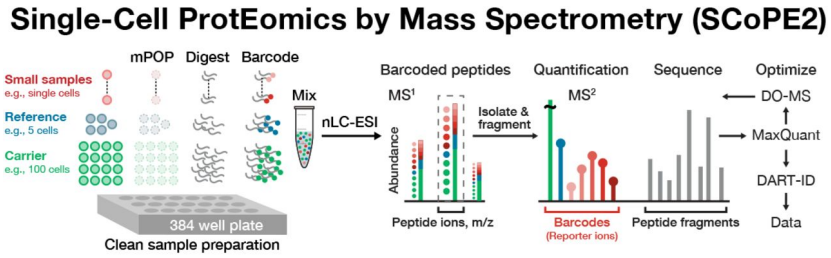
Single Cell Transcriptomics

Newsletter August 2019

Paper 1

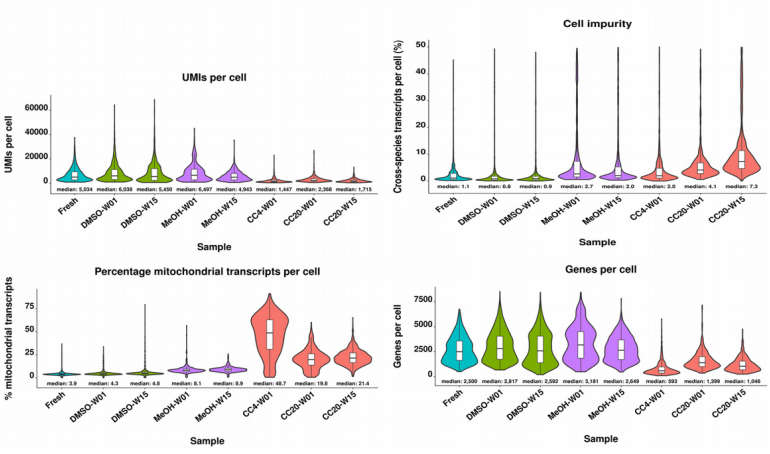
Specht H. et al. [High-throughput single-cell proteomics quantifies the emergence of macrophage heterogeneity] *bioRxiv pre-print*, 2019.



Proteomics approaches start to catch-up genomics in the field of single cell analysis. Comparing to CyTOFF (https://www.fluidigm.com/products/helios) biased by antibodies and finite multiplexing capabilities, protein mass-spectrometry (MS) is devoid of such limitations. However, sensitivity did not allow (wide) implementation of MS in single cell analysis before. Group of Nikolai Slavov previously developed Single-Cell ProtEomics by Mass Spectrometry (ScoPE-MS) and recently improved it (SCoPE2) to increase throughput and lower the costs (20-39$ per cell). Improvements included miniaturization and automation of sample preparation, changes in LC and parameters of MS. SCoPE2 also uses data-driven analytics to optimize instrument parameters for sampling more ion copies per protein. In proof-of-principle experiments lasting *85 hours* of instrument time, 2000 proteins were analyzed in *356 single monocytes and macrophages*, with *1000 proteins* identified per cell. Data allowed identification of cell types and proteome composition of cells, highlighting its utility for general research. Data also contained substantially lower quantities of counting errors comparing to scRNA-seq, mostly due to 1000 times higher abundance of proteins in the cell comparing to their corresponding mRNAs. However, variability existed in detected proteins between different technical sets of same cells, a phenomenon known for data-dependent acquisition (used for MS here), while the possible solution of this issue (use of specific list of peptides that can be targeted for analysis in each set) was already mentioned in the paper.

Paper 2

Wohnhaas C. T. et al. [DMSO cryopreservation is the method of choice to preserve cells for droplet-based single-cell RNA sequencing] *Scientific Reports*, 2019.



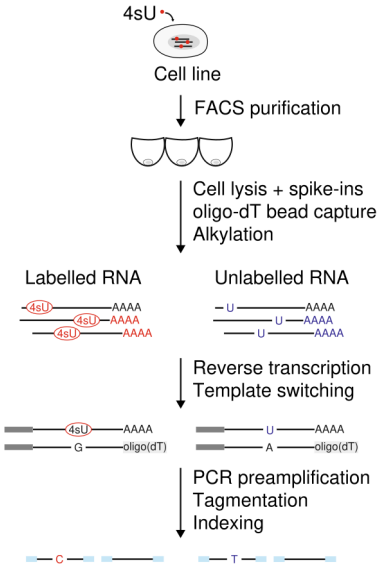
It is sometimes demanding (or even impossible) to perform cell/nuclei isolation steps with the downstream single cell library preparation same day. Different techniques of cell preservation exist, however the inherent transcriptomic changes associated with cell preservation and how they may bias further downstream analysis remain unknown. In this paper three preservation techniques (dimethyl sulfoxide based cryopreservation, methanol fixation and CellCover reagent) were compared to fresh cells. Cryopreservation proved to be the most robust protocol, maximizing both cell integrity and low background ambient RNA. Importantly, gene expression profiles from fresh cells correlated most with those of cryopreserved cells. Such similarities were consistently observed across the tested cell lines ($R \geq 0.97$), monocyte-derived macrophages ($R = 0.97$) and immune cells ($R = 0.99$). In contrast, both methanol fixation and CellCover preservation showed an increased ambient RNA background and an overall lower gene expression correlation to fresh cells. Upon preservation cells changed gene expression to a different extent. After DMSO preservation only 5-25 genes significantly changed expression, however these included upregulated FOS and FOSB indicative of cell activation upon damage.

This study highlights the importance of careful consideration for the use of preservation techniques (and their specific types) for the droplet based single cell transcriptomics.

Paper group 3. Monitoring newly synthesized RNA in single cells

Hendrikset G.-J. et al. [NASC-seq monitors RNA synthesis in single cells] *Nat. Communications*, 2019.
Cao J. et al. [Characterizing the temporal dynamics of gene expression in single cells with sci-fate] *bioRxiv preprint*, 2019.

Technologies to measure newly synthesized RNA are important to investigate cell perturbations and state transitions altogether with the general RNA turnover measurement. In a recent paper, lab of Rickard Sandberg (father of Smart-seq2) from Sweden described a new technique focused on the measurement of expression of newly synthesized RNA in single cells. This is a big improvement comparing to the previously available techniques, such as indirect RNA velocity (based on amounts of mRNA and unspliced newly synthesized pre-mRNA) and direct RNA sequencing techniques limited only to the bulk populations of cells. Technique was named NASC-seq (new transcriptome alkylation-dependent single-cell RNA sequencing) and allows simultaneous measurement of the newly synthesized and pre-existing RNAs in single cells. It is based on incorporation of 4-thiouridine (4sU) into newly synthesized RNA with following alkilation of 4sU. This is followed by T-C conversion during RT, Smart-seq2 library construction and further sequencing. Newly synthesized RNA is detected by the presence of C instead of T (miss-match) in the reads. Technique was validated on K562 cells and detected newly synthesized RNA overlapped with previously reported RNAs with high turnover. NASC-seq was also validated on Jurkat T-cells, where cell activation was revealed by both rapid up- and down-regulation of genes. The newly synthesized and pre-existing transcriptomes after T-cell activation were shown to be distinct, confirming ability of NASC-seq to simultaneously measure both temporal fractions of RNAs in single cells.

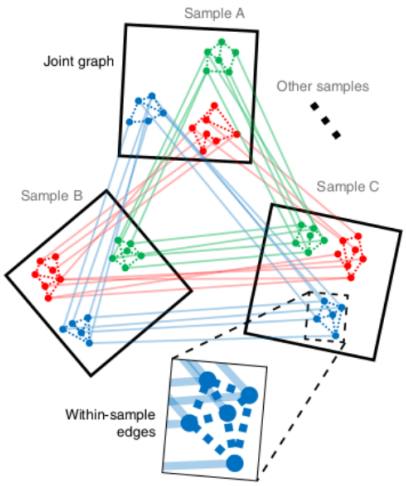


Interestingly enough, another paper by Jay Shendure group (USA, also in collaboration with Illumina) was released on bioRxiv describing a very similar approach for direct measurement of newly synthesized RNA in single cells. This time technique was called *sci-fate* while being considerably similar to the NASC-seq from Sandberg group. Sci-fate is also based on labeling of newly synthesized RNA using 4-thiouridine with the following thiol(SH)-linked alkylation reaction. In contrast to Hendrikset et al. paper, oligodT contained UMIs and cell-specific barcodes. T-C substitutions in reads were also used as a basis for identifying newly synthesized RNA. Similarly, technique was initially tested on cell lines and was confirmed to detect newly synthesized RNA. Further validation was also based on *in vitro* stimulation of cells (this time dexamethasone stimulation of adenocarcinoma-derived A549 cells) and was confirmed to detect newly synthesized RNA (together with remaining transcriptome) upon cell activation. Using computational approaches investigators went further and draw links between newly synthesized RNA and trascription factors expressed in the cells. Also, attempt was made to infer previous cell states from sci-fate data.

Paper group 4. Integration of single cell data between samples, technologies, tissues and even omics modalities

Barkas N., Petukhov V., et al. [Joint analysis of heterogeneous single-cell RNA-seq dataset collections] *Nat. Methods*, 2019.
Stuart T., et al. [Comprehensive Integration of Single-Cell Data] *Cell*, 2019.
Welch D. W., et al. [Single-Cell Multi-omic Integration Compares and Contrasts Features of Brain Cell Identity] *Cell*, 2019.

Current stage of single cell research reached a point where we faced the need to integrate data across different experimental approaches, samples, batches and even omics modalities. Thankfully, new tools were developed recently aimed to help with these tasks.



In a recent paper from Kharchenko and Khodosevich groups (*big congratulations to Viktor Petukhov!:*)) new tool named Conos (clustering on network of samples) was published (while previously available on github). It is based on a unified graph representation capturing relationships between cells in different samples. Initially it was verified on bone marrow and cord blood scRNAseq samples and confirmed to identify joint clusters of genuine cell populations across different samples. Conos proved to be robust in identifying common cell clusters even upon perturbations

(increased noise, missing cells) and across heterogeneous samples (being usual case in clinical research, etc). Moreover, cells of different states (tumor cells and nearby tissue) were not mixed together. Conos was tested using integration of Tabula Muris with other atlases and was effective at identifying common cell populations across samples measuring diverse tissues, as well as overcoming the differences of the three different scRNA-seq platforms used. Moreover, Conos can be used to integrate single cell data created from different molecular modalities (chromosome accessibility, scRNAseq, etc.). To note, we widely use Conos to integrate samples of different origins and experimental variations and we are always happy with the results.

Two other papers recently published in Cell target the same challenges Conos was designed to solve. These originate from Macosco and Satija groups also promise robust integration of single cell datasets even across different molecular modalities.

Next Single Cell Seminar

BREAK until September 27th

Contact: mykhailo.batiuk@bric.ku.dk