Systematic Comparison of High-throughput Single-Cell and Single-Nucleus Transcriptomes during Cardiomyocyte Differentiation

Alan Selewa[1], Ryan Dohn[1], Heather Eckart[1], Stephanie Lozano[1], Bingqing Xie[1], Eric Gauchat[1], Reem Elorbany[2], Katherine Rhodes[2], Jonathan Burnett[2], Yoav Gilad[1,2], Sebastian Pott[2*], Anindita Basu[1,3*]

[1] Department of Medicine, University of Chicago, [2] Department of Human Genetics, University of Chicago, [3] Center for Nanoscale Materials, Argonne National Laboratory

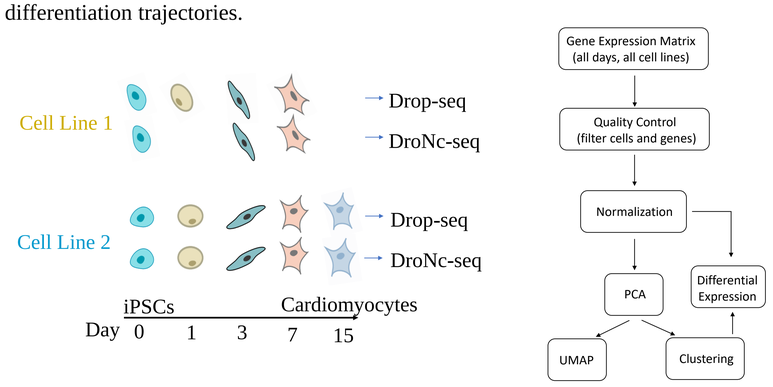
Introduction

- A comprehensive reference map of all cell types in the human body is necessary for improving our understanding of fundamental biological processes and in diagnosing and treating disease
- High-throughput single-cell RNA sequencing techniques have emerged as powerful tools to identify and characterize cell types in complex and heterogeneous tissues
- However, extracting intact cells from tissues and organs is often technically **challenging** or **impossible**, for example in heart or brain tissue
- **Single-nucleus RNA sequencing** provides an alternative way to obtain transcriptome profiles of such tissues
- To systematically assess the differences between high-throughput single-cell and single-nuclei RNA-seq approaches, we compared **Drop-seq** and **DroNc-seq**, two microfluidic-based 3' RNA capture technologies that profile total **cellular** and **nuclear RNA**, respectively

Methods

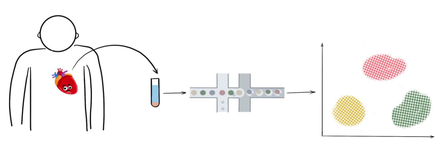
I. Comparison of Drop-seq and DroNc-seq

We performed Drop-seq and DroNc-seq, respectively, on induced pluripotent stem cells (iPSCs) undergoing differentiation into cardiomyocytes, following an established protocol from Bakken et al. Cells/nuclei were sampled at 5 time points (days 0, 1, 3, 7, 15) from the two cell lines. Technical replicates from each sampling underwent Drop-seq and DroNc-seq. This allowed us to compare Drop-seq and DroNc-seq with respect to read depth, transcriptome composition, cell types detected, and cellular differentiation trajectories.



II. DroNc-seq on adult male frozen primary heart tissue

We applied DroNc-seq to frozen human heart tissue to identify possible cardiac cell sub-types and non-cardiac cells within the tissue.



Results

I. Read Distributions in Drop-seq and DroNc-seq

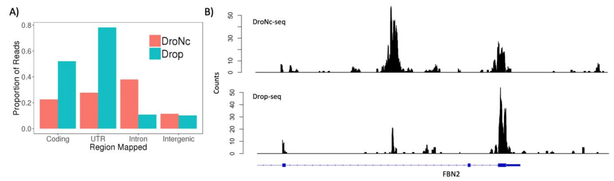


Figure 1: Comparison of distribution of reads across the genome between Drop-seq and DroNc-seq. A) Proportion of reads across various annotations. DroNc-seq samples more intronic reads compared with Drop-seq due to direct interrogation of nuclear RNA which contains more nascent RNAs than the cytosol. B) Distribution of reads at an example loci.

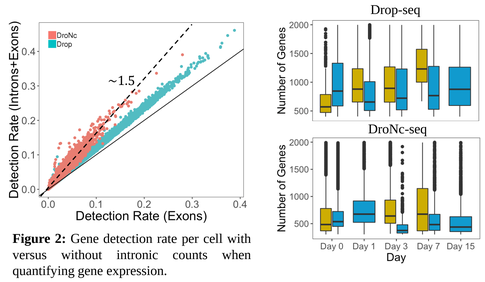


Figure 2: Gene detection rate per cell with versus without intronic counts when quantifying gene expression.

Figure 3: Comparison of genes detected (number of features with at least 1 count in a cell) between Drop-seq (top) and DroNc-seq (bottom). Overall, Drop-seq shows a higher number of genes and transcripts detected compared with DroNc-seq, reflecting the greater abundance of transcripts in the intact cell, compared with the nucleus alone.

II. Cell-type identification in Drop-seq and DroNc-seq data

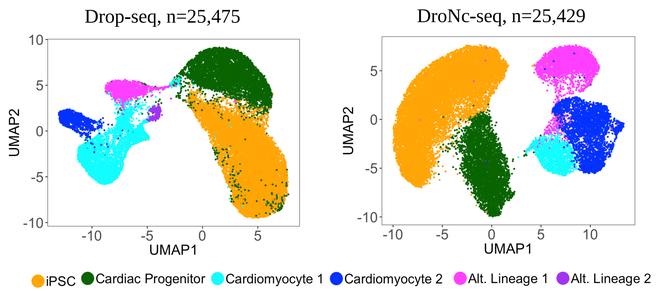
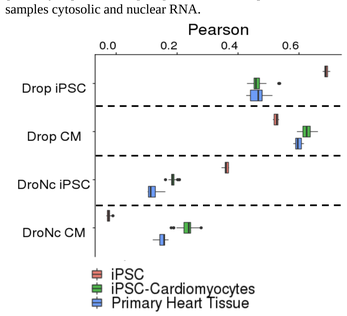


Figure 4: Cell type analysis and comparison between Drop-seq and DroNc-seq. Clustering was performed on the Euclidean nearest-neighbor graph, which was computed using the top 7 principle components for Drop-seq and DroNc-seq. Clustering results are visualized with UMAP. Differential expression was performed using a one-vs-all approach and resulting cluster labels were inferred from significantly differentially expression genes (given in Table 1).

Table 1. Summary of markers used to identify each cell-type and the cell-type proportion across all experiments

Markers	Cell type	Prevalence (Drop)	Prevalence (DroNc)
DPPA4	iPSC	48.9%	52%
EOMES APLNR	Cardiac Progenitor	23.3%	18.2%
MYH6 TNNT2	Cardiomyocyte 1	16.1%	5.6%
MYH6 TNNT2 AFP SERPINA 1	Cardiomyocyte 2	4.2%	12.7%
TTR FOXA2	Alternative Lineage 1	5.9%	11.3%
CD34 SCARF1 FLT1	Alternative Lineage 2	1.4%	0%

Figure 5: Correlation of iPSC and cardiomyocyte pseudo-bulk with matched bulk RNA-seq data of iPSCs, cardiomyocytes derived from iPSCs, and primary heart tissue. Drop and DroNc iPSCs correlate most with bulk iPSCs, and similarly for cardiomyocytes. The correlation is generally higher for Drop-seq data which is expected since bulk RNA-seq samples cytosolic and nuclear RNA.



Results

III. Reconstruction of differentiation trajectory

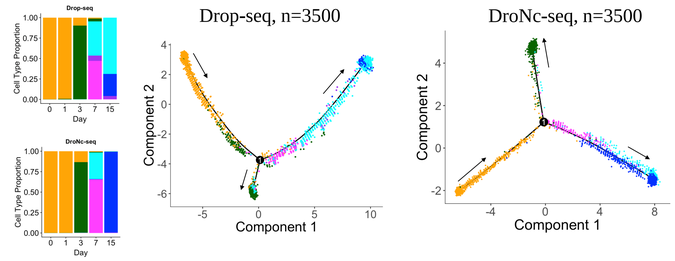


Figure 5: Comparison of differentiation trajectory in Drop-seq and DroNc-seq. The stacked bar-plots on the left show the cell-type proportion at each day in Drop-seq (top) and DroNc-seq (bottom). In both cases, we observe cardiomyocytes starting at day 7, as well as another cell-type representing an alternative lineage and/or immature cells that failed to differentiate by day 7. We applied Monocle to reconstruct the single cell differentiation trajectories (scatter plots on the right). Both methods detected a single branching point which partitions cardiomyocytes and immature cells that have yet to differentiate. The arrows indicate the direction of increasing pseudo time.

IV. Cell types in adult primary heart tissue

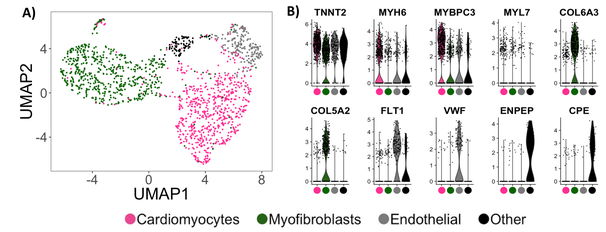


Figure 6: Application of DroNc-seq on human heart tissue. A) Cell type analysis visualized with UMAP. B) Distribution of marker genes identified with differential expression analysis. As expected, the majority of cells (~82%) were CMs and myofibroblasts.

Conclusions

- Primary differences between Drop-seq and DroNc-seq reflect the RNA in which they sample, i.e. cellular vs nuclear RNA. Mitochondrial and ribosomal RNAs were enriched in Drop-seq, while long non-coding RNAs were enriched in DroNc-seq
- Intron inclusion significantly increased the sensitivity of DroNc-seq, increased gene detection rate and generally improved cluster separation, in agreement with previous studies
- Expression profiles in Drop-seq and DroNc-seq confirmed the differentiation of iPSCs into CMs and revealed major cell types found within the *in vitro* differentiation model of iPSC-CMs
- Application of DroNc-seq to primary heart tissue revealed cardiomyocytes and myofibroblasts as well as a small population of endothelial cells
- This comparison of Drop-seq and DroNc-seq demonstrates the capability of DroNc-seq in dissecting the multicellular environment within a complex tissue such as the heart, which would otherwise not be possible with Drop-seq

Acknowledgements

This work was supported by the Chan-Zuckerberg Initiative pilot award #2017-174052. This work was performed, in part, at the Center for Nanoscale Materials, a U.S. Department of Energy Office of Science User Facility, and supported by the U.S. Department of Energy, Office of Science, under Contract No. DE-AC02-06CH11357. The computational resources were provided by the University of Chicago Research Computing Center. We thank Megan Rowton, Alex Guzzetta, and John Blischak for helpful comments on the manuscript.