

Sep 15, 2025

Bioinformatic methods for GRC scRNAseq poster

DOI

<https://dx.doi.org/10.17504/protocols.io.yxmvmmbpm9g3p/v1>

Hannah Aichelman¹

¹University of Rochester



Hannah Aichelman

University of Rochester

Create & collaborate more with a free account

Edit and publish protocols, collaborate in communities, share insights through comments, and track progress with run records.

Create free account

OPEN  ACCESS



DOI: <https://dx.doi.org/10.17504/protocols.io.yxmvmmbpm9g3p/v1>

Protocol Citation: Hannah Aichelman 2025. Bioinformatic methods for GRC scRNAseq poster . **protocols.io**
<https://dx.doi.org/10.17504/protocols.io.yxmvmmbpm9g3p/v1>

License: This is an open access protocol distributed under the terms of the **Creative Commons Attribution License**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

Protocol status: Working

We use this protocol and it's working

Created: September 15, 2025



Last Modified: September 15, 2025

Protocol Integer ID: 227312

Keywords: bioinformatic methods for grc scrnaseq, typical benchmarking with peripheral blood mononuclear cell, uniform bioinformatics pipeline, 10x genomics flex, grc scrnaseq, 10x genomics ocm, acute myeloid leukemia, derived acute myeloid leukemia, cell rna, scrna, resolution profiling of complex tissue, cells with low gene detection, diverse cell population, sequencing depth, low gene detection, illumina single cell, typical benchmarking, peripheral blood mononuclear cell, number of gene

Abstract

Background: Single-cell RNA sequencing (scRNA-seq) enables high-resolution profiling of complex tissues and has advanced our understanding of cellular heterogeneity. Many platforms are available, each with trade-offs in sensitivity, accuracy, and throughput.

Methods: We evaluated five scRNA-seq platforms—10X Genomics GEM-X, 10X Genomics Flex, 10X Genomics OCM, Parse Biosciences, and Illumina Single Cell—using manufacturer protocols and a uniform bioinformatics pipeline. Unlike typical benchmarking with peripheral blood mononuclear cells, we used patient-derived acute myeloid leukemia (AML) apheresis samples to evaluate performance on clinically relevant, heterogeneous material. While basic QC metrics were evaluated, we focus here on the value of increased sequencing depth compared to vendor recommendations.

Results: As expected, differences in basic metrics were observed among the methods, although all platforms recovered diverse cell populations. While differences in cluster resolution were observed, most clusters were represented across all methods. Unique clusters, comprised of cells with low gene detection, were associated with a single platform, requiring further investigation. Surprisingly, although number of genes detected increased significantly, higher sequencing depth did not affect clustering, independent of platform.

Conclusions: Platform performance differed by metric, underscoring that no single technology is universally optimal. Overall, we found limited value with increased sequencing depth, although further analyses remain to be done. Cell typing and differential expression analysis may benefit from more sequencing although we identified nearly identical clusters independent of sequencing depth. Ultimately, selection of an optimal scRNA-seq platform depends on budget, biological system, and specific experimental goals.

Data processing methods

- 1 Separate single cell runs were processed using Seurat v4.1.0 within R v4.1.1. Samples were imported and initially filtered for genes represented within at least 3 cells and cells containing at least 200 unique genes. Fully sequenced runs were down-sampled to manufacturer recommendations using the SampleUMI function on data import.

The PercentageFeatureSet function was used to quantify mitochondrial and ribosomal content, pattern matching against "^MT-" and "RPS*|RPL*" respectively. Doublets were determined using scDbtFinder v1.8.0 using the combination of assay and whether or not it was the full or recommended depth run. Doublets and cells containing greater than 20% mitochondrial content were filtered for downstream clustering.

Only samples sequenced across all assays were used for clustering. Samples were normalized using Seurat's SCTransform function and initial dimensional reduction performed using RunPCA with the first 20 dimensions. Samples were integrated using Harmony v0.1.0 and clustering was performed using Seurat's FindNeighbors, FindClusters, and RunUMAP functions on the Harmony reduction. A resolution of 0.3 was used for FindClusters.

Plotting methods

- 2 Violin plots were generated using Seurat's VlnPlot function. Stacked bar charts, knee plot, and saturation curves were generated using ggplot2 v3.3.5