Multiomic interrogation of the immune system at single-cell resolution

- Analyze full-length paired B- or T-cell receptors, cell surface protein expression, antigen specificity, and gene expression from a single cell with 10x Genomics Chromium Single Cell Immune Profiling

- Sequence V(D)J transcripts, antigen receptors, cell surface protein expression libraries, and 5' gene expression libraries with high accuracy on Illumina sequencing systems

- Reveal distinct mechanisms driving adaptive immune system response and gene regulation in healthy and disease states

In collaboration with

10x GENOMICS

illumına®

Introduction

The adaptive immune system is highly complex and dynamic, enabling our bodies to mount a response against infectious pathogens, including bacteria and viruses, and disease-causing cellular changes as in cancer. This complex process is mediated by the activation of B-cells and T-cells upon presentation of an antigen on the cell surface by the major histocompatibility complex (MHC). To recognize the vast range of possible antigens, B-cell and T-cell receptors require a diverse repertoire of recognition elements. This diversity is created through rearrangements in the variable (V), diversity (D), and joining (J) gene segments (V(D)J rearrangement) in the heavy immunoglobulin chain of B cells, the β chain of αβ T-cell receptors, and the δ chain of γδ T-cell receptors.[1] V(D)J rearrangement, along with other events during immune cell maturation, has the potential to generate > 10^{18} unique T-cell receptors in humans and a much more diverse B-cell repertoire.[2,3]

Next-generation sequencing (NGS) is proving to be a powerful immunogenomics research tool to map the vast repertoire of immune cells capable of recognizing a seemingly boundless array of targets.[4] For a more fully developed picture of immunological responses in infections, autoimmunity, cancer, and other pathologies, it is necessary to go beyond gene expression profiling of the B-cells and T-cells involved. Understanding how the adaptive immune system functions requires a multiomic approach that includes profiling cell phenotypes in the context of their cell surface protein markers (Figure 1).

This technical note outlines a protocol for profiling gene expression, B-cell or T-cell receptors, antigen specificity, and cell surface proteins (CSP) at single-cell resolution using 10x Genomics Chromium Single Cell Immune Profiling and Illumina sequencing systems. Generated data analyzed with 10x Genomics Cell Ranger and Loupe Browser can help to fully characterize the B-cells and T-cells involved in immunological responses.

Protocol overview

A workflow for multiomic profiling of the immune system typically involves initial sample preparation, labeling cDNA with a barcode for downstream detection, library preparation, sequencing, and data analysis (Figure 2). This protocol leverages the advanced Chromium System with Next GEM technology from 10x Genomics and proven Illumina sequencing technology.
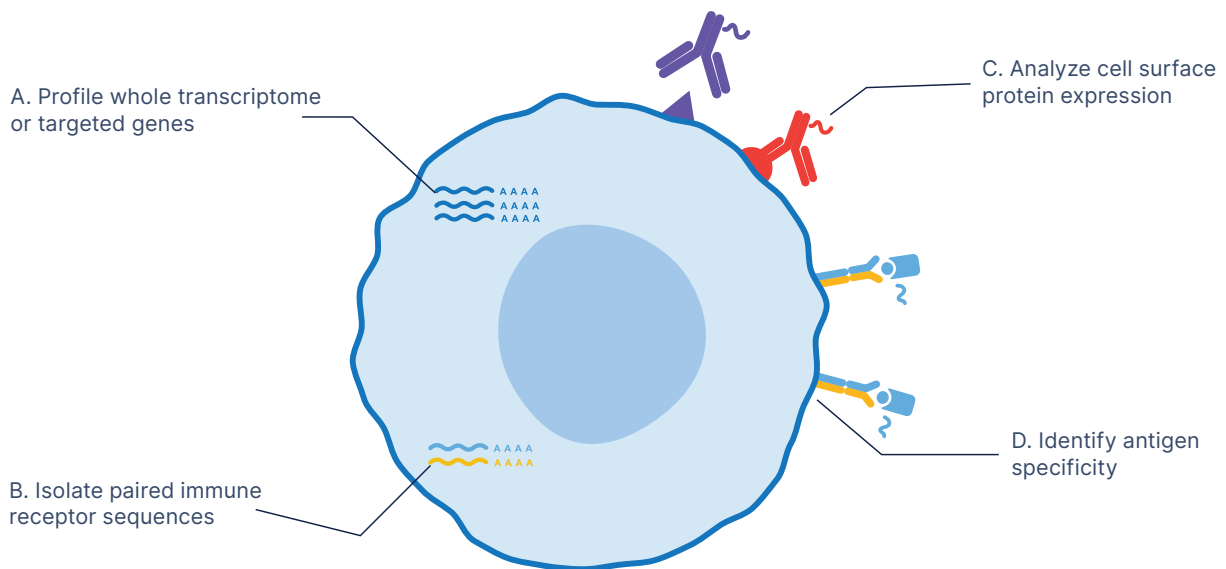

Figure 1: Fully characterize the adaptive immune response with multiomic profiling—(A) Analyze mRNA or targeted gene expression for hundreds to tens of thousands of cells per reaction using Chromium Single Cell Immune Profiling. (B) Identify distinct clonotypes through corresponding paired, full-length immune receptor sequences. (C) Measure up to hundreds of cell surface proteins at single-cell resolution with Feature Barcode technology. (D) Screen antigen specificity of B-cells or T-cells using Feature Barcode technology. Methods for C and D require compatible partner products (details on page 3).

The workflow begins with the Chromium Next GEM Single Cell 5' Kit v2, in which user-generated dissociated cells are partitioned individually in emulsions, or GEMs (Gel Bead-in-emulsion), within the Chromium instrument. This is followed by lysis and cDNA generation inside the GEM. Post reverse transcription, emulsions are broken and barcoded cDNA is amplified in bulk. Simultaneously, oligos conjugated to barcoded antibodies[*] specific to cell surface proteins are amplified. If measuring antigen specificity, then barcodes from MHC-peptide multimers are amplified at this step.[†] The amplified cDNA will generate sufficient material to enrich for full-length B-cell receptor (BCR) or T-cell receptor (TCR) sequences.

Starting with immune cells, users can generate as many as four library types, including gene expression, BCR, TCR, and 10x Genomics Feature Barcode libraries, for sequencing across various Illumina instruments, including the NovaSeq™ 6000, NextSeq™ 2000, NextSeq 1000, or NextSeq 500/550 Systems. Feature Barcode technology enables analysis of surface protein expression and antigen specificity from a single cell. Data analysis is performed using the Cell Ranger pipeline that allows users to link data from gene expression libraries with data from corresponding BCR or TCR libraries. Loupe Browser software makes it easy to visualize and explore single-cell immune profiling data.

---

* BioLegend TotalSeq reagents combine gene expression analysis with detection of hundreds of cell surface proteins at high resolution for ultrahigh parameter multiomic cytometry.

† Immudex dCODE Dextramers provide oligo-conjugated MHCs for T-cell profiling.

Prepare samples

Sample preparation starts with generating a suspension from cell culture, primary cells, or fresh tissue.[‡] High-quality single-cell suspensions are critical for good results. Demonstrated protocols for sample preparation from various sample types are available on the 10x Genomics Support website.[5]

Generate single-cell libraries

Once single-cell suspensions are isolated, they are ready for library preparation using the Chromium Next GEM Single Cell 5' Kit v2. The single-cell suspensions are loaded onto a microfluidic chip and run in the Chromium Controller, Chromium X/iX, or Chromium Connect. These systems partition individual cells into droplets with a single gel bead that contains a unique barcode. The droplets, or GEMs, are then incubated to attach barcodes to the mRNA which is reverse transcribed into cDNA (Figure 3).

Following this incubation, GEMs are broken and pooled fractions are recovered and purified. These fractions are taken through a pre-amplification PCR step to amplify 10x Genomics barcoded, full-length cDNA fragments. Amplification generates sufficient material to construct multiple libraries from the same cells, eg, B-cell or T-cell libraries and 5' gene expression libraries.

---

‡ This assay has been validated on peripheral blood mononuclear cells (PBMCs), and various challenging cell types. This assay can be applied to other sample types including fresh tissue but is not compatible with nuclei isolated from fresh or frozen tissue.
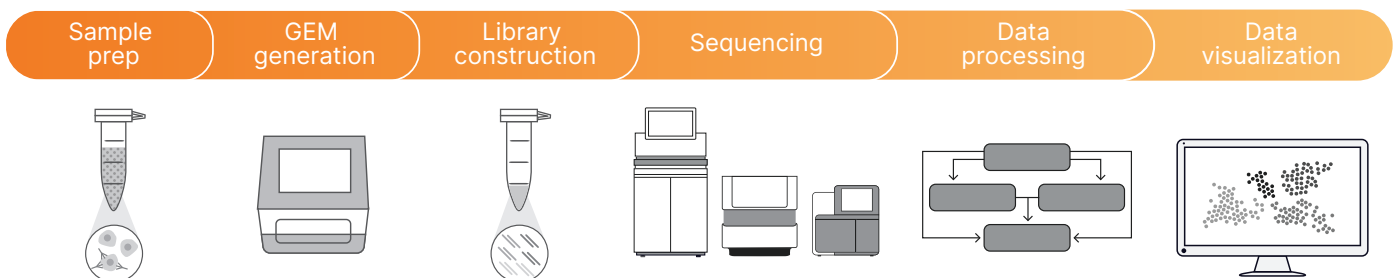

Figure 2: Workflow for multiomic profiling of the immune system—Cell suspensions undergo microfluidic single-cell partitioning and barcoding into GEMS on a Chromium system. Resulting V(D)J and gene expression libraries are sequenced on an Illumina instrument. Data are analyzed and visualized using Cell Ranger and Loupe Browser software, respectively.
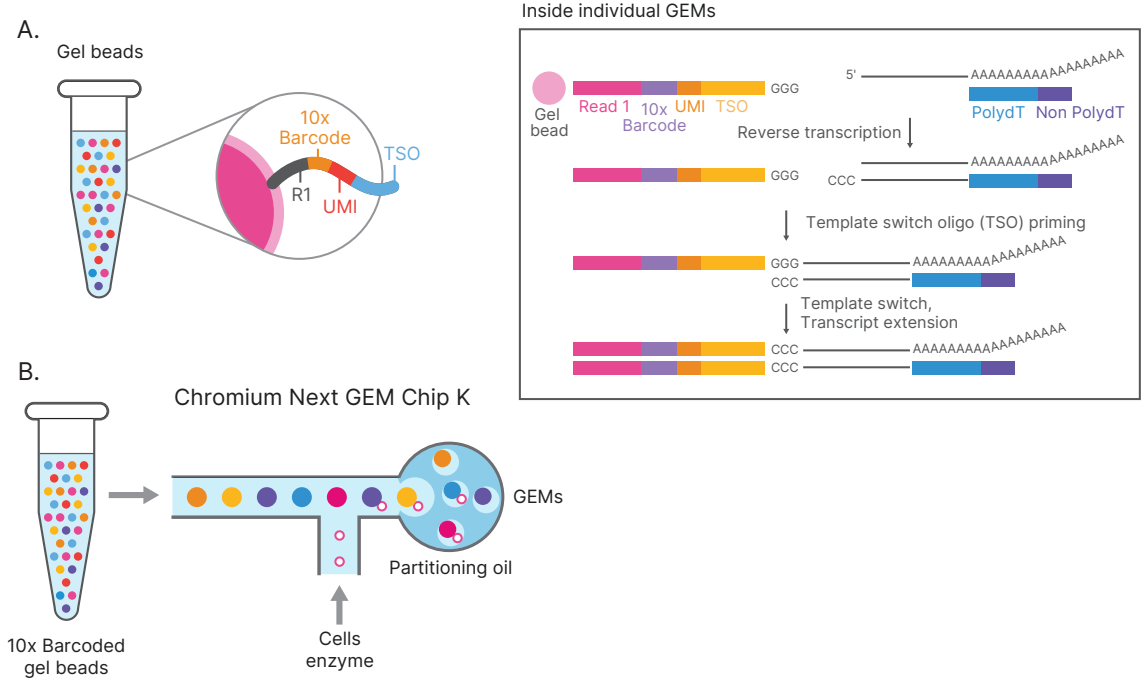
.

Figure 3: Single-cell suspensions are labeled with the 10x barcode—(A) Single cells are partitioned into individual GEMs on the 10x Chromium system. (B) Inside the GEM, mRNA is isolated and the GEM barcodes attached. The labeled mRNA is reverse transcribed and the resulting cDNA is ready for use in library preparation.

The amplified cDNA is used to amplify full-length V(D)J segments (10x barcoded) via PCR amplification with primers specific to either the BCR or TCR constant regions. If both B- and T-cells are expected to be present in the partitioned cell population, immunoglobulin (Ig) and TCR transcripts can be amplified in separate reactions using the same amplified cDNA material. Enzymatic fragmentation and size selection are used to generate variable length fragments that collectively span the V(D)J segments of the amplified BCR or TCR transcripts prior to library construction.

Amplified cDNA and V(D)J segments are used as input for library construction. The resulting barcoded single-cell immune profiling gene expression and V(D)J libraries are now ready for sequencing using Illumina NGS sequencing systems (Figure 4).

Sequence with Illumina instruments

To accommodate the sequencing output required for this application, we recommend sequencing the paired V(D)J, CSP, and gene expression libraries on a NovaSeq 6000, NextSeq 2000, NextSeq 1000, or NextSeq 500/550 System (Table 1). For V(D)J, CSP, and gene expression libraries, 10x Genomics recommends to spike in 1% PhiX. The spike-in acts as a control for the run and provides information regarding error rate metrics. Expected sequencing metrics for these libraries are available on the 10x Genomics Support website.
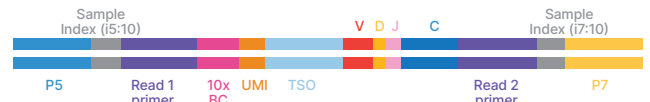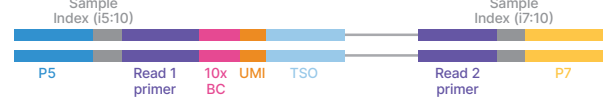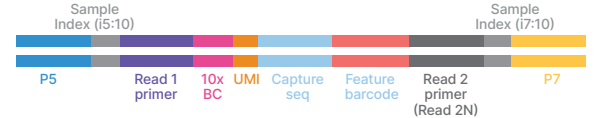
Figure 4: Configurations of single-cell libraries ready for sequencing—Chromium Single Cell 5' Gene Expression, V(D)J, and Cell Surface Protein Dual Index libraries comprise standard paired-end constructs compatible with Illumina sequencing systems. Chromium Single Cell 5' Gene Expression and V(D)J Single Index libraries (not shown) are also compatible with Illumina systems. BC, barcode; TSO, template switch oligo; UMI, unique molecular identifier

Sequencing depth and number of samples per sequencing run

Each standard assay run on the 10x Genomics Chromium Controller, Chromium X/iX, or Chromium Connect can target up to 10K cells, the high-throughput assay run on the Chromium X can target up to 20K cells.[6] Calculations regarding the number of libraries that can be pooled in a given sequencing run are based on an average of 5000 cells per sample (Table 1). Users will need to adjust the number of libraries per run based on the cell recovery.

The minimum sequencing depth is 20K reads pairs/cell for gene expression libraries, and 5000 read pairs/cell for V(D)J or Cell Surface Protein (CSP) libraries.[7] Adjust sequencing depth for the required performance or application. Specific sample types may require more reads for V(D)J (ie, plasma B cells) or gene expression (ie, cell lines) studies.

Choosing an Illumina sequencing system, illumina.com/systems/sequencing-platforms

Table 1: Example sample throughput for 10x Chromium Single Cell libraries on Illumina sequencing systems

10x library type	Sequencing depth (read pairs per cell)[a]	Library pooling ratio[b]	Average number of cells per sample[c]	Number of samples per run by system and flow cell type									
				NextSeq 550		NextSeq 2000			NovaSeq 6000			
				Mid output	High output	P1[d]	P2[d]	P3	SP	S1	S2	S4
Single reads per flow cell type				130M	400M	100M	400M	1.2B	650M	1.3B	3.3B	8B
V(D)J or CSP	5000[e]	n/a	5000	5	16	4	16	48	26	52	132[g]	320[g]
V(D)J + CSP	5000 (V(D)J)[e]	1	5000	2	8	2	8	24	13	26	66[g]	160[g]
	5000 (CSP)[e]	1	5000	2	8	2	8	24	13	26	66[g]	160[g]
V(D)J or CSP with 5' GEx	5000 (V(D)J or CSP)	1	5000	5[f]	8	4[f]	8	24	12	24	64	152[g]
	20,000 (GEx)[e]	4	5000	OR 1[f]	2	OR 1[f]	2	6	3	6	16	38[g]
V(D)J + CSP + 5' GEx	5000 (V(D)J)	1	5000	Sequencing system with higher throughput is required			4	16	8	16	40	100[g]
	5000 (CSP)	1	5000				4	16	8	16	40	100[g]
	20,000 (GEx)[e]	4	5000				1	4	2	4	10	25[g]

The number of single cell samples per sequencing runs is based on an Illumina PhiX control library at supported cluster densities/loading concentration. Actual performance parameters may vary based on sample type, sample quality, and clusters passing filter. See specifications page for each instrument for details. CSP, cell surface protein; GEx, gene expression; n/a, not applicable

a. Minimum read recommendations courtesy of 10x Genomics. Adjust sequencing depth for the required performance or application. The sequencing saturation metric and curve in the 10x Cell Ranger run summary can be used to optimize sequencing depth for specific sample types.

b. V(D)J, CSP, and 5' GEx libraries may be pooled for sequencing, taking into account the differences in depth requirements between the pooled libraries. Library pooling ratio: library can be brought to the same concentration and then pooled at a 1 :4 ratio (volume V(D)J:volume GEx or volume CSP:volume GEx). It is also possible to pool them at a 1:10 ratio volume V(D)J:volume GEx. In that case, the number of samples per run must be adjusted.

c. See "sequencing depth" and "number of samples per run" to understand the maximum number of cells that can be recovered, depending on which Chromium system is used.

d. P1 and P2 flow cells with the same sample throughput are available on the NextSeq 1000 System.

e. For V(D)J and CSP libraries, 5000 read pairs (or 10,000 individual reads) means 5000 clusters, or 5000 reads from Read 1 and 5000 reads from Read 2. For Gene Expression libraries, 20,000 read pairs (or 40,000 individual reads) means 20,000 clusters, or 20,000 reads from Read 1 and 20,000 reads from Read 2.

f. Users can run the indicated number of V(D)J/CSP libraries OR one gene expression library. If running two or three library types, use of a higher output flow cell is recommended.

g. Requires use of the NovaSeq XP workflow, which allows for individual lane loading.

Recommended read length

The recommended read length for sequencing 10x Genomics Dual Index V(D)J and Chromium GEx libraries is 26 bp for Read 1 (10x barcode, unique molecular identifier (UMI)), 10 bp for index 1, 10 bp for index 2, and 90 bp for Read 2 (Table 2). Shorter reads can lead to decreased application performance. In particular, Read 2 length is critical for spanning the V(D)J junctions. If CSP libraries are sequenced on their own, it is recommended to use read lengths of 26 × 10 × 10 × 25 bp given the lack of diversity after ~ 25 bases. If CSP libraries are sequenced with V(D)J and/or GEx libraries, recommended read lengths are 26 × 10 × 10 × 90 bp. Sample index reads must not be shorter than indicated; however, any read can be longer. Additional bases in sample index reads must be trimmed using Cell Ranger mkfastq or Illumina bcl2fastq prior to further analysis.

In most cases, the standard recommendations (5000 read pairs per targeted cell in a 26 × 90 bp read length configuration) for V(D)J and CSP Dual Index libraries are sufficient. Some sample types do not reach saturation at these sequencing depths and may require deeper sequencing. The most common cause for this is extreme variation in expression levels between cells in a sample; most sequencing data will represent high-expressing cells. If recovery of low-expressing cells is critical to the experiment design and the sample type is prone to high variation, use a higher sequencing depth.

Designing an experiment, 10x Genomics Support website

Sequencing configurations, 10x Genomics Support website

Analyze and visualize data

Analysis of gene expression, V(D)J, and CSP libraries is compatible with standard Illumina FASTQ generation pipelines. Following sequencing, the complete run folder of demultiplexed FASTQ files can be input into the Cell Ranger software. Cell Ranger produces count matrices for the gene expression and cell surface protein libraries and performs annotation and clonotype grouping for the V(D)J library. With the current version of Cell Ranger (6.1.1), users can carry out the analysis all at once with the multi pipeline (recommended) or separately using the count and V(D)J pipelines. Cell Ranger outputs web summaries containing QC metrics that can assess experimental and sequencing success.

Additional output files, such as .cloupe and .vloupe files, can be used for secondary analysis with Loupe Browser software. Here, users can remove poor quality cells/ outliers, recluster the data, carry out differential gene expression analysis, and overlay the V(D)J and cell surface protein information on the gene expression data. Third-party tools can be used for additional analysis.

Data highlights

Loupe V(D)J Browser allows users to analyze, search, and visualize V(D)J sequences and clonotypes from single-cell 5′ data produced by the Chromium platform (Figure 5). Use Loupe V(D)J Browser for quickly exploring the clonality and diversity of BCRs and TCRs at the single-cell level. Within Loupe V(D)J Browser, users can quickly determine the most common paired α and β TCR chains or heavy and light BCR chains within a sample, search for motifs within CDR3 regions, detect variants within V(D)J transcripts, and compare clonotype frequency between samples (Figure 6).

Table 2: Recommended read lengths for 10x Genomics libraries

Library type	Read 1 (bp)	i7 index (bp)	i5 index (bp)	Read2 (bp)
Purpose	10x Barcode, UMI	Sample index	Sample index	Insert or cDNA
V(D)J and GEx Single Cell 5' v2 Dual Index	26	10	10	90
CSP v2 Dual Index	26	10	10	25
CSP + V(D)J or CSP + GEx or CSP (Dual Index) + V(D)J + GEx	26	10	10	90
V(D)J and GEx Single Cell 5' (v1/v1.1) Single Index	26	8	0	91

A. Gene expression t-SNE
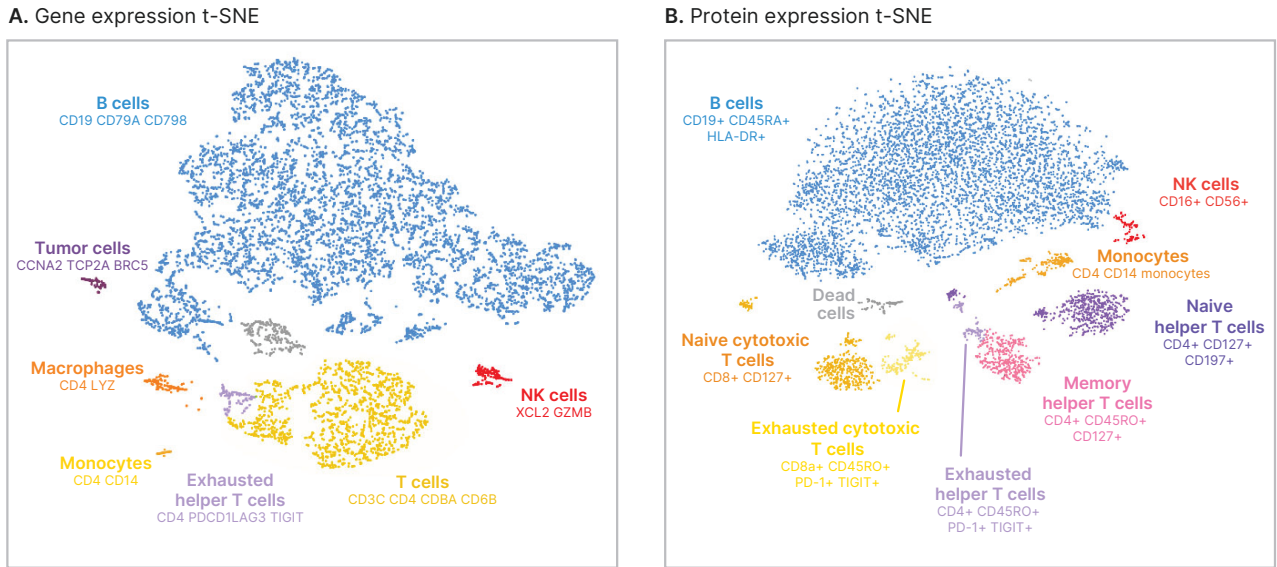
B. Protein expression t-SNE

Figure 5: Multiomic data provides complementary insights into diverse immune cell populations in a complex cancer sample—(A) t-SNE plot of 7859 cells from a dissociated human melanoma sample. Each cell is represented by a single dot. Broad categories of immune cell populations were identified through gene expression signatures, including cell populations that were not identified by protein markers such as macrophages (orange) and a small population of tumor cells (purple). (B) t-SNE projection of the same sample cluster based on cell surface protein expression using a panel of 16 markers. Analysis of cell surface protein expression enabled further classification of T cells into naïve, memory, and exhausted subpopulations.

A. Antigen binding specificities
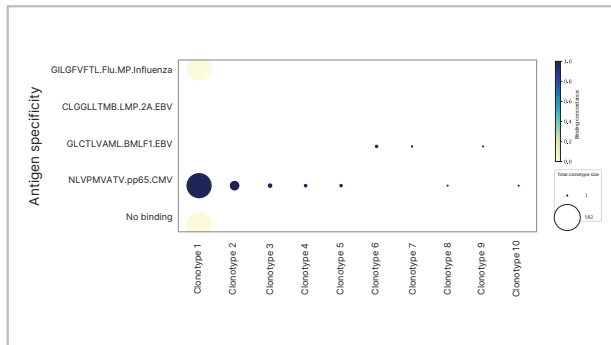
B. TCR clonotype frequency
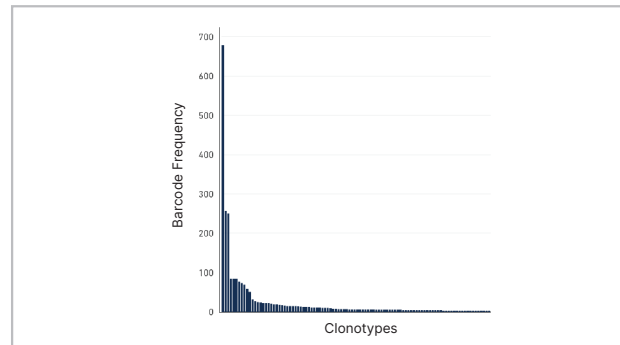
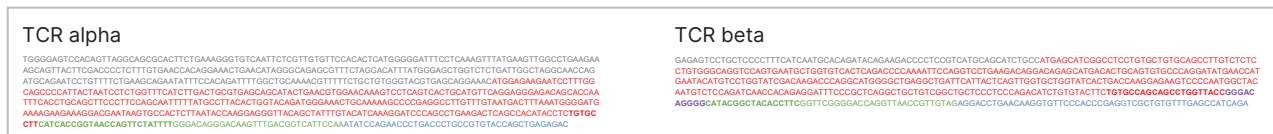**C.** Full-length sequences of clonotype 1: TCR alpha and beta chains

Figure 6: Characterization of antiCMV T-cells from a CMV+ donor by antigen-binding specificity, clonotype frequency, and paired, full-length T-cell receptor sequences—(A) The ten largest antigen-binding clonotypes are plotted along with their antigen specificity and binding concordances, as defined by dCODE Dextramer Reagent binding. "No binding" is defined as a clonotype showing no binding to peptide–MHC (pMHC) on the y-axis. The presence of a circle indicates that at least one member of the clonotype was specific for a particular pMHC. Circle size indicates the total number of cells for the given clonotype. Circle color indicates the proportion of cells within the clonotype that bind the specific pMHC, with darker colors having higher binding concordance. (B) TCR clonotypes are ranked by frequency. The top CMV-specific clonotype has 678 cells. (C) Shown are paired, full-length TCRα and β chain V(D)J sequences. V(D)J nucleotides are color coded as follows: 5' UTR (gray), V (red), D (purple), J (green), C (blue), and CDR3 (bold).

Access expert support

For sequencing Chromium Single Cell Immune Profiling libraries, the Illumina and 10x Genomics teams collaborate to ensure you are fully supported throughout the workflow. Contact 10x Genomics Support (support@10xgenomics. com) for assay and analysis questions, and Illumina Support (techsupport@illumina.com) for sequencing questions. The teams are also equipped to handle more complex issues together.

Summary

This protocol enables simultaneous profiling of gene expression, BCRs, TCRs, antigen specificity, and cell surface proteins at single-cell resolution. Using a multiomic approach can help reveal cellular mechanisms driving adaptive immune system response and gene regulation, including differences in healthy and disease states.

Learn more

NovaSeq 6000 System, illumina.com/systems/ sequencingplatforms/novaseq

NextSeq 1000 and NextSeq 2000 Systems, illumina.com/ systems/sequencing-platforms/nextseq-1000-2000

NextSeq 500/550 System, illumina.com/systems/ sequencingplatforms/nextseq

10x Genomics Chromium Single Cell Immune Profiling, 10xgenomics.com/products/single-cell-immune-profiling

References

1. Illumina. Immunology Research Review. illumina.com/content/ dam/illumina-marketing/documents/publications/ scientific-publication-review-immunology-research.pdf. Accessed October 28, 2021.
2. Vahedi G, Takahashi H, Nakayamada S, et al. STATs shape the active enhancer landscape of T cell populations. *Cell*. 2012;151(5):981-993. doi:10.1016/j.cell.2012.09.044
3. Venturi V, Price DA, Douek DC, Davenport MP. The molecular basis for public T-cell responses?. *Nat Rev Immunol*. 2008;8(3):231-238. doi:10.1038/nri2260
4. Woodsworth DJ, Castellarin M, Holt RA. Sequence analysis of T-cell repertoires in health and disease. *Genome Med*. 2013;5(10):98. Published 2013 Oct 30. doi:10.1186/gm502
5. 10x Genomics. Sequencing Requirements for Single Cell V(D) J. support.10xgenomics.com/single-cell-vdj/sequencing/doc/ specifications-sequencing-requirements-for-single-cell-vdj. Updated August 26, 2020. Accessed October 28, 2021.
6. 10x Genomics. What is the maximum number of cells that can be profiled?. kb.10xgenomics.com/hc/en-us/articles/ 360001378811-What-is-the-maximum-number-of-cells-that-can-be-profiled-. Accessed October 28, 2021.
7. 10x Genomics. Experimental Design for Immune Profiling. support.10xgenomics.com/single-cell-vdj/software/pipelines/ latest/design. Accessed October 28, 2021.

Ordering information

Sequencing reagents	Catalog no.
NovaSeq 6000 System reagents	
NovaSeq 6000 SP Reagent Kit v1.5 (100 cycles)	20028401
NovaSeq 6000 S1 Reagent Kit v1.5 (100 cycles)	20028319
NovaSeq 6000 S2 Reagent Kit v1.5 (100 cycles)	20028316
NovaSeq 6000 S4 Reagent Kit v1.5 (200 cycles)	20028313
NextSeq 1000 and NextSeq 2000 System reagents	
NextSeq 2000 P1 Reagents (300 cycles)	20050264
NextSeq 1000/2000 P2 Reagents v3 (100 cycles)	20046811
NextSeq 2000 P3 Reagents (100 cycles)	20040559
NextSeq 500/550 System reagents	
NextSeq 500/550 Mid Output Kit v2.5 (150 cycles)	20024904
NextSeq 500/550 High Output Kit v2.5 (150 cycles)	20024907

illumına®