# Single-cell RNA Sequencing of Human Embryonic Stem Cells Reveals Immediate and Deferred Transcriptomic Responses to ATRA Toxicity Across Cell Sub-populations.

D A Gallegos1, S Jeffay1, E S Hunter1, B N Chorley1

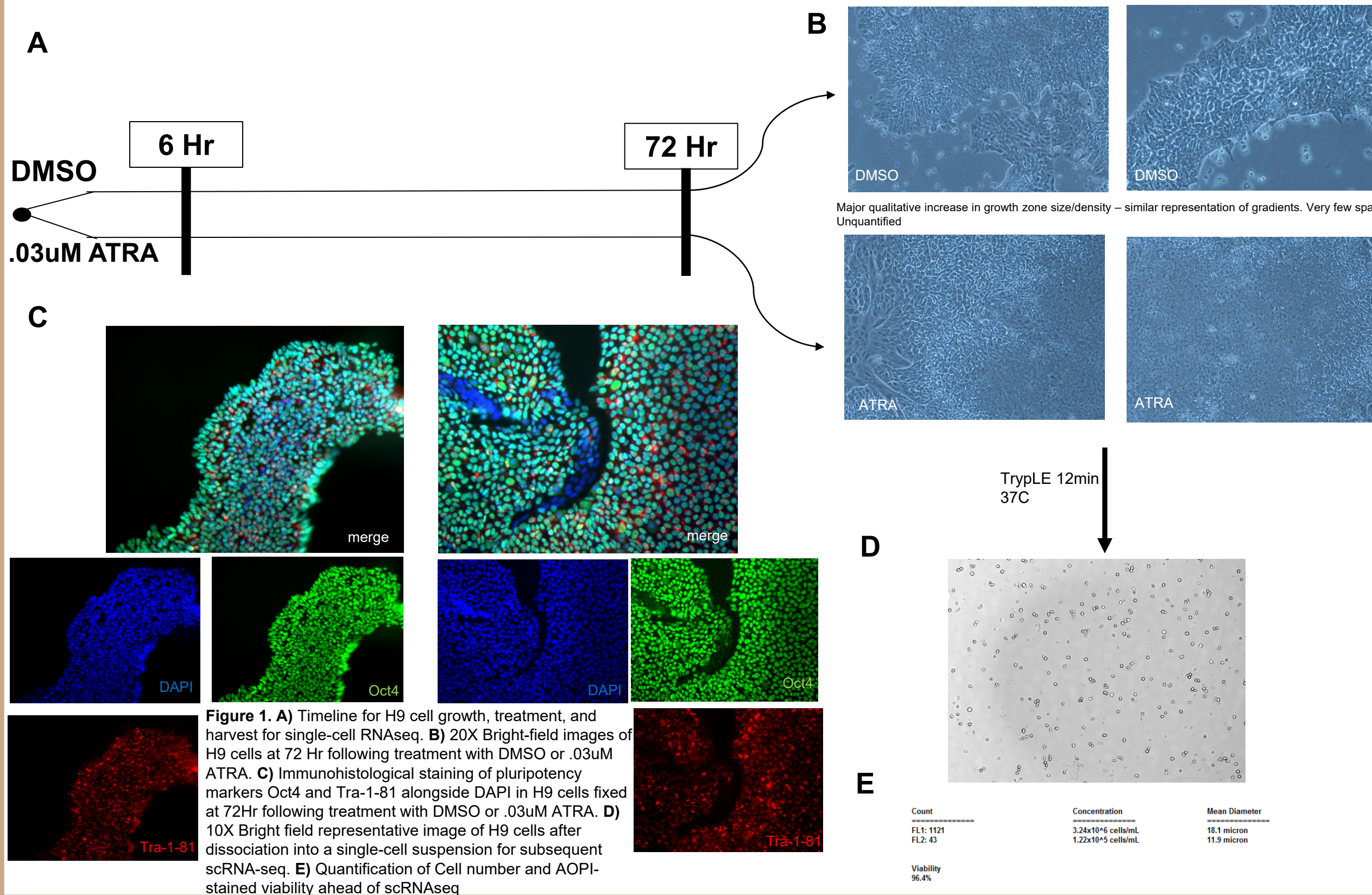1. Biomolecular and Computational Toxicology Division , CCTE, ORD, US EPA, RTP, NC, USA,

*This poster does not necessarily reflect EPA policy. Mention of trade names is not an endorsement or recommendation for use.*

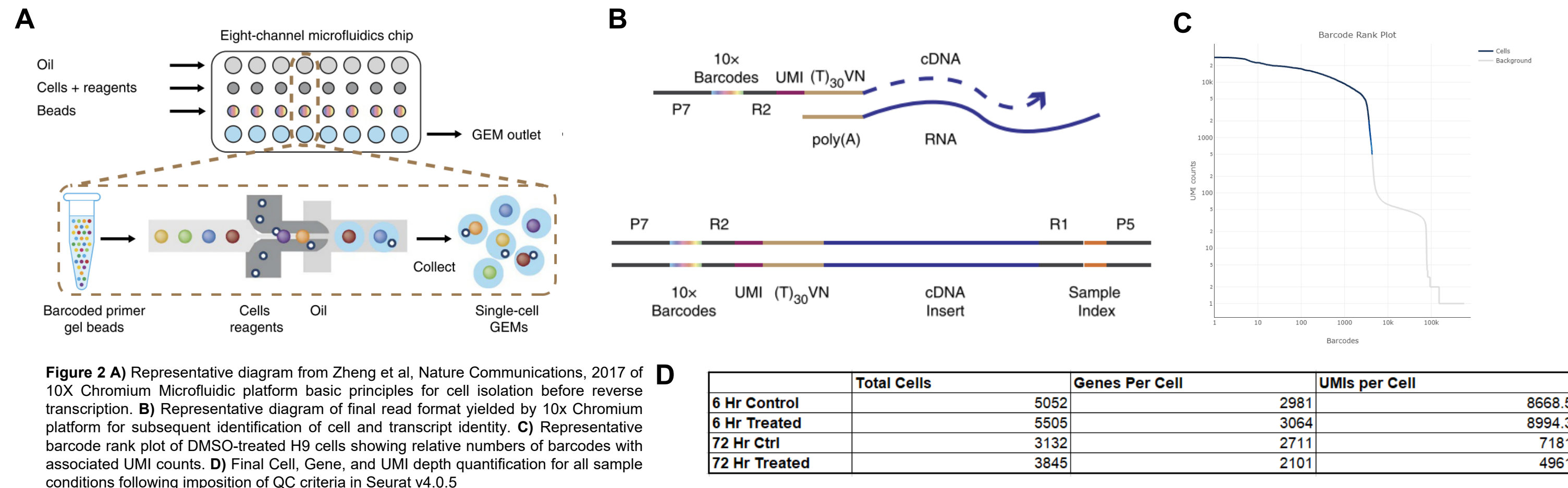David Gallegos I Gallegos.david@epa.gov I (919)-541-2623

## Abstract

Developmental disorders, such as heart malformations, affect 3% of children born in the U.S.. There is clear evidence that exposure to environmental chemicals can cause the development of many of these defects. Unfortunately, most commercial chemicals remain un- or under evaluated and there is a need for new approach methodologies (NAMs) that can accurately assess the impact on development at lower cost and in a higher throughput (HT) manner. Mouse embryonic stem cells (mESCs) studies have shown that less than 20% of *in vivo* developmental toxicants are able to be detected by mESC endpoints *in vitro*. More specific and sensitive *in vitro* assessments of early human development that leverage mechanistic understanding are needed to better model these early morphogenetic periods. To this end, we employed single-cell RNAseq (scRNAseq) using the 10XGenomics platform to assess our ability to survey direct, toxicant-induced effects on gene transcription at multiple post-exposure timepoints and putative cell subpopulations of H9 undifferentiated human embryonic stem cells (hESCs). In this proof-of-principle study, we measured scRNAseq profiles of immediate (6h) and persistent (72h) transcriptomic changes elicited by all-trans-retinoic-acid (ATRA) administered at a concentration near the lowest effect level (LOAEL) on hESCs (0.03μM). We identified putative subpopulations of cells within our *in vitro* system which may indicate very early cellular differentiation. We found a primary response gene program comprising direct metabolic responses to ATRA exposure as well as induction of transcripts involved in the regulation of hLh transcription factor DNA binding and developmental/differentiation processes. Following continued exposure (72h), we observed a qualitative phenotypic increase in cell proliferation, abundance, and morphological range as well as multi-faceted gene programs regulating multiple cellular processes including Wnt signaling, transcription factor binding, metabolic responses, and varied lines of differentiation/proliferation. These early results indicate that scRNAseq may help identify impacts of environmental chemical exposure on mechanisms of cellular differentiation and development. Such higher resolution and HT applications will lead to greater understanding of the impact of environmental chemicals on development and children's health. *This abstract does not reflect US EPA policy.*
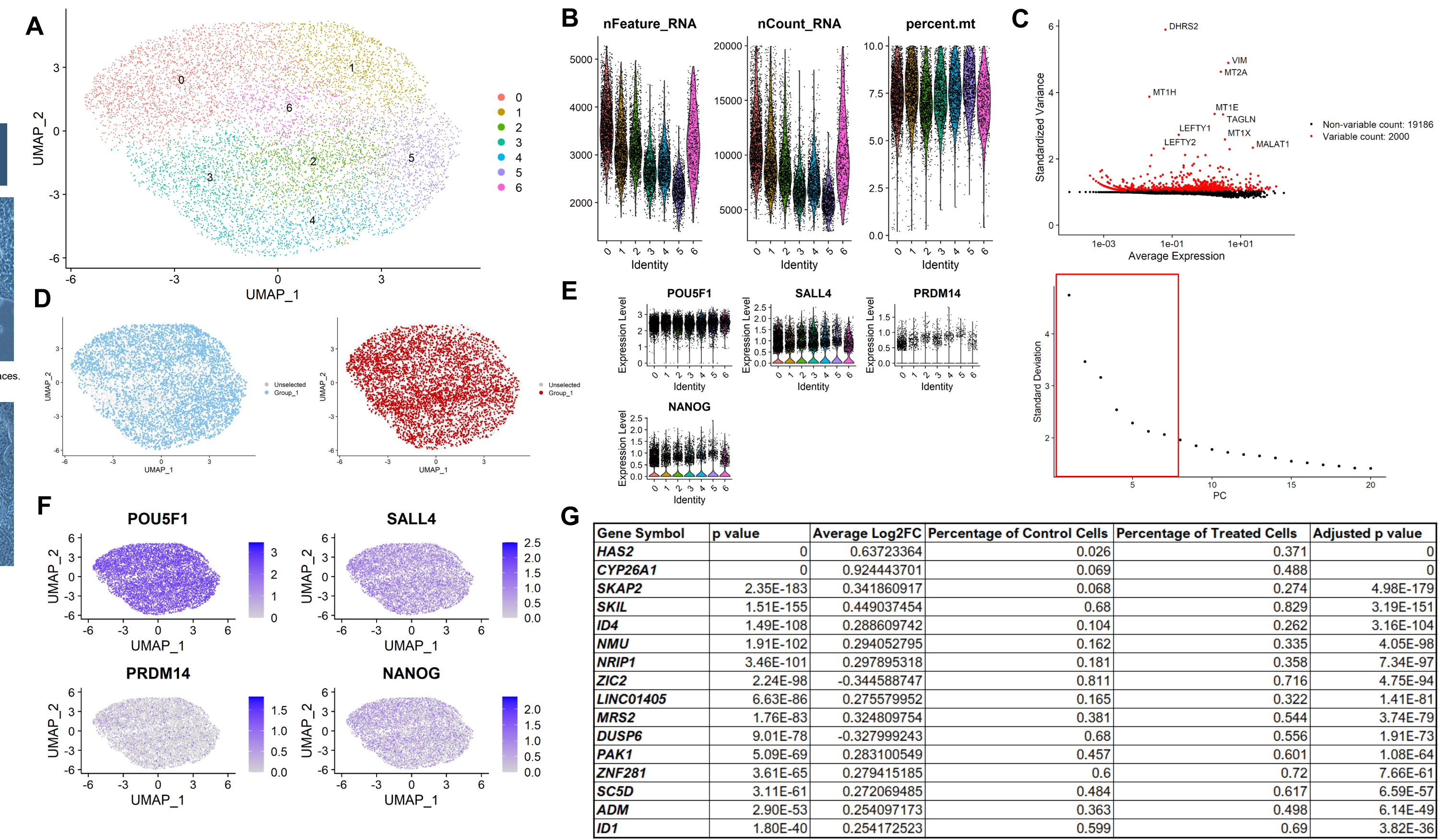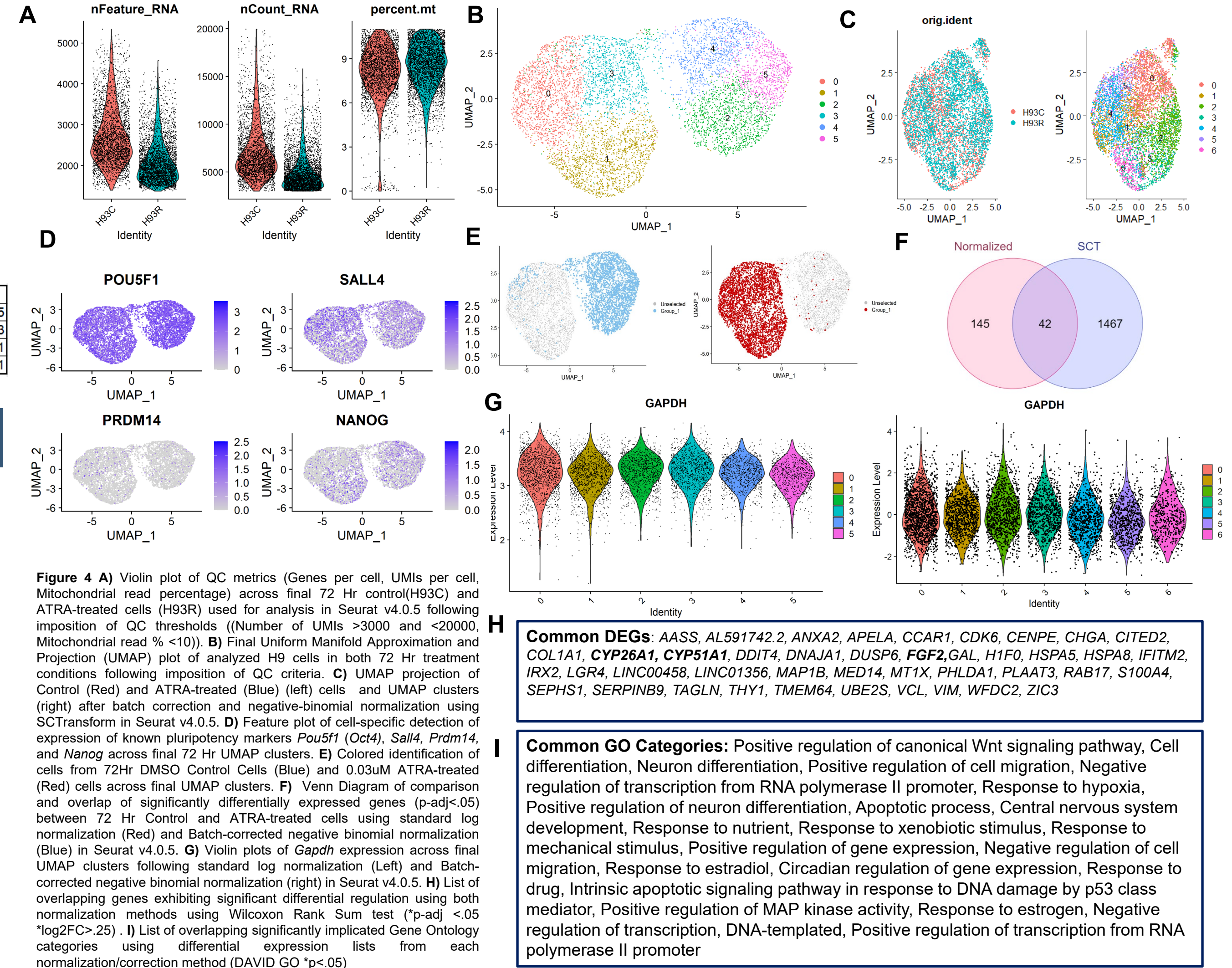
## Treatment Paradigm and Growth Phenotype



**Figure 1. A)** Timeline for H9 cell growth, treatment, and harvest for single-cell RNAseq. **B)** 20X Bright-field images of H9 cells at 72 Hr following treatment with DMSO or .03uM ATRA. **C)** Immunohistological staining of pluripotency markers Oct4 and Tra-1-81 alongside DAPI in H9 cells fixed at 72Hr following treatment with DMSO or .03uM ATRA. **D)** 10X Bright field representative image of H9 cells after dissociation into a single cell suspension for subsequent scRNA-seq. **E)** Quantification of Cell number and AOPI-stained viability ahead of scRNAseq.

## Mid-Output scRNAseq Performed For Proof-of-Principle Assessment at 6 & 72 Hr



**Figure 2 A)** Representative diagram from Zheng et al, Nature Communications, 2017 of 10X Chromium Microfluidic platform basic principles for cell isolation before reverse transcription. **B)** Representative diagram of final read format used by 10x Chromium platform for subsequent identification of cell and transcript identity. **C)** Representative barcode rank plot of DMSO-treated H9 cells showing relative numbers of barcodes with associated UMI counts. **D)** Final Cell, Gene, and UMI depth quantification for all sample conditions following imposition of QC criteria in Seurat v4.0.5

| | Total Cells | Genes Per Cell | UMIs per Cell |
|---|---|---|---|
| 6 Hr Control | 5052 | 2981 | 8668.5 |
| 6 Hr Treated | 5505 | 3064 | 8994.3 |
| 72 Hr Ctrl | 3132 | 2711 | 7181 |
| 72 Hr Treated | 3845 | 2101 | 4961 |

## 6 Hr Post-Exposure scRNA-seq Detects Early Transcriptional Response to ATRA



**Figure 3 A)** Final Uniform Manifold Approximation and Projection (UMAP) plot of analyzed H9 cells in both treatment conditions following imposition of QC criteria (Number of UMIs >3000 and <20000, Mitochondrial read % <10) in Seurat v4.0.5. **B)** Violin plot of QC metrics (Genes per cell, UMIs per cell, Mitochondrial read percentage) across final UMAP cell clusters of cells harvested at 6 Hr. **C)** Volcano plot of variance contributed by top 2000 most variable genes in hr dataset (top) and elbow plot of Variance explained by each component in PCA through PC 20 (bottom). Red box indicates the 6 PCs used for subsequent dimensionality reduction and UMAP generation. **D)** Colored identification of cells from 6Hr DMSO Control Cells (Blue) and 0.03uM ATRA-treated (Red) cells across final UMAP clusters. **E)** Violin plot of normalized expression of known pluripotency markers *Pou5f1* (Oct4), *Sall4, Prdm14,* and *Nanog* across final UMAP clusters. **F)** Feature plot of cell-specific detection of expression of known pluripotency markers *Pou5f1* (Oct4), *Sall4, Prdm14,* and *Nanog* across final UMAP clusters. **G)** Significantly Differentially-expressed genes detected between final 6 Hr DMSO-treated and .03uM ATRA treated cells using Wilcoxon Rank Sum test (*p-adj < .05 *log2FC>.25)

| Gene Symbol | p value | Average Log2FC | Percentage of Control Cells | Percentage of Treated Cells | Adjusted p value |
|---|---|---|---|---|---|
| HAS2 | 0 | 0.63723364 | 0.026 | 0.371 | 0 |
| CYP26A1 | 0 | 0.924443701 | 0.069 | 0.488 | 0 |
| SKAP2 | 2.35E-183 | 0.341860917 | 0.068 | 0.274 | 4.98E-179 |
| SKIL | 1.51E-155 | 0.449037454 | 0.68 | 0.829 | 3.19E-151 |
| ID4 | 1.49E-108 | 0.288609742 | 0.104 | 0.262 | 3.16E-104 |
| NMU | 1.91E-102 | 0.294052795 | 0.162 | 0.335 | 4.05E-98 |
| NRIP1 | 3.46E-101 | 0.297895318 | 0.181 | 0.358 | 7.34E-97 |
| ZIC2 | 2.24E-98 | -0.344588747 | 0.811 | 0.716 | 4.75E-94 |
| LINC01405 | 6.63E-86 | 0.275759952 | 0.165 | 0.322 | 1.41E-81 |
| MRS2 | 1.76E-83 | 0.324809754 | 0.381 | 0.544 | 3.74E-79 |
| DUSP6 | 9.01E-78 | -0.327999243 | 0.68 | 0.556 | 1.91E-73 |
| PAK1 | 5.09E-69 | 0.283100549 | 0.457 | 0.601 | 1.08E-64 |
| ZNF281 | 3.61E-65 | 0.279415185 | 0.6 | 0.72 | 7.66E-61 |
| SC5D | 3.11E-61 | 0.272069485 | 0.484 | 0.617 | 6.59E-57 |
| ADM | 2.90E-53 | 0.254097173 | 0.363 | 0.498 | 6.14E-49 |
| ID1 | 1.80E-40 | 0.254172523 | 0.599 | 0.69 | 3.82E-36 |

## 72 Hour Post-ATRA Treatment Reveals Evidence of Persistent Transcriptional Alterations With Implications on Differentiation



**Figure 4 A)** Violin plot of QC metrics (Genes per cell, UMIs per cell, Mitochondrial read percentage) across final 72 Hr control(H93C) and ATRA-treated cells (H93R) used for analysis in Seurat v4.0.5 following imposition of QC thresholds ((Number of UMIs >3000 and <20000, Mitochondrial read % <10)). **B)** Final Uniform Manifold Approximation and Projection (UMAP) plot of Control (Red) and ATRA-treated (Blue) cells in both 72 Hr treatment conditions following imposition of QC criteria. **C)** UMAP projection of Control (Red) and ATRA-treated (Blue) (left) cells and UMAP clusters (right) after batch correction and negative-binomial normalization using SCTransform in Seurat v4.0.5. **D)** Feature plot of cell-specific detection of expression of known pluripotency markers *Pou5f1* (Oct4), *Sall4, Prdm14,* and *Nanog* across final 72 Hr UMAP clusters. **E)** Colored identification of cells from 72Hr DMSO Control Cells (Blue) and 0.03uM ATRA-treated (Red) cells across final UMAP clusters. **F)** Venn Diagram of comparison and overlap of significantly differentially expressed genes (p-adj<.05) between 72 Hr Control and ATRA-treated cells using standard log normalization (Red) and Batch-corrected negative binomial normalization (Blue) in Seurat v4.0.5. **G)** Violin plots of *Gapdh* expression across final UMAP clusters following standard log normalization (Left) and Batch-corrected negative binomial normalization (right) in Seurat v4.0.5. **H)** List of overlapping genes exhibiting significant differential regulation using both normalization methods using Wilcoxon Rank Sum test (*p-adj < .05 *log2FC>.25) . **I)** List of overlapping significantly implicated Gene Ontology categories using differential expression lists from each normalization/correction method (DAVID GO *p<.05)

**Common DEGs:** AASS, AL591742.2, ANXA2, APELA, CCAR1, CDK6, CENPE, CHGA, CITED2, COL1A1, **CYP26A1, CYP51A1,** DDIT4, DNAJA1, DUSP6, **FGF2,** GAL, H1FO, HSPA5, HSPA8, IFITM2, IRX2, LGR4, LINC00458, LINC01356, MAP1B, MED14, MT1X, PHLDA1, PLAAT3, RAB17, S100A4, SEPHS1, SERPINB9, TAGLN, THY1, TMEM64, UBE2S, VCL, VIM, WFDC2, ZIC3

**Common GO Categories:** Positive regulation of canonical Wnt signaling pathway, Cell differentiation, Neuron differentiation, Positive regulation of cell migration, Negative regulation of transcription from RNA polymerase II promoter, Response to hypoxia, Positive regulation of neuron differentiation, Apoptotic process, Central nervous system development, Response to nutrient, Response to xenobiotic stimulus, Response to mechanical stimulus, Positive regulation of gene expression, Negative regulation of cell migration, Response to estradiol, Circadian regulation of gene expression, Response to drug, Intrinsic apoptotic signaling pathway in response to DNA damage by p53 class mediator, Positive regulation of MAP kinase activity, Response to estrogen, Negative regulation of transcription, DNA-templated, Positive regulation of transcription from RNA polymerase II promoter

## Conclusions and Future Directions for Hazard Assessment

- A single-cell transcriptomic, *in vitro*, approach was used to identify immediate and persistent mechanistic pathways involved in the toxicological response to ATRA in individual cells

- Even with differentiation prevented, sc-RNAseq is able to resolve transcriptional heterogeneity among ESCs that can enhance assessment in the context of downstream cell differentiation

- Future single cell RNA-seq and imaging analysis can provide increased resolution to capture complex cell-type specific mechanisms of developmental disruption. Such data will better establish early key events involved in developing adverse outcome pathways.



Cao et al, Science, 2017

Guo et al, Stem Cell Reports, 2019