Targeted epigenetic tools for the control of variability in cardiac differentiation of patient-derived iPSCs

Julien Morival[1], Lily Widyastuti[2], Michael Zaragoza[2], and Timothy Downing[1] Department of Biomedical Engineering, UC Irvine[1]; Departments of Pediatrics & Biological Chemistry, School of Medicine, UC Irvine[2]

Background and significance

The advancements in somatic cell reprogramming and subsequent differentiation into a cell type of interest has provided scientists with a method to create accurate disease models from patient cells. This technique can be of significant advantage when trying to recapitulate a model of cells that are normally difficult to attain in patients, like cardiac and brain tissues (Trounson 2012). Moreover, comparing these diseased cells to controls could indicate potential causes of the disease to researchers. This process of redefining cell state and identity, however, involves a complex and convoluted set of gene regulations controlled by several mechanisms like the epigenome. It is therefore unsurprising that the desired cell type output is made up of variable low and immaturity yields (Huang 2012, Yoon 2006). One of the main reasons for this problem has been attributed to epigenetic memory and cell plasticity, wherein iPSCs retain some of the epigenetic landscape inherent to their cell type of origin. Attempts so far to resolve this have focused on using genome-wide approaches, like histone deacetylase and DNA methylation inhibitors (Huang 2012). These efforts, however, still have limited success, in part due to their lack of specificity (Christman 2002). There is therefore a need to develop new techniques in order to increase yield and maturity of patient-derived cells and allow for more accurate disease models to be made.

The Zaragoza lab developed a series of induced pluripotent stem cells (iPSCs) from skin biopsy fibroblasts of patients with lamin-associated cardiomyopathies. These cell lines, however, showed variability in the efficiency of cardiac differentiation. We hypothesize that DNA methylation may be playing a role in hindering cell plasticity and the differentiation process due to epigenetic memory. *We propose to identify variations in the methylome across lamin-associated cardiomyopathy patient and control cell lines, in order to understand the underlying mechanisms that regulate yield, maturity, and disease-state.*

Preliminary data

Table 1. Fibroblast and iPS cell line pairs and cardiomyocyte differentiation efficiencies (N = 10).

iPSC/Fibroblast	Source	Genotype[a]	Sex	Age at Skin Biopsy (years)[b]	Differentiation efficiency (% cTNT + cells)[c]
PATIENT-1	Patient A1	+/-	F	38	75
PATIENT-2	Patient A2	+/-	M	62	62
PATIENT-3	Patient A3	+/-	F	70	42
PATIENT-4	Patient C1	+/-	M	51	27
PATIENT-5	Patient C3	+/-	M	29	P
CONTROL-1	Control A1	+/+	F	49	60
CONTROL-2	Donor 3	+/+	M	51	60
CONTROL-3	Control A3	+/+	F	68	57
CONTROL-4	Control C1	+/+	M	60	49
CONTROL-5	Control C3	+/+	F	26	P

[a] Genotype: +/+ homozygous normal; +/- heterozygous *LMNA* mutation.
[b] Age: average age ± SD of CONTROL (50.8 ± 16) vs. PATIENT (50 ± 17) is not significantly different $p > 0.05$, (t-test)
[c] Differentiation efficiency: Differentiation trial pending (P).

Creation of iPSCs from human fibroblasts: Ten matched pairs of fibroblasts and iPSC lines will be used (Table 1). Cell lines were obtained, created, and characterized in the Zaragoza Lab under approved protocols (IRB HS#2011-8030, IRB HS# 2014-1253, & hSCRO#: 2014-1043). Five PATIENT iPSC lines were derived from dermal fibroblasts cultured from skin biopsies obtained from affected individuals of two study families (A & C). These include three patients (A1, A2, and A3) heterozygous for *LMNA* splice-site (c.357-2A>G) (Zaragoza 2016) and two patients (C1 and C3) heterozygous for *LMNA* missense (p.Arg335Trp) mutation (Zaragoza 2017). The five CONTROL iPSC lines were

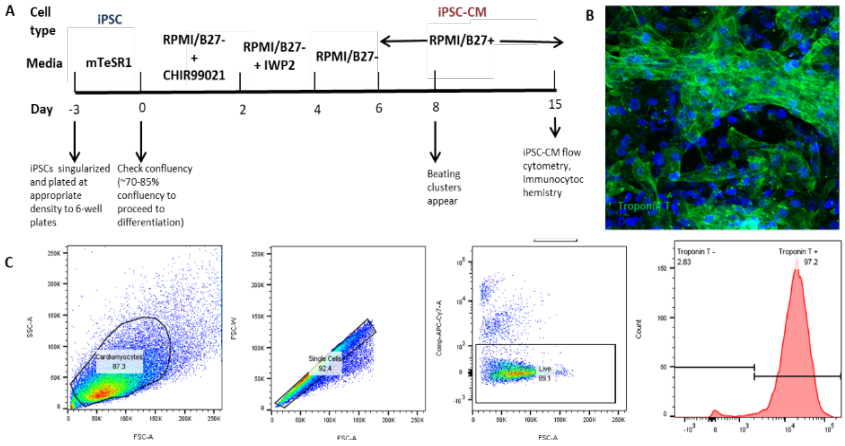

Figure 1: iPS-derived cardiomyocytes. (A) Differentiation and characterization workflow (**B**) Representative image of immunocytochemistry staining for cardiac-specific cell marker Troponin T (green) and DAPI (blue). **(C)** Representative dot plot from flow cytometry. Gating strategy excluded doublets and dead cells.

derived from dermal fibroblasts from unaffected, mutation-negative family members (A1, A3, C1 and C3) or healthy, unrelated individuals (Donor 3). Fibroblast lines were reprogrammed using non-integrating techniques (CytoTune-iPS Sendai Kit, Life Tech) (Fusaki 2009). For each line, one to three clones were characterized for pluripotency, differentiation capability, and chromosomal stability. By Sanger sequencing of all 12 *LMNA* exons in fibroblast DNA, presence or absence of mutation was confirmed.

iPSC-derived cardiomyocytes generation and evaluation of differentiation efficiency: iPSC-derived cardiomyocytes (CMs) were generated by Wnt signaling modulation (Lian 2013) (Figure 1A). On day 15, CMs were characterized by staining for cardiac Troponin T protein (Figure 1B) and flow cytometry to quantify the percentage of cardiac Troponin T positive (cTnT+) (Figure 1C). Differentiation efficiencies varied for PATIENT and CONTROL iPSCs (Table 1).

Results:

Identification of differentially methylated regions (DMR) responsible for cell state variability

Table 2. iPSC RRBS data statistics

ID	Source	iPSC gDNA conc (ng/uL)	Mapping %	Mapped reads	Mean CpG methylation	Mean read depth	# of CpGs
C1	Control A1	3980	82.30%	14005546	70.40%	3.3186	9163391
P1	Patient A1	2730	80.00%	13239664	69.30%	4.94019	7178314
C2	Donor 3	774.9	81.80%	15412513	68.80%	3.70308	9555966
P2	Patient A2	540.3	74.40%	13889429	69.80%	4.10576	8408397
C3	Control A3	418.3	79.90%	12818488	66.70%	5.13147	7131640
P3	Patient A3	1340	77.30%	11897293	68.90%	3.96827	7348060
C4	Control C1	325.8	74.70%	12541395	70.80%	4.09152	8152915
P4	Patient C1	1960	81.10%	12307525	69.90%	3.29007	8286874
C5	Control C3	2880	80.90%	11418775	69.10%	4.67716	6665970
P5	Patient C3	1154.1	81.20%	11057593	70.40%	4.38496	6794765

Previous studies have highlighted the detrimental impact of epigenetic memory on reprogramming and differentiation (Kim 2010). In order to determine the importance of DNA methylation in the context of the patient-derived cardiomyocytes, reduced representation bisulfite sequencing (RRBS-seq) was performed. We collected genomic DNA from corresponding fibroblasts and iPSC pairs (Table 1) and quantified the collected genomic DNA (gDNA) using Nanodrop spectrophotometer (Table 2). The resulting gDNA yield a concentration in the ranges of 525 ng/ul to 2880 ng/ul. Next, the resulting gDNA were subjected to reduced representation bisulfite sequencing (RRBS), so as to capture CpG-rich regions.

DNA collected from both iPSCs and fibroblasts was subjected to standard RRBS library preparation protocol, checked for quality, and sent out for sequencing at the Genomics High Throughput Facility at UCI. Collected raw fastq reads were subsequently trimmed, and put through Bismark, a widely utilized software for mapping bisulfite-converted sequence reads and determining cytosine methylation levels.

Table 3. Dynamic methylation data statistics

Dynamic CpGs	Hyper CpGs	Hypo CpGs	total DMRs	Ambiguous DMRs	custom DMRs	Hyper DMR	Hypo DMR	Genes Associated
657	203	454	601	3	598	175	423	984

From this analysis, we generated BED files, containing methylation ratios for each captured CpG (Table 2). These ratios are simply calculated at each CpG by dividing the number of reads with a methylated C at that location by the total number of reads that map there. This data was then run through methylKit, an R-based software for cytosine methylation analysis, which returned a non-biased clustering of the samples. As expected, fibroblasts and iPSCs separated from each other, however, patient and control samples were not as distinguishable, especially in iPSCs. However, PCA analysis revealed that 3 samples (iPSC control 1, iPSC patient 3, and iPSC patient 4) clustered away from the rest of the iPSCs (Fig. 2A). Interestingly, differentiation efficiencies of these three samples (60%, 42%, and 27%) are low (Table 1). In order to elucidate the impact of DNA methylation on differentiation efficiency, we determined dynamically methylated cytosines between the three identified samples and the remaining seven. These CpGs were chosen as long as they met the conditions of coverage ≥5, q-value ≤ 0.01, and a difference in methylation percentage ≥30%. Positive differences (>0; $CpG\%_{low\ differentiation} > CpG\%_{higher\ differentiation}$), were termed hypermethylated, while negative differences (<0; $CpG\%_{low\ differentiation} < CpG\%_{higher\ differentiation}$), were termed hypomethylated. 657 CpGs were identified as dynamic, of which 203 were hypermethylated, and 454 were hypomethylated. Dynamic CpGs were found to be located mostly within introns (39%) and intergenic (44%) (Fig. 2B).

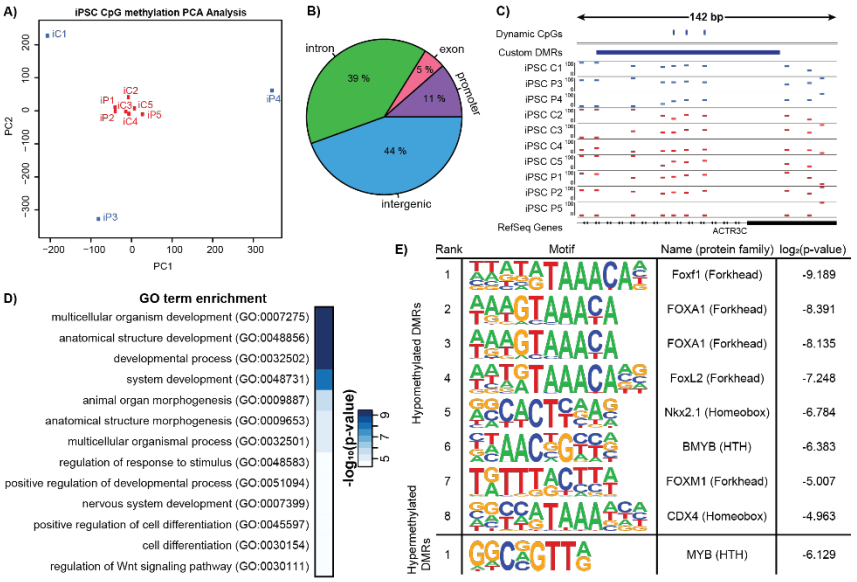

Figure 2: RRBS analysis of iPSCs based on differentiation efficiency A) PCA analysis of all iPSC RRBS samples. **B)** Percentage distribution of dynamic CpGs in genomic feature. **C)** Genome track of DNA methylation (0-100%) for all iPSC samples, including dynamic CpGs, and custom tiles created from them. **D)** GO term enrichment for DMRs. **E)** TF binding sites enriched in hypomethylated (top) and hypermethylated (bottom) DMRs.

To perform region-based analyses, we generated tiles to attain differentially methylated regions (DMRs) similarly to previously defined methods (Ziller 2013) These tiles were created by grouping any dynamic CpGs within 500bp of each other. Any tile with a size <100bp, had their genomic interval extended in both directions until they reached 100bp (Fig. 2C). Only non-ambiguous DMRs containing dynamic CpG methylation difference in the same direction (i.e. all negative, or all positive) were kept, and defined as hyper or hypomethylated accordingly. 598 DMRs were generated and classified for iPSCs (Table 3). Using the GREAT analysis from Stanford, we identified 984 genes associated with these DMRs. In order to perform GO term analysis, we generated a customized background file by performing the same tiling method as before on all CpGs acquired in our RRBS iPSC samples. Using this background, we found that of the 13 enriched terms, 11 were related to developmental and differentiation processes, and one was related to Wnt pathway regulation (Fig. 2D). These terms, along with the Wnt pathway's involvement in cardiac differentiation, suggest that DNA methylation may indeed be responsible for the lower differentiation in these iPSC samples. We additionally performed HOMER analysis to identify overrepresented transcription factor binding sites within our DMRs. Hypomethylated DMRs enriched for homeobox and forkhead protein families, while hypermethylated DMRs enriched for MYB, a protein involved in hematopoiesis regulation. Together, these results reinforce the fact that the identified genes may be dysregulated due to differential methylation, and therefore lead to differences in differentiation efficiency.

Future Directions

In the future, we aim to apply the findings of our current study to improve our cardiomyocyte differentiation protocols. By using a more targeted approach to activate or silence genes and transcription factor binding sites identified, we hope to obtain a more robust and more reproducible differentiation efficiency in the process of cardiomyocyte generation. Additionally, by using DNA from fibroblast cells, we will perform RRBS to similarly identify differentially methylated regions between patient and control samples, and their associated genes, in order to study the effect of DNA methylation on disease-state.

References

Cheedipudi, S. M., Matkovich, S. J., Coarfa, C., Hu, X., Robertson, M. J., Sweet, M. E., … Marian, A. J. Genomic Reorganization of Lamin-Associated Domains in Cardiac Myocytes is Associated with Differential Gene Expression and DNA Methylation in Human Dilated Cardiomyopathy. *Circulation Research*. 2019. https://doi.org/10.1161/CIRCRESAHA.118.314177

Christman, J. K. 5-Azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy. *Oncogene*. 2002;21:5483–5495. https://doi.org/10.1038/sj.onc

Dubois NC, Craft AM, Sharma P, et al. SIRPA is a specific cell-surface marker for isolating cardiomyocytes derived from human pluripotent stem cells. *Nat Biotechnol*. 2011;29(11):1011–1018. Published 2011 Oct 23. doi:10.1038/nbt.2005

Fusaki N, et al., Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. Proceedings of the Japan Academy Series B-Physical and Biological Sciences, 2009. 85(8): p. 348-362.

Gu, H., Smith, Z. D., Bock, C., Boyle, P., Gnirke, A., & Meissner, A. (2011). Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nature Protocols*, *6*(4), 468–481. https://doi.org/10.1038/nprot.2010.190

Huang, C., & Wu, J. C. Epigenetic Modulations of Induced Pluripotent Stem Cells: Novel Therapies and Disease Models. *Drug Discov Today Dis Models*. 2012;*9*(4):1–29. https://doi.org/10.1158/0008-5472.CAN-10-4002.BONE

Kim K, Doi A, Wen B, et al. Epigenetic memory in induced pluripotent stem cells. *Nature* 2010 467:285-290

Lian X, Zhang J, Azarin SM, et al. Directed cardiomyocyte differentiation from human pluripotent stem cells by modulating Wnt/β-catenin signaling under fully defined conditions. Nat Protoc. 2013 Jan;8(1):162-75.

Liu, X. S., Wu, H., Ji, X., Stelzer, Y., Wu, X., Czauderna, S., … Jaenisch, R. Editing DNA Methylation in the Mammalian Genome. *Cell*. 2016;*167*(1):233–247. https://doi.org/10.1016/j.cell.2016.08.056

Madakashira, B. P., & Sadler, K. C. DNA methylation, nuclear organization, and cancer. *Frontiers in Genetics*. 2017;*8*(76):1–7. https://doi.org/10.3389/fgene.2017.00076

Trounson, A., Shepard, K. A., & DeWitt, N. D. Human disease modeling with induced pluripotent stem cells. *Current Opinion in Genetics and Development*. 2012;*22*(5):509–516. https://doi.org/10.1016/j.gde.2012.07.004

Yoon, B. S., Yoo, S. J., Lee, J. E., You, S., Lee, H. T., & Yoon, H. S. Enhanced differentiation of human embryonic stem cells into cardiomyocytes by combining hanging drop culture and 5-azacytidine treatment. *Differentiation*. 2006;(74): 149–159. https://doi.org/10.1111/j.1432-0436.2006.00063.x

Zaragoza MV, et al., Exome Sequencing Identifies a Novel LMNA Splice-Site Mutation and Multigenic Heterozygosity of Potential Modifiers in a Family with Sick Sinus Syndrome, Dilated Cardiomyopathy, and Sudden Cardiac Death. PLoS ONE, 2016. 11(5).

Zaragoza MV, et al., Heart-hand syndrome IV: a second family with LMNA-related cardiomyopathy and brachydactyly. Clinical Genetics, 2017. 91(3): p. 499-500.

Zhou, W., Dinh, H. Q., Ramjan, Z., Weisenberger, D. J., Nicolet, C. M., Shen, H., … Berman, B. P. DNA methylation loss in late-replicating domains is linked to mitotic cell division. *Nature Genetics*. 2018;*50*(4): 591–602. https://doi.org/10.1038/s41588-018-0073-4

Ziller, M., Gu, H., Muller, F., Donaghey, J., Tsai, L., … Meissner, A. Charting a dynamic DNA methylation landscape of the human genome. *Nature*. 2013;*500*: 477-481.