Scupa: immune cell polarization assessment using the single-cell foundation model

Wendao Liu, Zhongming Zhao

MD Anderson Cancer Center UTHealth Houston Graduate School of Biomedical Sciences
McWilliams School of Biomedical Informatics at UTHealth Houston

Introduction

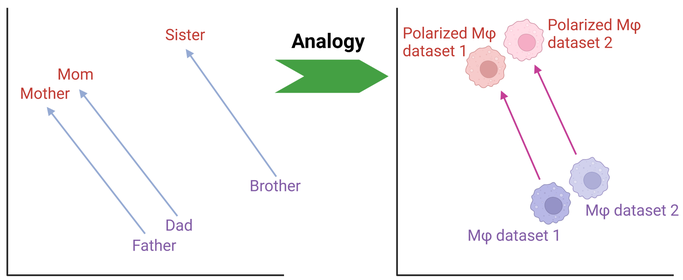
Immune cell polarization refers to the process by which immune cells adopt different functional states in response to various signals from the environment. The most famous example is macrophage polarization. M1 macrophages are pro-inflammatory and involved in defense against pathogens and tumor cells. M2 macrophages are anti-inflammatory and involved in tissue repair and remodeling.

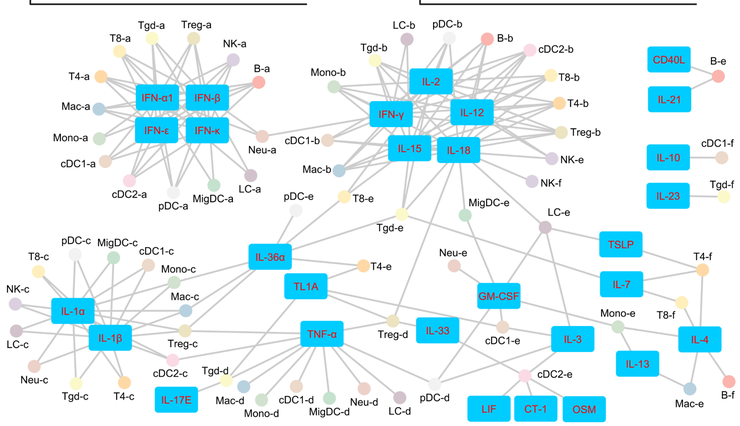
A recent study presented **Immune Dictionary**, a compendium of single-cell immune responses induced by 86 cytokines (Cui et al. 2024). The comprehensive cytokine-driven immune cell polarization states serve as a valuable reference for assessing immune cell polarization in other scRNA-seq studies.

The pretrained, transformer-based **single-cell foundation models** have demonstrated superior performance in multiple applications. We propose that it is possible to assess the immune cell polarization in scRNA-seq data using cell embeddings from the single-cell foundation model.

Word embeddings
Converting words to numeric vectors
Large language models
(e.g. GPT, BERT, LLaMA)
1. The embeddings of semantically similar words are close
2. Word relationships can be expressed in analogy vector operations

Cell embeddings
Converting cells to numeric vectors
Single-cell foundation models
(e.g. scGPT, scFoundation, UCE)
1. The embeddings of same cell types across datasets are close
2. Cell polarization can be expressed in analogy vector operations

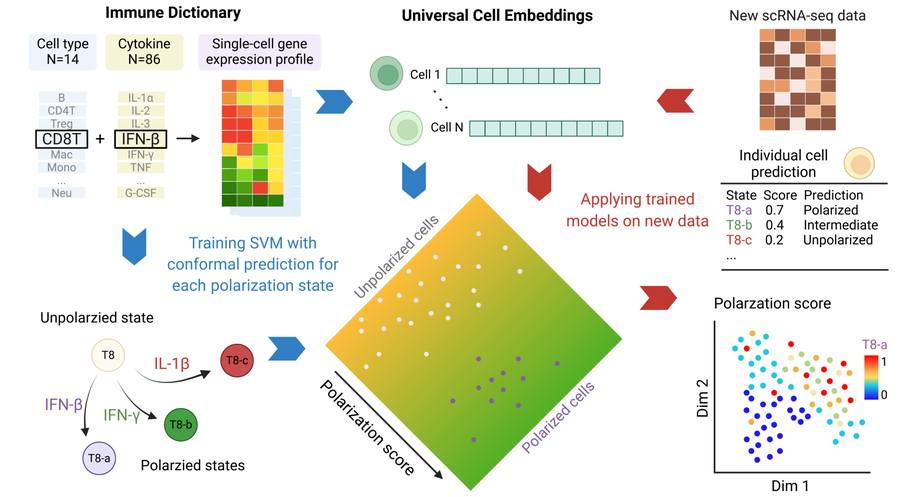




Cytokine-driven immune cell polarization states

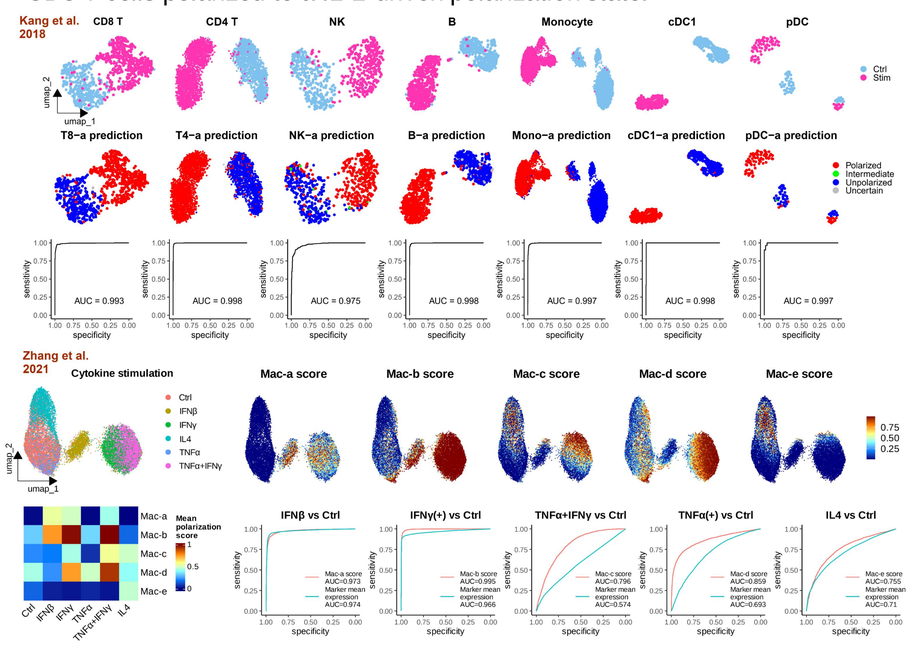
Method

We developed a R package: Scupa for Single-cell unified polarization assessment. We generated Universal Cell Embeddings (**UCE**) for the Immune Dictionary, and trained support vector machines to predict cell polarization. For new scRNA-seq data, Scupa outputs polarization scores and conformal predictions for individual cells.



Validation

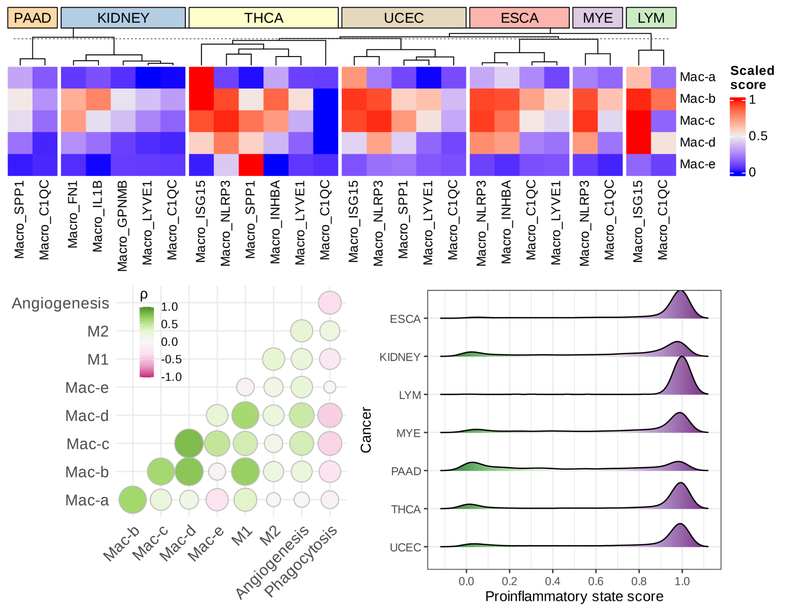
We validated Scupa using three independent cytokine-treatment datasets.
1. Human PBMCs treated with IFN-β *in vitro*: distinguishing cells between control and stimulated samples.
2. Human macrophages treated with four cytokines *in vitro*.
3. Mouse spleens treated with IL-2 *in vivo*: significantly higher percentage of CD8 T cells polarized to a IL-2-driven polarization state.



Application in cancer

We applied Scupa to a pan-cancer tumor infiltrating myeloid cell dataset and compared different polarization state scores in macrophage clusters.

Proinflammatory state score is defined based on Mac-b/c/d states to measure macrophage proinflammatory activity.



Conclusion

Scupa is a method for assessing immune cell polarization in scRNA-seq data, relying on Universal Cell Embeddings and using Immune Dictionary as reference. Scupa has been validated in *in vitro* and *in vivo* datasets and can be applied to immune cell scRNA-seq data.

Acknowledgements

Grant support: NIH R01LM012806, CPRIT RP240610.

Wendao Liu was supported by CPRIT BIG-TCR Predoctoral Training Program (RP210045) and John and Rebekah Harper Fellowship in Biomedical Sciences.

References

Liu W, Zhao Z. Scupa: single-cell unified polarization assessment of immune cells using the single-cell foundation model. Bioinformatics. 2025。
Cui A et al. Dictionary of immune responses to cytokines at single-cell resolution. Nature. 2024.
Rosen Y et al. Universal Cell Embeddings: A Foundation Model for Cell Biology. BioRxiv. 2023.