Toxicogenomics Reveals Dynamic Baseline Changes and Overlapping Effects of Epigenotoxicant Exposure on Stem Cell Differentiation and Tissue Development Networks

Hobbs CA[1], Sinha, N, Chepelev, N[1], Gagné R[2], Christy, N[1], Yauk, CL[2], Recio, L[1]

[1]Integrated Laboratory Systems, Inc., Research Triangle Park, NC, USA
[2]Health Canada, Ottawa, Ontario, Canada

ABSTRACT

Pluripotency and differentiation processes are regulated by networks of genes controlled, at least in part, by the stem cell epigenome. Disruption of this finely tuned regulatory circuit by exposure to drugs and environmental chemicals can lead to adverse health effects. A biomarker panel capable of detecting a chemical's potential to impact the stem cell epigenome could have utility for hazard identification. Induced pluripotent stem (iPS) cells are similar to embryonic stem cells in chromatin structure and gene regulation. We first used the ACS-1007™ (ATCC-HYR0103) human iPS cell line and next generation sequencing (RNA-seq) to examine spontaneous gene expression changes in iPS cells seeded as single cells in 96-well dishes and cultured for up to 4 days in three independent experiments. Notably, there were dynamic changes in the transcriptome, especially on Days 3-4, largely related to cellular homeostasis, cell death and survival, and differentiation. Next, iPS cells were exposed to three chemicals known to affect histone post-translational modifications—valproic acid (histone deacetylase inhibitor), garcinol (histone acetyltransferase inhibitor), or 3-deazaneplanocin-A (DZNep; histone methyltransferase inhibitor). After 24 hours of exposure, cells from three replicate plates were harvested for mRNA isolation and concurrent assessment of cell viability. RNA sequencing demonstrated substantial dose-dependent changes in gene expression that correlated with increasing cytotoxicity. Bioinformatic analyses demonstrated that over multiple doses of all three chemicals there was considerable overlap in the most affected biological functions, including processes related to stem cell differentiation and tissue development. Moreover, comparison of genes differentially expressed in response to at least two doses of all three chemicals identified 542 shared genes, implying that these chemicals affect some common pathways and genes. Our study suggests that a toxicogenomic approach may prove useful for developing a predictive signature of chemically-induced epigenetic responsiveness in human stem cells. (Supported by NIEHS: 1R43ES023526-01; 2R44ES023526-02)

INTRODUCTION

Stem cells can both self renew and generate progeny that are capable of following different differentiation pathways. In recent years, technology has evolved to artificially derive pluripotent stem cells from a non-pluripotent cell by inducing a forced expression of specific genes (1). Although not perfectly identical, these "induced pluripotent" (iPS) cells are very similar to embryonic stem (ES) cells in many aspects, including chromatin structure and gene expression programs (2).

Since pluripotency and stem cell differentiation processes are driven by networks of genes largely regulated by epigenetic mechanisms, iPS and ES cells are attractive models to study the effects of drugs and environmental toxicants (potential developmental toxicants) on the human epigenome. As a means of expanding the scope of toxicological evaluation of chemicals, we are working to develop an assay to evaluate the effects of chemicals directly on specific histone modifications at a panel of developmentally-relevant genes in human pluripotent stem cells.

For this study, we used the ATCC-HYR0103 (ACS-1007™) iPS cell line (derived from normal human adult male liver fibroblasts) and three potent inhibitors of histone-modifying enzymes:

Valproic acid: a drug used to treat epilepsy and bipolar disorder; inhibitor of histone deacetylases (HDACs)

Garcinol: a natural inhibitor of p300/CBP and PCAF histone acetyltransferases, found in the rinds of *Garcinia indica* fruit

3-Deazaneplanocin A (DZNep): a drug which acts as a S-adenosylhomocysteine synthesis inhibitor; inhibitor of histone methyltransferases, particularly EZH2, a component of the Polycomb Repressive Complex 2 which regulates genes important for differentiation and development

OBJECTIVES

1) Evaluate the potential for spontaneous changes (e.g., differentiation) following cell seeding in 96-well dishes to possibly mask or confound effects of chemical exposure.

2) Identify and compare differentially expressed genes following exposure to the canonical inhibitors of histone-modifying enzymes, valproic acid, garcinol, and DZNep.

METHODS

Cell Culture:

ACS-1007™ human iPS cells derived from liver fibroblasts from a normal adult male were purchased from ATCC (ATCC-HYR0103). The cells were passaged and maintained on Matrigel (Corning) in mTeSR1 medium (StemCell Technologies) as cell aggregates in 6-well dishes. ROCK inhibitor Y27632 (ATCC) was included in the medium at 2 µM on days of passaging. Cells were monitored daily and differentiated cells removed by aspiration. For the assays, cells were dissociated using TrypLE™ (Life Technologies) into single cell suspensions and seeded onto Matrigel in 96-well dishes at ~$1X10^5$ cells/well.

For the first study, cells were fed daily with mTeSR1 and harvested for four consecutive days (corresponding to 0, 24, 48, or 72 hours of chemical exposure in an assay regimen); three independent experiments were performed.

For the second study, cells were treated with valproic acid, garcinol, or DZNep at various concentrations for 24 hours in three 96-well plates (*i.e.*, biological replicates). An additional plate of cells was used to assess viability for each chemical treatment using the CellTiter-Fluor Viability assay (Promega) according to manufacturer's directions.

For both studies, at time of harvest medium was removed and cells were lysed with Trizol (Ambion) or Qiazol (Qiagen) and stored at -80ºC. Cells were pooled from ≥16 wells and mRNA was prepared using RNeasy Plus (Qiagen) with DNase purification. RNA integrity was verified using an Agilent Bioanalyzer 2100.

Ion Proton™ Sequencing:

Next-generation sequencing was performed using an Ion Proton™ sequencer (Life Technologies, Carlsbad, CA). Samples were spiked with ERCC RNA Spike-in Mix (Ambion®). Poly-A RNA enrichment (DynaBeads® mRNA DIRECT Micro Kit) was performed for each sample on 1 µg of total RNA. The Ion Total RNA-Seq Kit v2 was used to fragment and prepare the cDNA libraries from poly-A enriched samples. The 3' and barcode primers provided in the Ion Xpress™ RNA-Seq Barcode 1-96 Kit were used to amplify the cDNA (each PCR product receiving its own unique barcode). Libraries were amplified using the Platinum® PCR SuperMix, High Fidelity (Life Technologies, Carlsbad, CA). Aliquots of each library were pooled together for a total final concentration of 50 pM. Emulsion PCR, enrichment, and chip loading were done on an Ion Chef™ instrument, using Ion P1™ chips (version 3) and Ion PI™ Chef Kits (Life Technologies). Chips were then used in semi-conductor sequencing by the Ion Proton™ sequencer using the Ion Proton™ HI-Q™ Sequencing Kit (Life Technologies) to obtain ≥20 million reads per sample.

Bioinformatics:

The Proton™ Torrent Server version 4.4.3 interpreted the sequencing data and generated FASTQ files for each barcoded sample. Reads were aligned to the reference genome (GRCh38/v84) using Star (2.3.0)(3) and Bowtie (2.1.0)(4). Following alignment, gene counting was performed with HT-Seq count (0.6.1) (http://www-huber.embl.de/users/anders/HTSeq/doc/count.html) with m parameter set to "intersection-nonempty" using the Ensembl GTF annotation (GRCh38v75). The table of counts was then imported in R and genes with a total count less than one read per million reads were eliminated. The EdgeR (3.6.7)(5) package was used for the analysis by normalizing with TMM (embedded in EdgeR)(6) and calculating differentially expressed genes using the generalized linear models. Genes with a false discovery rate *p*-value greater than $p<0.05$ and having a fold change between -1.5 and 1.5 were excluded from the subsequent analyses using Venny (7) to compare overlapping gene sets and Ingenuity Pathway Analysis™ (IPA™, Ingenuity Systems, Redwood City, CA) to identify the biological pathways, functions, and processes that were affected by chemical exposure.

RESULTS

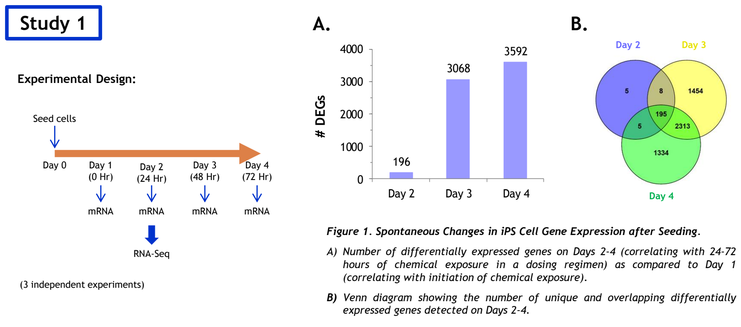

Figure 1. Spontaneous Changes in iPS Cell Gene Expression after Seeding.

A) Number of differentially expressed genes on Days 2-4 (correlating with 24-72 hours of chemical exposure in a dosing regimen) as compared to Day 1 (correlating with initiation of chemical exposure).

B) Venn diagram showing the number of unique and overlapping differentially expressed genes detected on Days 2-4.

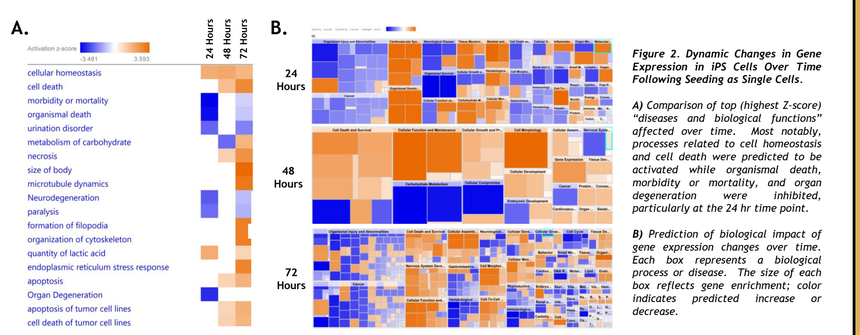

Figure 2. Dynamic Changes in Gene Expression in iPS Cells Over Time Following Seeding as Single Cells.

A) Comparison of top (highest Z-score) "diseases and biological functions" affected over time. Most notably, processes related to cell homeostasis and cell death were predicted to be activated while organismal death, morbidity or mortality, and organ degeneration were inhibited, particularly at the 24 hr time point.

B) Prediction of biological impact of gene expression changes over time. Each box represents a biological process or disease. The size of each box reflects gene enrichment; color indicates predicted increase or decrease.

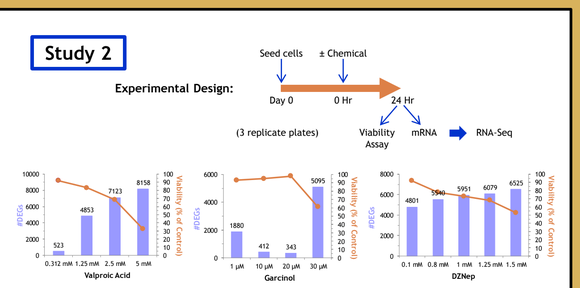

Figure 3. Dose Response of DEGs Following iPS Cell Exposure to Valproic Acid, Garcinol, and DZNep. Overall, the number of DEGs increased with chemical concentration and extent of cytotoxicity. However, exposure to DZNep resulted in a large number of DEGs even at a minimally cytotoxic concentration.

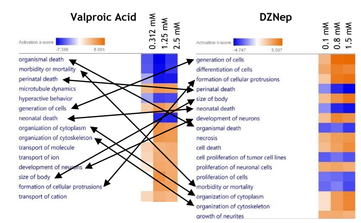

Figure 4. Comparison of Biological Functions Most Impacted by Chemical Exposure Reveals Substantial Overlap and Similar Directionality.

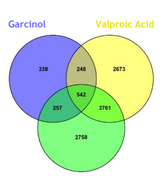

Figure 5. Identification of 542 Consensus Genes Impacted by Exposure of iPS Cells to the Three Chemicals. Venn diagram comparing genes that were differentially expressed in response to at least two concentrations of each chemical.

Top 5 Pathways	*p* -value	Overlap/ #Molecules
Canonical Pathways		
Hepatic Fibrosis/Hepatic Stellate Cell Activation	2.48E-04	7.1%
Glutamate Receptor Signaling	2.65E-04	12.3%
nNOS Signaling in Skeletal Muscle Cells	2.74E-04	26.7%
Bladder Cancer Signaling	7.23E-04	9.2%
Melatonin Degradation II	2.89E-03	50%
Molecular and Cellular Functions		
Cellular Movement	2.34E-03 - 5.17E-09	122
Cellular Growth and Proliferation	2.76E-03 - 1.22E-08	178
Cellular Assembly and Organization	2.89E-03 - 3.80E-07	99
Cellular Function and Maintenance	2.87E-03 - 3.80E-07	138
Cell Death and Survival	2.99E-03 - 4.98E-07	158
Physiological System Development and Function		
Cardiovascular System Development and Function	2.34E-03 - 5.17E-09	99
Organismal Development	2.76E-03 - 1.22E-08	161
Skeletal and Muscular System Development and Function	2.89E-03 - 3.80E-07	85
Embryonic Development	2.87E-03 - 3.80E-07	113
Hematological System Development and Function	2.99E-03 - 4.98E-07	48
Diseases and Disorders		
Cancer	2.78E-03 - 1.06E-12	436
Organismal Injury and Abnormalities	2.99E-03 - 1.06E-12	447
Gastrointestinal Disease	2.89E-03 - 3.36E-12	393
Cardiovascular Disease	2.89E-03 - 1.52E-08	111
Reproductive System Disease	2.13E-03 - 2.94E-07	264

Figure 6. Bioinformatic Pathway Analysis of the 542 Consensus Genes. Categories with the highest predictive confidence are listed. Note the high number of DEGs related to cancer and other diseases, as well as tissue development.

Summary and Conclusions

- ATCC-HYR0103 cells undergo dynamic changes in gene expression following seeding as single cells in 96-well dishes. Substantially more gene expression changes were observed on days 3 and 4 following seeding as compared to day 2.
  - Genes exhibiting altered expression included some involved in cellular homeostasis, possibly in response to the shock of passaging and/or culturing as single cells.
  - Processes related to cell death and survival were most affected but some genes indicative of spontaneous differentiation were also differentially expressed.
- Exposure to the histone modifying chemicals, valproic acid, garcinol, and DZNep led to substantial changes in gene expression in a dose-dependent fashion, generally correlating with extent of cytotoxicity.
  - There was considerable overlap in the biological functions most affected, including processes related to stem cell differentiation and tissue development.
  - Comparison of genes differentially expressed in response to at least two doses of all three chemicals identified 542 consensus genes, implying that these chemicals affect some common pathways and genes.
- Our study suggests that a toxicogenomic approach will be useful for developing a predictive signature of chemically-induced epigenetic responses in stem cells.

REFERENCES

1) Takahashi, K. *et al.* (2007) Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* **131**:861-872.

2) Guenther, M.G. *et al.* (2010) Chromatin Structure and Gene Expression Programs of Human Embryonic and Induced Pluripotent Stem Cells. *Cell Stem Cell* **7**:249-257.

3) Dobin, A. *et al.* (2013) STAR: Ultrafast Universal RNA-Seq Aligner. *Bioinformatics* **29**:15-21.

4) Langmead, B. and Salzberg, S.L. (2012) Fast Gapped-Read Alignment with Bowtie 2. *Nature Methods* **9**:357-359.

5) Robinson, M.D. *et al.* (2010) EdgeR: a Bioconductor Package for Differential Expression Analysis of Digital Gene Expression Data. *Bioinformatics* **26**:139-140.

6) Robinson, M.D. and Oshlack, A. (2010) A Scaling Normalization Method for Differential Expression Analysis of RNA-Seq Data. *Genome Biol* **11**:R25.

7) Oliveros, J.C. (2007-2015) VENNY. An Interactive Tool for Comparing Lists with Venn Diagrams. http://bioinfogp.cnb.csic.es/tools/venny/index.html

(Supported by NIEHS: 1R43ES023526-01; 2R44ES023526-02)