ISCVAM - Interactive Single Cell Visual Analytics tool for Multiomics

Thanh N Nguyen*[1], Zhihua Chen[1], Yijun Tian[2], Paulo C Rodriguez[3], Liang Wang[2], Xiaoqing Yu[1], Y. Ann Chen+[1,4]

1. Department of Biostatistics and Bioinformatics, 2. Department of Tumor Biology and 3. Department of Immunology, H. Lee Moffitt Cancer Center and Research Institute, Tampa, FL, 4. Huntsman Cancer Institute at the University of Utah
+Ann.Chen@hci.utah.edu

Background

In order to investigate tumor microenvironment (TME) heterogeneity at unprecedently resolution, the visual analytics of multimodal single cell platform is urgently needed for harnessing and synthesizing multiple sources of information.

ISCVAM Methods & Functionalities

ISCVAM was built with react.js for the web frontend, and node.js in the backend, with a data bridge to portable HDF5 storage. Commonly used R packages e.g., Seurat and Signac, were used for data processing and analyses.
The analyzed R object was converted into h5 format for integrative visualization on ISCVAM.

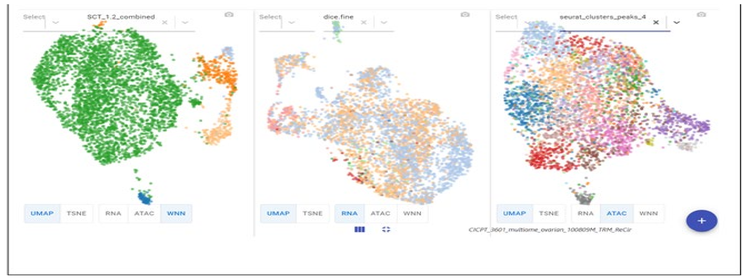
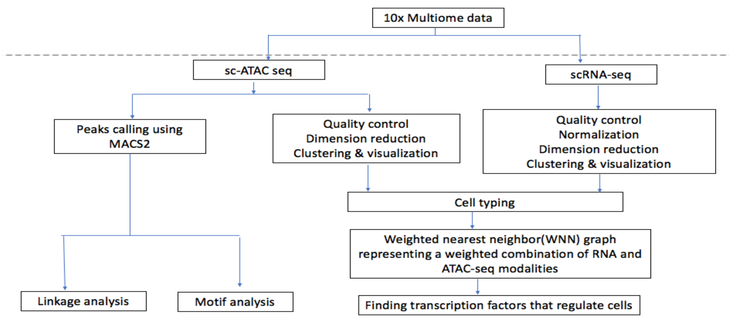


Figure 1: ISCVAM user interface. Visualization of data in multiple modalities (RNA, ATAC, WNN (weighted nearest neighbor) analysis) and different dimension reduction projections (tSNE & UMAP) with different referent panels in singleR package (e.g. MONACO, DICE) in independent panels in ISCVAM assists the identification of rare cell populations with unique RNA & ATAC peak expression profile.

Analysis pipeline



Proof-of-principle study

Discovery data

- One ovarian cancer patient with two samples of sorted 4,080 cells from tissue resident memory (TRM)-like and recirculating (ReCir) cells CD8+T cells (Anadon, Yu et al. 2022) was used as discovery data.
- This multiome dataset with paired single cell ATAC and gene expression was analyzed following the workflow, and the R object was converted into h5 format for interactive visualization on ISCVAM.
- We identified a rare cell population, MAIT, which was not reported in the original report
- We investigated the epigenetic landscape of MAIT (figure 5)

Internal validation

- Additional 3 patients with 6 samples from the same ovarian cohort

External validation

- ~11K cells from a PBMC sample (from 10X Genomics).

Proof-of-principle study (cont.)

In the discovery data, after debatching two samples using RNA assay and grouping cell clusters based on their marker genes, a total of 5 clusters were identified as seen in Figure 2A. Some exhaustion gene markers, such as CTLA4, ITGA2, CXCL13, are highly expressed in cluster 6, as reported in Anadon et al.
Cluster 11 has some specific gene markers, not shared with other clusters, including TLE1 and SLC4A10. We annotated this cluster 11 as MAIT cells, including some of the well-known markers of MAIT also expressed such as KLRB1, ZBTB16.

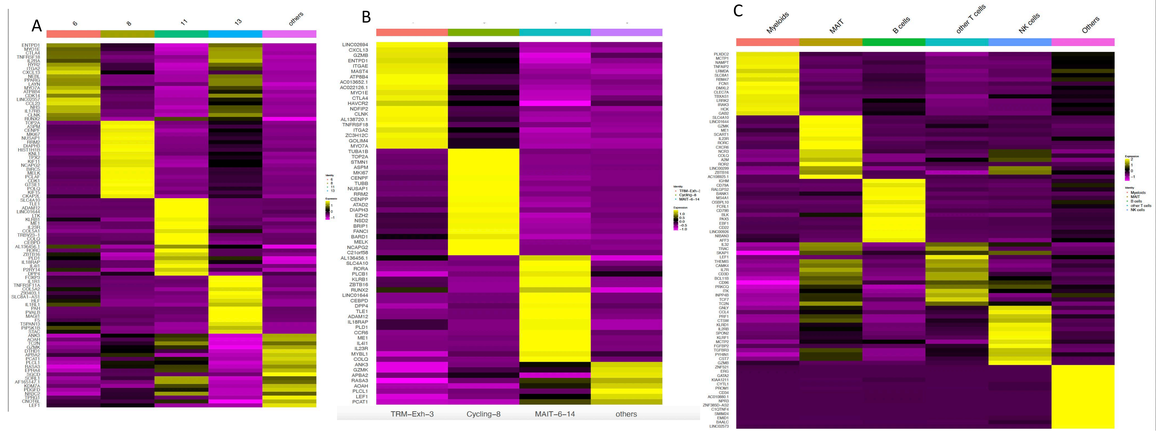


Figure 2: RNA markers heatmaps of 3 datasets from left to right: discovery dataset, internal validation and external validation. RNA markers of MAIT such as TLE1, SLC4A10 are expressed across all 3 datasets

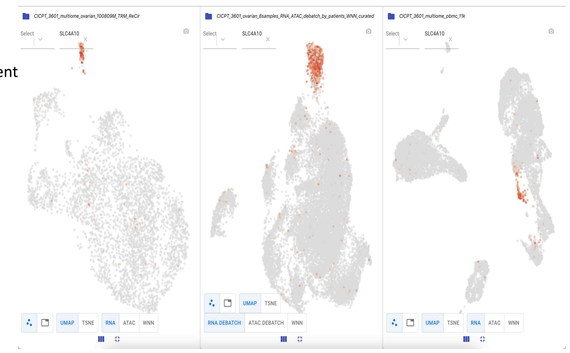


Figure 3: SLC4A10 gene expression across 3 different datasets visualized in ISCVAM

To identify the gene-peak relationship, we performed linkage analysis for SLC4A10 using Signac package. Briefly, LinkPeak function computes the correlation between gene expression and accessibility at nearby peaks.

Dataset	correlation score	Linked peak	zscore	pvalue
Discovery	0.0502	chr2-161426992-161427395	2.8577	0.00213373
	0.1313	chr2-161905529-161906149	9.3182	5.92E-21
Internal validation	0.3215	chr2-161901527-161906621	15.0021	3.56E-51
	0.2897	chr2-161911452-161917197	9.5261	8.17E-22
	0.1985	chr2-161900078-161900935	9.3931	2.91E-21
	0.1662	chr2-161907392-161909889	7.4164	6.02E-14
	0.1291	chr2-161708371-161709844	4.4796	3.74E-06
	0.1079	chr2-161898573-161899164	4.8290	6.86E-07
	0.0932	chr2-161919357-161920226	6.5425	3.03E-11
	0.0898	chr2-161570705-161571379	6.8350	4.10E-12
External validation	0.2597	chr2-161912196-161912507	23.9609	3.55E-127
	0.1667	chr2-161905346-161906011	10.2650	5.06E-25
	0.1306	chr2-161624831-161625031	15.5625	6.55E-55
	0.0961	chr2-161166641-161167253	4.8305	6.81E-07
	0.0583	chr2-161280713-161281178	6.5145	3.65E-11

Table 1: linkage analysis of gene SLC4A10 across 3 datasets. The linked peaks are in the nearby regions of chromosome 2

To find the transcription factor (TF) that may mediate MAIT cells, we performed motif enrichment analysis by first identified differentially accessible peaks, then identified enriched motifs in these top peaks using the function findMotifs in Signac package.

In the discovery dataset, among the enriched peaks of MAIT, we found one of our linked peak of gene SLC4A10, which is chr2-161,905,529-161,906,149 (table 1). Among the TF in this peak region, we noticed TF RORC is highly expressed with high specificity in MAIT group.

Our finding consistent with Cogswell et al. that MAIT cells can also express TF RORC making them distinct from other conventional CD4+ and CD8+ T cells.

The same analysis was performed on 2 validation datasets and showed the consistent results.

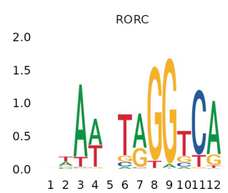


Figure 4: motif sequence logo of TF RORC

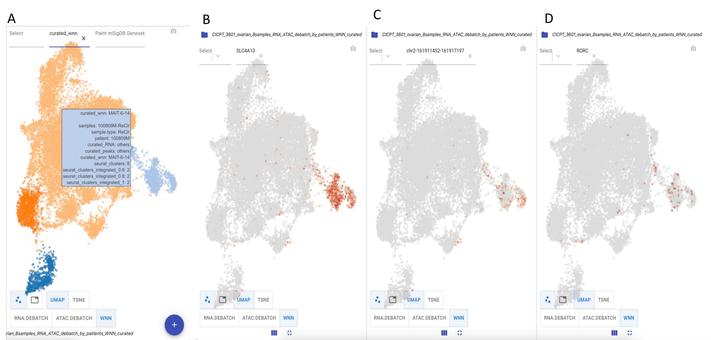


Figure 5: Trio regulatory relationship of internal validation data on ISCVAM
Figure 5A is cells painted by their cell type curation, the light blue cluster is annotated as MAIT cells with gene SLC4A10 is highly expressed in panel B, the expression of one of the linked peaks is shown in panel C, and the expression level of the enriched TF RORC (within this peak region) is shown in panel D.

Conclusion

We built ISCVAM as a visualization tool that:
- Extends to multimodality → explore regulatory relationships
- Incorporate external datasets → support in silico validation
- Multiple clustering resolutions → discover rare cell population

References

- Smalley et al. Single-Cell Characterization of the Immune Microenvironment of Melanoma Brain and Leptomeningeal Metastases. Clinical Cancer Research (2021)
- Anadon et al. *Ovarian cancer immunogenicity is governed by a narrow subset of progenitor tissue-resident memory T cell.* Cancer Cell (2022)
- Hao et al. *Integrated analysis of multimodal single-cell data.* Cell (2021)
- Stuart et al. *Single-cell chromatin state analysis with Signac*. Nature Methods (2021).
- Aran et al. *Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage*. Nat. Immunol. (2019).
- Cogswell et al. *MAIT Cells: Partners or Enemies in Cancer Immunotherapy? Cancers (*2021)

Acknowledgement

- Moffitt DRP Core Innovation Grant
- Moffitt Cancer Center, Quantitative Science Discovery Accelerator Award