

Stem cell research coupled with nanotechnology offers unique opportunities for biomimicry and in vitro studies of multicellular development from single cells or embryoids to tissues, organoids and organs. Most recently, human stem cells have been applied towards a range of prominent diseases such as Parkinson’s, cardiomyopathy, spinal injuries and age-related macular degeneration. Organoids derived from stem cells often employ EBs when modeling organ development. There continues to be a growing reliance on EBs and other 3D cultured tissues in bioengineering applications that include disease studies, tumor formation, drug screening and regenerative medicine.

EB technology remains important to the stem cell field. The multilineage capacity of human embryonic stem cells (hESCs) and induced pluripotent stem cells (iPSCs) is evaluated in vivo by teratoma formation but also in vitro initiating from two-dimensional (2D) stem cell colonies or 3D EBs. The typical self-assembly of EBs in free suspension static cultures is by nature non-homogeneous in size and shape and can be complicated by the fusion of multiple EBs. In mouse studies, single cell suspension formed EBs templated in lithography-enabled poly(ethylene glycol) (PEG) and polydimethylsiloxane (PDMS) hydrogel well arrays revealed more uniform gene expression of differentiation biomarkers. These studies also revealed differences on lineage differentiation related to EB sizes. In another study, to increase the efficacy of cardiogenesis, human stem cell EBs were formed from single cells seeded into polyurethane (PU) cuboidal microwells to form uniform 2D colonies then released into free suspension for uniform EB formation. These EB templating studies in addition to
commercially available high throughput cell aggregate tools \[46\], reveal that the uniformity and size of EBs are important parameters that impact differentiation efficiency.

In pluripotent stem cell differentiation, embryoid bodies (EBs) provide a three-dimensional multicellular precursor in lineage specification. The internal structure of EBs is not well characterized yet is predicted to be an important parameter to differentiation. Here we use custom SU-8 moulds to generate transparent lithography-templated arrays of polydimethylsiloxane (LTA-PDMS) for high throughput analysis of human embryonic stem cell (hESC) EB formation and internal architecture. EBs formed in 200 and 500 µm diameter microarray wells by use of single cells, 2D clusters, or 3D early aggregates were compared. We observe that 200 µm EBs are monocystic versus 500 µm multicystic EBs that contain macro-, meso- and micro-sized cysts. In adherent differentiation of 500 µm EBs, the multicystic character impairs the 3D to 2D transition creating non-uniform monolayers. Our findings reveal that EB core structure has a size dependent character that influences its architecture and cell population uniformity during early differentiation.

What additional factors contribute to influence size performance of EBs is unknown. To address this we engineered custom sized lithography templated arrays of PDMS (LTA-PDMS) and used these to generate uniform EBs from hESCs of two different sizes for comparison. We evaluated EB formation by single cells, 2D clusters, or 3D early aggregates to ensure that any differences we observed were not an artifact of EB formation method. We observe a single hollow core or cyst in our 200 µm LTA-PDMS EBs, that is in contrast to 500 µm LTA-PDMS EBs that are multicystic and exhibit varied cyst morphology. We describe these cysts as large macrocysts, polarized mesocysts and microcysts. Importantly, the cystic architectures in 200 or 500 µm EBs were independent of formation method. Cystic complexity impacted early differentiation, with 200 µm EBs
readily transforming from 3D to 2D monolayer, while the multicystic character in 500 µm EBs generated mixed 3D/2D complexity during adherent early differentiation. Our work reveals new morphological information linking EB core complexity to size and also its importance to optimizing differentiation with hESCs that go through EB precursors.

**Materials and Methods**

**Lithographic fabrication and generation of polydimethylsiloxane grids and microarrays for a modular, multifunctional engineered platform (LTA-PDMS).**

The L-Edit program (Tanner EDA, Monrovia, CA) was used to design a 2D pattern for the transparency masks (Fineline Imaging; Colorado Springs, CO) and subsequent photolithography. SU-8 2050 spin coated moulds with micropatterned features were generated on 4" silicon wafers and processed as manufacturer's instructions (Su-8 2050 and 2075 series resists; Microchem.com). For the micro grids, we generated cylindrical wells with diameters of 200, 300, 400 and 500 µm and depth of either 200 or 400 µm. Based on 3D culture generation, both 200 and 500 µm wide wells were able to generate 3D stem cell cultures, or EB formation, cell retention during washing steps, and 3D culture retrieval for further analysis or downstream applications, such as differentiation or post-patterning within the rest of the LTA-PDMS platform. Well diameters of 300 and 400 µm generated similar 3D culture artefacts as the 500 µm EBs in our preliminary analysis and so were not characterized further.

In addition to the LTA-PDMS grids, we also developed modular LTA-PDMS arrays, which comprise of a common device base, which can be paired with different loading device covers, or other modifications, for 3D culture post-patterning experiments, or 2D
and 3D culture chambers. The Su-8 moulds for the arrays were done the same way as the ones for the grids. Each LTA-PDMS array has 36 identical wells that have surface area of 1 mm² and are 200 µm deep. Each well has a build-in unique feature that allows for alignment of the device for tracking of live cultures over a time-course, and then matching the corresponding chambers with the generated data during immunocytochemical, or other, analysis.

Polydimethylsiloxane (PDMS) was prepared according to manufacturer instructions (Sylgard; Dow Corning, Midland, MI), with 10 parts polymer to 1 part cross-linking agent and mixed thoroughly. The SU-8 mould was treated with 1ml of Sigmacote (Sigma-Aldrich, St. Louis, MO) before spin coating PDMS to improve castings release. To generate the actual PDMS device, a spin step at 10 rpm for 1 minute was used to apply a uniform thickness PDMS layer, then degassed under vacuum for 15 minutes. Alternatively, 5ml of PDMS mixture was applied onto the wafer and degassed under vacuum for 15-30 minutes, until all visible trapped bubbles pop. Cross-linking of the PDMS-SU-8 mould was done by baking at 60°C for 90 minutes, followed by slow cool to room temperature. Both methods generate devices with mechanical robustness that is sufficient to allow for cell seeding under various conditions, as well as survive common UV, plasma, and ethanol sterilization conditions. The PDMS cast was then carefully peeled off to generate the lithography template array (LTA)-PDMS grids. The mould comprising of multiple grids was sterilized in 70% ethanol overnight at room temperature then stored until use. Immediately before preparation for cell loading, each LTA-PDMS grid was UV sterilized for 30 minutes.
Human ESC maintenance, 2D passage, non-templated EB formation and directed early germ layer commitment.

Human ESC line WA09 (WISC Bank, WiCell, Madison, Wisconsin) \cite{23} was maintained in feeder-free conditions on tissue culture treated dishes coated with 1:100 Matrigel (10 mg/ml; BD Biosciences, San Jose, CA) diluted into Hank's Buffered Saline Solution (Gibco HBSS; Life Technologies, Grand Island, NY), or on non-tissue culture treated petri dishes coated with Vitronectin-XF (StemCell Technologies, Vancouver, Canada). Stem cells were maintained in either mTeSR2 or mTeSR-E8 complete media (StemCell Technologies, Vancouver, Canada) and mechanically passaged between days 5-7. On day two following passaging slow release PLGA beads loaded with bFGF2 (StemBeads FGF2; Stem Culture Incorporated, Rensselaer, NY) were added with fresh mTeSR2 media. Media changes were done every three days with StemBeads FGF2 and mTeSR2. Preparation of non-templated EBs from hESC colonies was done by mechanical passaging to low adhesion 35 mm dishes in mTeSR2 media in the presence of 10 µM Rock inhibitor (Sigma-Aldrich, St. Louis, MO) over three days and with daily media changes. For directed multilineage differentiation we used the Human Pluripotent Stem Cell Functional Identification Kit (RnD Systems, Minneapolis, MN) and for early endoderm differentiation comparison we also used the STEMdiff Definitive Endoderm kit (StemCell Inc., Vancouver, BC).
EB formation in the LTA-PDMS grids from single cells or mechanically passaged intermediates.

Single cell suspensions were generated from 2D stem cell cultures using Gentle Dissociation Agent (StemCell Inc., Vancouver, BC) for 10 minutes at 37°C, 5% CO₂, followed by slow trituration. Plasma treated LTA-PDMS arrays (200 µm and 500 µm wells) were primed with 20 µl of warm mTesR2 media, loaded with single cell suspensions, and then kept stationary at 37°C, 5% CO₂ for 10min to allow cells to settle. Fresh mTeSR2 media with 10 µM ROCK inhibitor was added and the media changed every other day. ROCK inhibitor was discontinued after day two.

We used mechanical passaging to generate 2D cell clusters and subsequent 3D early aggregates. Aggregates are preformed in free suspension for two days after mechanical passaging. A 120 µl suspension in mTesR2 media was used to pipet load three 500 µm size grids of 100 wells per array and provides 300 EBs. A 30 µl suspension was used to load four 200 µm size grids of 100 wells per array and provides 400 EBs. Alternatively a short 30-minute vacuum draw can be used to coat wells with media, without plasma treating. Stem cell clusters or aggregates were allowed to settle at 37°C, 5% CO₂ for 10 minutes, then the grids washed gently to remove excess cells from the surface. To wash, 2 mls of mTeSR2 media was added at the edge of the dish and allowed to overflow into the grids. From this point the formation of EBs and their internal architecture was tracked in the transparent LTA-PDMS arrays. Bulk removal of EBs from wells was done by placing the array in a 1.5 ml microcentrifuge tube and gentle tapping or slow mixing with 200 µl mTeSR2 or mTeSR-E8 media.
**Microscopy and immunocytology**

Phase images were collected on a Nikon 80i epifluorescence microscope with a PLAN 10X 0.30 NA DL objective and a cooled QICam CCD camera. Fluorescent imaging was done on a Leica SP5 Laser Scanning Confocal Microscope with HC PL FLUOTAR 10X 0.30 NA or HCX PL APO CS 20X .70 NA objectives and on a Zeiss AxioObserver Z1 Inverted Microscope with Colibri LED illumination, 100X oil 1.45 NA PlanFLUAR or 63X Plan-Apochromat 1.4 NA oil DIC objectives, and Hamamatsu ORCA ER CCD camera using Zeiss Axiovision Rel 4.8 acquisition software. Images were compiled using Adobe Photoshop (Adobe Systems Inc., San Jose, CA) and Microsoft PowerPoint (Microsoft Corp., Redmond, WA) software. Immunocytology of 2D cell cultures and 3D EBs was done following modified fixation in 1% glutaraldehyde [47] or by -20°C methanol for 15-20 minutes, followed by gentle washing in HBSS (Life Technologies, Grand Island, NY) or by methanol fixation. Nonspecific binding was blocked by 3 x 20 minute incubations in 1% BSA-HBSS, followed by an additional HBSS wash. Pluripotency primary antibodies used (1:1000 each) were anti-Oct4A C-10 (Santa Cruz Biotechnology, Dallas, TX) and anti-SSEA4 (Millipore, MA). Differentiation lineage antibodies (1:100 each; Human Pluripotent Stem Cell Functional Identification Kit RnD Systems, Minneapolis, MN) were anti-OTX2 (ectoderm), anti-SOX17 (endoderm), and anti-Brachyury (mesoderm). Secondary antibodies used (1:1000) were AlexaFluor 488 and AlexaFluor 594 (A-11001, A-11037, Invitrogen, Carlsbad, CA) and were used with Hoechst in 3% BSA-HBSS at 20°C for two hours with a final HBSS wash. Mounting was performed in ProLong Gold at 20°C overnight in the dark before imaging or storing at 4°C.
Results

LTA-PDMS EBs resemble non-templated EBs in gross structure, pluripotency and size-dependent behavior.

Custom PDMS microarrays, referred to as LTA-PDMS, consisting of 100 wells with diameter of either 200 or 500 μm were generated for uniform EB formation and analysis. Templated EB formation in LTA-PDMS used mechanically passaged 2D clusters or 3D early aggregates. The individual uniformly sized EBs were tracked during early differentiation by immunocytology and phase microscopy (Figure 1). Excess non-seeded clusters and aggregates were retained and allowed to form EBs in free suspension as controls. Representative images of individual EB formation tracked in the transparent LTA-PDMS 200 and 500 μm wells starting with 2D clusters are shown (Figure 1A). The ability to form EBs and their gross morphology were comparable between LTA-PDMS or non-templated EBs. The individual tracking of EBs over time and the ability to monitor morphological changes during formation of individual EBs (Figure 1A) was only possible using microarrays and provides an additional set of parameters for comparative EB analysis. For example, we observed that EBs retained longer in the wells (Figure 1A, 200 μm, day 5) grow out of wells or collapse inward indicative of a less dense core. The early 3D aggregates generated for EB templating are close to 200 μm size and rapidly fill wells within 3 days of seeding. The advantage of using 3D early aggregates is both in ability to form EBs and time required for formation. We observed that, similar to free suspension controls, the 2D stem cell clusters do not always form EBs. We achieved generally 36% (200 μm) and 47% (500 μm) efficiency of EB formation in the LTA-PDMS arrays using 2D clusters. Whereas by comparison, using 3D early aggregates we observed 91% (200 μm)
or 95% (500 μm) efficiency of EB formation in the LTA-PDMS arrays. Standard deviation in all of these cases was less than 4% (N>300). We next examined directed early differentiation along three germ layers of LTA-PDMS EBs for multilineage potential (Figure 1B and 1C). EBs were allowed to attach to 1:100 Matrigel coated glass using coverslips in circular dishes or in multi-well chambers. Germ layer adherent differentiation was done using the RnD Systems Human Pluripotent Stem Cell Functional Identification Kit over the recommended period for each germ layer. ImmunocytoLOGY staining (Figure 1B) with Hoechst and antibodies against OTX2 (ectoderm), SOX17 (endoderm), and Brachyury (mesoderm) revealed expected staining.

We observed that similar to control EBs, the templated EBs differentiated along all germ lineages. However, EB size affects transition time to monolayer formation (Figure 1B see standard and reversed images; and Figure 1C). The 200 μm EBs more rapidly transition from 3D to a uniform 2D monolayer. By contrast, 500 μm EBs retain a gradient of 3D central and 2D outermost features during what we refer to as the 3D/2D transition. This is particularly evident when comparing day 3 of three-day time lapse tracking of differentiating EBs (Figure 1C). Large non-templated EBs behave similarly to the 500 μm LTA-PDMS EBs. Dumbbell shaped structures formed from fusion of differently sized EBs in free suspension cultures, reflect the size characteristics of their fused parts. That is, the smaller end transitions more quickly from 3D to 2D (see Figure 1C, Ectoderm differentiation). Our findings indicate that the LTA-PDMS and non-templated EBs form with similar gross structure and exhibit multi-lineage potential. In addition use of microarrays for comparative analysis of EB differentiation reveals longer retention of 3D/2D mixed morphology with larger EBs.
Figure 1. Templated EBs are pluripotent but reveal size-dependent differentiation profiles. (A) Time course for EB formation in 200 or 500 μm wells seeded with 2D clusters. 200 μm EBs form in 3 days while larger 500 μm EBs require 5 days by this method. (B) Pluripotency of EBs. Immunocytochemistry of early germ layer biomarkers during multi-lineage differentiation of 200 μm, 500 μm and free suspension (non-templated) control EBs. Images shown are day 3 post-differentiation (RnD kit) for Ectoderm (OTX2), Endoderm (SOX17) and Mesoderm (Brachyury-T). Inversion of the merged image highlights the dispersal pattern of cells during differentiation. Templated EBs differentiate similarly to controls, including the mixed 3D/2D character present with 500 μm EBs. (C) Time course comparison of 200 μm, 500 μm and control EBs during directed differentiation (RnD kit) along Ectoderm, Endoderm and Mesoderm. By day 3 the 200 μm EBs transition readily from 3D spheroids into 2D cell monolayers, while 500 μm and larger free suspension EBs retain mixed 3D/2D morphologies with visible internal cyst architecture. Scale bars are 200 μm.
The LTA-PDMS 500 µm EBs display multiple cysts as part of a complex internal architecture.

LTA-PDMS formation of EBs revealed a less dense core for 200 µm EBs and possible complexity of 500 µm EB core (Figure 1). Previous studies in both human and mouse EBs also indicate a more tightly compact EB surface versus the core \(^{[48-51]}\). Surprisingly limited information is available on the core architecture of EBs or how EB size impacts this architecture as well as early differentiation when initiated from EBs. We used confocal microscopy and immunocytoology to investigate the core architecture of uniformly sized LTA-PDMS EBs formed with mechanically passaged 2D clusters. We observed either a single monocyst in 200 µm EBs (Figure 2A) or multiple cysts in the 500 µm EBs (Figure 2B to 2D; supplemental movies). In Figure 2B a 20X confocal image is shown for day 4 500 µm EBs undergoing early endoderm differentiation (illustrated in Figure 2C) along with multiple 63X confocal slices through that image stack. Multiple cysts of varying size are observed as well as some cysts that exhibit a multicellular polarized structure. These structures persist during early differentiation in multiple protocols we applied, as well as in other lineages (Figure 2D) and across different EB formation methods (Figure 3 and Figure 4).
Figure 2. Multiple cysts present in the core of 500 μm EBs contribute to mixed 3D/2D morphology during early differentiation. Cysts in LTA-PDMS EBs formed from mechanically passaged intermediates. 

(A) Phase and confocal images show the monocystic structure of undifferentiated day 3 formed 200 μm EBs, generated using 2D cell clusters. Phase image is loaded LTA-PDMS EB. Top panel: single EB in microarray and confocal section stained for Hoechst (nucleus, blue), pluripotency marker SSEA4 (surface, red), Actin (surface, green) and merged image. Bottom panel: four confocal sections of an extracted EB under a coverslip stained with Hoechst (nucleus, blue), SSEA4 (surface, red) and tight junction protein ZO1 (green). Scale bars are 200 μm. 

(B) The multiple cyst structure present in 500 μm EBs is retained during early endoderm differentiation. Two representative LTA-PDMS 500 μm EBs differentiated by the STEMdiff (top) or by RnD kit (bottom). Z-stack images of two representative cysts are shown for each EB (Merged: Hoechst, blue; ZO1, green; SOX17, red). Scale bars are 200 μm (left) and 40 μm in zoomed images (right). 

(C) Schematic of EB differentiation from non-attached EB to mixed 3D/2D adherent cells. 

(D) Differentiating 500 μm EB (RnD; mesoderm) with characteristic presence of multiple cysts (monochrome image of Hoechst staining). Cartoon schematic of multiple cysts enclosed within a large cavity is shown in the second panel.
We previously observed rapid formation of uniform monolayers during multi-lineage adherent differentiation of 200µm EBs versus 500µm EBs (Figure 1) and observed mono versus multiple cyst features dependent on size (Figure 2). We next examined the adherent interface of 500µm EBs during early differentiation. Confocal imaging of 500 µm EBs formed with 3D early aggregates that are undergoing endoderm differentiation (STEMdiff and RnD kits) is shown in Figure 3A and 3B. Slices through a confocal image stack from adherent layer to top of the differentiating EBs illustrate the variability between each layer in cell-cell organization, including gaps, and different 3D versus 2D features. Similar findings were observed independent of protocol, however colony height varied (Figure 3C, N=30 EBs each). The LTA-PDMS 500µm EBs differentiated by the RnD kit were 70-100µm in height (average 84.33±8.92µm) and those differentiated by the STEMdiff kit were 20-80µm in height (average 54.90±16.18 µm).
Figure 3. The colony profile of differentiating 500 μm EBs is a complex architecture of mixed 3D/2D morphology. (A) Representative confocal sections through differentiating 500 μm EBs at day 5 of STEMdiff endoderm differentiation (Merged: Hoechst, blue; ZO1, green; SOX17, red). Sections shown were chosen as the most informative slices to reveal the 3D/2D transition profile. EB1 sections at 6, 11, 18, 29, 51 μm; EB2 sections at 9, 13, 18, 23, 27 μm; EB3 sections at 11, 17, 31, 45, 48 μm. (B) Representative confocal sections through differentiating 500 μm EBs at day 4 of RnD endoderm differentiation (Merged: Hoechst, blue; ZO1, green; SOX17, red). EB1 (sections at 10, 23, 31, 52, 67 μm); EB2 (13, 27, 35, 49, 71 μm); EB3 (9, 23, 42, 56, 77 μm). Scale bars are 200 μm. See also Supplemental Movies. (C) Compared height distribution retained during early endoderm differentiation of 500 μm diameter EBs (n=30).
Three distinct cyst types contribute to core complexity of 500 µm LTA-PDMS EBs independent of EB formation methods.

The complexity of 500 µm vs 200 µm EBs in cyst number is independent of type of mechanically passaged cells used, that is 2D clusters or 3D early aggregates (Figures 1, 2, and 3). To determine whether multiple cysts present in 500 µm EBs are artifacts of fusion of mechanically passaged intermediates we generated LTA-PDMS EBs from single cells (Figure 4A) and performed immunocytology and confocal imaging (200 µm, Figure 4B; 500 µm, Figure 4C). The formation time to generate the 500 µm EBs from single cells is double that (12 days versus 6 days) of 2D clusters or 3D cell aggregates. However, the cyst character of 200 µm (monocystic) and 500 µm (multiple cysts) EBs formed from single cell suspensions was similar to what we observed by EB formation from mechanically passaged intermediates.

To examine the cell-cell contacts within the cyst structures in 500 µm EBs, we used the biomarker ZO1, required for maturation of tight junctions [52]. ZO1 and Hoechst staining revealed three types of EB cysts that we refer to as macro-, meso- and micro- cysts based on cyst size and surrounding cell organization (illustrated in Figure 2E and Figure 4D). The macrocyst is a large single cavity that is 10 to 20% of the EB volume, and is present in both 200 and 500 µm EBs. The mesocyst is 40 to 150 µm in diameter and is surrounded by tightly packed cells with a polarized appearance due to concentrated ZO1 towards the core. The microcysts have a core of 10 to 40 µm that is typically surrounded with loosely packed cells. Measurements are in longer diameter when cysts are oval. Confocal immunocytology on 40 LTA-PDMS 500 µm EBs revealed that 28% have an empty macrocyst (Figure 4E), whereas most common (72%, Figure 4F) is a large macrocyst containing meso- and micro- cysts. These results identify new structures not previously
characterized in EBs. Together with our previous analysis here, our findings indicate that the internal core structure of 500 µm EBs exhibits a multicystic complexity that exceeds the monocystic structure of 200 µm EBs and occurs independent of formation method.
Figure 4. Templated 500 µm EBs reveal three distinct cyst types in two architectural arrangements. Cysts in LTA-PDMS EBs formed from single cell suspensions. (A) Left two panels are 200 µm and 500 µm LTA-PDMS grids loaded with single cell suspensions of H9 hESCs at day 0. Scale bar is 1000 µm. Right panels are 200 µm EBs in wells at day 4 (top) and 500 µm EBs in wells at day 12 (bottom). Scale bars are 200 µm. (B) Pluripotent 200 µm EB (Merged: Hoechst, blue; ZO1, green) generated from single cells, showing large internal cyst structure (arrow) Scale bar is 200 µm (N=200). (C) Single cell-generated 500 µm EB revealing cyst structures (yellow boxes) similar to 500 µm EBs formed using mechanical dissection intermediates (glutaraldehyde fixation). (D) Schematic depicting the polarized meso- and non-polarized micro-cyst structures observed in 500 µm EBs. (E) Confocal section of an empty macrocyst in a single cell-generated 500 µm LTA-PDMS EB (day 4 STEMdiff endoderm differentiation; methanol fixation). Boxed areas are typical internal meso- and micro- cyst structures: microcysts (box 1, top two images); and mesocysts (box 1, bottom image and box 2). (Merged: Hoechst, blue; ZO1, green). Scale bars: 200 µm in the full EB image, 100 µm in the boxed 1 and 2 images. (F) Confocal section showing a macrocyst with internal meso- and micro- cysts in a 500 µm LTA-PDMS EB generated from single cells (day 4 STEMdiff endoderm differentiation; methanol fixation). Shown boxed are: a microcyst (box 1, top); and mesocysts (box 1, middle and bottom, and box 2). (Merged: Hoechst, blue; ZO1, green). Scale bars: 200 µm in the full EB image, 100 µm in box 2 and finally 50 µm in Box 1.
Discussion

Embryoid bodies are important intermediates in stem cell differentiation and expected to be particularly effective for \textit{in vitro} 3D organoid development. EB formation methods vary, and whether optimal parameters in terms of size and morphological features exist remains unclear. Here we evaluated multiple EB parameters under a single uniform platform allowing comprehensive comparative analysis of EB formation from three different stem cell sources, including time tracked formation, morphological features, and efficacy of custom sized EBs in early adherent differentiation by two protocols. Our study revealed new structural features of EBs that are multicystic complexity of the core, impacted by size that affects uniformity of early differentiating cultures. These findings emphasize the importance of controlling EB size and uniformity. The resiliency of the cyst structures to form regardless of method of EB formation suggests that inherent unknown factors influence self-assembly of the EB architecture.

We compared hESC EB formation from single cells or mechanically dissected 2D clusters or 3D early aggregates that are three commonly used protocols. We observe that 200 µm EBs are uniform and monocystic by multiple methods of formation. In contrast, 500 µm EBs exhibit internal multicystic complexity that is comprised of macro-, meso- and micro- cysts. Upon tri-lineage differentiation analysis, 500 µm EBs retain complex mixed 3D/2D morphological character of the colony preventing early formation of a uniform adherent monolayer, versus 200 µm EBs that readily transition from 3D to 2D monolayer in a consistent manner. Our findings indicate that use of 200 µm EBs offers uniformity in differentiation for high throughput studies and consistent results. Size, rather than method of EB formation appears to be most critical for influencing the internal architecture of EBs.
How the cystic structure of EBs forms is unknown. Whether the microcysts we identified act as precursors to the more tightly packed mesocysts and whether these contribute to the macrocyst requires further investigation. Such an underlying organizational timeframe could reflect components of early developmental cues present during embryonic gastrulation or be a distinct \textit{in vitro} behavior. It has been reported for mouse embryoid bodies that the presence of an internal cyst in EBs increases differentiation efficiency, while conversely, EBs that do not generate such a cavity express reduced differentiation ability \cite{48,51,53}. Parallels between cystic EBs can also be drawn with early \textit{in vivo} embryonic stages, where the embryo must undergo a cavitation step in order to further develop. Continued analysis of EB stem cell biology combined with informative bioinformatics and utilizing advanced materials for tissue bioengineering are needed. Understanding how EB size influences its internal structure has immediate application to improve differentiation protocols where EB intermediates contribute to tissue and organoid bioengineering.
Chapter Conclusions

3D culture in general, and embryoid bodies (EBs) in particular, remain an important intermediate to bioengineering of stem cells for research discovery and biomedical applications. Custom microarray templating of EBs aided by engineered devices such as the here developed LTA-PDMS platform offer a powerful means to carefully evaluate critical parameters of EB formation, 3D multicellular structure and use of EBs in multi-lineage differentiation protocols and organoid formation. There is a complex structure in 500 µm EBs that is absent smaller 200 µm EBs, and which results in mixed 2D/3D islands during differentiation. However there is still evidence in the literature which may favor use of larger EBs for development of tubular-based structures such as the heart. Using the LTA-PDMS grids, we were able to identify new cell-cell organization features that appear as cysts in human EBs, with a surrounding rosette formation, whose developmental importance is yet unknown. The cysts were observed independent of EB formation method by comparative analysis of three methods. EB size and core complexity therefore are critical parameters in generation of consistent results for stem cell research that include optimization of differentiation protocols, tissue bioengineering and development and disease studies.
Chapter References:


Chapter 3:
Derivation of Ethnically Diverse Human Induced Pluripotent Stem Cell Lines

State of the Field

The promise of stem cells to recapitulate aspects of normal or impaired development, mature cellular function and disease states provides hope for new insights applicable to regenerative medicine, cell and tissue therapies, and drug studies in human healthcare. Stem cell research is already providing important information on monogenetic and metabolic diseases by utilizing patient iPSCs for disease mechanism studies surrounding the affected cell types, their interactions within the diseased tissues, and responses to target drugs. Expanding the availability of iPSC lines that reflect age, gender, or ethnic group is expected to further refine our understanding of phenotypically different responses to disease onset or drug treatments versus healthy controls. Thus, derivation and characterization of new high quality ethnically diverse-induced pluripotent stem cells (ED-iPSC) lines of low passage number and confirmed pluripotency is critical for the advancement of healthcare and biomedical research towards next generation patient-driven regenerative medicine. In order to achieve this, derived ED-iPSC lines need to undergo extensive core quality control, assuring that any newly derived lines are free of mycoplasma, endotoxins, bacteria, yeast, mold and viruses and analyzed for normal karyotype in addition to comprehensive pluripotency evaluation.
The ED-iPSC lines need to also be maintained in vitro as xeno-free cultures, if possible to avoid possible cross-contamination with animal membrane proteins. This can be done either on human foreskin fibroblasts (HFF) or on feeder-free extracellular xeno-free matrices. Due to the inherent variability and low efficacy of current reprogramming strategies to generate patient-specific iPSC lines, pluripotency analysis of any such line need to be done jointly by teratoma formation in vivo, as well as evaluation of embryoid body (EB) formation and subsequent initial tri-lineage commitment to early differentiation in vitro. Such xeno-free ED-iPSC lines will be an important resource for stem cell research that evaluates ethnic origin contributions for biomedical discovery towards a multitude of clinical applications, such as next generation drug discovery and personalized regenerative medicine.

**Hypothesis**

The human genome with all its ethnic variations is known to contribute to medical treatments and therefore likely to underline differences in human development, aging, disease, repair, which makes it an exciting area of research and clinical study. Unfortunately, development of ethnically diverse stem cell lines, which will be are critical to start piecing together these variations, is limited and has not kept pace with other advances in stem cell research. Thus, we predict that availability of well-characterized ethnically diverse stem cell lines xenofree ED-iPSC lines will represent a valuable resource for future research, novel drug discovery and clinical investigations in regenerative medicine.
Strategy

Here we will derive xeno-free ethnically diverse-human induced pluripotent stem cell (ED-iPSC) lines from fibroblasts obtained from individuals of African American, Hispanic-Latino, Asian, and Caucasian self-reported ethnic origin and will characterize the lines under a previously developed and validated LTA-PDMS uniform platform for comparative analysis. Derived ED-iPSC lines will be kept at low passage number and evaluated in vivo by teratoma formation and in vitro by high throughput microarray analysis of EB formation and early differentiation for tri-lineage commitment to endoderm, ectoderm and mesoderm. The validated lines will then be made available to researchers as a high quality stem cell resource that fills an important niche is the current biomedical and healthcare research landscape.
Acknowledgements and Contributions:

I performed experiments and analysis with microarrays, EB formation, xenofree maintenance on feeder free, and multi-lineage commitment assays and compiled figures, assisted by Zach Olmsted. Jose Cibelli and Janet Paluh designed and supervised experiments, analysis and manuscript and figure preparations. Eun-Ah Chang performed experiments and analysis for derivation, pluripotency, karyotype and teratomas. Steven Suhr and Jiesi Luo performed semi- and q-RT-PCR. Copyright permission for the purposes of this dissertation can be found at: http://www.nature.com/srep/journal-policies/editorial-policies

Parts of this work has been published in:

Abstract

The human genome with all its ethnic variations contributes to differences in human development, aging, disease, repair, and response to medical treatments and is an exciting area of research and clinical study. The availability of well-characterized ethnically diverse stem cell lines is limited and has not kept pace with other advances in stem cell research. Here we derived xenofree ethnically diverse-human induced pluripotent stem cell (ED-iPSC) lines from fibroblasts obtained from individuals of African American, Hispanic-Latino, Asian, and Caucasian ethnic origin and have characterized the lines under a uniform platform for comparative analysis. Derived ED-iPSC lines are low passage number and evaluated in vivo by teratoma formation and in vitro by high throughput microarray analysis of EB formation and early differentiation for tri-lineage commitment to endoderm, ectoderm and mesoderm. These new xenofree ED-iPSC lines represent a well-characterized valuable resource for future research and drug discovery or clinical investigations.
Introduction

The promise of stem cells to recapitulate aspects of normal or impaired development, mature cellular function and disease states provides hope for new insights applicable to regenerative medicine, cell and tissue therapies, and drug studies in human health care. Stem cell research is already providing important information on monogenetic and metabolic diseases by utilizing patient iPSCs for disease mechanism studies surrounding the affected cell types (1-3). Analysis of early or late onset monogenetic diseases includes neurodegeneration, such as the role of SMN1 in early motor neuron death in a spinal muscular atrophy patient (4) or late onset familial Parkinson’s disease that is exacerbated by mutation of LRRK2 in dopaminergic neurons of disease patients (5,6). Disorder of carbohydrate metabolism in type I diabetes with insufficient production of insulin hormone may also be approachable through stem cell therapies (7). By expanding the availability of iPSC lines that reflect age, gender or ethnic group further refinement in understanding phenotypically different responses to disease onset or drug treatments versus healthy controls is expected.

Here we present details of the derivation and characterization of new high quality ethnically diverse-induced pluripotent stem cells (ED-iPSC) lines of low passage number and of African American, Hispanic-Latino, Caucasian, and Asian ethnicity. Derived ED-iPSC lines have undergone core quality control to be free of mycoplasma, endotoxins, bacteria, yeast, mold and viruses and analyzed for normal karyotype. The ED-iPSC lines were maintained in vitro on either human foreskin fibroblasts (HFF) or on feeder-free extracellular xenofree matrix. Pluripotency analysis was done by teratoma formation in vivo as well as evaluation of embryoid body (EB) formation and subsequent tri-lineage
commitment to early differentiation in vitro. Comparative in vitro analysis employed high throughput custom lithography template microarrays for uniform EB formation (8). These xenofree ED-iPSC lines are an important new resource for stem cell research that evaluates ethnic origin contributions for biomedical discovery towards clinical applications.

**Materials and Methods**

**In vitro culture of primary human fibroblasts and lentivirus reprogramming**

Human fibroblasts for iPSCs derivation were obtained from Coriell Institute (Camden, New Jersey; Table 1) and reprogrammed using a single polycistronic vector using four-factor 2A (4F2A) doxycycline (DOX)-inducible lentivirus encoding mouse cDNAs for Oct4, Sox2, Klf4, and c-Myc separated by three different 2A peptides (P2A, T2A, and E2A, respectively). The lentiviral plasmids are p20321 (TetO-FUW-OSKM) and p20342 (FUW-M2rtTA) (Addgene, Cambridge, MA) originally developed by Carey et al (9). Lentiviral particles (4F2A and M2rtTA) were packaged in HEK 293T cells. The primary fibroblast cells were co-transfected using the lentivirus construct, psPAX and pCMV-VSVG vectors by calcium phosphate co-precipitation. Viral supernatants from cultures packaging each of the two viruses were pooled, filtered through a 0.45 μm filter and concentrated by ultracentrifugation and stored at −80°C.

The 5 human fibroblast lines were transduced by viral particles in xenofree human fibroblast culture medium (10) in the presence of polybrene (8 μg/mL). Forty-eight hours after infection, less than 15% of fibroblasts tested immunopositive for viral-derived OCT4. The procedure was carried out in 1 well of a 6-well plate with cells at 70% confluence to allow for cell growth after viral infection an appearance of stem cell colonies. The medium
was replaced two days after infection, and then daily, with xenofree hES medium plus doxycycline (1μg/ml) formulated to maintain stem cell pluripotency. After 35 days of culture, small cell clumps distinguishable from the fibroblast morphology appeared. Those that formed cell colonies with hESC-like morphology were mechanically isolated and passed on to mitotically inactivated xenofree human foreskin feeder cells (ATCC PCS-201-010). Overall reprogramming efficiency by this method was calculated to be 0.002 ~ 0.004 %. The iPSC colonies were expanded for several passages under xenofree conditions without doxycycline and evaluated for expression of markers of pluripotency by quantitative RT-PCR (qRT-PCR) and immunocytotherapy.

**Real-time PCR genomic analysis**

Quantitative PCR analysis was done by isolation of total RNA from the hESC or iPSC lines and parental fibroblast lines and purification using the NucleoSpin RNA XS Total RNA isolation kit (Clontech). Reverse transcription (RT) was performed in a 20ul reaction volume using Superscript II (Invitrogen) and the cDNA reaction was diluted to a 300ul working stock volume. Primers for use in qPCR were first validated by maximally amplifying cDNA from a range of samples to confirm that a single PCR reaction product was produced and that the amplicon was of the predicted length. For validation, 10ul of cDNA from H9 hESCs (WA09, Wicell, Madison, WI), control fibroblasts (line A-2), and two of iPSC lines (A-2.2.1 & A-2.2.2) for each primer set was amplified for 36 cycles (95°C 30s, 55°C 30s, 72°C 30s). For endogenous and transgene expression, 5ul of cDNA from each iPSC lines for each primer set was amplified for 32 cycles and resolved on a 3% nusieve agarose gel and visualized by ethidium bromide staining. Quantitative PCRs contained 10 ng of cDNA, 400 nM of each primer, and SYBR Green PCR Master Mix.
Each sample was analyzed by triplicate by an ABI PRISM 7000 sequence detection system. Data was analyzed using the system’s software. The expression of gene of interest was normalized to GAPDH in all cases and compared with hESCs.

**Mycoplasma analysis of ED-iPSC lines**

We used the MycoAlert™ PLUS Assay (Lonza, USA) mycoplasma detection kit essentially as manufacturer's instructions. Briefly, after centrifugation (1500 rpm, 5min) of cell supernatant during passage of suspension iPSC cultures, the supernatants were transferred into luminescence compatible tubes (Corning). The viable mycoplasma was lysed to allow enzymes to react with MycoAlert™ PLUS substrate, catalyzing the conversion of ADP to ATP. The level of ATP in the sample both before (reading A; ATP background) and after (reading B) the addition of MycoAlert™ PLUS substrate was assessed using a luminometer (Victor³, Perkin-Elmer, Waltham, Massachusetts, USA), so that a ratio B/A was obtained. Reading B assesses the conversion of ADP to ATP and is a monitor of contaminated samples. If the ratio of B/A is greater than 1 the cell culture was considered to be contaminated by mycoplasma. For control samples, the MycoAlert TM assay positive and negative control set was used.

**Xenofree stem cell maintenance, passaging in 2D, in 3D as uniform embryoid bodies and directed early differentiation of ED-iPSC and hESC lines.**

Ethnically diverse-induced pluripotent stem cell (ED-iPSC) lines maintained on human foreskin fibroblast feeders were transferred to feeder-free conditions in non-tissue culture treated dishes coated with xenofree vitronectin or 1:100 Matrigel (10 mg/ml; BD
Biosciences, San Jose, CA) diluted into Hank's Buffered Saline Solution (Gibco HBSS; Life Technologies, Grand Island, NY). Cells were maintained in mTeSR2 complete media (StemCell Technologies, Vancouver, Canada) and mechanically passaged between days 5 and 7. Media was replaced on day 1 after the first passage of the series and cells grown overnight. On day 2, slow release bFGF2 beads (20 microliters of PLGA beads loaded with bFGF2; StemBeads; Stem Culture Incorporated, Rensselaer, NY) were added with fresh mTeSR2 media. Media changes were done every 3 days with StemBeads FGF2 and mTeSR2. Preparation of uniform sized EBs from iPSCs colonies was done in custom lithography template microarrays (LTA) generated in-house. Chemical dissociation of the stem cell colonies into single cell suspension was done before and loading of the cells into LTA- polydimethylsiloxane (PDMS) grids in mTeSR2 media in the presence of 10 μM Rock inhibitor (Sigma-Aldrich, St. Louis, MO) at day 0. Stem cells were maintained in grids for five days with media changes every two days. For directed multi-lineage early differentiation we used the Human Pluripotent Stem Cell Functional Identification Kit (RnD Systems, Minneapolis, MN).

**Immunocytoology**

For immunocytology of biomarkers in iPSC colonies, cells were prepared by two methods. Cells were fixed using 4% paraformaldehyde in PBS for 15 min at room temperature and blocked by incubating cells for 90 min in a solution containing 3% normal donkey serum and permeabilized by 0.1% Triton-X 100 for 10min before antibody addition. Incubations with the primary antibodies of anti-Nanog (Santa Cruz) and anti-SSEA4 (Santa Cruz) were done at 4°C overnight, followed by incubation with a secondary antibody conjugated with Alexa 647 or Alexa 488 (Abcam, Cambridge, MA). After rinsing
with phosphate buffered saline (PBS), the DNA was stained with bisbenzimide and cells imaged using a digital camera connected to a Nikon TE-2000 inverted microscope.

Phase imaging for in vitro differentiated samples was done on a Nikon 80i epifluorescence microscope using a PLAN 10X 0.30 NA DL objective and images captured with a cooled QICam CCD camera. Fluorescent images were obtained on a Leica SP5 Laser Scanning Confocal Microscope using either HC PL FLUOTAR 10X 0.30 NA or HCX PL APO CS 20X .70 NA objectives and also on a Zeiss AxioObserver Z1 Inverted Microscope with Colibri LED illumination, using a 100X oil 1.45 NA PlanFLUAR or 63X Plan-Apochromat 1.4 NA oil DIC objectives. Images were captured with a Hamamatsu ORCA ER CCD camera and Zeiss Axiovision Rel 4.8 acquisition software. Figures were compiled using Adobe Photoshop (Adobe Systems Inc., San Jose, CA) and Microsoft PowerPoint (Microsoft Corp., Redmond, WA) software. The immunocytology of 2D cell cultures or three dimensional EBs was done by first fixing cells for 10 minutes at room temperature in 4% paraformaldehyde and stored overnight in PBS + 0.1% Tween20 at 4°C. Immediately before incubation with antibodies, the cells were permeabilized with PBS + 0.5% Triton X-100 for 1 hour at 4°C. Nonspecific binding was blocked by a 20 minute incubation in 1% BSA in HBSS and followed by a single HBSS wash. Antibodies used for gauging pluripotency recognized Oct4A C-10 (Santa Cruz Biotechnology, Dallas, TX) and anti-SSEA4 (Millipore, MA) (1:1000 each). Analysis of lineage commitment to differentiation was done using antibodies to OTX2 (ectoderm), SOX17 (endoderm), and Brachyury (mesoderm; 1:100 each) provided in the Human Pluripotent Stem Cell Functional Identification Kit (RnD Systems, Minneapolis, MN). Secondary antibodies were either AlexaFluor 488 or AlexaFluor 594 (A-11001, A-11037, Invitrogen, Carlsbad, CA). Nuclei were stained with Hoechst (1:1000 in HBSS) at 4°C overnight and followed by
washing one hour in HBSS at 4°C. Samples were mounted in ProLong Gold antifade reagent (Life Technologies, Grand Island, NY) at 20°C overnight in the dark before imaging immediately or storing at 4°C.

**In vivo teratoma formation assay**

Approximately 2 million ED-iPSCs were injected subcutaneously in the flank region of NOD scid gamma (NSG) mice (The Jackson Lab, Bar harbor, ME). After 12-24 weeks, teratomas were formed from 10 iPSC lines (Figure 3), and tumors were excised & fixed in 10% normal buffered formalin (NBF) overnight. The samples were processed for histology by the Division of Human Pathology at MSU. Hematoxylin- and eosin (H&E)-stained sections were examined under a microscope.

**Results**

**Derivation of human ED-iPSC lines from African American, Hispanic-Latino, Caucasian, and Asian ethnicities**

In evaluating genetic contributions to human disease, the generation of induced pluripotent stem cells (iPSCs) constitutes one of the most exciting scientific breakthroughs in the last 20 years. Although multiple sources of pluripotent Caucasian lines exist, there is limited availability of ethnically diverse (ED) iPSC lines. To provide a new high quality resource of ED-iPSCs for biomedical research, we obtained primary human fibroblasts from African American (AA), Hispanic-Latino (HL) and Asian (AS) origins as well as Caucasian parental lines (Coriell Institute, Camden, NJ) and used these to derive ED-iPSC lines (Table 1). To reprogram fibroblasts, polycistronic lentiviral plasmid vectors with
tet-inducible expression TetO-FUW-OSKM and FUW-M2rtTA (9; plasmids 20321 and 20342, Addgene) were used and fibroblasts positive for viral derived Oct4 identified by immunocytochemistry (Figure 1a-c). The polycistrionic cassette encodes four factors that are Oct4, Sox2, Klf4, and c-Myc mouse genes whose expression was tracked by semi-quantitative RT-PCR. Over a time course of day 35 to 75, cells grown in 2D culture under xenofree conditions on human foreskin fibroblasts (HFF) were monitored for the appearance of stem cell like clustered colonies (Figure 1d-f). Derived stem cell lines were maintained on HFF and standard quality control tests done to ensure safety of the iPSCs in regard to absence of contaminating mycoplasma or other biological agents. Karyotyping on G-banded metaphases of the iPSCs was outsourced (Table 2, n=20 cells per iPSC line; Cell Line Genetics; Madison, WI). Fourteen of the 17 ED-iPSC lines displayed normal karyotyping by this method and were further analyzed. The remaining lines were stored and may be useful for analysis of phenotypes associated with specific chromosomal regions.

Analysis of ED-iPSC lines for pluripotency by genetic and cell biological biomarker profiles and teratoma analysis.

The ED-iPSC lines we generated from reprogramming of fibroblasts of African American, Hispanic-Latino, Caucasian and Asian ethnicities were analyzed by a platform of tests that included immunocytochemistry for pluripotency markers Nanog and SSEA-4 (Figure 2a) and semi-quantitative RT-PCR analysis of gene expression during reprogramming of fibroblasts to pluripotency (Figure 2b and 2c).
Figure 1. Derivation of ethnically diverse (ED)-iPSC lines from primary human fibroblasts. (a) Addgene plasmid 20321 (TetO-FUW-OSKM and 20342; FUW-M2rtTA) is a single polycistronic vector using four-factor 2A (4F2A) doxycycline (DOX)-inducible lentivirus encoding mouse cDNAs for Oct4, Sox2, Klf4 and cMyc separated by three different 2A peptides (P2A, T2A, and E2A, respectively; Carey et al., 2009). (b) Schematic of the process of reprogramming and culture conditions. (c) 48-hours after infection, less than 15% of fibroblasts were immunopositive for viral-derived Oct4 by immunocytochemistry (scale bar is 50 μm). (d, f) After 35 days, infected ED fibroblasts generated stem cell-like colonies that were isolated for derivation of ED-iPSC lines. (e) Original cell morphology of parental ED-fibroblasts. Images in (d-f) are phase contrast (scale bars are 200 μm).
Figure 2. Characterization of gene expression in reprogrammed ethnically diverse (ED)-iPSC lines. (a) Representative immunocytology biomarker analysis of pluripotency in 9 ED-iPSC lines. DNA is in blue by Hoechst, Nanog is in red and SSEA-4 is in green (scale bar is 50 μm). (b) Pluripotency-related genes are expressed from the endogenous loci in 26 ED-iPSC lines, while the virally delivered transgene is predominantly silenced revealed by semi-quantitative RT-PCR. Primers were validated in H9 ESCs, parental fibroblasts (A-2), and two of iPSC lines (A-2.2.1 & A-2.2.2); collagen (COL), fibrinonectin (FNC), descortin (DEC), transgene; 2A-S(Sox2), 2A-O(Oct4), and 2A-M(C-myc) and endogeneous gene; hOCT (Oct4), hSOX (Sox2), hNAG (Nanog) and hREX (human REX). (c) Semi-quantitative expression of the silenced three transgenes and expressed two endogenous pluripotency genes shown in 26 iPSC lines. GAP (GAPDH) is used as a loading control for each lane. (d) Expression of fibronectin (FNC) and collagen type 1α2 (COL) in individual parental fibroblasts (A-2, A-3, C-1, F-3, and H-3) by quantitative RT-PCR. No FNC or COL expression is observed in the ED-iPSCs or H9 hESCs. Expression for OCT4, SOX2, NANOG, and REX1 levels is increased in ED-iPSCs relative to H9 hESCs. PCR reactions were normalized against internal controls (GAPDH) and plotted relative to expression levels in H9 hESCs.
Primers sets for RT-PCR were designed to recognize several classifications of target. These are, Set 1: an internal control housekeeping gene GAPDH (GAP) expressed at similar levels in most cell types and allows normalization of samples. Set 2: genes associated with fibroblast phenotype. Targets included collagen type 1α2 (COL), fibronectin (FNC), and decorin (DEC). Set 3: the lentiviral transgenes that are expressed to induce pluripotency but shut down when iPSCs have fully reprogrammed to pluripotency. To ensure that only mRNAs expressed from the transgenes -- and not endogenous genes -- were recognized, the three primer sets were designed with one primer recognizing the 2A elements unique to the viral vector and a second primer targeting the mouse Sox2 (2A-S), the mouse Oct4 (2A-O), or the mouse cMyc (2A-M) transgenes. Set 4: Endogenous human genes that should reactivate in iPSCs. The primers selected included human Oct 4 (hOCT), SOX2 (hSOX), nanog (hNAG) and REX1 (hREX). The evaluation of primer sets is shown in Figure 2b, agarose gels of PCR products in Figure 2c, and quantitative RT-PCR data shown by histogram in Figure 2d. The gold standard test of pluripotency is generation of teratomas in vivo, as shown in Figure 3. Our analysis is consistent with successful reprogramming to generate human iPSCs from ethnically diverse origins.
Figure 3. Teratoma formation by ED-iPSCs in Nod scid gamma (NSG) mice. (a) Representative image series of hematoxylin-eosin (H-E) stained sections from a formalin-fixed teratomas produced from 9 ED-iPSC lines. Each line formed mature, cystic teratomas with tissues representing the three embryonic germ layers including squamous epithelium, neuroectodermal tissues and melatonin formation (ectoderm), respiratory epithelium and intestinal glandular epithelium (endoderm) and cartilage, bone and connective tissues (mesoderm). Scale bars are 0.2 cm.
Microarray Embryoid Body formation and tri-lineage commitment of ED-iPSCs in vitro.

The ability of the iPSCs to form embryoid bodies (EBs) and commit to tri-lineage differentiation was compared in vitro. For uniform evaluation we generated custom 200 µm microwells by photolithography and coating with poly-dimethylsiloxane (PDMS) (Figure 4a). The formation of ED-iPSC EBs in vitro was done by first dissociating 2D colonies to single cells or by mechanical passaging to generate 2D clusters then loading wells as previously described (8). EB formation including a characteristic internal cyst occurred with similar efficiency by all lines (Figure 4b). To evaluate ability of iPSCs to commit to multi-lineage early differentiation, we used a commercially available tri-lineage germ layer commitment kit (RnD Systems, Minneaopolis, MN) with EBs over a five day period then stained cells for biomarkers corresponding to each germ layer that included, Otx2 for ectoderm, Brachyury for mesoderm, and Sox17 for endoderm (Figure 4c). We observed that over 94% of the cells stain positively for the corresponding germ layer marker. Minimal cell death was observed during the differentiation, consistent with high quality pluripotent stem cell lines.
Figure 4. In vitro tri-lineage commitment to differentiation by ED-iPSCs. (a) Schematic of the process flow to generate LTA-PDMS grids used for high throughput embryoid bodies (EBs) templating (top image), a PDMS mold of arrays prepared on a 4 inch wafer (middle), and 5 day time course of EBs formation in 200 μm LTA-PDMS grid wells (bottom image; (N>400; scale bar is 200 μm). (b) EBs characterization by Hoechst, immunocytology for pluripotency markers Sox2 or SSEA-4, and Actin (phalloidin) staining. The F3.5.2 ED-iPSCs EBs shown was templated in a 200 μm well to day 5. Scale bar is 200 μm. (c) Representative images of differentiated 200 μm ED-iPSCs EBs using a Human pluripotent stem cell identification kit. Left column image shows bright field of pluripotent ED-iPSCs 2D colonies. Differentiated ED-iPSCs were fixed for ICC analysis at day 4 for the ectoderm and endoderm lineages and at day 3 for the mesoderm lineage. Antibodies used for immunocytology were indicated at the top of each column. Scale bars are 200 μm. (d) Colocalization of Hoescht nuclear stain and germ layer specific markers, listed in each image, to the nucleii of differentiated iPSC cells. Negative control is secondary antibody only. Representative images for germ layer commitment are shown for the iPSC line A2.2.2 at day 4 for ectoderm and endoderm and at day 3 for mesoderm. Scale bars are 200 μm.
Discussion

Here we generated high quality xenofree ethnically diverse-human induced pluripotent stem cell (ED-iPSC) lines from fibroblasts obtained from individuals of African American, Hispanic-Latino, Caucasian and Asian origin to provide an important resource to benefit research and clinical discovery. The significance of our study is that ED-iPSC lines were derived by same viral vector, single method, and same culture conditions throughout the same place, instead of quite diverse iPSCs derivation methods and different individual laboratories. When we compared our data with previously characterized H9 hESCs, the growth pattern and characterization of ED-iPSCs are similar to hESCs. The availability of these cell lines will be useful to evaluate the contributions of ethnic backgrounds to cell or drug based therapies that have a genetic component to disease onset or resistance, an aspect that is expected to be important as personalized medical procedures advance.

Chapter Conclusions

Our comprehensive in-depth characterization of these ethnically diverse-iPSC (ED-iPSC) lines derived and characterized under a uniform platform includes karyotype, teratoma analysis, gene expression analysis by qRT-PCR, and in vitro analysis under xeno-free conditions and using high throughput microarrays to uniformly generate embryoid bodies for optimal comparative evaluation of cell line potential for tri-lineage commitment to differentiation.
Chapter References:


### Tables

**Table 1. Ethnically diverse parental fibroblasts and induced Pluripotent Stem Cell Lines.**

<table>
<thead>
<tr>
<th>Sources</th>
<th>Catalogue no</th>
<th>Nature</th>
<th>Sex</th>
<th>Age</th>
<th>Ethnicity</th>
<th>Symbols</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coriell</td>
<td>AG08498</td>
<td>Skin (foreskin)</td>
<td>male</td>
<td>1 year</td>
<td>Asian</td>
<td>A-2</td>
</tr>
<tr>
<td>Coriell</td>
<td>AG22284</td>
<td>Skin (foreskin)</td>
<td>male</td>
<td>1 day</td>
<td>Asian</td>
<td>A-3</td>
</tr>
<tr>
<td>Coriell</td>
<td>GM00970</td>
<td>Skin (foreskin)</td>
<td>male</td>
<td>26 years</td>
<td>Caucasian</td>
<td>C-1</td>
</tr>
<tr>
<td>Coriell</td>
<td>GM22268</td>
<td>Skin (foreskin)</td>
<td>male</td>
<td>1 day</td>
<td>African American</td>
<td>F-3</td>
</tr>
<tr>
<td>Coriell</td>
<td>GM22186</td>
<td>Skin (foreskin)</td>
<td>male</td>
<td>1 day</td>
<td>Hispanic Latino</td>
<td>H-3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Asian</th>
<th>Asian</th>
<th>Caucasian</th>
<th>African American</th>
<th>Hispanic Latino</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-2.1.1</td>
<td>A-3.1.1</td>
<td>C-1.1.1</td>
<td>F-3.2.2</td>
<td>H-3.1.1</td>
</tr>
<tr>
<td>A-2.2.1</td>
<td>A-3.2.1</td>
<td>C-1.2.1</td>
<td>F-3.3.1</td>
<td>H-3.2.1</td>
</tr>
<tr>
<td>A-2.2.2</td>
<td>A-3.3.1</td>
<td>C-1.3.1</td>
<td>F-3.3.2</td>
<td>H-3.3.1</td>
</tr>
<tr>
<td>A-2.2.3</td>
<td></td>
<td>C-1.5.1</td>
<td>F-3.4.1</td>
<td>H-3.4.1</td>
</tr>
<tr>
<td>A-2.3.1</td>
<td></td>
<td></td>
<td>F-3.4.2</td>
<td>H-3.5.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>F-3.5.2</td>
<td>H-3.5.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>F-3.6.1</td>
<td></td>
</tr>
<tr>
<td>iPSC lines</td>
<td>Passage number</td>
<td>Karyotype</td>
<td>iPSC lines</td>
<td>Passage Number</td>
</tr>
<tr>
<td>------------</td>
<td>----------------</td>
<td>--------------------</td>
<td>------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Asian</td>
<td></td>
<td></td>
<td>Asian</td>
<td></td>
</tr>
<tr>
<td>A-2.1.1</td>
<td>p 9</td>
<td>46, XY</td>
<td>F-3.3.2</td>
<td>p 6</td>
</tr>
<tr>
<td>A-2.2.1</td>
<td>p 12</td>
<td>46, XY</td>
<td>F-3.5.2</td>
<td>p 7</td>
</tr>
<tr>
<td>A-2.2.2</td>
<td>p 7</td>
<td>46, XY</td>
<td>F-3.6.1</td>
<td>p 7</td>
</tr>
<tr>
<td>Asian</td>
<td></td>
<td></td>
<td>Hispanic-Latino</td>
<td></td>
</tr>
<tr>
<td>A-3.1.1</td>
<td>p 11</td>
<td>46, XY</td>
<td>H-3.1.1</td>
<td>p 8</td>
</tr>
<tr>
<td>A-3.2.1</td>
<td>p 8</td>
<td>46, XY</td>
<td>H-3.5.1</td>
<td>p 9</td>
</tr>
<tr>
<td>A-3.3.1</td>
<td>p 7</td>
<td>46, XY</td>
<td>H-3.5.2</td>
<td>p 9</td>
</tr>
<tr>
<td>Caucasian</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-1.2.1</td>
<td>p 7</td>
<td>46, XY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-1.5.1</td>
<td>p 7</td>
<td>46, XY</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A-2.3.1</td>
<td>p 7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C-1.1.1</td>
<td>p 9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C-1.3.1</td>
<td>p 7</td>
</tr>
</tbody>
</table>
Chapter 4:
Human Embryoid Body Transcriptomes Reveal Maturation Differences Influenced by Size and Formation in Custom Microarrays

State of the Field

Pluripotent stem cells offer unique and unparalleled opportunities in human tissue engineering through the ability to direct differentiation along desired lineages to generate specific cells, tissues or multicellular organoids. In addition to cell replacement therapies, biomedical advances will benefit from stem cell studies that provide a new understanding of normal processes of development and repair or of malfunction of human tissues or organs as a consequence of perturbation by aging, injury or disease. The ability to manipulate human stem cells towards biomedical applications in vivo, or for on-chip therapeutic devices in vitro is challenging. It requires expanding our knowledge of environmental, chemical, and physical influences on stem cells in a niche as well as during in vitro manipulations, and including studies with an increasing diversity in stem cell lines in regard to age, sex and ethnicity.

Transcriptome data along with bioinformatics analysis provides a comprehensive means to cross-compare pluripotent cell types such as hESC, hiPSC and human embryonic carcinoma cells (hECCs) as well as multiple independently generated pluripotent lines. Such analysis compliments and extends in vitro multi-lineage differentiation studies along with teratoma or more direct in vivo functional studies. Open
source comparative transcriptome platform resources include Pluritest and Cellnet. Differences in transcriptome profiles of pluripotent stem cells exist and are expected in part to reflect inherent variation in lines generated by different labs, applying different reagents and protocols across the research community. Continued comparative analysis will allow variances and similarities in gene expression to be defined along with their relevance to specific biomedical applications.

Frequently used in pluripotent stem cell differentiation strategies are embryoid body (EB) intermediates that are naturally forming 3D stem cell aggregates that retain pluripotent potential. Pluripotent stem cells grown in two-dimensional culture have the capability to spontaneously form 3D EBs when mechanically passaged and released into non-adherent media conditions, as demonstrated for mouse and human 2D cell clusters. Biomedical scaled up applications with human stem cells are alternatively exploring use of dissociated to single cells followed by seeding them in custom microarray templates to enhance cell interactions and 3D assembly of iPSC or hESC EBs or by applying shear forces in a bioreactor.

Embryonic stem cell EBs share some similarities with 3D embryonic-like structures such as the presence of a single internal cyst-like core structure in naturally forming 200 micron-sized EBs. However not all methods of EB formation generate a less dense core structure, such as in the use of high shear in bioreactors to generate EBs that are then used for differentiation. In addition, increased EB size can result in a multicystic core as revealed by use of custom high throughput photolithography-templated microarrays. Current human EB transcriptome data have shown gene activation involved with early differentiation and mRNA levels in their analysis of naturally forming EB intermediates
however there is no data on the effect of EB size, and method of EB formation, on differentiation.

**Hypothesis**

Since stem cell differentiation strategies and optimization for generating lineage-specific cells and tissues most frequently rely on a three-dimensional embryoid body (EB) intermediate, we propose to evaluate the transcriptome differences in individual uniformly generated EBs. Previous work from our group has shown that formation of EBs of 200 or 500 micron size reveal distinct morphological differences that are single or multicystic cores, respectively, independent of method of formation from single cells or two-dimensional (2D) clusters. Our LTA-PDMS microarray platform allow for high throughput uniform formation of EBs of custom diameters, which is ideal to generate the EBs for the proposed precise downstream analysis.

**Strategy**

We will use our LTA-PDMS array generated EBs to obtain 3D cultures under a standardized platform for transcriptome analysis, which will compare EB size and the method of EB formation from single cells or mechanically passaged 2D clusters. This analysis will evaluate RNA expression and transcriptome in EBs formed from the human embryonic stem cell (hESC) line WA09 and from ethnically diverse human induced pluripotent stem cell lines (ED-iPSC) of African American and Hispanic Latino ethnicity recently derived in our laboratory. This will be the first comprehensive study on EB transcriptomes including multiple size parameters, EB formation methodologies, and
ethnicities. Our detailed bioinformatics analysis will include comparative expression of pluripotency genes as well as up-regulated gene ontology (GO) pathways. We anticipate that this information will be broadly useful to the stem cell and bioengineering communities in optimization of tissue engineering with pluripotent stem cells and understanding sources of variation.

**Acknowledgements and Contributions:**

I performed experiments for lithography, stem cell growth, maintenance, EB generation, microscopy and RNA harvesting, and assisted manuscript writing and figures. Janet Paluh designed and supervised experiments, analysis, and manuscript and figure preparations. Mike Buck performed and supervised RNAseq bioinformatics; Maria Tsompana performed bioinformatics and RNAseq sample processing. Zach Olmsted performed experiments for stem cell growth, maintenance, and assisted manuscript figures.

Copyright permission for the purposes of this dissertation can be found at: [http://www.aspbs.com/jnn/JNN%20Copyright%20Transfer%20Form.pdf](http://www.aspbs.com/jnn/JNN%20Copyright%20Transfer%20Form.pdf)

**Parts of this chapter are published in:**

Abstract

Stem cell differentiation strategies and optimization for generating lineage-specific cells and tissues most frequently rely on a three-dimensional embryoid body (EB) intermediate. We previously applied nanotechnology tools of photolithography to generate custom microarrays that allow high throughput uniform formation of EBs of custom size for precise downstream analysis. Formation of EBs of 200 or 500 micron size revealed distinct morphological differences that are single or multicystic cores, respectively, independent of method of formation from single cells or two-dimensional (2D) clusters. Here we utilize photolithographic array generated EBs to obtain 3D cultures under a standardized platform for transcriptome analysis to compare EB size and the method of EB formation from single cells or mechanically passaged 2D clusters. Our analysis evaluates RNA expression in EBs formed from the human embryonic stem cell (hESC) line WA09 and from ethnically diverse human induced pluripotent stem cell lines (ED-iPSC) of African American and Hispanic Latino ethnicity recently derived in our laboratory. This is the first comprehensive study on EB transcriptomes including multiple size parameters, EB formation methodologies, and ethnicities. Our analysis indicates upregulation of genes involved in wound healing for mechanically passaged cells and of genes for embryonic tube formation in 500 micron multicystic EBs. We propose that EB maturation may be a longer process than previously realized. The type or extent of maturation possible may be influenced by EB size, with larger EBs capable of more extensive remodeling as revealed by multicystic morphology and initiation of early tube formation pathways while retaining pluripotency status. We anticipate that this information will be broadly useful to the stem cell and bioengineering communities in optimization of tissue engineering with pluripotent stem cells and understanding sources of variation.
Introduction

Pluripotent stem cells offer unique and unparalleled opportunities in human tissue engineering through the ability to direct differentiation along desired lineages to generate specific cells, tissues or multicellular organoids. In addition to cell replacement therapies, biomedical advances will benefit from stem cell studies that provide a new understanding of normal processes of development and repair or of malfunction of human tissues or organs as a consequence of perturbation by aging, injury or disease. The ability to manipulate human stem cells towards biomedical applications in vivo\(^1\), or for on-chip therapeutic devices\(^2\) is challenging. It requires expanding our knowledge of environmental, chemical, and physical influences on stem cells in a niche as well as during in vitro manipulations, and including studies with an increasing diversity in stem cell lines in regard to age, sex and ethnicity\(^3\).

Transcriptome data along with bioinformatics analysis provides a comprehensive means to cross-compare pluripotent cell types such as hESC, hiPSC and human embryonic carcinoma cells (hECCs)\(^4,5,6\) as well as multiple independently generated pluripotent lines. Such analysis compliments and extends in vitro multilineage differentiation studies along with teratoma or more direct in vivo functional studies. Open source comparative transcriptome platform resources include Pluritest\(^7\) and Cellnet\(^8\).

Differences in transcriptome profiles of pluripotent stem cells exist and are expected in part to reflect inherent variation in lines generated by different labs, applying different reagents and protocols across the research community. Continued comparative analysis will allow variances and similarities in gene expression to be defined along with their relevance to specific biomedical applications.
Frequently used in pluripotent stem cell differentiation strategies are embryoid body (EB) intermediates that are naturally forming 3D stem cell aggregates that retain pluripotent potential. Pluripotent stem cells grown in two-dimensional culture have the capability to spontaneously form 3D EBs when mechanically passaged and released into non-adherent media conditions, as demonstrated for mouse⁹,⁰,¹¹ and human¹² 2D cell clusters. Biomedical scaled up applications with human stem cells are alternatively exploring use of dissociated to single cells followed by seeding them in custom microarray templates to enhance cell interactions and 3D assembly of iPSC¹³ or hESC¹⁴ EBs or by applying shear forces in a bioreactor¹⁵. ESC-EB shares some similarities with 3D embryonic-like structures¹⁶,¹⁷ such as the presence of a single internal cyst-like core structure in naturally forming 200 micron-sized EBs. However not all methods of EB formation generate a less dense core structure, such as in the use of high shear in bioreactors to generate EBs that are then used for differentiation. In addition, increased EB size can result in a multicystic core¹⁵ as revealed by use of custom high throughput photolithography-templated microarrays. Current human EB transcriptome data¹⁸,¹⁹,²⁰ have shown gene activation involved with early differentiation and mRNA levels in their analysis of naturally forming EB intermediates however there is no data on the effect of EB size, and method of EB formation, on differentiation.

This study applies nanotechnology tools of photolithography to generate uniformly-sized EBs for comparison by transcriptome analysis. This is the first comprehensive whole transcriptome analysis of EBs comparing EB size (200 or 500 microns) as well as EB formation method that is by single cells or 2D cell clusters. Our detailed bioinformatics analysis includes comparative expression of pluripotency genes as well as up-regulated gene ontology (GO) pathways. In addition we compare EB transcriptomes for the hESC
line WA09 and newly derived ethnically diverse induced pluripotent stem cell lines for Hispanic-Latino and African American ethnicities\(^3\). We observe slight but significant elevation of multiple pluripotency genes in 500 micron EBs. Significant differences in GO pathways were also present, reflecting influences of EB formation and EB size and morphology. Our findings provide new insights into EB maturation. They offer a long-awaited explanation as to the unpredictable EB-differentiation behavior observed between protocols and laboratories, and should be taken into account by the stem cell bioengineering community when designing stem cell differentiation strategies.

**Materials and Methods**

**Maintenance of human pluripotent stem cell ESC and ED-iPS lines**

The ethnically diverse (ED)-iPS cells\(^3\) and hESC line WA09 (WISC Bank, WiCell, Madison, Wisconsin) were maintained in mTeSR2 complete media (Stem Cell Technologies, Vancouver, Canada) on StemAdhere (Stem Cell Technologies, Vancouver, Canada) coated non-tissue culture treated dishes. Cells were enzymatically passaged between days 5-7 using the Gentle Cell Dissociation Agent (Stem Cell Technologies, Vancouver, Canada). Media was replaced one day after the first passage and cells were grown overnight with addition of 10µl/mL slow release hFGF2 (StemBeads FGF2; Stem Culture Incorporated, Rensselaer, NY) in fresh mTeSR2 media. Media was then completely changed every two days, with fresh hFGF2 beads added.
Formation of uniformly sized hESC and ED-iPS EBs in custom microarrays

Generation of 200 or 500 micron uniform EBs was done by templating in custom microarrays of polydimethylsiloxane (PDMS) as previously described\textsuperscript{14}. Essentially, to load microarrays from single cells we chemically dissociated 2D stem cell colonies using the Gentle Dissociation Agent (Stem Cell Technologies, Vancouver, Canada). Alternatively we generated small 2D clusters, by mechanical passaging for loading into microarrays. The loaded microarrays were gently washed after loading with 2 mL of fresh HBSS (Hank’s Balanced Salt Solution, Fisher Scientific, Grand Island, NY) to remove excess cells not loaded into the wells. After the wash, loaded microarrays were incubated in mTeSR2 media containing 10 μM Rock inhibitor (Sigma-Aldrich, St. Louis, MO) for two days. The media was changed every two days, until EBs reached the uniform 200 or 500 μm diameter. No ROCK inhibitor was added after the first two days. On average, EB formation from 2D clusters required 3 to 5 days (200 μm), 7 to 10 days (500 μm). EB formation from single cells required 5-7 days (200 μm) or 10 to 14 days (500 μm). Once formed, the template EBs were readily removed by liquid expulsion with a p1000 micropipette. Individual EBs from each group to be analyzed were then collected manually for processing by suction with a p200 micropipette tip under a 10x microscope magnification.

RNA isolation

RNA isolations for bioinformatics analysis of the hESC and ED-iPS cell lines were done using the Ambion PureLink Mini RNA isolation kit (Life Technologies). Eluted total RNA from the stem cell EBs was stored at -80°C until ready to ship for analysis. For RNA-
seq analysis, total RNA was processed for sequencing using the TruSeq RNA Sample Preparation Kit (Illumina). Samples were four-plexed and 50bp single-end sequenced on an Illumina HiSeq2500 producing 30 to 50 million sequence reads to generate bioinformatics data.

**RNA-Seq data processing**

Raw sequencing reads were mapped to the *Homo sapiens* genome (*hg19 build*) using Tophat – version 2.0.7\(^{60}\), count files generated with HTSeq-count\(^{61}\). Significantly enriched genes (FDR <0.001) were then determined using DESeq2\(^{62}\). Enriched genes were then tested for annotation enrichment with DAVID\(^{63}\) and clustered using unsupervised hierarchical clustering with complete linkage and Euclidean distance. We identified the following GO terms to be of interest during our analysis: GO: 0060562~epithelial tube morphogenesis; GO: 0035295~tube development; GO: 0021915~neural tube development; GO: 0001841~neural tube formation; GO: 0001843~neural tube closure; GO: 0060606~tube closure; GO: 0001838~embryonic epithelial tube formation. The generated GEO access number is GSE74792.

**Results**

**Generation and transcriptome profiling of custom sized EBs**

To control EB size and formation we engineered the stem cell microenvironment, by applying nanotechnology tools of photolithography to produce custom templated microarrays. The designed Su-8 stencil was used as a template to make a polydimethylsiloxane (PDMS) microarray platform for generating uniform EBs of 200 to
500 microns from hESC and ED-iPSC lines for analysis (Table 1). Single cells or 2D-clusters of cells were seeded into microarrays as previously described\textsuperscript{14} and allowed to form EBs over multiple days and grow to fill the well dimensions before immunocytological (ICC) and transcriptome analysis. Representative images of microarray formed EBs from the hESC WA09 line as well as two ED-iPSC lines and strategy for generating microarrays are shown in Figure 1. Images reveal the typical single cyst internal structure of 200 µm EBs (Figure 1A) versus the multicyst internal structure of 500 µm EBs (Figure 1B) and include brightfield and staining for OCT 4 and SSEA4 markers of pluripotency and the tight junction protein ZO1. Microarray templating was used to ensure size and uniformity of EBs and compared to spontaneously formed EBs in free-suspension (Figure 1C). The process flow for the lithography generated grids used in uniform EB formation is illustrated in Figure 1D and loading of grids in Figure 1E. Differentiation potential of 200 versus 500 micron EBs is previously described\textsuperscript{14}. 


Figure 1. Three-dimensional formation of uniform stem cell hESC and hiPSC EBs in custom photolithography templated arrays. (A) Representative images of 200 µm EBs generated from the WA09 hESC line as well as from ED-iPSC lines stained for the pluripotency markers Oct4 and SSEA4 and for the tight junction marker ZO1. Brightfield images show the characteristic cyst found in EBs of that size. Images in the last column show representative images of 200 µm diameter EBs forming within a loaded 200 LTA-PDMS grid (top) and of 500 µm diameter EBs recovered from a 500 LTA-PDMS grid (bottom). The insert image is of a 500 µm diameter EB before removal from the LTA-PDMS grid. (B) ICC stained images for the nuclear marker Hoechst show the multicystic nature of larger EBs. (C) Brightfield Image shows representative images of free suspension EBs that vary in size and shape. (D) Process flow starting with blank silicon wafer through the generation of LTA-PDMS grids for high throughput uniform EB creation. (E) Generation of high throughput uniform EBs starting with either single cell suspension, or with 2D cell clusters. Scale bars in the figure are 200 µm.
Transcriptome variation is present in 200 versus 500 micron hESC WA09 EBs independent of formation method

To determine gene expression differences between EB size and morphology, we used RNA-seq followed by bioinformatics analysis. Overall, we observed a dramatic difference in transcriptomes of 200 versus 500 micron regardless of formation method. We found 1654 differentially expressed genes at a false discovery rate (FDR) <0.01 (less than 1% false positives) and 661 genes at FDR <0.001 (less than 0.1% false positives). All further analysis focused on the 661 differentially expressed genes at a FDR of <0.001. The majority of the 661 genes (476/661) were upregulated in the 500- micron compared to the 200 micron EBs. To understand the gene expression relationship and consistence across all experimental replicates we used hierarchical clustering analysis (Figure 2) and examined the gene class enrichments for up- or down-regulated genes. Table 2 indicates the top genes up- or down-regulated along with established roles. Functional distinctions are further discussed in following sections. See Supplementary Table 1 for the full list of genes (GEO access number GSE74792).
Figure 2. Transcriptomic comparison of 200 and 500 micron EBs from WA09 hESC formed by two methods. RNA-seq, Heat Map and Jaccard analysis. (A) 200 or 500 micron sized embryonic bodies (EBs) formed via 2D clusters or by templating of single cell (SC) suspensions from WA09 were compared by RNA-seq. The top 50 genes from 661 total differentially expressed genes at a FDR <0.001 were hierarchically clustered and the top 50 genes were divided into (B) downregulated or (C) upregulated tables. Sequence counts at each gene was standardized for visualization by log_2 (gene counts/gene mean). See Table 2 for top genes and Supplementary Table 1 for the full list of genes.
Pluripotency gene expression is independent of EB size and morphology or formation method

Previous reports have identified multiple genes contributing to pluripotency as well as the information on the relative levels of expression of pluripotency genes\textsuperscript{4,21-25} to each other. We compared pluripotency gene expression levels for our 200- or 500-micron EB hESC samples generated using dissociated single cells (Table 3) at the end point of uniform EB formation for each method. We first focused on genes that have been previously reported as important for inducing and maintaining pluripotency in embryonic and induced stem cell lines. This includes expression of the Yamanaka pluripotency factors OCT4, SOX2, KLF4, and c-MYC\textsuperscript{26}, plus the two additional factors NANOG and LIN28A/B identified by James Thomson\textsuperscript{27}. In addition, we evaluated several genes that have been more recently implicated in pluripotency, that include the factors UTF1, SALL4, NR5A2, TBX3, ESRRB, DPPA4\textsuperscript{28,29,30}, and REX-1 (zfp42)\textsuperscript{3,31}. The FDR for all genes examined was <0.001. We observed small but insignificant up-regulation of KLF4 (1.34 fold), Sox2 (0.68), UTF1 (0.97), Lin28B (0.49), and SALL4 (0.45) and downregulation of NR5A2 (-0.92), although again insignificant. Overall, we observe no significant gene expression differences in pluripotency markers dependent on EB size and method of formation in our pluripotent stem cell samples.

Expression of early embryonic tube forming genes in 500 micron multicystic EBs

We previously identified distinct morphological features present in large 500 micron EBs versus 200 micron EBs (Figure 1B; See also Reference 14) that is the presence of multiple cysts that are reminiscent of tubes or neural rosettes. To provide information on these structures we performed bioinformatics analysis on 200 versus 500 micron EBs.
GO analysis reveals up-regulation of genes associated with embryonic tube formation (Table 4 and Materials and Methods). Specifically, genes involved in early tube formation that are upregulated include DVL2, MIB1, TSC1, TSC2, LAMA5, PTK7, PBX1, CELSR1, SCRIB, ZIC2, HECTD1. The proteins encoded by genes for LAMA5, PTK7, and SCRIB is a scaffold protein involved in planar cell polarization processes prior to and during early tube formation in both the neural tube and in cardiogenesis. TSC1 and TSC2 and CELSR1 are genes that are associated primarily with cavitation driven tube formation. MIB1 has been shown to be involved in apoptosis driven tube formation. The general tube formation regulator PBX1 is also upregulated. The proteins encoded by DVL2, ZIC2, and HECTD1 have been shown to be involved in later stages of tube formation, specifically during tube closures.

Additional upregulated apoptosis mediating genes such as AGTR2 and BMF may indicate a role for cell death in remodeling of the internal multicyst formation in 500 micron EBs. The list of genes includes functional roles in cytoskeletal remodeling and cell adhesion, such as E-cadherin and plakins plus various tight junction proteins. Up-regulation of factors in signaling pathways for TGF-β, Wnt, and EGF were also observed and in early organ development are implicated in guiding early tube formation for multiple organ systems including trachea, kidneys, and lungs. Genes that are involved in extracellular matrix remodeling such as zinc finger proteins and other calcium ion receptors and transporters are also upregulated. Genes that may suggest a primed state also are significantly up or down regulated and include USP9X and NYNRIN (up) block protein degradation, EXOSC6 (down) targets RNA for degradation, HES3 (up) inhibits Helix-loop-Helix proteins, H1F0 and WDR82 (up) histone compacting genes, BCORL1 (up) transcriptional repressor, and CCDC129 and CCDC96 (down) coiled-coil domain proteins.
As well, we observe upregulation in both ECM remodeling genes and in early priming transcriptional factors implicated in differentiation. These include for example, genes involved in neurogenesis (SYT7, CNTFR, EFNB1, L1CAM), skeletal muscle (HRC, ACTN3), pituitary gland (TEF, GSX1), pancreas (MAPK8IP1), germ cells (YBX2), and cardiogenesis (SPNS2, PLEC, FOXH1).

**Expression of wound responsive genes in EBs generated by mechanical passaging.**

The formation of EBs is typically done by mechanical passaging of colonies into 2D clusters or by use of single cells in microarrays or bioreactors. Here we compared mechanical passaging versus templating of single cells for 200 micron EBs (**Figure 3**). We observed that compared to single cell EB generation, mechanical passaging into 2D clusters for EB formation activated genes within the GO wound healing (**Table 5 – GO:0009611~Response to wounding**) that include PROCR, GATM, SERPINE1, DSP, PROS1, and PLAU. Roles for these genes that may be relevant include extracellular matrix remodeling, cytoprotection (PROCR and PROS1\(^{51,52,53}\)), as well as ECM remodeling and degradation (PLAU\(^{54,55}\)) and cell-cell tight junction maintenance (DSP\(^{56,57}\)). SERPINE1\(^{58}\) and GATM\(^{59}\) are regulatory genes that are associated with monolayer repair and cell migration. The timeline for EB formation from 2D clusters is 3-5 days for 200 micron EBs and 7-10 days for 500 micron EBs, indicating that these genes if induced by mechanical wounding persist in the newly formed EB despite no noticeable variation in EB morphology versus single cell formed EBs. See **Supplementary Table 2** for the full list of 661 differentially expressed genes (GEO number GSE74792).
Figure 3. Transcriptomic comparison of WA09 hESC 200 micron embryonic bodies derived from cell clusters (2D) or single cells (SC). Embryonic bodies derived via 2D or single cell were compared by RNA-seq. (A) The 334 differentially expressed genes at a FDR <0.001 were hierarchically clustered and the top 50 genes were divided into downregulated (B) or upregulated (C) tables. Sequence counts at each gene was standardized for visualization by log₂ (gene counts/gene mean). See Table 3 and Table 4 for top genes and also Supplementary Table 2 for the full list of genes.
Discussion

Nanotechnology and microtechnology provides critical tools for customized in vitro analysis of stem cells\textsuperscript{64-67}. Here we applied photolithography to generate custom microarrays of defined size to control EB formation and size uniformity for comparative transcriptome analysis. Our findings indicate that manipulation of stem cells for EB formation, as well as EB size, each significantly alter transcriptome profiles which should be considered when differentiation methodologies utilize EBs (Figure 4A and 4B). Thus EBs formed by different strategies may not provide comparable results, consistent with concerns for more uniformity in protocols in the stem cell field\textsuperscript{68,69,70}. The expression of pluripotency genes was consistent and independent of method of formation or EB size (Figure 4C). Gene ontology (GO) analysis revealed an upregulation of genes for wound healing when mechanical passaging is used to form EBs from 2D clusters, not present when single cell dissociation is used. Comparison of multicystic 500 micron EBs versus monocystic 200 micron EBs indicates upregulation of GO pathways for early embryonic tube formation and organogenesis, including but not limited to neural tube formation. This analysis fills important gaps in our understanding of the EB stem cell intermediate and has significance for understanding potential variation and efficiencies in differentiation protocols employing EBs.
Figure 4. 500 micron EBs retain pluripotency while also expressing early embryonic tube formation pathway genes. Representative 3D renderings of monocyctic 200 and polycystic 500 micron EBs are shown in (A), arrows highlight the internal cystic structures. Simulated scale bar is 200 microns. (B) 3D renderings of early embryonic tube formation pathways that appear active in the 500 micron EBs, based on observed gene expression (listed genes). (C) ICC staining of a polarized cyst in a 500 micron EB showing levels of nuclear pluripotency marker Sox2, Hoechst nuclear stain, and ZO1 tight junction marker. Scale bar is 50 microns.
Tube formation has been shown to be a prerequisite for proper development in multiple organs and there are multiple factors that direct this process. The morphological processes of early tube formation share common features that include molecular events to establish apical membrane polarity. Genes that govern this process are upregulated in large 500 micron EBs compared to the smaller 200 micron EBs. Critical genes that are involved in tube formation include tight junction proteins and calcium ion receptors and transporters, which are involved in ECM remodeling to generate conditions that promote tube formation. Additionally, other genes involved in ECM remodeling have been shown to prevent the apical surfaces of the forming tube from sticking together. Within the cell, multiple genes that code for cytosolic transport proteins, which are involved in early tube formation cues are also consistently upregulated. These factors regulate the concentration of small soluble molecules in the ECM, which help generate the tight junction between the cells within the forming tube walls that may affect the size and shape of the tube being formed. In addition, genes that are involved in programmed cell death are also upregulated in the multicystic 500-micron EBs (Figure 4B). These genes assist formation of cavities, such as pre-amniotic cavities during embryonic development in vertebrates. Transcriptome analysis further indicates the expression of genes related to remodeling of ECM, cytoskeleton, cell-cell interactions, and signaling, consistent with the observed multicystic internal structure. The 500-micron EBs may be primed for formation of these more complex tissues and organoids71. Select genes that code for proteins found at early stages of multi-lineage differentiation are in some cases upregulated while later expressed proteins remain downregulated in 500-micron EBs. Together the combination of upregulated and downregulated genes in larger 500-micron EBs is intriguing and may in some cases provide unique advantages for differentiation when tube formation is
needed\textsuperscript{71}. In this regard it is interesting that stem cell differentiation for cardiogenesis has been proposed to occur with higher frequency by use of larger EBs in the range of 500 microns, but smaller, in the range of 200 microns, EBs were shown to be more enriched for beating cardiomyocytes\textsuperscript{72,73,74}.

**Chapter Conclusions**

Our findings reveal new insights into size-dependent variability in EBs as well as retention of signals arising from manipulation of stem cells for EB formation. Elevated expression of genes involved in wound healing when mechanical passaging is used is in retrospect not surprising and consistent with in vitro wounding models typically used in culture\textsuperscript{58,75}. However the retention of these signals up to 5 days during the formation of EBs in microarray templates indicates that EB maturation may be a longer process then realized. In addition, the type or extent of maturation possible is also reflected in EB size, with larger EBs capable of more extensive remodeling as revealed by multicystic morphology and initiation of early tube formation pathways while retaining pluripotency status.
Chapter References:


59. V. Wessagowit, R. Mallipeddi, J. A. McGrath, and A. P. South, Clinical and Experimental Dermatology. 29(6), 664-668 (2004)
64. A. Khademhosseini, R. Langer, J. Borenstein, and J. P. Vacanti, PNAS. 103(8), 2480-2487 (2005)


### Table 1: EB RNA Samples Used for Microarray Expression Analysis

<table>
<thead>
<tr>
<th>EB Formation Method</th>
<th>200 µm EB 2D cluster</th>
<th>200 µm EB single cell</th>
<th>500 µm EB 2D cluster</th>
<th>500 µm EB single cell</th>
</tr>
</thead>
<tbody>
<tr>
<td>hESC</td>
<td>WA09</td>
<td>WA09</td>
<td>WA09</td>
<td>WA09</td>
</tr>
<tr>
<td>ED-iPSC</td>
<td>F3.5.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ED-iPSC</td>
<td>H3.3.1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** ED= ethnically diverse; F= African American; H= Hispanic-Latino;

No samples for grey boxes. Two EBs analyzed for each sample.
Table 2: Genes up- or down-regulated in 500 micron EBs.

<table>
<thead>
<tr>
<th>Genes Downregulated in 500 micron EBs</th>
<th>Genes Upregulated in 500 micron EBs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gene Name</strong></td>
<td><strong>Gene Name</strong></td>
</tr>
<tr>
<td>PDXP</td>
<td>MOGS</td>
</tr>
<tr>
<td>Pyridoxine, Vitamin B6</td>
<td>Mannosyl-Oligosaccharide Glucosidase</td>
</tr>
<tr>
<td>Active form of vitamin B6</td>
<td></td>
</tr>
<tr>
<td>FCRL4</td>
<td>NBAS</td>
</tr>
<tr>
<td>Fc Receptor-Like 4</td>
<td>Neuroblastoma Amplified Sequence</td>
</tr>
<tr>
<td>Fc receptor-like glycoprotein</td>
<td></td>
</tr>
<tr>
<td>OR13C3</td>
<td>ISOC2</td>
</tr>
<tr>
<td>Olfactory Receptor, Fam 13, Sub-fam C, Mem. 3</td>
<td>Isochorismatase Domain Containing 2</td>
</tr>
<tr>
<td>Initiates neuronal response to perceive smell</td>
<td></td>
</tr>
<tr>
<td>SLC15A4</td>
<td>FOXH1</td>
</tr>
<tr>
<td>Solute Carrier Family 15, Member4</td>
<td>Forkhead Box H1</td>
</tr>
<tr>
<td>Proton oligopeptide cotransporter</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SH3KBP1</td>
</tr>
<tr>
<td>SH3-Domain Kinase Binding Protein 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SLC39A7</td>
</tr>
<tr>
<td>Solute Carrier Family 39, Member 7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MIB1</td>
</tr>
<tr>
<td>Mindbomb E3 Ubiquitin Protein Ligase 1</td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>NUMA1</td>
<td>Nuclear Mitotic Apparatus Protein 1</td>
</tr>
<tr>
<td></td>
<td>Aligns mitotic spindle in asymmetric cell division</td>
</tr>
<tr>
<td>SLC7A2</td>
<td>Solute Carrier Family 7, Member 2</td>
</tr>
<tr>
<td></td>
<td>Cationic amino acid transporter</td>
</tr>
<tr>
<td>PIP5K1A</td>
<td>Phosphatidylinositol-4-Phosphate 5-Kinase</td>
</tr>
<tr>
<td></td>
<td>Catalyzes phosphorylation of PtdIns4P</td>
</tr>
<tr>
<td>ZBED4</td>
<td>Zinc Finger, BED-Type Containing 4</td>
</tr>
<tr>
<td></td>
<td>Zinc finger protein</td>
</tr>
<tr>
<td>WDR82</td>
<td>WD Repeat Domain 82</td>
</tr>
<tr>
<td></td>
<td>Involved in mitosis to interphase transition</td>
</tr>
<tr>
<td>CAPNS1</td>
<td>Calpain, Small Subunit 1</td>
</tr>
<tr>
<td></td>
<td>Remodeling of cytoskeleton during cell cycle</td>
</tr>
<tr>
<td>NFATC1</td>
<td>Nuclear Factor Of Activated T-Cells</td>
</tr>
<tr>
<td></td>
<td>Gene expression in embryonic cardiac cells</td>
</tr>
<tr>
<td>GTF2H4</td>
<td>General Transcription Factor IIH, Polypeptide 4</td>
</tr>
<tr>
<td></td>
<td>Involved in nucleotide excision repair</td>
</tr>
<tr>
<td>USP9X</td>
<td>Ubiquitin Specific Peptidase 9, X-Linked</td>
</tr>
<tr>
<td></td>
<td>Chromosome alignment and segregation</td>
</tr>
<tr>
<td>PLEC</td>
<td>Plectin</td>
</tr>
<tr>
<td></td>
<td>Interlinks intermediate filaments and microtubules</td>
</tr>
<tr>
<td>PTMS</td>
<td>Parathymosin</td>
</tr>
<tr>
<td></td>
<td>Remodels chromatin structure</td>
</tr>
<tr>
<td>YAP1</td>
<td>Yes-Associated Protein 1</td>
</tr>
<tr>
<td></td>
<td>Regulates cell proliferation and apoptosis</td>
</tr>
<tr>
<td>UNC13B</td>
<td>Unc-13 Homolog B</td>
</tr>
<tr>
<td></td>
<td>Synaptic vesicle maturation during exocytosis</td>
</tr>
<tr>
<td>TEF</td>
<td>Thyrotrophic Embryonic Factor</td>
</tr>
<tr>
<td></td>
<td>Transcription factor for the TSHB promoter</td>
</tr>
<tr>
<td>ANKFY1</td>
<td>Ankyrin Repeat FYVE Domain Containing 1</td>
</tr>
<tr>
<td></td>
<td>Proposed effector of Rab5</td>
</tr>
<tr>
<td>CTNNA2</td>
<td>Catenin, Alpha 2</td>
</tr>
<tr>
<td>Gene</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>CNOT1</td>
<td>CCR4-NOT Transcription Complex, Subunit 1</td>
</tr>
<tr>
<td></td>
<td>Maintenance of embryonic stem cell identity</td>
</tr>
<tr>
<td>CERS4</td>
<td>Ceramide Synthase 4</td>
</tr>
<tr>
<td></td>
<td>Involved in the production of sphingolipids</td>
</tr>
<tr>
<td>EMD</td>
<td>Emerin</td>
</tr>
<tr>
<td></td>
<td>Formation of the nuclear actin cortical network</td>
</tr>
<tr>
<td>GRAMD4</td>
<td>GRAM Domain Containing 4</td>
</tr>
<tr>
<td></td>
<td>Mediator of E2F1-induced apoptosis</td>
</tr>
<tr>
<td>IGSF3</td>
<td>Immunoglobulin Superfamily, Member 3</td>
</tr>
<tr>
<td></td>
<td>Involved in cyst and duct formation</td>
</tr>
<tr>
<td>PTK7</td>
<td>Protein Tyrosine Kinase 7 (Inactive)</td>
</tr>
<tr>
<td></td>
<td>Involved in Wnt/planar cell polarity signaling</td>
</tr>
<tr>
<td>LRP4</td>
<td>Low Density Lipoprotein Receptor-Rel. Protein 4</td>
</tr>
<tr>
<td></td>
<td>Involved in canonical Wnt pathway regulation</td>
</tr>
<tr>
<td>EFNA3</td>
<td>Ephrin-A3</td>
</tr>
<tr>
<td></td>
<td>Migration, repulsion, and adhesion regulator</td>
</tr>
<tr>
<td>MACROD1</td>
<td>MACRO Domain Containing 1</td>
</tr>
<tr>
<td></td>
<td>Enhances ESR1-mediated transcription activity</td>
</tr>
<tr>
<td>EEF2</td>
<td>Eukaryotic Translation Elongation Factor 2</td>
</tr>
<tr>
<td></td>
<td>Ribosomal translocation in translation elongation</td>
</tr>
<tr>
<td>ORMDL3</td>
<td>Sphingolipid Biosynthesis Regulator 3</td>
</tr>
<tr>
<td></td>
<td>Negative regulator of sphingolipid synthesis</td>
</tr>
<tr>
<td>ARHGEF11</td>
<td>Rho Guanine Nucleotide Exchange Factor 11</td>
</tr>
<tr>
<td></td>
<td>Regulation of RhoA GTPase</td>
</tr>
<tr>
<td>NYNRIN</td>
<td>NYN Domain &amp; Retroviral Integrase Containing Nucleic acid binding</td>
</tr>
<tr>
<td>CLASP1</td>
<td>Cytoplasmic Linker Associated Protein 1</td>
</tr>
<tr>
<td></td>
<td>Alignment of chromosomes to the mitotic spindle</td>
</tr>
<tr>
<td>NAV2</td>
<td>Neuron Navigator 2</td>
</tr>
<tr>
<td></td>
<td>Neuronal development, cell growth and migration</td>
</tr>
<tr>
<td>Gene</td>
<td>Description</td>
</tr>
<tr>
<td>-----------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>KLHL15</td>
<td>Kelch-Like Family Member 15 Protein ubiquitination &amp; cytoskeletal organization</td>
</tr>
<tr>
<td>NMNAT2</td>
<td>Nicotinamide Nucleotide Adenylyltransferase 2 Catalyze an essential step in NADPpathway</td>
</tr>
<tr>
<td>NUFIP2</td>
<td>Nuclear Fragile X Protein Interacting Protein 2 RNA binder</td>
</tr>
<tr>
<td>RACGAP1</td>
<td>Rac GTPase Activating Protein 1 Controls cell growth and differentiation</td>
</tr>
<tr>
<td>HIPK1</td>
<td>Homeodomain Interacting Protein Kinase 1 Regulation and TNF-mediated cellular apoptosis</td>
</tr>
<tr>
<td>XPR1</td>
<td>Xenotropic And Polytropic Retrovirus Receptor 1 G-protein coupled signal transduction</td>
</tr>
<tr>
<td>PITPNM1</td>
<td>Phosphatidylinositol Transfer Protein RHOA activity and cytoskeleton remodeling</td>
</tr>
<tr>
<td>SLC12A7</td>
<td>Solute Carrier Family 12, Member 7 Electroneutral potassium-chloride cotransporter</td>
</tr>
</tbody>
</table>
Table 3: Comparison of pluripotency gene expression in 200 vs 500 micron hESC EBs.

<table>
<thead>
<tr>
<th>Pluripotency Gene</th>
<th>Log2 Fold Change</th>
<th>FDR Value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oct 4/POU5F1</td>
<td>+ 0.19</td>
<td>&gt;0.001</td>
<td>0.34</td>
</tr>
<tr>
<td>Myc/c-Myc</td>
<td>- 0.76</td>
<td>&gt;0.001</td>
<td>0.025</td>
</tr>
<tr>
<td>KLF4</td>
<td>+ 1.34</td>
<td>&gt;0.001</td>
<td>0.00023</td>
</tr>
<tr>
<td>Sox2</td>
<td>+ 0.68</td>
<td>&gt;0.001</td>
<td>0.0024</td>
</tr>
<tr>
<td>Nanog</td>
<td>- 0.37</td>
<td>&gt;0.001</td>
<td>0.21</td>
</tr>
<tr>
<td>UTF1</td>
<td>+ 0.97</td>
<td>&gt;0.001</td>
<td>0.0043</td>
</tr>
<tr>
<td>LIN28A</td>
<td>+ 0.43</td>
<td>&gt;0.001</td>
<td>0.029</td>
</tr>
<tr>
<td>LIN28B</td>
<td>+ 0.49</td>
<td>&gt;0.001</td>
<td>0.0027</td>
</tr>
<tr>
<td>SALL4</td>
<td>+ 0.45</td>
<td>&gt;0.001</td>
<td>0.00077</td>
</tr>
<tr>
<td>NR5A2</td>
<td>-0.92</td>
<td>&gt;0.001</td>
<td>0.0041</td>
</tr>
<tr>
<td>TBX3</td>
<td>+ 0.69</td>
<td>&gt;0.001</td>
<td>0.12</td>
</tr>
<tr>
<td>ESRRB</td>
<td>+ 0.58</td>
<td>&gt;0.001</td>
<td>0.16</td>
</tr>
<tr>
<td>DPPA4</td>
<td>- 0.11</td>
<td>&gt;0.001</td>
<td>0.51</td>
</tr>
<tr>
<td>Rex-1 (zfp42)</td>
<td>+ 0.10</td>
<td>&gt;0.001</td>
<td>0.67</td>
</tr>
</tbody>
</table>
Table 4: Tube formation genes in 200 versus 500 micron EBs.

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Function</th>
<th>Fold Change</th>
<th>(p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DVL2</td>
<td>Dishevelled Segment Polarity Protein 2 Regulates Wnt signaling pathways</td>
<td>+ 0.55</td>
<td>3.67E-07</td>
</tr>
<tr>
<td>MIB1</td>
<td>Mindbomb E3 Ubiquitin Protein Ligase 1 Ubiquitinates Notch protein ligands</td>
<td>+ 0.89</td>
<td>7.47E-17</td>
</tr>
<tr>
<td>TSC1</td>
<td>Tuberous Sclerosis 1 Negatively regulator of mTORC1 pathway</td>
<td>+ 0.66</td>
<td>1.89E-07</td>
</tr>
<tr>
<td>TSC2</td>
<td>Tuberous Sclerosis 2 Negatively regulator of mTORC1 pathway</td>
<td>+ 0.55</td>
<td>1.11E-05</td>
</tr>
<tr>
<td>LAMA5</td>
<td>Laminin, Alpha 5 Cell-cell interactions in embryonic development</td>
<td>+ 1.32</td>
<td>2.73E-06</td>
</tr>
<tr>
<td>PTK7</td>
<td>Protein Tyrosine Kinase 7 (Inactive) Wnt signaling mediated cell-cell interactions</td>
<td>+ 1.59</td>
<td>2.79E-12</td>
</tr>
<tr>
<td>PBX1</td>
<td>Pre-B-Cell Leukemia Homeobox 1 Transcription factor involved in cell patterning</td>
<td>+ 0.49</td>
<td>2.11E-05</td>
</tr>
<tr>
<td>CELSR1</td>
<td>Cadherin, EGF LAG G-Type Receptor 1 Cell-cell signaling during neural development</td>
<td>+ 1.04</td>
<td>1.77E-07</td>
</tr>
<tr>
<td>SCRIB</td>
<td>Scribbled Planar Cell Polarity Protein Regulator of polarized cell-based differentiation</td>
<td>+ 1.12</td>
<td>9.89E-06</td>
</tr>
<tr>
<td>ZIC2</td>
<td>Zic Family Member 2 Transcription factor in early CNS organogenesis</td>
<td>+ 1.23</td>
<td>1.05E-05</td>
</tr>
<tr>
<td>HECTD1</td>
<td>HECT Cont. E3 Ubiquitin Protein Ligase 1 Required for neural tube closure and others</td>
<td>+ 0.71</td>
<td>2.86E-05</td>
</tr>
</tbody>
</table>
Table 5: Wound responsive genes in single cell versus 2D intermediates generated 200 micron EBs

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Function</th>
<th>Fold Change</th>
<th>-log10 (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAGLN</td>
<td>Transgelin, Actin cross-linking, contributes to senescence</td>
<td>-2.46</td>
<td>23.64</td>
</tr>
<tr>
<td>ANXA3</td>
<td>Annexin A3, Regulates cell growth and signal transduction</td>
<td>-2.24</td>
<td>29.97</td>
</tr>
<tr>
<td>CXCL5</td>
<td>Chemokine (C-X-C Motif) Ligand 5, Involved in cell migration and invasion</td>
<td>-2.04</td>
<td>29.19</td>
</tr>
<tr>
<td>PROS1</td>
<td>Protein S (Alpha), ECM remodeling protein</td>
<td>-1.79</td>
<td>11.34</td>
</tr>
<tr>
<td>SERPINE1</td>
<td>Serpin Peptidase Inhibitor, Clade E, Serine protease inhibitor, ECM remodeling</td>
<td>-1.52</td>
<td>6.47</td>
</tr>
<tr>
<td>ANXA1</td>
<td>Annexin A1, Promotes membrane fusion and exocytosis</td>
<td>-1.42</td>
<td>5.51</td>
</tr>
<tr>
<td>H1F0</td>
<td>H1 Histone Family, Member 0, Found in late differentiated non-dividing cells</td>
<td>-1.36</td>
<td>8.97</td>
</tr>
<tr>
<td>PLAU</td>
<td>Plasminogen Activator, Urokinase, ECM modification, migration and proliferation</td>
<td>-1.32</td>
<td>7.00</td>
</tr>
<tr>
<td>HSPB8</td>
<td>Heat Shock 22kDa Protein 8, Regulates cell proliferation and apoptosis</td>
<td>-1.30</td>
<td>5.64</td>
</tr>
<tr>
<td>PRNP</td>
<td>Prion Protein, Neuronal development and synapse plasticity</td>
<td>-1.25</td>
<td>8.97</td>
</tr>
<tr>
<td>MYO1C</td>
<td>Myosin IC, Transcription factor for actin-based motors</td>
<td>-1.24</td>
<td>9.33</td>
</tr>
<tr>
<td>CER1</td>
<td>Cerberus 1, DAN Family BMP Antagonist, Regulates Nodal signaling during gastrulation</td>
<td>-1.22</td>
<td>5.01</td>
</tr>
<tr>
<td>Gene Symbol</td>
<td>Description</td>
<td>log2Ratio</td>
<td>FDR</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------</td>
<td>-----------</td>
<td>-----</td>
</tr>
<tr>
<td>PROCR</td>
<td>Protein C Receptor, Endothelial Binds to and enhances activated Protein C</td>
<td>-1.20</td>
<td>5.63</td>
</tr>
<tr>
<td>EMP3</td>
<td>Epithelial Membrane Protein 3 Cell proliferation and cell/cell interactions</td>
<td>-1.15</td>
<td>3.27</td>
</tr>
<tr>
<td>APOBEC3B</td>
<td>Apolipoprotein B MRNA Editing Enzyme Implicated in growth or cell cycle control</td>
<td>-1.15</td>
<td>3.59</td>
</tr>
<tr>
<td>CLDN3</td>
<td>Claudin 3 Regulates tight junction formation</td>
<td>-1.13</td>
<td>3.91</td>
</tr>
<tr>
<td>SCRN1</td>
<td>Secernin 1 Calcium dependent secretase in mast cells</td>
<td>-1.12</td>
<td>9.20</td>
</tr>
<tr>
<td>SERPINB6</td>
<td>Serpin Peptidase Inhibitor, Clade B M6 Regulates blood serine proteases in the brain</td>
<td>-1.10</td>
<td>5.89</td>
</tr>
<tr>
<td>SOCS2</td>
<td>Suppressor Of Cytokine Signaling 2 Regulates cytokine transduction of GH/IGF1</td>
<td>-1.08</td>
<td>5.26</td>
</tr>
<tr>
<td>BMP7</td>
<td>Bone Morphogenetic Protein 7 Involved in embryonic TGFβ pathways</td>
<td>-1.07</td>
<td>4.77</td>
</tr>
<tr>
<td>ACTA1</td>
<td>Actin, Alpha 1, Skeletal Muscle Involved in cell motility, structure and integrity</td>
<td>-1.07</td>
<td>4.09</td>
</tr>
<tr>
<td>HPN</td>
<td>Hepsin Regulates cell growth and cell morphology</td>
<td>-1.05</td>
<td>3.27</td>
</tr>
<tr>
<td>NMRK2</td>
<td>Nicotinamide Riboside Kinase 2 Regulates Laminin-based cell adhesion</td>
<td>-1.04</td>
<td>5.19</td>
</tr>
<tr>
<td>TMSB4X</td>
<td>Thymosin, Beta 4, X Chromosome Cell proliferation, migration &amp; differentiation</td>
<td>-1.03</td>
<td>6.25</td>
</tr>
<tr>
<td>ADM</td>
<td>Adrenomedullin Regulates hormone secretion &amp; angiogenesis</td>
<td>-1.02</td>
<td>7.27</td>
</tr>
<tr>
<td>S100A10</td>
<td>S100 Calcium Binding Protein A10 Cell cycle progression &amp; differentiation</td>
<td>-1.02</td>
<td>6.22</td>
</tr>
<tr>
<td>ANXA2</td>
<td>Annexin A2 Regulates cell growth and transduction</td>
<td>-1.01</td>
<td>3.88</td>
</tr>
<tr>
<td>TAGLN2</td>
<td>Transgelin 2</td>
<td>-1.01</td>
<td>5.90</td>
</tr>
<tr>
<td>Gene Name</td>
<td>Gene Function</td>
<td>Fold Change</td>
<td>-log10 (p-value)</td>
</tr>
<tr>
<td>-----------</td>
<td>---------------</td>
<td>-------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>HIST1H1A</td>
<td>Histone Cluster 1, H1a Regulator of individual gene transcription</td>
<td>2.71</td>
<td>25.79</td>
</tr>
<tr>
<td>NLRP2</td>
<td>NLR Family, Pyrin Domain Containing 2 Suppresses TNF/CD40-induced NFKB1 activity</td>
<td>1.69</td>
<td>17.99</td>
</tr>
<tr>
<td>ZNF883</td>
<td>Zinc Finger Protein 883 Involved in transcriptional regulation</td>
<td>1.57</td>
<td>7.69</td>
</tr>
<tr>
<td>SYT11</td>
<td>Synaptotagmin XI Involved in vesicular trafficking and exocytosis</td>
<td>1.54</td>
<td>9.5</td>
</tr>
<tr>
<td>TYW3</td>
<td>TRNA-YW Synthesizing Protein 3 Homolog Stabilizes codon-anticodon interactions</td>
<td>1.3</td>
<td>6.21</td>
</tr>
<tr>
<td>LMO3</td>
<td>LIM Domain Only 3 (Rhombotin-Like 2)</td>
<td>1.26</td>
<td>5.29</td>
</tr>
</tbody>
</table>

**Putative marker of differentiated smooth muscle**

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Function</th>
<th>Fold Change</th>
<th>-log10 (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLA-DQB1</td>
<td>MHC, Class II, DQ Beta 1 Plays a central role in the immune system</td>
<td>-1.00</td>
<td>3.79</td>
</tr>
<tr>
<td>GATM</td>
<td>L-Arginine:Glycine Amidinotransferase Embryonic muscle &amp; CNS development</td>
<td>-1.00</td>
<td>4.28</td>
</tr>
<tr>
<td>SERPINB1</td>
<td>Serpin Peptidase Inhibitor, Clade B, M1 Protects from inflammatory-based damage</td>
<td>-0.93</td>
<td>3.26</td>
</tr>
<tr>
<td>DSP</td>
<td>Desmoplakin Junction protein linking to intermediate filaments</td>
<td>-0.90</td>
<td>3.88</td>
</tr>
<tr>
<td>AK3</td>
<td>Adenylate Kinase 3 Maintains homeostasis of cellular nucleotides</td>
<td>-0.79</td>
<td>3.40</td>
</tr>
<tr>
<td>AMT</td>
<td>Aminomethyltransferase Critical component of glycine cleavage system</td>
<td>-0.75</td>
<td>3.43</td>
</tr>
<tr>
<td>FGFBP3</td>
<td>Fibroblast Growth Factor Binding Protein 3 Involved in heparin binding and function</td>
<td>-0.75</td>
<td>3.88</td>
</tr>
</tbody>
</table>

**Genes Upregulated in Single Cell Generated 200 micron EBs**
<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
<th>Fold Change</th>
<th>t-Value</th>
</tr>
</thead>
</table>
| SYT4   | Synaptotagmin IV
Involved in vesicular trafficking and exocytosis                  | 1.23        | 3.80    |
| CRYZ   | Crystallin, Zeta (Quinone Reductase)
Binds NADP & enhances mRNA stability in BCL2 | 1.21        | 5.78    |
| DOCK2  | Deducator Of Cytokinesis 2
Involved in cytoskeletal rearrangement                                       | 1.18        | 3.60    |
| HIST1H1B | Histone Cluster 1, H1b
Regulator of individual gene transcription                                    | 1.17        | 4.82    |
| ZNF619 | Zinc Finger Protein 619
Involved in transcriptional regulation                                          | 1.15        | 3.88    |
| TMEM132B | Transmembrane Protein 132B
Integral membrane component                                                   | 1.08        | 5.63    |
| P2RY1  | Purinergic Receptor P2Y
Receptor for extracellular ATP and ADP                                         | 1.08        | 3.43    |
| BTBD17 | BTB (POZ) Domain Containing 17
Transcription factor in early differentiation                                    | 1.00        | 3.60    |
| ZFP42  | ZFP42 Zinc Finger Protein (Rex-1)
Acquisition & maintaining of ES pluripotency                                    | 0.88        | 5.90    |
Chapter 5:
Distinct and Shared Determinants of Cardiomyocyte Contractility in Multi-Lineage Competent Ethnically Diverse iPSCs

State of the Field

The generation and study of human induced pluripotent stem cells (iPSCs) is proving instrumental in understanding normal development and disease mechanisms in the context of natural human genetic variation. Reprogrammed pluripotent cells exhibit variation amongst derived lines and in comparison to other pluripotent cell types such as embryonic stem cells (ESCs) or embryonal carcinoma cells (ECC). This complexity in the reprogramming process itself generates additional variability in gene expression profiles, even within replicate iPSC lines, and necessitates further functional analysis of differentiated cells and tissues. Natural variations based on sex, age and ethnicity are also known to influence therapeutic outcomes. Of vital interest to biomedical applications with human iPSCs is a better understanding of how developmental pathways are impacted by the initial reprogrammed pluripotent state of iPSCs and how this translates into functional competencies and disease models.

Ethnic genetic diversity can have a determining role in human disease by effecting disease prevalence or efficacy of drug therapy in patients. A limited number of ethnically varied genetic loci are known, and the impact that these loci have on disease occurrence and drug resistance is not well established. Stem cell biology offers an opportunity to investigate ethnic genetic contributions with potential for biomedical impact. To increase
the ethnic diversity in available stem cell resources that is reflective of the global human population, we recently reported derivation of new ethnically diverse human iPSC lines (ED-iPSC) well characterized by teratoma formation, qRT-PCR, and multi-lineage commitment. These cells are likely to be an important resource in biomedical research towards regenerative medicine and novel drug discovery, which are geared towards utilizing ethnicity and environment induced factors to optimize efficacy of developed treatments.

**Hypothesis**

We propose to cross-compare ED-iPSC lines by bioinformatics analysis of whole genome transcriptome and epigenetic profiles and rigorously examine under a single platform multi-lineage differentiation to multiple differentiated tissues. This will include ectoderm-derived pyramidal neurons, Type I astrocytes, and retinal pigment epithelial (RPE) cells, endodermal pancreatic progenitors, and mesoderm-derived smooth muscle and cardiomyocytes from established protocols initiated from uniformly sized EB intermediates. The bioinformatics evaluation of differentiation efficiency within and across ethnicities will provide the first comparative analysis of differential gene expression pathways in ethnically diverse differentiation pathways, with a focus on cardiomyocyte development. We predict that our work will help reveal both distinctive and shared pathways among the ethnic iPSCs evaluated for cardiogenesis and likely will include genes with known ethnic impact on cardiac therapy. These findings are expected to be broadly important to the stem cell community and to enhance our understanding of the genes and pathways involved in cardiac development.
Strategy

Identification of the interplay between key determinants in genetic pathways underlying tissue development and function is expected to benefit by inclusion of ethnically diverse populations. We previously derived ethnically diverse induced pluripotent stem cells (ED-iPSCs) of African American, Hispanic Latino and Asian self-designated ethnic origins with normal karyotype and pluripotency as evaluated by qRT-PCR of relevant biomarkers, *in vivo* teratoma formation, and *in vitro* tri-lineage commitment. Here we will evaluate multi-lineage differentiation into multiple cell types and use bioinformatics analysis of whole genome RNA-Seq and Chip-seq data within ED-iPSC replicates and within ethnically distinct lines to cross-evaluate gene expression. We will do bioinformatics comparisons between ethnic replicate lines and across ethnicities which likely will reveal both shared and distinct determinants. Our analysis will identify control transcription factors in genetic pathways for multi-lineage differentiation and specifically cardiogenic contractility, both in ethnicity-independent, and dependent fashion. The evaluation of these ED-iPSC lines will represent an ethnically informative new stem cell resource for understanding and optimizing multi-lineage differentiation necessary for therapeutic applications of human iPSCs.
Acknowledgements and Contributions:

I contributed to conception and design, collection and assembly of data, data analysis and interpretation, and manuscript writing. Zachary T. Olmsted aided in collection and assembly of data, and manuscript writing. Haluk Dogan contributed to assembly of data, data analysis and interpretation, and manuscript writing. Eda Gongorurler helped with data analysis and interpretation. Maria Tsompana contributed to collection and assembly of data, and data analysis and interpretation. Hasan H. Otu contributed to conception and design, collection and assembly of data, data analysis and interpretation, and manuscript writing. Michael Buck helped with conception and design, data analysis and interpretation, and manuscript writing. Eun-Ah Chang helped with collection and assembly of data. Jose Cibelli contributed to conception and design, provision of study material or patients, and data analysis and interpretation. Janet L. Paluh contributed to conception and design, financial support, administrative support, provision of study material or patients, data analysis and interpretation, and manuscript writing. Copyright permission for the purposes of this dissertation can be found at: http://www.nature.com/srep/journal-policies/editorial-policies

Parts of this work have been submitted for publication:

Abstract

The realization of personalized medicine through induced pluripotent stem cell (iPSC) technology remains at a pivotal stage in which expanded transcriptomics, epigenomics, and bioinformatics studies continue to play an important role in elucidating genetic pathways that direct tissue development and function. The diverse human population will also benefit from iPSC studies that enable contributing factors such as ethnicity, age or gender to be evaluated. Such studies are only beginning to be enabled as resources are developed. Previously we derived replicate iPSC lines of African American, Hispanic-Latino and Asian self-designated ethnically diverse (ED) origins with normal karyotype, qRT-PCR verified pluripotency biomarkers, teratoma formation and in vitro tri-lineage commitment. Here, bioinformatics of high quality ED-iPSC RNA-Seq and ChIP-seq pluripotency data sets are compared along with demonstrated differentiation to multiple neural, pancreatic, smooth muscle and cardiomyocyte multi-lineage cell types. In our analysis we observed functional differences within ED-iPSC replicate lines and across ethnicities to generate beating cardiomyocytes. Towards advancing cell therapeutic goals for cardiac health that use iPSC derived cardiomyocytes we applied bioinformatics analysis of transcriptome data and compared Asian and Hispanic-Latino ED-iPSC replicate beating/non-beating pairs. Our analysis identifies genes that may serve as potential predictors of beating outcomes and their relevant GO pathways by comparing ethnically distinct, shared or inversely regulated genes associated with beating or non-beating outcomes. We discuss our findings in terms of similarities to genome wide association studies, contributing disease pathways and prediction of functional cardiomyocytes from the reprogrammed epigenome.
Introduction

Human induced pluripotent stem cells (iPSCs; 1, 2) and their differentiated progenitors are providing increased opportunities in personalized medicine, including use in tissue engineering (3) and in studies to distinguish disease mechanisms from normal development (4-6). The growing number of potential therapeutic applications of iPSCs include high impact diseases, related to cardiac health (7) spinal muscular atrophy (8), neurodegeneration (9, 10) and Type I diabetes (11). Investigation into natural variations based on gender, age and ethnicity and how these contribute to disease and therapeutic outcomes are receiving increased attention that is possible now through inclusion of these parameters in new stem cell resources (12-16). Most challenging to iPSC technology is perhaps ethnic diversity and the additional complexity it brings in distinguishing reprogramming versus ethnic-specific contributions to the generation of multiple functional cells or tissues, which has not yet been investigated.

Reprogramming remains an inefficient and poorly understood process, in which epigenetic variation and differences in gene expression in iPSC lines reflects an underlying stochastic mechanism (17). Bioinformatics is playing an increasing role in evaluation of iPSC lines, including development of comparative bioinformatics models that evaluate pluripotency profiles of lines developed under different platforms (18) and investigation into dynamic changes within lines as cells move from pluripotency through differentiation stages (19-21). Also critically needed are larger scale bioinformatics studies of multiple lines that are derived, differentiated and analyzed under a uniform platform by multiple comprehensive strategies. The ED-iPSC lines analyzed in this study fit this need and include replicate lines across ethnicities derived and analyzed under a single platform.
that is based on additionally optimized protocols to increase accuracy of comparative analysis (13, 22, 23). By our previous validation of these ED-iPSCs we confirmed normal karyotype, qRT-PCR verified pluripotency biomarkers, teratoma formation as well as *in vitro* tri-lineage commitment. The current analysis extends these studies substantially to include comparative transcriptomics and epigenomics along with *in vitro* multi-lineage differentiation to six cell types of high interest for iPSC-derived cell and tissue therapeutic applications.

In this bioinformatics and cell biological study with ED-iPSCs we expand our knowledge of the reprogrammed epigenetic landscape, revealing stochastic differences that fall into two primary clusters, each capable of pluripotency as gauged by teratoma and qRT-PCR (13) as well as new data on multi-lineage differentiation into pyramidal neurons, CD44+/GFAP+ astrocytes, RPE cells, pancreatic progenitors, smooth muscle and beating cardiomyocytes. We used previously established protocols and initiated differentiation from custom templated uniformly sized EB intermediates for consistency in comparative analysis (22, 23). At the epigenetic level, ED-iPSC line reprogramming was evaluated by ChIP-seq for histone modifications H3K4me1 and H3K27ac. We generated RNA-Seq data for transcriptome analysis of the pluripotent state, used in comprehensive bioinformatics analysis to compare gene expression profiles and identify gene ontology (GO) pathways contributing to observed differences in beating and non-beating ED-iPSC derived cardiomyocytes within and across two ethnicities, that are Asian and Hispanic-Latino. Beating was optimized and then quantified in custom microarray wells to observe 3D aggregation, expression of NKX2.5 and GATA4, and expression and banded striations of Troponin-T. By evaluating gene expression in so-called beater versus non-beater ED-iPSC lines, within and across ethnicities, we identify determinants distinct to the ethnic
lines examined or shared between these ethnic lines. We further cross-compare similarly up or down regulated genes as well as inversely regulated genes. Finally we evaluate implicated GO pathways and discuss those with cardiac relevance, as well as disease associated cardiac pathways. Our findings are expected to be broadly important to the stem cell and biomedical communities by enhancing our understanding of the reprogrammed pluripotent environment and genes, pathways, and alternate regulation across ethnicities that may contribute to predicting efficacy in cardiac development.

**Materials and Methods**

hESC and ED-iPSC line maintenance, 2D/3D passaging, and microarray formation of uniform embryoid bodies.

The human ED-iPSCs and hESC line WA09 (WISC Bank, WiCell, Madison, WI) were maintained in mTeSR2/TeSR-E8 complete media (StemCell Technologies, Vancouver, Canada) on StemAdhere (StemCell Technologies, Vancouver, Canada) coated non-tissue culture treated dishes. Cells were enzymatically passaged between days 5-7 using the Gentle Cell Dissociation Agent (StemCell Technologies, Vancouver, Canada) with 10 µl/mL slow release bFGF2 (StemBeads FGF2; Stem Culture Incorporated, Rensselaer, NY) added at each mTeSR2/TeSR-E8 media change.

Generation of 200 µm uniform templated EBs (t-EBs) from chemically dissociated single cells was done in custom microarrays of polydimethylsiloxane (PDMS) as previously described (22). t-EBs formed within five days in the grids with media changes every two days and then gently removed by liquid expulsion with a p1000 micropipette (see also Fig 1a).
In vitro directed differentiation of hESC and ED-iPSC into astrocytes, pyramidal neurons, retinal pigment epithelial cells and pancreatic progenitors.

t-EBs were used or single cells suspension of hESC WA09 or ED-iPSCs that were seeded into microarray wells in mTesR2/TeSR-E8 media with ROCK inhibitor (StemCell Technologies; Vancouver, Canada) and incubated overnight on Matrigel or Vitronectin-XF coated TC-treated dishes. On day 1 of differentiation, the maintenance media was replaced with Neural Induction Media (StemCell Technologies, Vancouver, Canada) and neural stem cells were generated per the protocol over 7-10 days. The resultant rosettes were then manually selected and replated onto Matrigel-coated dishes in Neural Induction Media, supplemented with 10 µM ROCK inhibitor overnight.

For astrocyte generation, adhered neural stem cells were incubated in Astrocyte Media (Sciencell Research Laboratories, Carlsbad, CA) at 37°C and 5% CO2. Complete media changes were done every 3 days. Successful differentiation was confirmed by immunocytology to identify multiple astrocyte specific biomarkers. By day 7 (total day 17 in the protocol) in Astrocyte Media, the bulk of cells were GFAP and CD44 positive.

To generate neurons we followed a protocol for pyramidal neuron generation, essentially as described (24-26), starting with templated 200 micron EBs plated onto either Matrigel or Vitronectin-XF coated dishes. Immunocytology on multiple neuronal biomarkers was performed at critical media transition steps during the differentiation protocol.

For retinal pigment epithelial (RPE) cell generation, we used an adapted protocol (27, 28), starting with pluripotent cells at 80-95% confluence. mRPE media (DMEM +
20\% KOSR, 1x p/s, 1x NEAA, 1x GlutaMAX, 100\mu M \beta\text{-}Mercaptoethanol) supplemented with Nicotinamide (NIC; 10mM) and Activin-A (140ng/ml ;first 14 days) and Retinoic Acid (250nM) was changed every other day until characteristic RPE cell morphology was achieved, usually between days 21 and 55. Once the RPE sheet was formed, media was switched back to mRPE + NIC and changed every 3 days.

To generate pancreatic progenitors, t-EBs, or single cells suspension of either hESC WA09 or ED-iPSCs were seeded into Matrigel or Vitronectin-XF coated dishes. On day 1 of differentiation, we started the StemDiff endoderm differentiation kit (StemCell Technologies). At day 5, we switched to a protocol that had been shown to generate putative \(\beta\)-islet cells (29-31) essentially as described. Immunocytochemistry was done at day 14, day 21, day 28, and day 35 to assess differentiation. See Supplementary Figure S1 for all protocol outlines.

**In vitro directed differentiation of ED-iPSC to cardiomyocytes and smooth muscle cells.**

Templated embryoid bodies (t-EBs) were generated as described above, or single cell suspensions of either hESC WA09 or ED-iPSCs as an alternate starting material for differentiation. Both were seeded into Matrigel or Vitronectin-XF coated plates overnight, prior to starting the differentiation protocol. To initiate differentiation, we used a validated protocol (32, 33) essentially as described. Immunocytochemistry was performed at day 14, day 21, and day 28 to assess expression of the cardiac-specific biomarker Troponin-T. We additionally stained for the cardiac markers Nkx2.5 and GATA4 under further optimized conditions.
Generation of smooth muscle cells was done by protocol optimized from literature (34, 35), essentially as described. Cells were stained for alpha smooth muscle actin (α-SMA) at days 10, 14, 21, 28, and 35 to verify differentiation. See Supplementary Figure S1 for all protocol outlines.

**Optimized cardiomyocyte differentiation protocol for ED-iPSCs lines reveals a strict Gsk3 inhibition requirement.**

We hypothesized that a key step in our cardiac differentiation is the initial Gsk3 inhibition (CHIR99021), as has been reported previously (32, 33). The A2.2.2 and H3.1.1 ED-iPSC lines can generate banded Troponin-T at 12µM GSK3 inhibitor (CHIR99021) incubation for 24 hours. By contrast, the F3.5.2 line exhibits massive cell death within 24 hours in the differentiation culture at those conditions. Cell death is greatly reduced when the GSK3 inhibitor concentration is optimized for both concentration and exposure time. Based on bioinformatics data on the A2.2.2, H3.3.1 and F3.5.2 ED-iPSC lines, the GSK3 inhibitor concentration was dropped to 6µM, and the incubation time was further experimentally adjusted down to 6 hours, which generated banded Troponin-T+ cells and spontaneously beating areas, as well as induce beating in LTA-PDMS generated 3D cultures (Supplementary Movies 1-4).
Microscopy and immunocytology.

Phase images were obtained on a Nikon 80i epifluorescence microscope with a PLAN 10X 0.30 NA DL objective and a cooled QICam CCD camera. Fluorescent images were obtained on a Leica SP5 Laser Scanning Confocal Microscope with HC PL FLUOTAR 10X 0.30 NA or HCX PL APO CS 20X .70 NA objectives and on a Zeiss AxioObserver Z1 Inverted Microscope with Colibri LED illumination, 100X oil 1.45 NA PlanFLUAR or 63X Plan-Apochromat 1.4 NA oil DIC objectives, and Hamamatsu ORCA ER CCD camera using Zeiss Axiovision Rel 4.8 acquisition software. Images were compiled using Adobe Photoshop (Adobe Systems Inc., San Jose, CA) and Microsoft PowerPoint (Microsoft Corp., Redmond, WA) software. Immunocytology was done following modified fixation in 4% formaldehyde followed with a 0.5% PBST incubation overnight at 4˚C and then by a gentle wash in HBSS (Life Technologies, Grand Island, NY). Nonspecific binding was blocked by 3 x 10 minute incubations in 1% BSA-HBSS, followed by an additional HBSS wash.

Pluripotency primary antibodies used (1:1000 each) were anti-Oct4A C-10 (Santa Cruz Biotechnology, Dallas, TX). Specific cell type differentiation antibodies (1:1000 each) were used. For astrocyte differentiation, we used A2B5, CD44, GFAP, Vimentin and Nestin markers. Pyramidal neuron differentiation was screened with A2B5, MAP2, VGlut1, and Tuj1 markers. Retinal pigment epithelium (RPE) differentiation was evaluated by phase microscopy, and with the biomarkers Occludin and ZO1, which stain for epithelial tight junctions. Pancreatic differentiation was evaluated with PDX1 and Pax6 markers. Cardiomyocyte generation was stained with the Troponin-T, NKX2.5 and the GATA4 markers. Smooth muscle cells were stained for the α-smooth muscle actin marker. Secondary antibodies used (1:1000) were AlexaFluor 488 and AlexaFluor 594 (A-11001,
A-11037, Invitrogen, Carlsbad, CA) and were used with Hoechst in HBSS at 4°C overnight, followed with a final 1 hour HBSS wash at 4°C. Mounting was then performed in ProLong Diamond (Life Technologies, Grand Island NY) at 20°C overnight in the dark before imaging or storing in the dark at 4°C.

**Nucleic acid preparation for RNA-Seq and ChIP-seq.**

Total RNA isolation for transcriptome analysis of the eight ED-iPSC lines and the four initial fibroblast cell lines were done using the Ambion PureLink Mini RNA isolation kit (Life Technologies, Grand Island, NY). Total RNA integrity and library quality control was validated using the Agilent Technologies 2100 Bioanalyzer. TruSeq RNA Sample Preparation v2 low sample (LS) protocol (Illumina, San Diego, CA) was used for mRNA preparation, library construction, and purification. Constructed RNA-Seq libraries were purified with Agencourt AMPure XP beads and quantified using the Quant-iT PicoGreen ds DNA Assay Kit (Invitrogen, Carlsbad, CA) and the KAPA Library Quantification Kit (Kapa Biosystems, Wilmington, MA) using qPCR. Libraries were normalized based on qPCR values and pooled using the TruSeq SR Cluster Kit v3-cBot-HS (Illumina, San Diego, CA). Pooled samples were four-plexed and sequenced with the HiSeq 2500 v3 sequencer (Illumina, San Diego, CA) producing 30 to 50 million single-end 50 bp reads.

Chromatin extractions from ED-iPSC cell lines were done using the Chromatin Shearing Optimization Kit - Low SDS (Diagenode, Denville, NJ) and sonicated in the Bioruptor Plus (Diagenode, Denville, NJ). Reactions were carried out with the SX-8G IP-Star® Compact System (Diagenode, Denville, NJ) using the CHIP_16_IPURE_100_I protocol and purified with the MinElute PCR purification kit (QIAGEN, Valencia, CA). ChIP
DNA was quantified using the Quant-iT PicoGreen ds DNA Assay Kit (Invitrogen, Carlsbad, CA) and 0.8 to 5.7 μl were used for library construction with the ThruPLEX™-FD Prep Kit (Rubicon Genomics, Ann Arbor, MI). Prepared ChIP-seq libraries were quantified with the Quant-iT PicoGreen ds DNA Assay Kit (Invitrogen, Carlsbad, CA) and the KAPA Library Quantification Kit (Kapa Biosystems, Wilmington, MA) using qPCR. Library quality control was performed with the Agilent Technologies 2100 Bioanalyzer and normalized based on qPCR, then pooled using the TruSeq SR Cluster Kit v3-cBot-HS (Illumina, San Diego, CA). Pooled samples and reference inputs were four-plexed and sequenced with the HiSeq 2500 v3 sequencer (Illumina, San Diego, CA) producing 20 to 56 million single–end 50 bp reads.

**RNA-Seq analysis.**

Raw reads were analyzed with FASTQC (v. 0.11.3) for quality control ([Supplementary Figure S2](#)) (36). Overrepresented (e.g. adapter and similar technical) sequences remaining in the raw reads were assessed and subsequently removed using the fastq-mcf module (v 1.04) under the ea-utils suite (37). Low quality base threshold was set at 25. Following technical sequence and low quality base removal using fastq-mcf, reads that were shorter than 20bp were filtered out. Transcript quantification was done using the GRCh38 reference genome by RSEM (v. 1.2.22) with default parameters, which utilizes Bowtie2 (38) for read alignment (39). RSEM uses Expectation Maximization (EM) to calculate the Maximum Likelihood (ML) estimates of expression (40). Differential expression analysis was done using EBSeq (v. 1.10), which uses an empirical Bayes model (41). Gene expression levels were subject to median normalization with differential
expression $\geq 0.95$ (a target false discovery rate of 5%) were considered as significantly differentially expressed.

For each comparison, EBSeq model convergence was manually verified, clustering of samples and/or genes was done using the Unweighted Pair Group Method with Arithmetic-mean method and Pearson’s correlation as the distance measure (42). The Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.7 (43) was used for functional analysis of the gene lists interrogating Biological Process (BP), Molecular Function (MF), and Cellular Component (CC) Gene Ontology (GO) categories (44) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (45). Biologically relevant categories that are over-represented in the gene set and therefore may be of further interest were assessed using the Expression Analysis Systematic Explorer (EASE) score in the DAVID tool. We picked GO categories that have EASE scores of 0.05 or lower as significantly over-represented.

**ChIP-seq analysis.**

Genomic sites enriched for the H3K27ac and H3K4me1 post-translational histone tail modifications common to one or more of the eight investigated ED-iPSC cell lines, were isolated and merged with a 200 bp window using Bedtools v2.17.0 (46, 47), providing with a total number of 124,297 and 233,972 intervals respectively. Coverage was estimated for each of these intervals using Bedtools v2.17.0 and normalized using SAMtools v1.1 (48). Final standardized coverage values for all samples and enriched genomic sites were plotted as a heatmap with the R package Pheatmap version 0.7.7 (http://www2.uaem.mx/r-mirror/web/packages/pheatmap/index.html) for the standardized
H3K27ac and H3K4me1 data respectively. GO biological processes related to differentially marked genomic sites were identified using GREAT (49).

Results

Ectoderm and endoderm multi-lineage differentiation potential of ED-iPSC lines.

To extend our previous teratoma analysis, biomarker profiles and \textit{in vitro} lineage commitment studies, we used established protocols to generate pyramidal neurons, astrocytes, and retinal pigment epithelium (RPE) of ectoderm origin, as well as pancreatic progenitors arising from endoderm differentiation (\textbf{Fig. 1, Fig. 2}). Cell types were chosen based on broad clinical interest in stem cell-derived cell-based therapies for neurodegeneration, injured brain and spinal cord, photoreceptor regeneration with RPE, and diabetes (24-27, 51, 52-54). All derivations used templated embryoid body (EB) intermediates of uniform size and dimensions (t-EB; \textbf{Fig. 1a}; 22, 23). Neural stem cell (NSC) progenitors were generated essentially as described (Neural Induction Media; StemCell Technologies, Vancouver, Canada). Typical rosette morphology was identified after 7 days of differentiation using brightfield microscopy, and the presence of Nestin in the rosette cells was confirmed by immunocytology (ICC; \textbf{Fig. 1b}). The NSCs were then used to derive astrocytes, and pyramidal neurons. Astrocytes were generated from NSCs by commercial media (Astrocyte Media; Sciencell Research Laboratories, Carlsbad, CA). Positive expression of CD44, GFAP, and Vimentin were monitored in maturing astrocytes (\textbf{Fig. 1b, Fig. 2a}). In immature dividing astrocytes Nestin and A2B5 were present but decrease over time and plateau with maturation as expected. ED-iPSC lines behaved
similarly during astrocyte differentiation (N>3 replicates). The implemented astrocyte protocol produces primarily CD44+/GFAP+ astrocytes.
Figure 1. Overview of differentiation strategies applied to ED-iPSC lines. Figure legend is on next page.
Figure 1. Overview of differentiation strategies applied to ED-iPSC lines. Multi-lineage differentiation to ectoderm, endoderm and mesoderm specialized cell types. (A) ED-iPSC lines were maintained as 2D feeder-free cultures, then differentiated from uniformly sized embryoid bodies (EB) formed in custom lithography patterned well arrays (200 µm t-EBs). The t-EBs were plated onto Matrigel or Vitronectin-XF for differentiation. Scale bars are 200 µm. (B) Early Ectoderm cell types. Differentiation of plated t-EBs to astrocytes, neurons, and retinal pigment epithelial (RPE) cells was done from neural stem cell (NSC) precursors. Neural rosettes appeared by day 7 upon addition of STEMdiff Neural Induction Medium. Early differentiating neural populations expressed intermediate filament markers Nestin (green; shown ED-iPSC F3.6.1) and Vimentin (green; shown ED-iPSC F3.5.2). Functional astrocytes expressed GFAP (green; shown ED-iPSC H3.3.1) and CD44 (green; shown ED-iPSC A2.2.2) by day 31 of differentiation. Differentiated neurons expressed β-III-tubulin (Tuj1; red; shown is ED-iPSC A2.2.2), MAP2 (red; shown is H9 hESC), VGlut1 (red; shown is ED-iPSC F3.5.2) and TBR1 (red; shown ED-iPSC F3.5.2). Representative differentiation to RPE cells (shown ED-iPSC F3.5.2), stained for Occludin (green) and ZO1 (red) tight junction markers showing cobblestone morphology. Scale bars are 50 µm. (C) Definitive endoderm, representative differentiation to pancreatic progenitor cells (shown ED-iPSC A2.2.2). Scale bars are 50 µm. (D) Mesoderm cell types. Representative differentiation to cardiomyocytes (shown is ED-iPSC A2.2.2) and smooth muscle cells (shown ED-iPSC F3.5.2). The hESC line WA09 (H9) when compared followed similar protocols. Scale bars are 50 µm.

Directed differentiation of pyramidal neurons from NSCs was done as previously described (24-26; Fig. 1b, Fig. 2b). Time-course and ICC revealed no differences in neuronal differentiation efficiency amongst ED-iPSC lines or when compared to hESC line WA09 (WiCell, Madison, WI) over multiple rounds (N>3; Fig 2b). The neural specific biomarker Tuj1 (β-III-tubulin) and neuronal specific MAP2, both present in mature neurons, were used to track the progress of differentiation along with two pyramidal specific biomarkers TBR1 and VGlut1 (Fig. 1b). During differentiation, the Tuj1 and MAP2 biomarkers appeared along with typical neuronal morphology of dendrites and elongated axons as expected. By ICC, 70-80% of the differentiated populations of ED-iPSCs were pyramidal neurons as revealed by specific biomarkers. A representative comparative ICC time-course of differentiation to pyramidal neurons for African-American, Hispanic-Latino and Asian ED-iPSC lines is shown (Fig. 2b). Differentiation into RPE cells from pluripotent t-EBs used established protocols (27, 28; Fig. 1b) with no differences observed among the ED-iPSC lines.
Differentiation of ED-iPSCs to pancreatic PDX1+/PAX6+ progenitors, which can generate β-islet cells to address pancreatic diseases, was done using established protocols (29-31). ICC was used to track PAX6 and the mature endoderm biomarker PDX1 (Fig. 1c, Fig. 2c). PAX6 localizes specifically to the perinuclear region consistent with maturing pancreatic cells (Fig. 2c, second panel). Similar staining was present across ED-iPSC lines and was comparable to differentiated hESC WA09 cells. In addition to the representative three lines shown, differentiation efficacy of ED-iPSC lines F3.6.1, A2.1.1 and H3.1.1 for astrocytes, pyramidal neurons, and pancreatic progenitors revealed no observable differences. RPE differentiation was tested on the three ED-iPSC lines shown in Fig. 2 (A2.2.2, F3.5.2, and H3.3.1). Our studies show no apparent differences between ED-iPSC lines examined and for differentiation towards evaluated ectodermal or endodermal cells.
Figure 2. Multi-lineage differentiation of ED-iPSC lines to astrocytes, pyramidal neurons and pancreatic progenitors. Figure legend is on next page.
Figure 2. Multi-lineage differentiation of ED-iPSC lines to astrocytes, pyramidal neurons and pancreatic progenitors. Comparative arrays of ED-iPSC lines A2.2.2, F3.5.2 and H3.3.1 showing representative images of differentiation to astrocytes, pyramidal neurons, and pancreatic progenitors following protocols outlined in the text, with detailed differentiation protocol schematics under each corresponding image sets. The hESC H9 control line was also compared for astrocytes and pyramidal neurons. Time frames shown all initiate from the pluripotent stage. (A) Astrocyte differentiation. Astrocyte-specific glial fibrillary acidic protein (GFAP) and surface marker CD44 are shown in images acquired at days 17, 20, 24 and 31 of differentiation. (B) Pyramidal neuron differentiation. Neuron biomarkers are β-III-tubulin (Tuj1, red) and microtubule associated protein 2 (MAP2, green) are shown in images acquired at days 31, 38, 45 and 52 of differentiation. (C) Early pancreatic progenitors. Expression of pancreatic-specific markers PAX6 and PDX1 are shown in images acquired at 14, 21 and 28 days of differentiation, along with a higher magnification image of PAX6 and PDX1 (day 21). Scale bars are 50 μm.

**ED-iPSC lines similarly generate smooth muscle but differ in ability to generate beating cardiomyocytes.**

Multiple cell types of clinical interest arise from mesoderm including cardiomyocytes, smooth muscle, and cells of hematopoietic lineage amongst others. We examined differentiation to smooth muscle and cardiomyocytes from mesoderm committed Brachyury-T positive lines A2.2.2, F3.5.2, and H3.3.1 (13; Fig.1d, Fig. 3a, 3b) using established protocols (32-35; Sciencell Research Laboratories, Carlsbad, CA) and compared to hESC line WA09 (H9, Fig. 3c). All ED-iPSC lines tested readily generated smooth muscle (Fig. 3a). In contrast, the ability to generate 3D aggregated cardiomyocytes with banded Troponin-T, required for beating activity, differed in the ED-iPSC lines (Fig. 3b). We observed functional differences within ethnic replicate line sets for ED-iPSC Asian lines (A2.1.1 beating versus A2.2.2 non-beating) and Hispanic-Latino lines (H3.1.1 beating versus H3.3.1 non-beating). The two African American ED-iPSC replicate lines examined revealed weak assembly of Troponin-T (F3.5.2) or diffuse non-banded Troponin-T (F3.6.1) in cardiomyocytes that did not beat. Cardiomyocyte maturation was followed over a time course of 28 days. ICC staining for the cardiac specific Troponin-T biomarker was generally found in regions of the 2D culture that have spontaneously assembled 3D cell clusters. Troponin-T in such 3D aggregated, beating
cardiomyocytes has typical banded striations, as observed in ED-iPSC lines A2.2.2 and H3.1.1 as well as hESC WA09 (H9). We also observed an intermediate Troponin-T phenotype (F3.5.2) and non-banded/diffuse Troponin-T (A2.1.1, F3.6.1, H3.3.1, H3.5.2), each with no beating. In summary, the ED-iPSC lines, though behaving similar in smooth muscle differentiation as well as ectoderm and endoderm differentiation, revealed differences in generating beating cardiomyocytes (**Fig. 3b**, N=3) that can be characterized by the inability to assemble 3D clusters with banded Troponin-T.
Figure 3. Mesoderm lineage differentiation of ED-iPSC lines to smooth muscle and cardiomyocytes. Comparative arrays of ED-iPSC lines A2.2.2, F3.5.2 and H3.3.1 showing representative images of differentiation to smooth muscle and cardiomyocytes, and compared to hESC H9 following protocols outlined in the text. Time frames shown all initiate from the pluripotent stage. Representative images are shown. (A) Smooth muscle. Cells are stained for Hoechst (blue), alpha smooth muscle actin (α-SMA, green). Acquired images show days 10, 14, 21, 28, and 35 of differentiation. (B) Cardiomyocytes. Expression of the cardiomyocyte-specific marker Troponin-T (green) is shown at days 14, 21, and 28. ED-iPSC lines with unpolymerized Troponin-T did not develop a beating phenotype. (C) Differentiation of the H9 hESC pluripotent line into smooth muscle and cardiomyocytes. (D) Detailed differentiation protocol schematics, top one illustrates smooth muscle differentiation, while bottom one shows cardiomyocyte generation protocol. Scale bars are 50 μm.
Cardiomyocyte differentiation protocols are sometimes optimized by varying GSK3 or IWP2 inhibitor concentrations, as well as time of exposure to these inhibitors (32, 33, 55). Optimization of GSK3 inhibition specifically has been demonstrated to assist Troponin-T banding and contractility. The differentiation protocol was optimized for each non-beating line (Fig. 4a) by systematically reducing the concentration of the GSK3 inhibitor CHIR99021 (Tocris, Bristol, UK) and time of exposure to 6 hours, from conditions optimal for hESCs (12 µM CHIR, 24-25 hrs) that were used in Fig 3b. The optimization was done in 2D cultures and then quantified in our custom 3D templated 100 well microarrays (Fig. 4, Supplementary Movies 1-3).
Figure 4. Chemical and physical optimization of ED-iPSC cardiomyocyte protocol. Optimization formation of banded Troponin-T in differentiated ED-iPSC cardiomyocytes was done in 2D cultures along with ability to form 3D aggregates using custom lithography microwells. Representative images are shown. (A) Optimizing cardiac differentiation of ED-iPSC lines A2.2.2, F3.5.2, and H3.3.1, and compared to the H9 hESC line. The GSK3 inhibitor (CHIR99021) optimal concentration and exposure time were individually determined for each line. Asterisk (*) denotes the optimal concentration of CHIR99021 exposure at 6 hours for each line. Detailed differentiation protocol schematic is shown under the image set. Scale bars are 50 μm. (B) Comparison between traditional 2D culture and 3D patterned culture of derived cardiomyocytes. Lithography templated microarrays efficiently and easily generate beating 3D cultures of cardiomyocytes. White arrows point to 3D clusters, while black arrows mark empty space within the microwell. (C) Derived ED-iPSC and H9 hESC cardiomyocytes by our optimized protocol retain banded Troponin-T expression and beating phenotype when seeded and grown within patterned microarray wells. (D) Recovered 3D cardiomyocyte culture from the F3.5.3 ED-iPSC line and stained for Hoechst (blue) and Troponin-T (red) shows the characteristic banded morphology typical of beating cardiomyocytes. (E) Histogram of loaded LTA-PDMS grid wells versus generated beating 3D cardiomyocyte clusters in representative ED-iPSC lines and the control H9 hESC line at optimized differentiation conditions. Scale bars are 200 μm.
To quantitatively compare ability of ED-iPSC lines to self-assemble in 3D and generate beating cardiomyocytes, we dissociated optimized cardiomyocyte 2D cell cultures into single cell suspensions and seeded custom microarray wells at high to low density (Fig. 4b, Materials and Methods). The microarray cultures were maintained for one week and 3D self-assembly and development of a beating phenotype was monitored by phase microscopy, followed by ICC confirmation of cardiac biomarkers (Fig. 4, Fig 5).

The importance of a 3D microarchitecture for beating and robust Troponin-T expression was evident in over-seeded microarrays in which a 2D cell layer was present between wells that did not stain for Troponin-T, while 3D cell aggregates remaining within individual wells stained for Troponin-T and exhibited banding with a beating phenotype (Fig. 4b). In summary, the microenvironment of the wells has several advantages, allowing us to compare in this high throughput platform multiple parameters that are 3D self-assembly and Troponin-T expression (Fig. 4b, 4c), Troponin-T banding in extruded clusters (Fig. 4d) along with other biomarkers (Fig. 5, NKX2.5 and GATA4), and monitoring of beating phenotype (Supplementary Movie 4: F3.5.2 derived cardiomyocytes). Similarly, we can monitor other cardiac biomarkers in the microarray clusters, such as NKX2.5 and GATA4 (Fig. 5.). We observed that optimization is most effective for ED-iPSC line F3.5.2 as expected, and of limited effect for H3.3.1 that displays diffuse Troponin-T by ICC (Fig. 3b). To further understand functional cardiomyocyte outcomes we applied comparative bioinformatics of ED-iPSCs transcriptomes (Fig. 6 and Fig. 7).
Figure 5. Cardiac biomarker expression of GATA4 and NKX2.5 in ED-iPSC derived cardiomyocytes in 2D and 3D cultures. Expression of early cardiogenic biomarker expression was evaluated in 2D and 3D culture, including templated 3D. Representative ED-iPSC lines A2.2.2, F3.5.2, and H3.3.1 stained for the cardiac specific markers NKX2.5 (red), GATA4 (green) and the Hoechst nuclear stain (blue) are shown under mixed 2D and 3D culture conditions or templated 3D culture in custom lithography generated PDMS microwells. Scale bars are 200 μm.
Transcriptome reprogramming of ED-iPSC lines is evident in gene expression analysis.

Jaccard analysis comparison between the ED-iPSC lines (Fig. 6a) based on enriched peaks for H3K27ac and H3K4me1 and analysis of differentially expressed genes allowing hierarchical clustering of ED-iPSCs versus parental fibroblasts (Fig. 6b) was done. The ED-iPSC lines are distinct from their initial fibroblasts with dramatically remodeled transcriptomes, suggesting complete reprogramming into an embryonic-like pluripotent phenotype. Raw RNA-Seq average read count and percent of high quality bases (bases with a quality score > 25) per sample were 36.85M and 97.39%, respectively. After trimming and filtering there were 36.73M reads and 98.12% high quality bases on average per sample (See Supplementary Fig. 2 – RNA-Seq average read count and quality of ED-iPSCs). For each of the initial fibroblast cell lines labeled A2 (Asian), A3 (Asian), F3 (African American), and H3 (Hispanic), corresponding ED-iPSC lines were generated in biological replicates, labeled (A2.1, A2.2), (A3.1, A3.3), (F3.5, F3.6), and (H3.1, H3.3).

RNA-Seq analysis generated expression data for 45,989 transcripts. The lowest 2% of the expression values, in all 12 samples (i.e. 2\textsuperscript{nd} percentile of 12x45,989 signal values), was considered as the background signal value. Genes with an expression value less than the background in all 12 samples were considered “undetected” and filtered out from the downstream analysis leaving a total of 34,439 genes. Hierarchical clustering of the samples using all 34,439 genes imply that the ED-iPSCs cluster together and are separate from parental fibroblasts (Fig. 6c). Two clusters of ED-iPSCs are present. ED-iPSC replicate lines, derived from the same ethnic parental fibroblast, Asian A2.1/A2.2 and Hispanic Latino H3.1/H3.3 fall within the same cluster group. In contrast, the
African American ED-iPSC lines F3.5.2/F3.6.1 or Asian lines A3.1.1/A3.3.1 are split between the two clusters. Based on our comprehensive functional analysis, ED-iPSC lines within either of the two clusters are effectively pluripotent.

**Two distinct epigenetic architectures support differentiation into beating cardiomyocytes.**

To further compare ED-iPSC line reprogramming at the epigenetic level, we performed ChIP-seq for histone modifications H3K4me1 and H3K27ac (Fig 6d). H3K4me1 is found at active and primed enhancers while H3K27ac is found at active enhancers and active promoters. To further characterize active and primed chromatin states the sites containing H3K4me1+/H3K27ac+ or H3K4me1+/H3K27ac- were evaluated respectively (Fig. 6e and Fig. 6f). Epigenomic analysis of the ED-iPSC lines, as with the transcriptome analysis, reflects the stochastic nature of the reprogramming process. Even within replicate ED-iPSC pairs derived from the same fibroblast source, the ED-iPSC lines do not always cluster together, whereas different ethnic ED-iPSC lines may cluster. For example, the African American ED-iPSC line F3.6.1 and Hispanic-Latino line H3.1.1 cluster together for both modifications H3K4me1 and H3K27ac, as well as overlap of these modifications, while the replicate lines F3.5.2 and H3.3.1 fall within a separate cluster. The grouping differences for F3.6.1 and H3.1.1 are driven in part by regulatory regions associated with genes involved in extracellular matrix (ECM) organization, collagen fibril organization and cardiovascular development.
Figure 6. Hierarchical clustering and epigenomic profiling of ED-iPSC and initial fibroblast lines. Figure legend is on next page.
Figure 6. Hierarchical clustering and epigenomic profiling of ED-iPSC and initial fibroblast lines. (A) Jaccard analysis comparison between the ED-iPSC lines based on enriched peaks for H3K27ac and H3K4me1 ChIP-seq data. The Jaccard index is defined as the ratio of the combined length of the genomic regions within both samples (intersect) and in at least one sample (union). The A2.1.1, A2.2.2, A3.1.1, A3.3.1, F3.5.2, F3.6.1, H3.1.1, and H3.3.1 represent the eight investigated ED-iPSC cell lines. (B) Heat map of hierarchical clustering using union of genes differentially expressed between ED-iPSCs and corresponding initial fibroblast cells. (C) Dendrogram showing the hierarchical clustering of replicate lines within ethnicity and between the ethnically diverse lines (D). Comparison of ED-iPSC at the epigenomic level by clustering ChIP-seq coverage at all genomic locations for H3K27ac and H3K4me1. The number of sites for each clustered group is shown with the red to blue color gradient represents coverage value expressed as reads per million (rpm) for each sample. Significantly enriched GO-Biological process were determined by GREAT and is listed below the heatmap. Enriched GO-Biological Processes were determined for active regions defined as H3K4me1+/H3K27ac+ (E) and for primed regions defined as H3K4me1+/H3K27ac- (F).

Although clustering differences are present, this does not reflect incomplete remodeling for pluripotency. ED-iPSC lines H3.1.1 as well as A2.2.2 for example, differentiated into all cell types tested in this study including beating cardiomyocytes, and all of these lines were previously validated by teratoma formation, qRT-PCR of pluripotency biomarkers, and tri-lineage commitment (13). These multiple epigenomic states, a result of stochastic complexity in reprogramming (17), identify genes and pathways as starting points for pluripotency and multi-lineage differentiation. Bioinformatic assessment strategies such as Pluritest (18) are beginning to be used to compare pluripotency profiles, however, we were unable to apply Pluritest here due to the current incompatibility of our HiSeq 2500 data files with that software. Our current study provides insights into both the impact of the reprogrammed pluripotent state as well as its effect on downstream multi-lineage differentiation.

Differential gene expression analysis and GO pathways defining beating and non-beating outcomes for Asian and Hispanic-Latino ED-iPSC replicate lines.

Significantly differentially expressed genes were found for the following comparisons: between initial fibroblasts, between ED-iPSCs, ED-iPSCs vs. initial
fibroblasts, and beating vs. non-beating ED-iPSC biological replicates (See Supplementary Data File 1 – Differentially expressed gene comparisons). The results are summarized in Supplementary Table 1. On average, the observed difference between the ED-iPSC cell lines was less than the difference between the initial fibroblast cells. The highest difference was observed between the ED-iPSCs and the corresponding fibroblast cells, as expected. The topology of sample clustering remained unchanged when the union of differentially expressed genes (n=8,503) between ED-iPSCs and corresponding fibroblasts was used (Fig. 6b).

Of particular interest were the replicate ED-iPSC lines in which one (A2.2.2 and H3.1.1) led to beating cells, while the other (A2.1.1 and H3.3.1) did not. We focused on in order to understand the genes and corresponding functional mechanisms that could explain the observed difference in obtaining functional beating cardiomyocytes we focused on the replicate ED-iPSC lines that behaved differently in that regards (Fig. 7). We compared genes that were differentially expressed between beating and non-beating ED-iPSC lines; A2.2.2 versus A2.1.1 and H3.1.1 versus H3.3.1. There were 354 genes up-regulated and 107 genes down-regulated in the beating ED-iPSC line compared to the non-beating replicate for the A2 line (Fig. 7a). We called these 461 genes (354+107) “beating-associated” genes for the A2 line. Similarly, there were 974 “beating-associated” genes for the H3 line. There were 51 genes (30+21) that showed concurrent direction of expression in the beating cells irrespective of the ethnicity of parental fibroblast. We labeled these as “ethnicity-independent” beating-associated genes. Figure 7a presents the heatmap of differentially expressed genes in the Venn diagrams, along with the cardiac related over-represented GO categories and pathways.
We identified functional categories that are uniquely or commonly overrepresented in A2 and H3 dependent beating-associated gene lists. The functional categories that are commonly overrepresented in ethnicity-dependent gene lists are likely to show non-overlapping active regions for these categories, as the identity and the change of expression (beating versus non-beating) of the genes in these lists are not the same. On the other hand, categories that are uniquely associated in one of the ethnicity-dependent beating-associated gene lists show the distinct mechanisms leading to the same phenotype in different ethnic lines. Functional groups I (461 genes) and III (974 genes) are beating-associated in the A2 or H3 lines, respectively, including up or down-regulated genes. Functional group II (51 genes) represents ethnic independent beating associated genes. We also analyzed differential expression to evaluate inversely regulated genes in comparing A2 versus H1, shown in Fig. 7b. Significantly overrepresented GO functional categories in all four gene lists can be found in Supplementary Data File 2. Finally, we generated Venn diagrams to compare overlap in identified GO pathways that included 380 Asian specific, 420 Hispanic-Latino specific, and 28 ethnicity independent categories (Fig. 7c). Our analysis is the first to evaluate differential gene expression pathways of generated functional cardiomyocytes with ethnic diversity contributions.
Figure 7. Comparative analysis of ED-iPSC beating and non-beating Asian and Hispanic-Latino replicate lines. Figure legend is on next page.
Figure 7. Comparative analysis of ED-iPSC beating and non-beating Asian and Hispanic-Latino replicate lines. (A) Genes concurrently up-/down-regulated between beating and non-beating lines in different ethnic groups are shown in the Venn diagrams. Heatmaps of the three different beating-associated gene lists are shown; I: A2 dependent, II: ethnicity independent, III: H3 dependent. Cardiac related GO categories that are uniquely overrepresented in the gene lists I, II, and III are shown next to each heatmap. Functional categories that are commonly overrepresented in A2 and H3 dependent beating associated gene lists are also shown. Sample genes in the overrepresented categories are listed in boxes next to the category graphs. (B) Analysis of oppositely regulated genes between the A2 and H3 lines, with the attendant heatmap, GO categories and representative gene list. (C) Venn diagram comparing overlap in identified GO categories.

Discussion

Successful therapeutic intervention in cardiac health is influenced by ethnic diversity, however stem cell resources used to generate and evaluate cardiomyocytes for cardiac therapies have typically been limited in ethnic diversity (56, 57). Previously we derived ethnically diverse (ED) iPSCs from fibroblasts sources obtained from self-designated African American, Hispanic-Latino and Asian ethnicities, establishing a pluripotent resource evaluated by teratoma formation, qRT-PCR verified gene pluripotency profiles and multi-lineage differentiation commitment (13). Here by multi-lineage differentiation of the ED-iPSCs to ectoderm, mesoderm and endoderm specific cells of potential therapeutic interest, we demonstrate uniform differentiation potential with the exception of functional cardiomyocytes. To benefit cell therapeutic goals for cardiac health that use iPSC derived cardiomyocytes we applied bioinformatics analysis to RNA-Seq data of Asian and Hispanic-Latino ED-iPSCs to identify potential predictors of beating outcomes. Our work identifies differentially expressed genes and GO pathways shared between ethnicities, across ethnic comparisons, or reverse regulated across ethnicities.

The ED-iPSC replicate lines differentiate similarly into ectoderm derived pyramidal neurons, astrocytes, and RPE cells as well as endoderm derived pancreatic precursors. Following commitment to Brachyury+ mesoderm, cells similarly formed smooth muscle
and in the cardiac lineage expressed transcription factors Nkx2.5 and Gata4, but differed in ability to form 3D beating aggregates in custom microarray wells and express banded (striated) troponin-T. Our bioinformatic analysis focused on the Hispanic-Latino (H) and Asian (A) replicate lines H3.1.1/H3.3.3 and A2.2.2/A2.1.1 as beater/non-beater pairs. Since cardiac differentiation is driven through regulation of TGF-β signaling mediated by Nodal, Activin and BMP ligands (32, 33, 55, 58) we evaluated genes in this pathway. In the A2.2.2/A2.1.1 beating/non-beating pair Nodal (2.0X) and BMP10 (8X) were upregulated in ED-iPSCs for beating cells. Nodal is also up-regulated in the H3.1.1 beating versus H3.3.1 non-beating ED-iPSC lines. In contrast, BMP10 (~6X) was elevated in the non-beating H3.3.1 ED-iPSC line versus beating H3.1.1 line. BMP10 is implicated in trabeculation, or 3D remodeling, of the embryonic heart and early cardiomyocyte growth and survival but its levels are not an effective predictor of cardiomyocyte beating across both of these ethnicities. The HAND1 transcription factor is important for morphogenic changes of cardiac looping, chamber septation and ventricular development and its misregulation is implicated in cardiomyopathies (59). HAND1 is elevated across ethnicities in beating A2.2.2 (~2.6X) and H3.1.1 (~2X) ED-iPSC lines versus non-beating A2.1.1, or H3.1.1 lines. Since its expression helps to maintain the differentiating cardiac precursor state, it must be later down-regulated to drive terminal cardiomyocyte differentiation. MYL4 which encodes the Myosin Light Chain 4 protein is present in embryonic muscle and regulates muscle contractions through cyclic intracellular Ca^{2+} dependent interactions with actin-rich thin filaments. MYL4 is upregulated in both beating lines H3.1.1 (~6X) and A2.2.2 (~5X) as compared to the non-beating H3.3.1 and A2.1.1 lines. In addition to NODAL, HAND1, and MYL4, other genes that were similarly up-regulated in beaters
versus non-beaters include ACTC1, COL11A1, APOA4, APOB, and MYL7, MYH14, and GUCY1A3 and downregulation of CNTN4. Inversely regulated genes that here were not good predictors of cardiomyocyte beating outcomes included BMP10, SERPINE1, GATA4, COL8A1, and CILP amongst others.

The gene ontology, GO, functional categories represented in each ethnic beating/non-beating pair, shown in Figure 7, reveal different emphasis. For example, Group I Asian A2 includes early development and differentiation, while Group III, Hispanic-Latino, includes later cardiac tissue development. Shared genes distinguishing beaters from non-beaters across these ethnicities, referred to as ethnic independent groups II and IV in Figure 7, include cellular infrastructure categories such as cytoskeletal remodeling, cell migration and transport, morphogenesis, focal adhesion, matrix remodeling and ECM composition, as well as functional myosin complex generation. The inability to readily form 3D aggregates and Troponin-T banding correlates with non-beating activity in our experimental analysis. In this regard GO categories and genes involved in 3D tissue structure such as ECM remodeling, ECM-receptor interactions, actin binding and actin filament movement pathways were of interest. Cardiogenesis is known to have a strong requirement for cell-cell interactions and ECM remodeling, which can be enhanced by engineered devices and control of 3D culture formation (60-65). A pathway implicated in mouse neonatal heart maturation (66) is the cholesterol transporter pathway, which is upregulated across ethnicities in our ED-iPSC Asian and Hispanic-Latino lines. Of well-studied canonical pathways in cardiogenesis, such as cardiac cell differentiation, heart looping, and heart morphogenesis, these were upregulated in the beating lines, while other pathways such as BMP regulation, tube development, cardiac muscle growth and cardiac tissue development categories included inversely regulated genes across
ethnicities when comparing beating/nonbeating. The functional pathways revealed in the ED-iPSCs represent differences in staging preceding differentiation to cardiomyocytes.

An increasing number of studies over the last decade are using genome wide association studies (GWAS) to identify risk factors in ethnic populations (67), including African Americans (68), Hispanic-Latino (69), and Europeans and South Asians (70). In addition, gene-disease associations include β1-adrenergic receptor polymorphism linked to β-blocker response (71), proteasome subunit a6, PSMA6, linked to myocardial infarctions in Japanese (72) and methylthioadenylsine phosphorylase, MTAP, and cyclin-dependent kinase inhibitor 2B, CDKN2B, linked to heart disease in Chinese Han populations (73). Currently over 40 putative risk factors have been compiled from numerous recent studies (74-80), that include the following genes: PPAP2B, ANKS1A, TCF21, ZC3HC1, ABO, CYP17A1, CNNM2, NT5C2, ZNF259, COL4A1, COL4A2, HHIPL1, CYP46A1, ADAMTS7, RALL, PEMT, RAD51, SRR, UBE2Z, SMCR3, PSRC1, CELSR2, SORT1, CDKN2B, SMARCA4, MRPS6, CYP17A1, AS3MT, APOA1, APOC3, APOA4, APOA5, CHRNA3, CHRNA5, SH2B3, ATXN2, TRAFD1, SMG6, TSR1, SGSM2, ABCC8, INS, MTAP, PSMA6, ADRB1, IDDM19, IDDM2, C1LA4, SLC41A2, MRS2, KCNJ11, CNNM1, NIPA1, CALPN10, PPARA, and CYP11B2. Of these genes, those differentially expressed between our beaters/nonbeaters include: PEMT, CDKN2B, APOA1, in Asian lines and APOA4, IDDM19, PPARA and ABCC8 in Hispanic-Latino lines. While it is unclear yet the extent to which stem cell based transcriptome studies such as this one, on small populations of cells may help to provide additional indicators, the analysis is expected to be informative for cardiomyocyte-relevant injuries such as in myocardial infarction (81). Finally, cardiac linked disease pathways of diabetes (Type 1
and Type 2) and cholesterol metabolism that are observed in larger GWAS studies, are also represented in GO pathways in our analysis of genes differentially expressed in beaters/non-beaters.

**Chapter Conclusions**

This study extends characterization of high quality ED-iPSC replicate lines generated and characterized under a common platform (13, 22, 23) with additional comprehensive comparative bioinformatics analysis of epigenome and transcriptome data along with optimized multi-lineage cell differentiation studies. Our generated ED-iPSC lines, RNA-Seq and ChIP-seq datasets, and bioinformatics analysis represent a new ethnically diverse resource for continued analysis by the stem cell biomedical community and may find additional usefulness in bioengineered platforms for ethnically relevant drug testing and metabolomics.
Chapter References:


Supplementary Table 1. List of differentially expressed genes. Row $i$, column $j$ represents genes up-regulated in sample $j$ compared to sample $i$.

<table>
<thead>
<tr>
<th></th>
<th>A2</th>
<th>A3</th>
<th>F3</th>
<th>H3</th>
<th>A2.1.1</th>
<th>A2.2.2</th>
<th>A3.1.1</th>
<th>A3.3.1</th>
<th>F3.5.2</th>
<th>F3.6.1</th>
<th>H3.1.1</th>
<th>H3.3.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2</td>
<td>446</td>
<td>578</td>
<td>665</td>
<td>2122</td>
<td>2617</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td>536</td>
<td>204</td>
<td>586</td>
<td>2199</td>
<td>2584</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F3</td>
<td>569</td>
<td>241</td>
<td>564</td>
<td>2174</td>
<td>2026</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H3</td>
<td>892</td>
<td>795</td>
<td>793</td>
<td>2537</td>
<td>2664</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2.1.1</td>
<td>2431</td>
<td></td>
<td>384</td>
<td>282</td>
<td>87</td>
<td>2154</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2.2.2</td>
<td>2106</td>
<td></td>
<td>128</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3.1.1</td>
<td>2382</td>
<td></td>
<td></td>
<td>481</td>
<td>109</td>
<td>697</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3.3.1</td>
<td>1782</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>F3.5.2</td>
<td>2159</td>
<td></td>
<td></td>
<td>242</td>
<td>170</td>
<td></td>
<td>702</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F3.6.1</td>
<td>1727</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>H3.1.1</td>
<td>1777</td>
<td></td>
<td></td>
<td>1882</td>
<td>377</td>
<td>226</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>565</td>
</tr>
<tr>
<td>H3.3.1</td>
<td>1767</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>480</td>
</tr>
</tbody>
</table>
Supplementary Figure S1. ED-iPSC lines differentiation protocols. Detailed protocol outlines used to generate (A) astrocytes, (B) pyramidal neurons, (C) retinal pigment epithelial (RPE) cells, (D) pancreatic progenitors, (E) smooth muscle cells, and (F) cardiomyocytes, either using the standard protocol, or the ED-iPSC optimized one.
Supplementary Figure S2. RNA-Seq average read Counts and quality of ED-iPSCs raw data. Analysis of ED-iPSC read counts and raw output show high quality RNA-Seq data.

Supplementary Data File 1. Differentially expressed gene list.

Differentially expressed genes in fibroblast parent lines compared to ED-iPSCs, ED-iPSCs vs. initial fibroblasts, and beating vs. non-beating ED-iPSC biological replicates.
Supplementary Data File 2. Go and pathway lists.
Full list of differentially expressed GO terms and pathways in the ED-iPSC replicate lines.

Supplementary Movie 1: Mixed cardiomyocyte culture with beating and non-beating regions.
H9 hESC derived cardiomyocytes grown on Matrigel coated dish show random non-beating 2D regions and beating 3D self-assembled regions within the same culture by day 15. Movie is at 1X speed, phase acquisition.

Supplementary Movie 2: Multi-well array of aggregated beating cardiomyocytes.
High density seeded H9 hESC 2D culture derived cardiomyocytes were dissociated into single cells and seeded within the microwell array. Cells form 3D beating clusters when allowed to self-assemble within patterned microwells by day 5. Movie is at 1X speed, phase acquisition. Well diameter is 200 μm. Beating clusters we observed in 396/400 wells in three separate experiments.

Supplementary Movie 3: Zoom-in of multi-well array of aggregated beating cardiomyocytes.
High density seeded H9 hESC 2D culture derived cardiomyocytes were dissociated into single cells and seeded within the microwell array. The cells preferentially form 3D beating clusters compared to inter-well areas, when allowed to self-assemble within patterned microwells by day 5. Movie is at 1X speed, phase acquisition. Well diameter is 200 μm. Beating clusters we observed in 790/800 wells spread over three separate experiments.
Supplementary Movie 4: Aggregated cardiomyocytes beat in microwell arrays.

2D culture F3.5.2 derived cardiomyocytes were dissociated into single cells and loaded into a microwell array. Cells preferentially self-assemble into beating 3D clusters in templated wells within 5 days. Movie is at 1X speed, phase acquisition. Well diameter is 200 μm.
Chapter 6:

Thesis Overview

Conclusions

Pluripotent stem cells hold tremendous potential for advancement of medical treatments that will enhance the quality of human life. To fulfill this promise, several key gaps must be addressed. There is a need for high quality stem cell lines of ethnic diversity, an expanded bioinformatic understanding of gene expression during development, and for innovative advances to generate and analyze complex multicellular architectures as \emph{in vitro} models to biomimic tissue functions. My research in human stem cell biology addresses current deficiencies in each of these important areas.

I morphologically and functionally characterized new ethnically diverse (ED) induced pluripotent stem cell (iPSC) lines by multi-lineage differentiation and contributed to generation of whole genome bioinformatics analysis for the panel of ED-iPSC lines. In Aims 2 and 3, I focused on biomimicry of cardiomyocyte tissue microarchitectures during cardiogenesis and for neuron-glia interactions during cortical development. These goals combined human stem cell biology, tissue specific functional assays, bioinformatics, lithography generated devices and bioengineering through use low and high throughput platforms to generate a highly detailed profile, including optimized parameters for differentiation, of each evaluated ED-iPSC line that would be useful in future differentiation-based projects and for organoid formation. Outside collaborative efforts
have been important to this work and include Dr. Jose Cibelli, Michigan State University – Reprogramming, Dr. Michael Buck, SUNY Buffalo – Bioinformatics, and Dr. Hasan Out, University of Nebraska-Lincoln – Bioinformatics.

**Specific Aim 1:** Evaluation of pluripotency in newly derived ethnically diverse pluripotent stem cells as a biomedical resource. Through a collaborative effort iPSC lines were derived by reprogramming human fibroblasts from ethnically diverse donors. My analysis of these lines for pluripotency included *in vitro* multi-lineage commitment and terminal differentiation assays with functional and cell biological analysis. This work required optimization of established protocols, bioinformatics analysis, and both 2D cultures and 3D microenvironments using lithography-templated devices. The human embryonic stem cell (hESC) line WA09, served as a differentiation control and a benchmark to compare efficacy of differentiation protocols during optimization. In conclusion, I have generated comprehensively characterized high quality ED-iPSCs to use as a biomedical resource along with datasets on whole genome, transcriptome and epigenome and multi-lineage in vitro differentiation.

**Specific Aim 2:** Determination of key morphological and transcriptional parameters in human cardiomyocyte tissue engineering including ethnicity. Ethnic genetic background is known to contribute to cardiac disease and treatment efficacy. Cardiomyocyte protocols validated in Aim 1 were used or optimized to understand underlying biological, bioinformatics and microenvironment contributions to cardiomyocyte formation in 2D colonies, or 3D microarchitectures. This included mesoderm commitment and later cardiac-specific markers expression, as well as
acquisition of functional features of such as banded tropomyosin, cardiac-specific transcription regulators, and beating. There is still a question whether the cardiomyocytes generated in this work are fetal, or fully mature, which can have impact on the translational projects they would be used for in the future. For example, it is known that contractility and electrophysiology differs between fetal and adult cardiomyocyte, which may cause arrhythmia if implanted in a patient. While finding out their nature would be important, either type are valuable for research and can be used to advance biomedical projects and drug screening. To conclude, I have advanced the cardiomyocyte field by analysis of contractility and identification of genes providing predictive information on beating function across ethnicities.

**Specific Aim 3: Biomimicry of the human cerebral cortex microarchitecture during development and aging.** Minimal regional models of neural tissue, based on stem cell derived neurons and support glia cells were generated and used to analyze neuron-glia interactions in controlled and engineered microenvironments. Both 2D co-cultures and 3D microarchitectures will be generated and apply biocompatible microstrands, lithography-templated microarrays to drive self-assembly, assisted self-assembly, or directed patterning to study neural cortical development and plasticity. I have established new methodologies for patterning neural cortical microarchitectures of ED-iPSC derived neurons and astrocyte co-cultures for cell-based applications in CNS related disease, aging, or trauma.
Future Directions

Optimize and elaborate on 3D culture devices and the differentiation platform to biomimic target organ features, which will improve tissue analog generation and in vivo compatibility. A possible approach would be to re-engineer the loading covers for the LTA-PDMS patterning platform by adding a fluidic channel that would allow for cells and 3D cultures to float directly into the patterning holes, rather than relying on somewhat random loading probability in the current generation of devices. This will improve loading efficiency and potentially reduce wasted cell cultures and 3D cultures that is currently an issue that needs to be addressed to fully load the devices in a more reliable and cost-effective manner. Coupling a microfluidic solution with bio-functionalization of the surface will likely be needed in the next generation of devices.

Careful balancing of functional coatings would be very valuable when it comes to generating tunable surfaces that would be an active part of the device purpose, be it cell attachment, differentiation, or evaluation of cell and tissue state. The logical next step would be improve on the current state of the platform, where the engineered aspects mostly function to allow passive cell adhesion, while differentiation and other desired functional changes are left to more classical methods such as soluble small molecules and direct cell-cell interactions with limited directed input from the user.

An exciting future direction to elaborate on engineered devices would be ones that are specifically tailored to the tissue analog, based on analysis of the actual organ or tissue. This approach to functionalize the surfaces with cell-specific factors was tested (Fig. 1a-c) with the astrocyte specific surface marker CD44 covalently patterned onto the surface, then the areas between the lines was passivated to prevent cell attachment.
Expanding on such engineered solutions involving nano-scale and micro-scale fabrication at CNSE would make use of the world-class expertise available in multiple areas, such as rapid prototyping, polymer chemistry, lithography-inspired design, bio-conjugation, high-throughput patterning, etc., that need to come together in such next-generation devices. Figure 1 illustrates the workflow for a next generation device that would need to incorporate many of the techniques listed above. Starting with a concept (Fig. 1a), through test structure evaluation (Fig. 1b) and finally onto a proof-of-principle (Fig. 1c) demonstration of the engineered device, including how controlled integration with the ED-iPSC derived neural cell cultures could be achieved. Fig. 1d shows some of the promising devices that can utilize such engineered approaches towards creating a guided organ development, for instance towards a cortical layered structure in vitro.

Use ED-iPSC derived cardiomyocytes, neurons, and other differentiated cell types to study development and diseases, with a focus on ethnic-specific conditions based on single nucleotide polymorphisms (SNPs). Recent reports and drug studies suggest that there are strong correlations between epigenetic background and susceptibility to disease. Now that the ED-iPSC panel of lines has been characterized fully, there are several exciting approaches that the lines and the attendant bioinformatics data can be used to study cardiomyopathies and neurodegenerative disease in particular that are influenced by SNPs. If the ED-iPSC lines have the SNPs that have been implicated in certain conditions, then differentiating them into cell lines of that organ, specifically in regards to cardiomyopathies, neurodegenerative diseases such as Alzheimer’s and Parkinson’s, macular degeneration, and glaucoma-related cell degeneration will generate cells that are predisposed to express the pathology associated with the disease. This will open up the possibility to generate disease models for
ethnically-associated conditions. Additionally, if the cells happen to have the more common healthy phenotype, then we can use the CRISPR and TALEN methods to induce the disease-associated SNP into the line, thus generating clonal line pair for that disease and the healthy control, which would be even more useful when it comes to studying genetic-based diseases or modelling the organ conditions in the disease state, both with differentiated cells from the major affected organ, but also by differentiating cell cultures that are found in organs that are more tangentially related to the disease, but could still be impacted.
**Figure 1. Strategies to generate next-generation hybrid neural cell devices.** (a) Antibody-based platform for cell patterning, based on antibody-antigen specificity, covalently attached to surfaces that have been CVD silanized for specific antibody orientation. (b) Visualization of patterned lines and dots made by micro-contact printing of CD44 antibody. Cells in the last two images do not recognize the constant region that is facing up, so they do not pattern. (c) CD44 antibody patterned in lines onto glass surface with the variable region facing up allows astrocyte cells to recognize it and preferentially attach and align along the lines. Surfaces between the lines was passivated with different compounds known to prevent cell attachment. (d) Possible engineering solutions that can be combined with micro contact printing and other means to functionalize surfaces in future projects involving neural and other co-cultures.
Develop cardiac organoids, utilizing micro- and nano-engineered scaffolds and devices to improve cell differentiation and functionality done in-line with bioinformatics and electrophysiology analysis. With self-organization likely to play a key role in terminal cardiac differentiation, an improved engineered answer that standardizes 3D formation during cardiogenesis would further improve cardiomyocyte generation in vitro and allow us to study early developmental processes and simulate cardiomyopathies such as infarcts, valve damage, etc. Bioinformatics analysis and electrophysiology measurements on cardiac cells derived from ED-iPSC lines would add to our knowledge of cardiogenesis and expand on the factors that are involved in this complex process. Comparing the epigenetic state of pluripotent ED-iPSC lines and at key points during differentiation would build up a more complete bioinformatics picture of the process, which will help with evaluating cell lineage commitment and can further optimize differentiation protocols, based on the unique background of a newly generated iPSC line.

Introducing electrophysiology measurements on the ED-iPSC derived cardiomyocytes would add another layer of evaluation of pluripotency to the already extensive array of characterization tools that the panel of lines was put through. Developing the ability to do functional measurements on cardiomyocytes derived in the lab would also open the door to do further, more involved projects that would apply our unique pluripotent lines to functional tissue engineering, drug evaluation, and disease studies.

Rapid prototyping of cardiac devices to simulate architecture in the heart could also be coupled with a fabricated platform that incorporates electrophysiological readout within the design. We can leverage our optimized techniques to generate functional cardiomyocyte cell populations from our ED-iPSC lines to incorporate cell culture with
engineered platforms to design *in vitro* models of *in vivo* cardiac architecture that are present within the heart, such chambers, valves, and tissue structure. This will be helped by the fact that cardiac cultures readily form 3D cell construct when given the right conditions as part of their differentiation pathway *in vitro*. Having the high quality cardiomyocyte generation capability then only leaves engineering the non-biological component of the hybrid device. For example by providing micro-strands onto which cardiomyocytes can attach, simulating muscle fibers, or 3D heart tissue construct that would mimic cardiac tissue *in vivo*. Another approach would be incorporate cell culture with a model of a heart valve scaffold to generate an *in vitro* model of a functional valve for use in drug testing, or disease modeling involved with cardiomyopathies.

An example of rapid prototyping for a potential heart valve hybrid device is illustrated in **Figure 2**. Starting with a 3D concept design (**Fig. 2a**) of the device and the required moulds to create it. Once the 3D concept of the mould has been finalized, 3D printed prototypes can be generated very rapidly (**Fig. 2b**), and then the actual engineered device can be cast out of PDMS (**Fig. 2c**) and prepared for functionalization and eventual integration with the cell component of the platform. Any modifications based on experimental feedback from experiments would be easy to incorporate due to the extremely fast engineering and production speed of the process flow, which can be as low as 1.5-2 days between iterations. This makes is very easy and quick to iterate devices until the most appropriate design and cell integration is achieved for the stated purpose, be it as a novel drug testing platform, or for an *in vitro* cardiomyopathy disease model.
Figure 2. Rapid Prototype Process for Engineered in Vitro Heart Valve Hybrid Device. (a) 3D design for a hybrid device heart valve (top) and 3D design for a two piece mould adapted from the concept. (b) Rapid prototyping of multiple generations of the mould using a stereolithography 3D printer. (c) Heart valve compared to PDMS casting of heart valve scaffolds from three generations of devices. The whole process from 3D concept through fabrication to PDMS scaffold ready to be seeded with cells takes on average 2 days per iteration. Scale bars are 1000µm.
Bioengineered devices to simulate ED-iPSC derived retinal pigment epithelium (RPE) layer assembly and model diseases such as macular degeneration in vitro. An exciting aspect in the development of characterization protocols was the ability to differentiate multiple tissue specific cell populations, one of which were RPE cells. RPE cells arise from cells committed to the ectoderm lineage during development, but they are separate from the neural stem cell progenitor lineage, which gives rise to neurons and astrocytes, the other two cell types that were generated in this project. As such RPEs are valuable to show that our derived NSC cultures are both stable and pure, as they did not give rise to RPEs when under differentiation conditions. On their own, RPE cells derived from our ED-iPSC lines could also be used to study ethnicity-dependent disease progression and serve as a platform to evaluate the efficacy of novel drugs on a wider range of ethnic and epigenetic backgrounds.

Unfortunately, current protocols take a significant time to generate RPE cells, up to 4 months for pure cultures (Fig. 3a). In order to make them more viable as a research tool, then, we need to improve their generation and maintenance by reducing the differentiation time (Fig. 3b), while maintaining high purity and the morphology and gene expression typical for this cell type. Functional RPE cultures have a unique requirement to maintain a specific morphology (Fig. 3c), and expression of tight junction markers, due to their function in vivo. If one of those parameters is not lost, then the culture is likely not going to be very effective as an evaluation tool, or as a tissue transplant source. Implementing the suggested protocol modifications, and adding engineered platforms that simulate more faithfully the specific ECM of this tissue would improve RPE generation and provide a better platform for novel drug testing and disease simulation in vitro.
**Figure 3. Differentiation of ED-iPSC lines into retinal pigment epithelium (RPE) cells.** (a) Schematic of the differentiation protocol that was used. (b) Modifications and optimizations that were done to improve the protocol efficacy and generate RPE cells that were xeno-free. (c) First image shows a brightfield section of a confluent cell sheet, showing the typical tightly packed cuboidal structure associated with RPE cells. ICC images show staining for the tight junction markers expressed in RPEs. Scale bars are 200μm in the brightfield image and 100μm in the ICC images.
Jointly develop 3D cultures and devices to improve cell differentiation and functionality in-line with bioinformatics and electrophysiology towards layered cortical CNS constructs for drug testing, disease modeling, epigenetic-optimized care and biomimetic machine learning. Our preliminary results suggest that pure 2D and 3D cultures of neurons likely won’t generate desirable neural network complexity without a support cell population, much as is the case in vivo (Fig. 4a). Thus, we needed to co-culture them with support cells, such as astrocytes, to promote cell survival and neural network formation. Functionalized scaffolds and surfaces to promote neural support cell attachment to promote neuron survival and patterning will likely be critical in mimicking cortical tissue, specifically the complex layers in vitro. Achieving this will open up exciting avenues for future projects where bioengineered scaffolds can be used to introduce a further layer of control over the derived neural co-cultures for a myriad of purposes.

Neurodegenerative diseases such as Alzheimer’s and Parkinson’s have been tracked to neuronal death and pose as attractive targets to develop in vitro models that can be generated with the help of bioengineered cell-device interface. Lithography-based devices offer an unparalleled control over the features that can be produced, ranging from the millimeter and larger scale all the way down to the sub-micron ranges. Our ability to introduce specifically patterned proteins within the devices can be leveraged towards creating an environment that closely simulates the conditions in the diseased brain, in regards to the protein expression within the ECM matrix surrounding the neural tissue culture analog. Additionally, the cortical tissue analogs can be used to simulate in vitro brain response to next-generation neural prosthetics, such as those for Parkinson’s, which
rely on deep brain stimulation to delay, or reverse, the effects of the disease without the need to use animal models in the early stages of clinical testing.

Cortical architecture mimicry is an important step towards developing an *in vitro* model of the tissue to be used as a platform for novel drug testing and tissue models of disease. Unfortunately, derivation of functional neurons is ineffective and takes several months. We were able to cut about two weeks off the differentiation time by starting with stable neural stem cell progenitors (NSCs). Within the stem cell field, it is unknown yet whether protocols that are shorter compared to natural development length should be any less relevant or useful. Further, if we generate templated neurospheres prior to differentiation, this further improves neuronal generation within the formed 3D cultures. In general, the requirement for self-organization during differentiation appears to be common and essential within the tissue analogs that replicate complex organs, such as the cortex and the heart. Thus improved engineered approaches not only will eliminate variability in complex tissue analog generation, but also promote desired alignment within the tissue, based on the high degree of control that micro-engineered and functionalized scaffolds could provide.

For instance, aligned micro-strands (*Fig. 4b and 4c*) can be used to simulate the seven layer architecture in the cortex, by promoting support cells migration and neuron attachment to glial support infrastructure. LTA-PDMS based next generation microfluidic devices can be designed to establish long term 3D neuronal cultures to allow them time to self-assemble into a layered cortical-like structure. Currently impeller-based bioreactors are used to stimulate formation of 3D aggregates for EBs or for neural layered cortical organoids. Whether or in what ways mechanical shear stress benefits, or hinders cortical assembly is not yet known. Micro-engineered devices would be a useful addition to the
currently available technologies, due to the relatively less reagents that would be used to generate the 3D tissue analogs. They would also provide a tunable approach that would allow for direct control over the conditions that the cells are exposed to, which in turn would allow such devices to more easily simulate various brain environments, from healthy to specific diseased ones, such as Alzheimer’s or stroke, and from early on in development to late, thus also allowing for development of hybrid devices to model aging of the brain cortex under different conditions.

High-throughput analysis of neural networks would also be possible, given the proper engineering approaches, combined with optimized ED-iPSC derived neural cell populations of neurons and support cells. Engineered surfaces and microfluidics-based MEMs devices (Fig. 1d) would allow for such as neuron to neuron communication and neuron to support cells interactions at the multiple levels, from the whole tissue analysis to the interactions between several cells, and potentially as fine as the single cell and sub-cell levels. Precise patterning of proteins such as neurotransmitters, cell surface factors, or attraction and repellant factors that have been identified to be critical for neuronal network formation and pruning during maturation in vivo would allow creation-by-design neural populations mimicking the cortical development in vitro to be generated with high fidelity in a relatively easy manner. This would have implications in developing novel drug testing platforms, elucidate brain development and neuronal lifecycle, as well as help with the design of novel fields such as brain mimicry in chip design and machine learning.
Figure 4. Engineered solutions towards tunable neural co-culture to simulate the cortex in vitro. (a) Comparison between ED-iSPC derived cortical neurons with and without support cells shows the need for co-culture in neuronal survival. Middle image shows neuron-only culture tends to self-aggregate when grown in 3D. Bottom image shows Astrocyte cells preferentially interact with microfibers that simulate ECM arrangements in vivo. Scale bars are 100μm. (b) Co-culturing with support cells promotes long-term neuron survival, function, and interactions (white arrows) that mimic the in vivo state in the cortex. Scale bars are 50μm. (c) Engineered device shows proof of principle for directed neural co-culture by using aligned microfibers to simulate cortical ECM. Neurons and support astrocyte cells align on the microfibers in morphology that is very close to what the cells look like in vivo. Scale bars are 1000μm.
**Discussion**

Predicting and understanding how a specific patient would react to a life critical drug or a regenerative treatment would be one of the single most important steps towards achieving true personalized healthcare. The goal is to not only treat the patient more effectively, but also to cut down on current medical costs. Addressing the need to have diverse stem cell lines available would be both beneficial for biomedical research, and would also shed light on the impact of epigenetic variance in human populations as it relates to disease treatments. To truly validate *in vitro* models relative to a diverse human population, we would also need to be able to evaluate these models with diverse stem cell lines.

Adding to their importance in regenerative medicine and patient-specific treatments, ED-iPSCs would also reduce the reliance on animal models in the initial stages of drug discovery. According to the FDA, currently the vast majority of candidate drugs fail at the *in vivo* test stages. This is mostly due to toxicity or adverse side effects, and is seen even if they had passed the pre-clinical stages with flying colors. Developing more appropriate platforms for those initial tests before the animal test stage would save money, time, and be more ethical. In order to overcome the gap between *in vitro* and *in vivo*, a combined approach that incorporates functional differentiation with engineered solutions that take into account tissue- and system-specific multicellular mechanical, physical, and chemical parameters would be needed.

Even the most “simple” organs in Nature are marvels of engineering. The intricate associations between cell types, ECM, environmental cues, responses to stimuli, plasticity, adaptation, and more are essential. Ignoring any aspect of these would at best
slow down progress due to erroneous or incomplete findings, or at worst it will fail to accurately generate the desired easy to use and high throughput organ-on-a-chip test beds that would be critical in next-generation biomedical research, drug development and regenerative healthcare. Engineering knowledge and intricate systems design has long been the bedrock of semiconductor research and development. They have fueled the current revolution in electronics and to a large degree are responsible for modern society. If anything manmade could come even close to the complexity of our bodies and organs, it would be without a doubt the FET-based 3D inteconnect chip.

Leveraging know-how in design, optimization, and consistent improvements to aid in the next generation of biological advances, especially combined with high quality bioinformatics available, would translate the success of the semiconductor field into making the next big opportunity possible, which should be the biomedical revolution. In this thesis, I combine engineered devices with classical stem cell biology to achieve high-throughput reproducible results in multi-lineage differentiation and show how bioinformatics can be used as both an analysis tool, but also as a feedback in differentiation protocol optimization. Successfully combining all three would be crucial to reach the ultimate goals in bioengineering: in vitro organs-on-a-chip that are comparable, or better, to animal models for drug discovery and disease modeling, and functional implantable tissue analogs for patient-specific regenerative medicine. My research has laid such a foundation for the areas of cardiac health and CNS-based neural disease.