Cancer cells produce cytokines and chemokines that attract a diverse population of immune cells, including macrophages, neutrophils, and lymphocytes, although other cell types may also be present. However, persistent activation of the immune system and failure of the inflammatory response to resolve can lead to chronic inflammation, which promotes tumor growth.

The intricate interplay between tumor and immune cells in the microenvironment leads to the production of a wide variety of cytokines and growth factors that foster tumor cell proliferation, survival, and metastasis. The complex nature of this communication highlights the significant impact that immune cells have in the tumor microenvironment, with both pro-tumoral and anti-cancer roles.

Recent studies have shown that accounting for the heterogeneity of immune cell infiltration can result in more sensitive survival analyses and more accurate tumor subtype predictions. Ongoing research is focused on the role of infiltrating lymphocytes and other immune cells in the tumor microenvironment.

Myeloid cells such as macrophages, monocytes, dendritic cells, neutrophils, basophils, and eosinophils are frequently found in the stroma of various tumors. Their presence can affect the effectiveness of different types of cancer treatments.

Recent advances in single-cell RNA sequencing and flow cytometry have enabled the identification of various immune cell populations. As a result, numerous methodologies have been proposed to infer the proportions, or deconvolve, individual cell types from bulk RNA-seq samples.

However, several studies have highlighted the crucial role of factors such as data transformations, scaling/normalization, cell type composition, cell type specific marker selection and choice of methodology on the accuracy of deconvolution results.

We will evaluate the impacts of these different factors on the most recent deconvolution methods: MuSiC2, EPIC , BisqueRNA, deconvSeq, SCDC, FARDEEP, DeconvRNASeq (list subject to changes).

Our approach will start by identifying relevant single-cell RNA-seq datasets to extract cell type-specific expression values and cell-type proportions. We will search for these datasets on platforms such as Gene Expression Omnibus (GEO), The Single Cell Portal, or in specific past research. Alternatively, we may use deconvolution tools that provide both single cell and bulk RNA-seq datasets.

Next, we will obtain bulk tumor data for a specific cancer type from the Cancer Genome Atlas data portal (TCGA). Accurately determining immune cell-type proportions in bulk samples can be challenging, so we will generate pseudo-bulk mixtures with known compositions. Additionally, we will explore the possibility of creating a tool that uses signature of genes highly correlated with immune infiltration levels in different cancer types and differentially expressed in both types of datasets, rather than relying on the expression of immune cells.

To evaluate the performance of each deconvolution method, we will measure correlation coefficients and root-mean-square error (RMSE) between the estimated immune cell-type compositions and the known compositions from TCGA datasets or he pseudo-bulk mixtures. Finally, we will also statistically validate the significance of our results.

References

Aran, Dvir, et al. “XCell: Digitally Portraying the Tissue Cellular Heterogeneity Landscape.” *Genome Biology*, vol. 18, no. 1, Dec. 2017, p. 220. *DOI.org (Crossref)*, https://doi.org/10.1186/s13059-017-1349-1.

Dong, Meichen, et al. “SCDC: Bulk Gene Expression Deconvolution by Multiple Single-Cell RNA Sequencing References.” *Briefings in Bioinformatics*, vol. 22, no. 1, Jan. 2021, pp. 416–27. *DOI.org (Crossref)*, https://doi.org/10.1093/bib/bbz166.

Huang, Ruichao, et al. “Combining Bulk RNA-Sequencing and Single-Cell RNA-Sequencing Data to Reveal the Immune Microenvironment and Metabolic Pattern of Osteosarcoma.” *Frontiers in Genetics*, vol. 13, Oct. 2022, p. 976990. *PubMed Central*, https://doi.org/10.3389/fgene.2022.976990.

Lei, Haoyun, et al. “Semi-Deconvolution of Bulk and Single-Cell RNA-Seq Data with Application to Metastatic Progression in Breast Cancer.” *Bioinformatics*, vol. 38, no. Supplement_1, June 2022, pp. i386–94. *DOI.org (Crossref)*, https://doi.org/10.1093/bioinformatics/btac262.

Li, Pei-Heng, et al. “Recent Developments in Application of Single-Cell RNA Sequencing in the Tumour Immune Microenvironment and Cancer Therapy.” *Military Medical Research*, vol. 9, Sept. 2022, p. 52. *PubMed Central*, https://doi.org/10.1186/s40779-022-00414-y.

Menden, Kevin, et al. “Deep Learning–Based Cell Composition Analysis from Tissue Expression Profiles.” *Science Advances*, vol. 6, no. 30, July 2020, p. eaba2619. *PubMed Central*, https://doi.org/10.1126/sciadv.aba2619.

Patrick, Ellis, et al. “Deconvolving the Contributions of Cell-Type Heterogeneity on Cortical Gene Expression.” *PLOS Computational Biology*, edited by Qing Nie, vol. 16, no. 8, Aug. 2020, p. e1008120. *DOI.org (Crossref)*, https://doi.org/10.1371/journal.pcbi.1008120.

Qi, Zongtai, et al. “Single-Cell Deconvolution of Head and Neck Squamous Cell Carcinoma.” *Cancers*, vol. 13, no. 6, Mar. 2021, p. 1230. *DOI.org (Crossref)*, https://doi.org/10.3390/cancers13061230.

Racle, Julien, et al. “Simultaneous Enumeration of Cancer and Immune Cell Types from Bulk Tumor Gene Expression Data.” *ELife*, vol. 6, Nov. 2017, p. e26476. *DOI.org (Crossref)*, https://doi.org/10.7554/eLife.26476.

Wang, Jiebiao, et al. “Bayesian Estimation of Cell Type–Specific Gene Expression with Prior Derived from Single-Cell Data.” *Genome Research*, vol. 31, no. 10, Oct. 2021, pp. 1807–18. *DOI.org (Crossref)*, https://doi.org/10.1101/gr.268722.120.

Zhang, Shaoshi, et al. “Deciphering Tumour Microenvironment of Liver Cancer through Deconvolution of Bulk RNA-Seq Data with Single-Cell Atlas.” *Cancers*, vol. 15, no. 1, Dec. 2022, p. 153. *DOI.org (Crossref)*, https://doi.org/10.3390/cancers15010153.