Linking two worlds: RNA and protein at single cell resolution

High-throughput single cell analysis lets researchers see the functions of complex biological systems. Detecting key proteins at the cell surface together with gene expression adds an additional layer of information. Reading the transcriptome and the proteome together at the single cell level reveals how one relates to the other, and how both perform cellular roles and functions. Drug development, research into infectious diseases and oncology are just some of the applications this new knowledge empowers.

Applications

Quantification

Perform protein level measurements

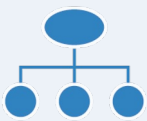
Protein variants

Research into post-translational modifications and splicing isoforms

Correlations

Investigate relationships between RNA and proteins

Clustering



Explore clearly distinguished cell subtypes

From surface protein to sequencing ready read-out

To incorporate proteomics data with the existing single cell transcriptomics GEXSCOPE technology, protein information can be converted into a nucleic acid sequence that can be sequenced in parallel with the transcriptome (Stoeckius et al., 2017). Commonly, antibody-oligonucleotide conjugates (Ab-oligo, Figure 1) are used: antibodies detect specific cell surface proteins while the oligonucleotides allow compatibility with next generation sequencing (NGS) systems.

The conjugated oligonucleotide sequence consists of a unique short sequence, the antibody-specific barcode, which enables identification of the target protein. The oligonucleotide also contains a 3' poly (A) sequence, which can be captured with poly (T) probes. In addition, a universal PCR handle is used for the generation of NGS libraries. These antibodies can be used in combination with Singleron's GEXSCOPE Single Cell RNA Library Kits to enable robust detection of surface protein markers and mRNA from the same single cells.

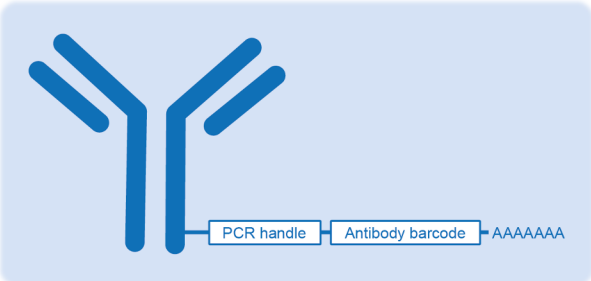


Figure 1: **Antibody conjugated to oligonucleotide.** Ab-oligo containing PCR handle, antibody barcode and poly (A) tail.

Transcriptome and proteome research in one workflow

For the combined detection of whole transcriptome and cell surface proteins, cells are first incubated with Ab-oligos. Following wash steps to remove unbound antibodies, stained cells are then processed on the Singleron SCOPE-chip®, a portable, microfluidic chip that isolates single cells into hundreds of thousands of microwells. Barcode Beads carrying poly (T) probes are then added to the microwells on the SCOPE-chip. Thereafter, the cellular mRNA together with Ab-oligos are captured onto the Barcode Beads allowing for cell-specific barcoding of each mRNA and Ab-oligo. Parallel mRNA and Ab-oligo NGS libraries are then generated. The Ab-oligo NGS libraries are used to determine the identity and quantity of the targeted proteins from individual cells (Figure 2, 3).

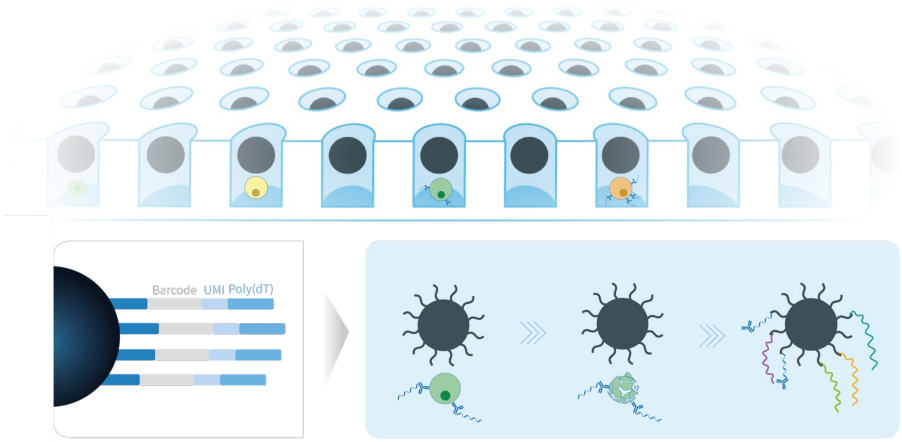


Figure 2. **SCOPE-chip technology.** The SCOPE-chip captures and partitions Ab-oligo stained single cells into hundreds of thousands of microwells. Barcoding Beads with attached capture oligonucleotides, carrying a cell label (Barcode), a unique molecular identifier (UMI) and a poly (T) probe, are then added to the microwells. After cell lysis, the mRNA and the Ab-oligos are captured by binding of the poly (A) tail to the poly (T) sequence of the captured oligonucleotides. Barcoding Beads are thereafter collected from the microwell chip and used for subsequent library generation.

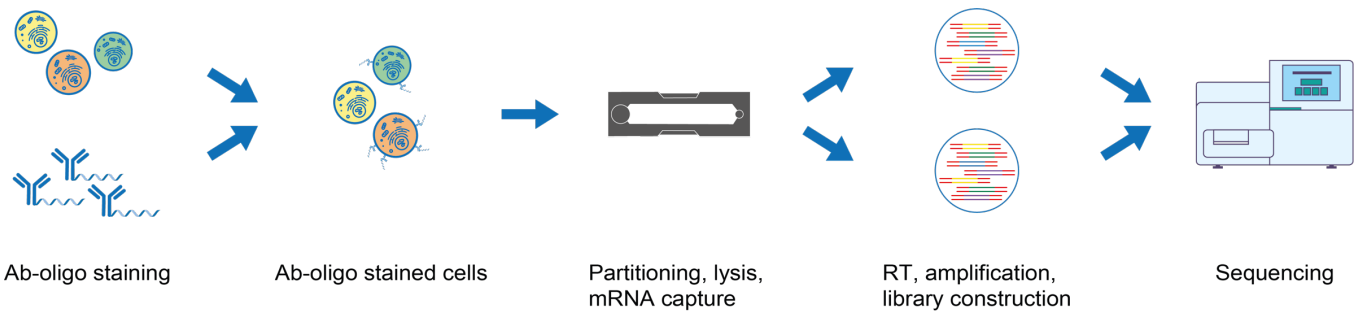


Figure 3. **Combined single cell transcriptomics and proteomics workflow.** Single cell suspension are incubated with Ab-oligos and excess of unbound antibodies are washed off. Cell suspension with bound Ab-oligos are loaded onto the SCOPE-chip for single cell partitioning, cell lysis, mRNA capture and molecular labeling. After reverse transcription of the captured oligos, barcoded cDNA is amplified and used to construct two parallel libraries, one for transcriptome and one for Ab-oligos. The NGS library is sequenced on a compatible sequencer and analyzed using bioinformatical tools.

Application: Single cell characterization of peripheral blood mononuclear cells with an antibody cocktail

Compatibility of GEXSCOPE Single Cell RNA Library Kit with surface protein detection through Ab-oligos was demonstrated using the BD AbSeq Immune Discovery Panel (IDP). The IDP antibody cocktail contains antibodies against 30 human immune markers. Together, the GEXSCOPE Single Cell RNA Library Kits and BD AbSeq IDP panel were used to detect single cell surface proteins and transcriptome of peripheral blood mononuclear cells (PBMCs).

Commercially available, cryopreserved PBMCs were used in this experiment. The IDP antibodies were reconstructed and used in accordance with manufacturer's instructions (BD Biosciences). One million PBMCs were resuspended in cell staining buffer and stained with the antibody cocktail. Following three washes to remove unbound antibody conjugates, stained cells were diluted and loaded onto four SCOPE-chips. This was followed by library preparation and sequencing.

Pre-processing and mapping were performed with CeleSCOPE v1.10.0, using *Homo sapiens* genome GRCh38 assembly as a reference. Downstream analysis was performed using R package Seurat v4.3.0 (Hao et al., 2021). A total of 34,179 cells were identified with a median of 1,028 genes per cell. Clusters are annotated using the SynEcoSys database, followed by manual verification. Moreover, each cluster was subsequently checked with BD panel's corresponding surface protein markers. Cell types (plateletes, plasma cells, erythrocytes progenitors) representing a low percentage of the sample, and cell types that are not targeted in the BD Panel were filtered out. Final data had a total number of 30,991 cells.

Increasing resolution using multi-omics data

PBMCs were clustered and annotated using mRNA data or mRNA and protein data together. Relying on transcript abundance, it was possible to determine five major cell types with T cells and monocytes being the most dominant cell types. Eight different T cell subtypes have been identified. Using transcript and protein abundance together, a better resolution of cells was achieved not only between the major cell types but also within cell populations (Figure 4, Figure 5).

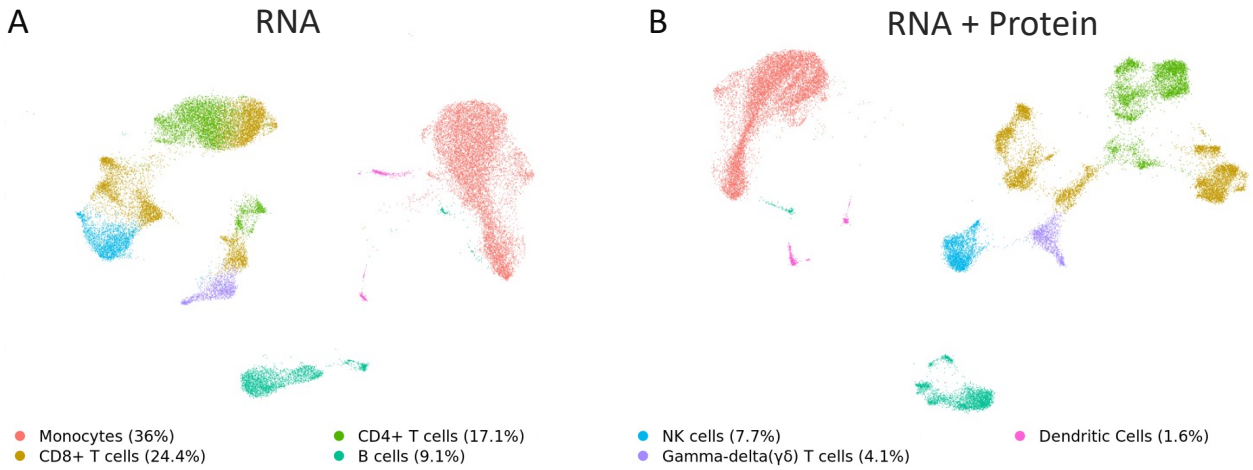


Figure 4. **UMAP clustering of PBMCs.** Cells were colored based on cell type. A) UMAP coordinates were plotted for all PBMCs based on mRNA data only. B) UMAP coordinates were plotted for all PBMCs based on mRNA and protein data.

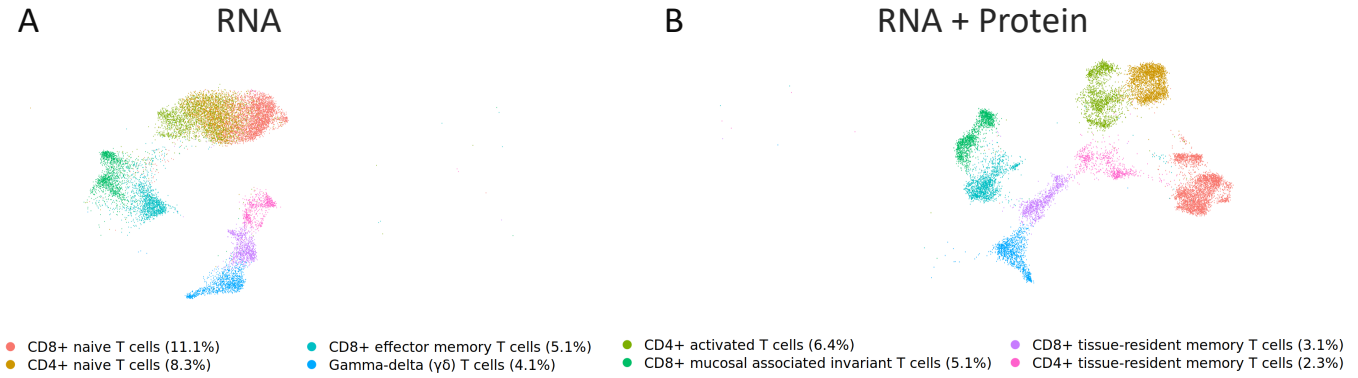


Figure 5 **UMAP clustering of T cells.** Cells were colored based on cell type. A) UMAP coordinates were plotted for all T cells based on mRNA data only. B) UMAP coordinates were plotted for all T cells based on mRNA and protein data.

High sensitivity detection at low transcript abundance

The correlation between mRNA and the respective proteins can differ significantly depending on various factors like distinct regulation mechanisms. The abundance and lifespan of mRNA molecules are subjected to changes. It might be challenging to reliably reflect levels for transcripts with low expression levels or fast mRNA turn-over. In these cases, evaluation of protein abundance can provide deeper insights into the cell identity and/or ongoing molecular processes.

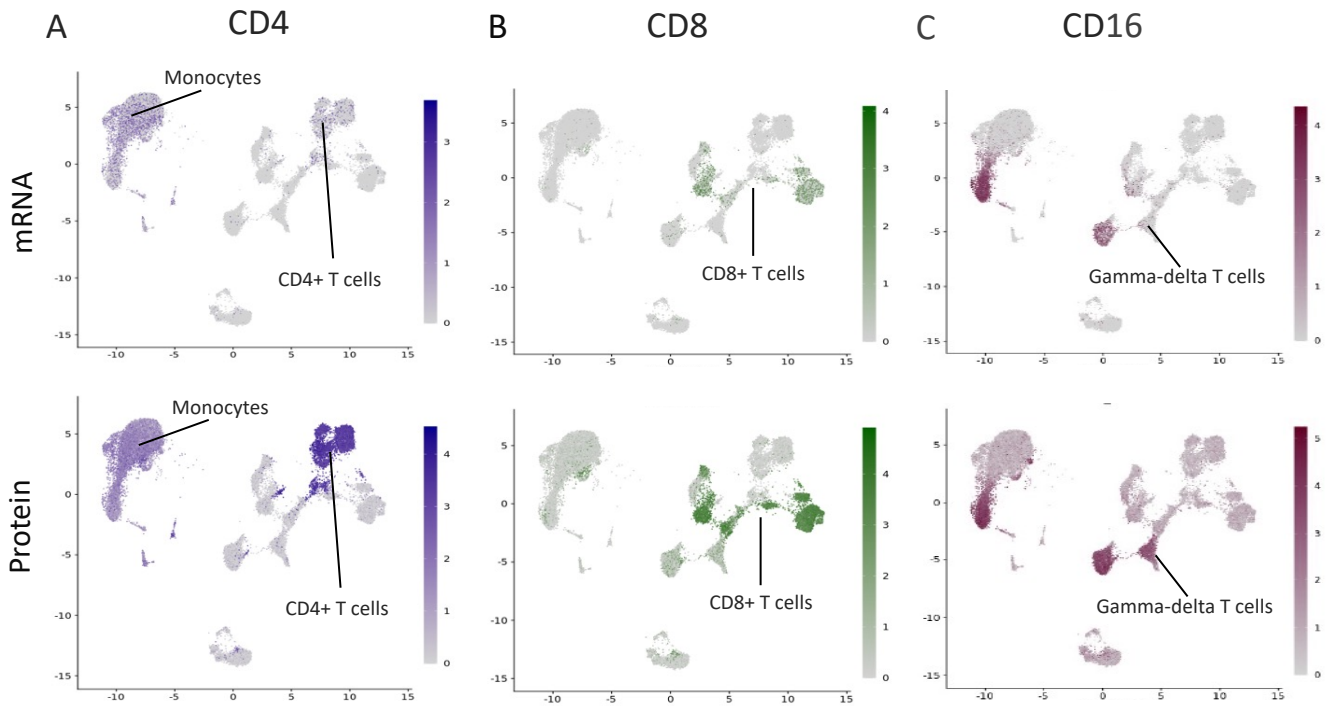


Figure 6. **Detection of surface proteins with low mRNA expression.** UMAP projection and cell types are the same as in Figure 4B. Color code corresponds to expression of mRNA (upper row) or protein (lower row). Expression levels of CD4 (A), CD8 (B) and CD16 (C) are shown.

CD4 transcript abundance is generally low in CD4+ T cells and monocytes. However, high levels of CD4 protein in CD4+ T cells and moderate levels in monocytes have been observed at a protein level (Figure 6A). Similarly, examination of CD8 abundance revealed its higher expression in CD8+ T cells compared to CD8 expression at the mRNA level (Figure 6B). CD16 is used as a marker for non-classical monocytes, natural killer cells and gamma-delta T cells. For gamma-delta T cells, detection at a protein level showed more abundance than at the mRNA level (Figure 6C).

References and Resources

- Stoeckius, M., et al. (2017). Simultaneous epitope and transcriptome measurement in single cells. Nat Methods 14, 865–868. https://doi.org/10.1038/nmeth.4380
- BD Biosciences, BD AbSeq Immune Discovery Panel Brochure, https://www.bdbiosciences.com/en-de/products/reagents/single-cell-multiomics/abseq/abseq-immune-discovery-panel
- Hao, Y., et al., (2021). Integrated analysis of multimodal single-cell data. *Cell*, *184*(13), 3573-3587.e29. https://doi.org/10.1016/j.cell.2021.04.048

Acknowledgement: We thank BD for providing the antibodies and their advice for downstream data analysis.

Singleron Biotechnologies GmbH

Tel.: +49 221 16824777 Email: info@singleronbio.com Address: Gottfried-Hagen-Str. 60, 51105 Cologne, Germany