Single-cell RNA-seq analysis report

Example deliverable · public dataset (10x Genomics PBMC 3k)
Prepared by OMICSDESK · 2026-07-31

What this document is. This is the report format a client receives, produced from a public dataset so that it can be shared. Every number was computed by the pipeline that generated it; the script is published at omics.sinogenomics.com/code and re-running it reproduces this document.

1 · Summary

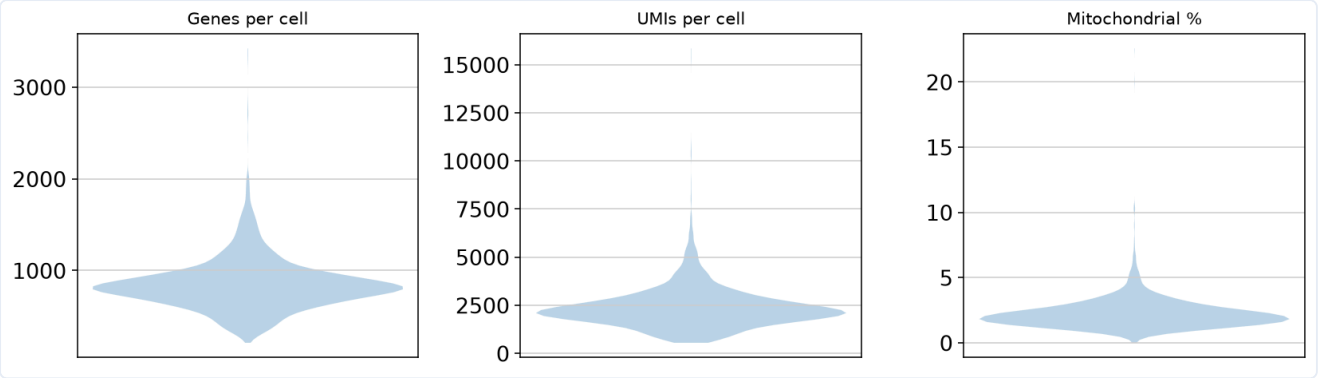
2,700	2,638	817	2.03%
cells in raw matrix	after QC (−2.3%)	median genes/cell	median mitochondrial

6	0.0%
clusters	unassigned cells

2,638 cells passed quality control and resolved into 6 populations, all of which could be assigned to known cell types by two independent methods. No population was left unannotated.

2 · Quality control

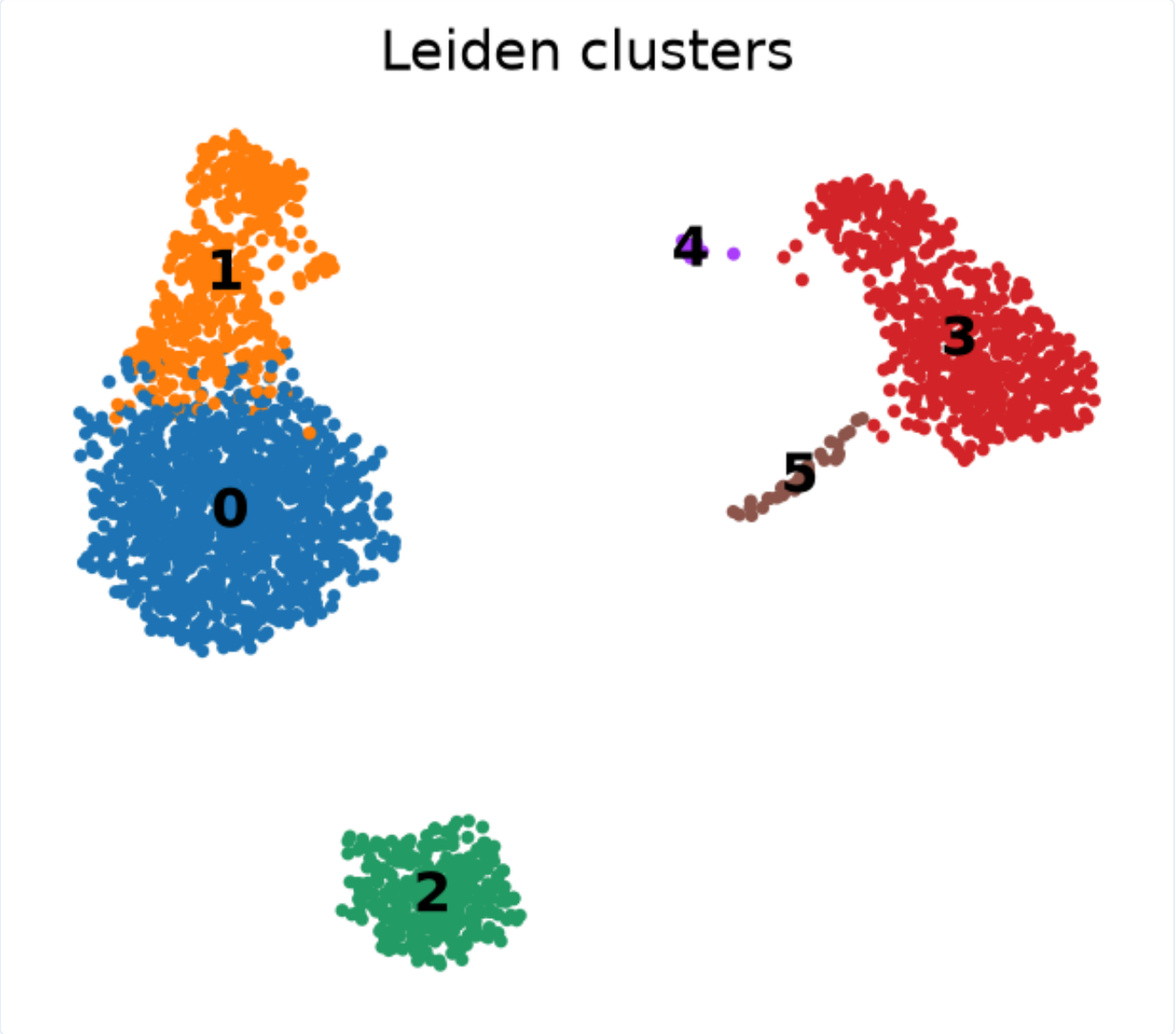
Per-cell distributions before filtering. Cells with fewer than 200 detected genes, more than 2,500 detected genes or more than 5% mitochondrial reads were removed.



METRIC	VALUE
Cells before / after QC	2,700 → 2,638 (2.3% removed)
Genes retained	13,714 of 32,738
Median genes per cell	817
Median UMIs per cell	2,197

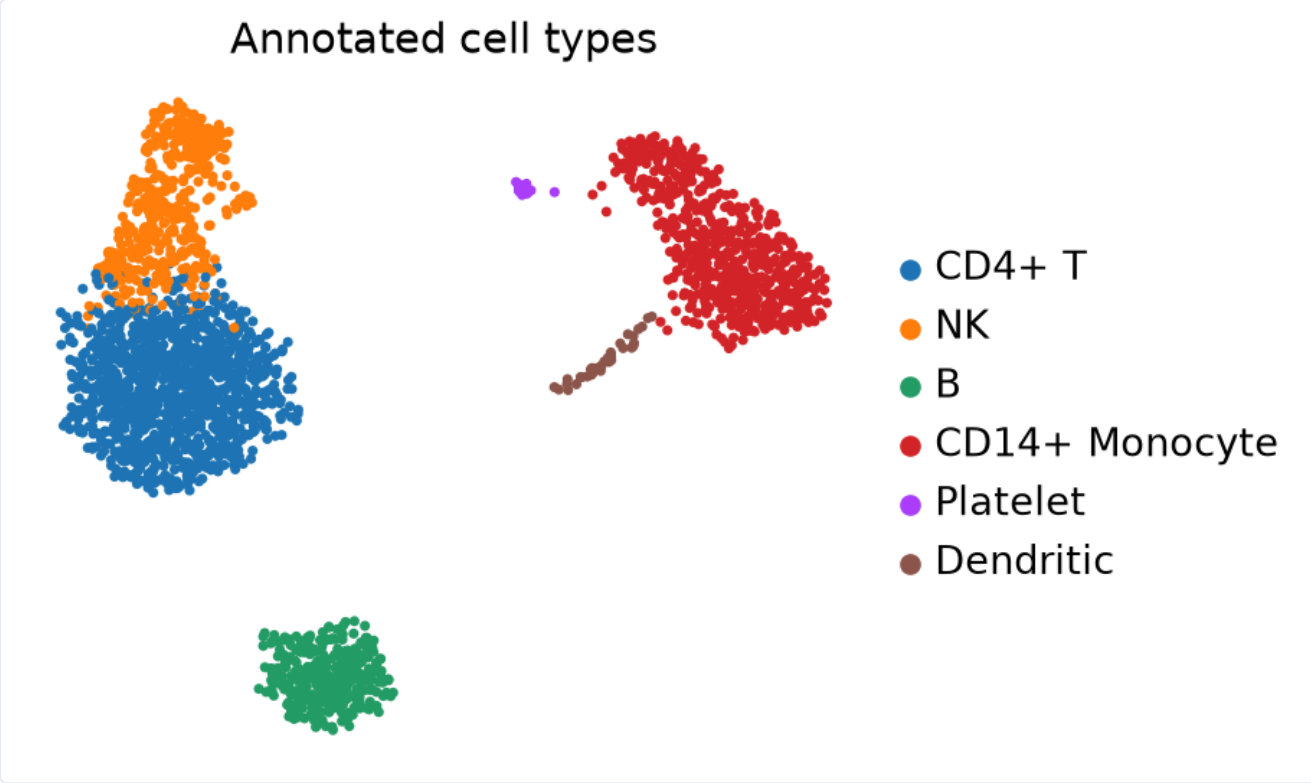
Median mitochondrial fraction	2.03%
Highly variable genes selected	1,838

3 · Clustering



4 · Cell-type annotation

Clusters were annotated from canonical marker panels and then checked against CellTypist Immune_All_Low (public), run without access to the manual labels. The two agree on 6 of 6 clusters at lineage level.

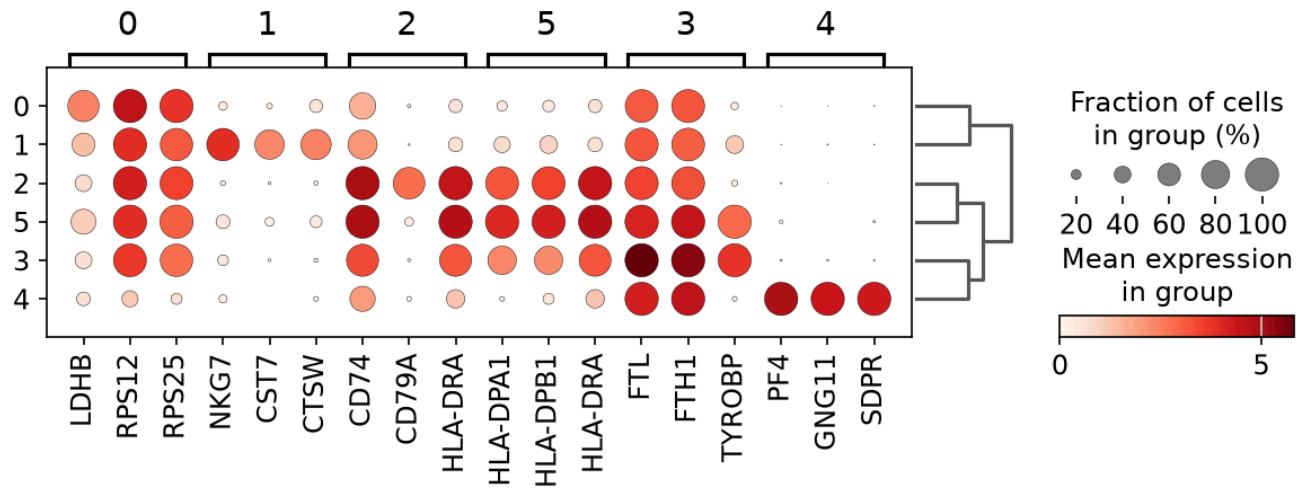


CELL TYPE	CELLS	SHARE
CD4+ T	1,193	45.2%
CD14+ Monocyte	635	24.1%
NK	422	16.0%
B	340	12.9%
Dendritic	35	1.3%
Platelet	13	0.5%

Independent cross-check

CLUSTER	CELLS	MARKER-BASED	AUTOMATED (BLIND)	CONSISTENCY
0	1,193	CD4+ T	Tcm/Naive helper T cells	100.0%
1	422	NK	CD16+ NK cells	100.0%
2	340	B	B cells	100.0%
3	635	CD14+ Monocyte	Classical monocytes	100.0%
4	13	Platelet	Megakaryocytes/platelets	100.0%
5	35	Dendritic	DC	100.0%

5 · Marker genes



CLUSTER	TOP MARKERS
0	LDHB, RPS12, RPS25, RPS3, RPS27, RPS6
1	NKG7, CST7, CTSW, B2M, GZMA, CCL5
2	CD74, CD79A, HLA-DRA, CD79B, HLA-DPB1, HLA-DQA1
3	FTL, FTH1, TYROBP, LYZ, CST3, LST1
4	PF4, GNG11, SDPR, PPBP, NRGN, GPX1
5	HLA-DPA1, HLA-DPB1, HLA-DRA, HLA-DRB1, HLA-DQA1, CD74

6 · Methods (ready to paste)

Single-cell RNA-seq data were processed with Scanpy. Cells with fewer than 200 detected genes, more than 2,500 detected genes, or more than 5% of counts from mitochondrial genes were excluded, leaving 2,638 of 2,700 cells. Counts were normalised to 10,000 per cell and log-transformed. 1,838 highly variable genes were selected (mean expression 0.0125–3, dispersion > 0.5); total counts and mitochondrial fraction were regressed out and the data scaled to unit variance with values clipped at 10. Principal component analysis was computed on the scaled data and a neighbourhood graph built from the first 30 components (k = 10). Cells were embedded with UMAP and clustered with the Leiden algorithm (resolution 0.5), yielding 6 clusters. Cluster marker genes were identified with a Wilcoxon rank-sum test. Cell types were assigned from canonical markers and independently verified with CellTypist Immune_All_Low (public). Analysis code and package versions accompany this report.

7 · What was delivered

- This report (PDF and HTML).
- Publication-ready figures at 300 dpi, each regenerable from the script.
- Results tables: per-cluster markers, cell-type composition, per-sample QC metrics.

- The processed object (h5ad / Seurat) for downstream work.
- The analysis script and environment specification.

> **Statement of limitations.** This example uses a single healthy-donor sample, so no between-group comparison was possible. For studies comparing conditions, differential expression is performed on pseudobulk profiles aggregated per subject rather than per cell, and cell-type abundance is tested at subject level — see omics.sinogenomics.com/sample-pseudobulk and /sample-abundance.
