Tissue-Driven Metabolic Specialization in Mononuclear Phagocytes: An ImmGen-Based Analysis

IL181.007 - Immunology

Prof. Stan

15th February 2025

GitHub Repository:

https://github.com/SyedHassan20/Tissue-Driven-Metabolic-Specialization-in-Mononuclear-Phagocytes-An-ImmGen-Based-Analysis

Background

Mononuclear phagocytes (MNPs), including macrophages, monocytes, and dendritic cells (DCs), are essential components of the immune system, integrating innate immune recognition with adaptive responses. Their function is tightly regulated at the molecular level through receptor-mediated signaling pathways, transcriptional networks, and metabolic programming, all of which shape their roles in antigen presentation, immune surveillance, and cytotoxicity (Guilliams et al., 2018).

MNP activation is initiated through pattern recognition receptors (PRRs) such as Toll-like receptors (TLRs) and C-type lectin receptors, which recognize pathogen-associated molecular patterns (PAMPs) (Jang et al., 2015). Upon ligand binding, these receptors trigger downstream signaling cascades involving key transcription factors such as NF-κB, interferon regulatory factors (IRFs), and AP-1. These pathways regulate cytokine production, antigen presentation via MHC molecules, and the induction of metabolic programs that sustain immune function (Jang et al., 2015). Dendritic cells, for instance, increase MHC-II expression and costimulatory molecules (CD80/CD86) to initiate T cell activation, while macrophages exhibit phenotypic plasticity, transitioning between inflammatory and reparative states based on cytokine exposure (Tze et al., 2011).

Metabolism plays a central role in defining MNP function. Glycolysis is rapidly upregulated in activated macrophages to support biosynthetic and energy demands, whereas oxidative phosphorylation and fatty acid oxidation dominate in tissue-resident macrophages that maintain homeostasis (Soto‐Heredero et al., 2020; Wculek et al., 2023). Cholesterol metabolism in dendritic cells is essential for their migratory function, as cholesterol biosynthesis supports DC movement to lymph nodes. Experimental inhibition of cholesterol synthesis using simvastatin reduces DC migration, underscoring the necessity of this pathway for immune surveillance (You & Chi, 2023). Glutathione metabolism in macrophages regulates inflammatory responses, particularly through cysteinyl leukotriene production. Inhibition of glutathione synthesis significantly reduces leukotriene secretion in tissue-resident macrophages, highlighting its role in macrophage-mediated immune signaling (Chen et al., 2023).

To understand how metabolic diversity influences MNP specialization, my paper utilizes principal component analysis (PCA) and uniform manifold approximation and projection (UMAP) to analyze transcriptomic data from the ImmGen MNP Open Source dataset.PCA was used to evaluate metabolic variance and identify dominant sources of variation across tissues and cell types, revealing that metabolic programs exhibit strong tissue specificity. This supports the hypothesis that tissue microenvironments impose metabolic constraints on MNPs, shaping their functional identity more than cell-intrinsic factors. UMAP was employed to enhance these distinctions, capturing nonlinear relationships in transcriptional profiles and refining metabolic clustering. By applying these dimensionality reduction techniques, I provide a clearer resolution of tissue-driven metabolic specialization while also identifying finer metabolic substructures

within MNP populations. These findings contribute to understanding the molecular basis of metabolic adaptation in MNPs and its implications for immune function.

Understanding the metabolic underpinnings of MNP specialization is critical for deciphering their roles in immunity and disease. By integrating receptor-mediated signaling, transcriptional control, and metabolic regulation, this study provides a framework for exploring how tissue-specific cues shape immune cell metabolism and function.

Data Preprocessing

The dataset utilized in this study, GSE122108, originates from the Immunological Genome Project (ImmGen) and provides a comprehensive transcriptional profile of MNPs across multiple murine tissues. This dataset comprises ~400 samples representing macrophages, monocytes, and dendritic cells (DCs) isolated from diverse microenvironments, including the liver, lung, brain, spleen, and peritoneal cavity. High-throughput RNA sequencing (RNA-seq) was used to capture gene expression across ~50,000 genes, providing a high-resolution view of immune cell metabolism and functional specialization.

To improve computational efficiency and focus on biologically relevant transcriptional variation, 8,000 high-variance genes were selected based on variance calculations, ensuring the retention of genes contributing most to transcriptional heterogeneity across MNP populations. Preprocessing steps included log2 transformation (log2(data + 1)) to stabilize variance, conversion of raw counts to numeric values, and removal of NA values to ensure data integrity. The dataset was

further standardized using mean centering and scaling before applying dimensionality reduction techniques.

Principal Component Analysis (PCA) was performed on the transposed gene expression matrix, with each sample treated as a row. PCA was used to evaluate dominant sources of variance across tissues and cell types, and variance explained by the first five principal components was calculated to assess the dataset's structure. To complement PCA, UMAP was applied to refine distinctions between metabolic profiles, capturing nonlinear relationships in transcriptional data. Tissue and Cell type labels were extracted from GSE122108 metadata files.

Figures and Analysis

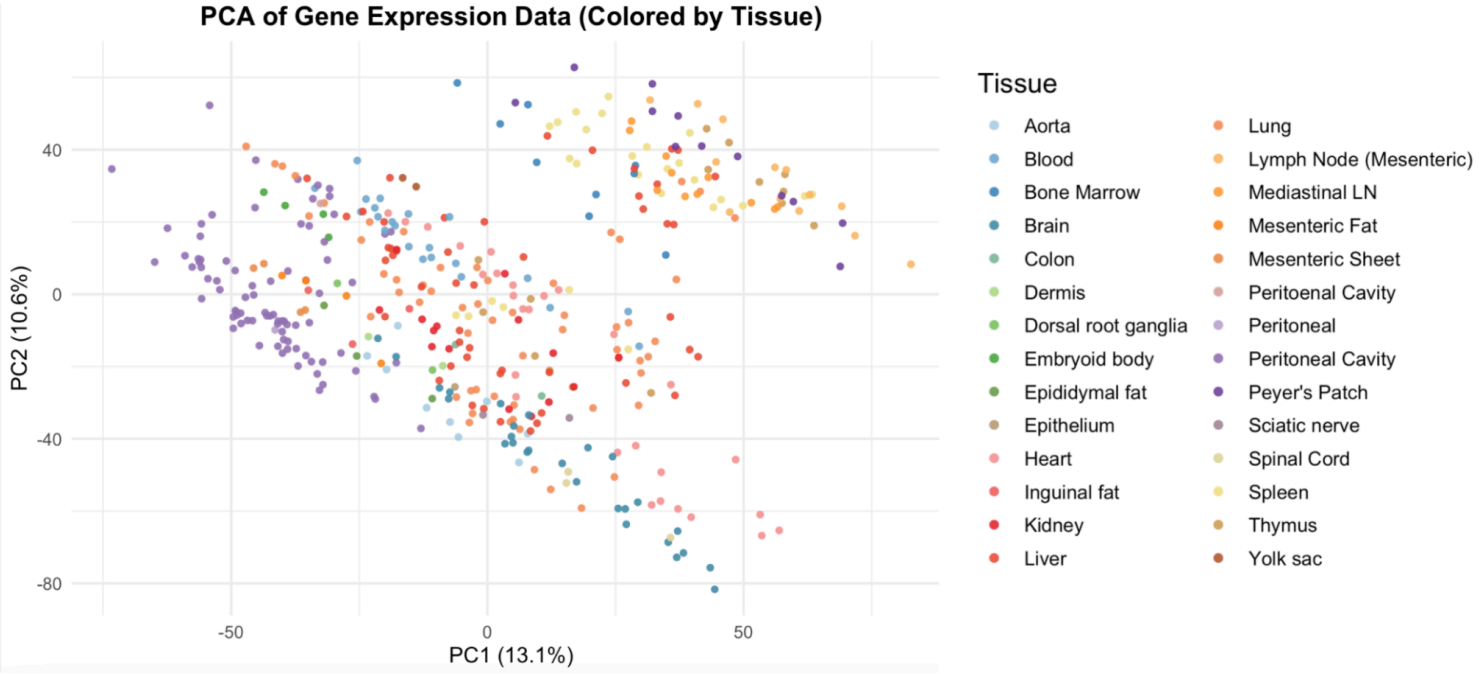


Fig. 1 shows the PCA of Gene Expression Data Across Different Tissues of ImmGen MNPs open-source data. This scatter plot displays the first two principal components (PC1: 13.1%, PC2: 10.6%) of gene expression data, with each point representing a sample and colored by tissue type. The PCA effectively reduces dimensionality while preserving variance, revealing tissue-specific clustering patterns, though some overlap suggests shared gene expression profiles. The color-coded legend enhances interpretability, though some colors are closely similar, which

may hinder differentiation. The axes are labeled with explained variance percentages, indicating the proportion of total variance captured by each component. The spread along PC1 suggests significant biological variation between tissue types, while the presence of outliers may indicate tissue heterogeneity or technical noise. A potential improvement could include convex hulls or ellipses to highlight tissue clusters more explicitly; however, when tried the hulls created more confusion.

This PCA plot suggests that metabolic programs exhibit notable tissue specificity, as samples tend to cluster by tissue of origin along PC1 (13.1% variance) and PC2 (10.6% variance). The application of standardization and gene filtering ensures that the dominant variance captured is biologically meaningful rather than driven by technical noise.

Distinct clustering of liver, brain, and lung MNPs indicates that tissue microenvironments likely impose metabolic constraints on resident immune cells, potentially influencing their reliance on oxidative phosphorylation, glycolysis, or lipid metabolism. The broader distribution of peritoneal cavity macrophages along PC2 suggests metabolic plasticity, possibly due to their dual exposure to both tissue-resident and circulating factors.

However, some overlap between tissues is observed, particularly in immune-related organs such as the spleen, bone marrow, and lymph nodes. This could reflect shared metabolic features among immune cell types or intra-tissue heterogeneity. While PCA provides an overview of tissue-driven metabolic variation, further dimensionality reduction techniques, such as UMAP, can help clarify finer distinctions in metabolic states across MNP populations

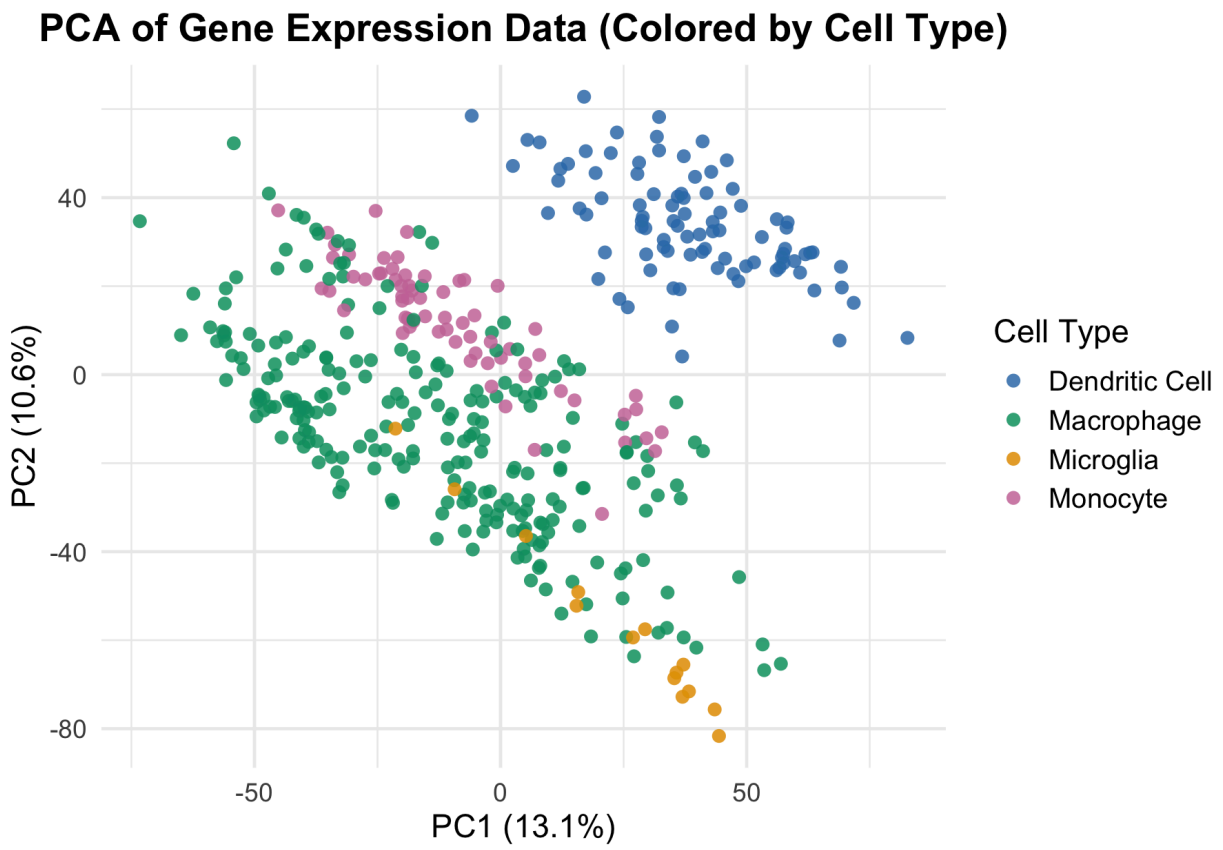


Fig. 2 shows Principal Component Analysis (PCA) of Gene Expression Data Across Mononuclear Phagocyte (MNP) Cell Types. This PCA plot visualizes the first two principal components (PC1: 13.1%, PC2: 10.6%, same as Fig. 1, of gene expression data, with each point representing a sample and colored by MNP subtype. Unlike the tissue-based PCA, this plot shows that cell types exhibit partial separation, with dendritic cells (blue) clustering distinctly, while macrophages (green) and monocytes (pink) show greater overlap. The clustering of microglia (orange) toward one region suggests a unique metabolic signature compared to other macrophages. The greater dispersion of macrophages along both PCs may indicate metabolic plasticity across tissue environments. This plot supports the idea that while tissue identity strongly shapes metabolic programs, cell-intrinsic factors also contribute to metabolic variation within MNPs. Further analyses, such as hierarchical clustering or pathway enrichment, could clarify the extent to which cell type versus tissue context drives these metabolic differences.

Although PC1 (13.1%) and PC2 (10.6%) explain the same proportion of variance as in the tissue-based PCA, the clustering pattern indicates that cell-type identity contributes less to metabolic variation than tissue type. Dendritic cells form a distinct cluster, suggesting fundamental differences in metabolic programming related to antigen presentation. However,

macrophages and monocytes show substantial overlap, indicating shared metabolic pathways such as lipid metabolism and phagocytosis. The less pronounced separation between cell types suggests that metabolic differences are more tissue-dependent, with cell-intrinsic programs being modulated by the local microenvironment. The more diffuse clustering pattern, particularly among macrophages, highlights metabolic plasticity, reinforcing the conclusion that tissue-driven metabolic adaptation is the dominant factor shaping MNP metabolism.

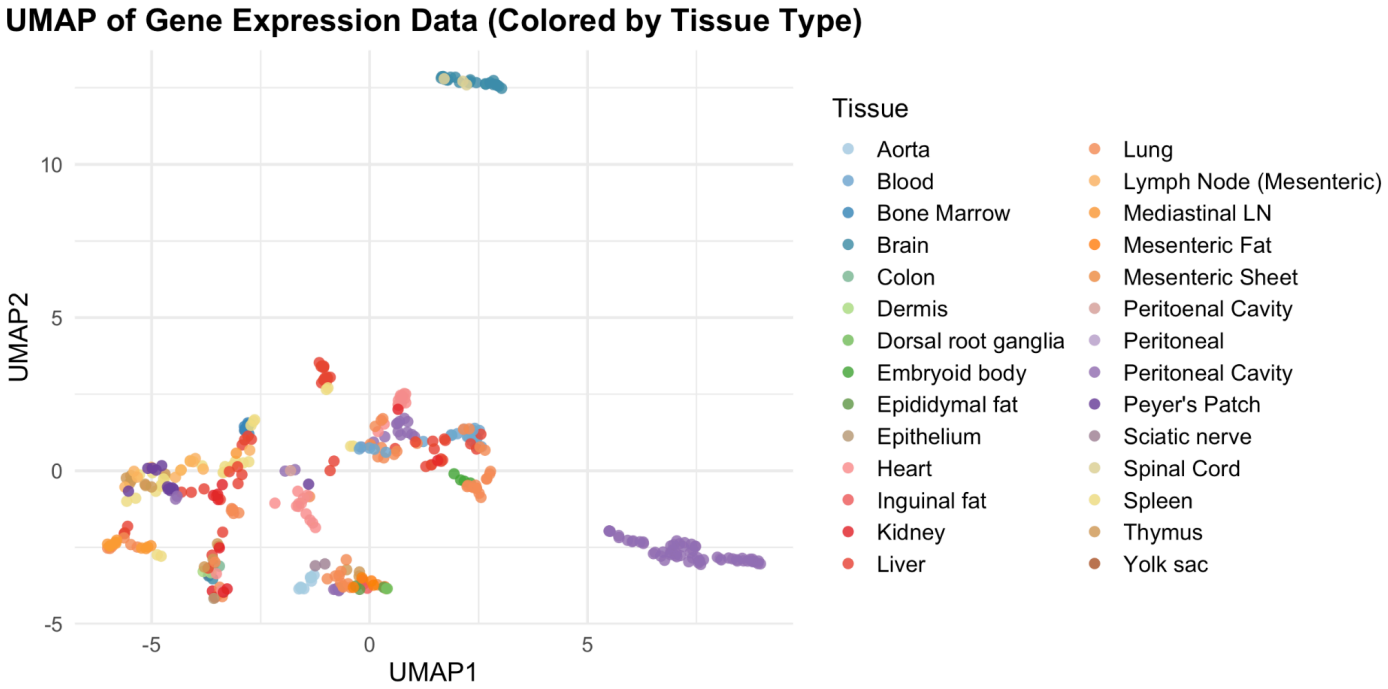


Fig. 3 shows the UMAP of Gene Expression Data Across Tissues. This Uniform Manifold Approximation and Projection (UMAP) visualization presents the global structure of gene expression data, with each point representing a sample colored by tissue type. Unlike PCA, UMAP preserves local and global relationships, revealing tissue-specific clusters with varying degrees of separation. Notably, some tissues, such as the liver and Peyer's patch, form distinct, well-separated clusters, while others, like mesenteric tissues and blood-derived samples, display partial overlap, suggesting shared gene expression features or transitional states. The presence of compact clusters suggests strong tissue-driven metabolic or transcriptional programs, while more diffuse regions may reflect either biological plasticity or methodological noise. This visualization highlights the influence of tissue microenvironments on gene expression patterns and suggests potential transitions between functionally related tissues.

UMAP provides a clearer view of tissue-driven metabolic specialization compared to PCA, revealing stronger clustering in tissues with highly conserved metabolic programs. The distinct separation of liver, lung, and brain MNPs suggests that metabolic constraints imposed by the tissue microenvironment strongly shape gene expression. The tightly clustered liver MNPs likely reflect the liver's essential role in lipid metabolism and detoxification, whereas the broader dispersion of peritoneal cavity MNPs highlights their metabolic heterogeneity, likely due to their exposure to both circulating and tissue-resident signals. The partial overlap between certain tissues suggests metabolic convergence in functionally related microenvironments. Compared to PCA, UMAP better preserves local structures, allowing for finer resolution of tissue-specific metabolic states and potential transitions between them. These findings support the hypothesis that while MNPs share core metabolic pathways, tissue-driven transcriptional programs exert a dominant influence, warranting further investigation into tissue-specific metabolic regulation.

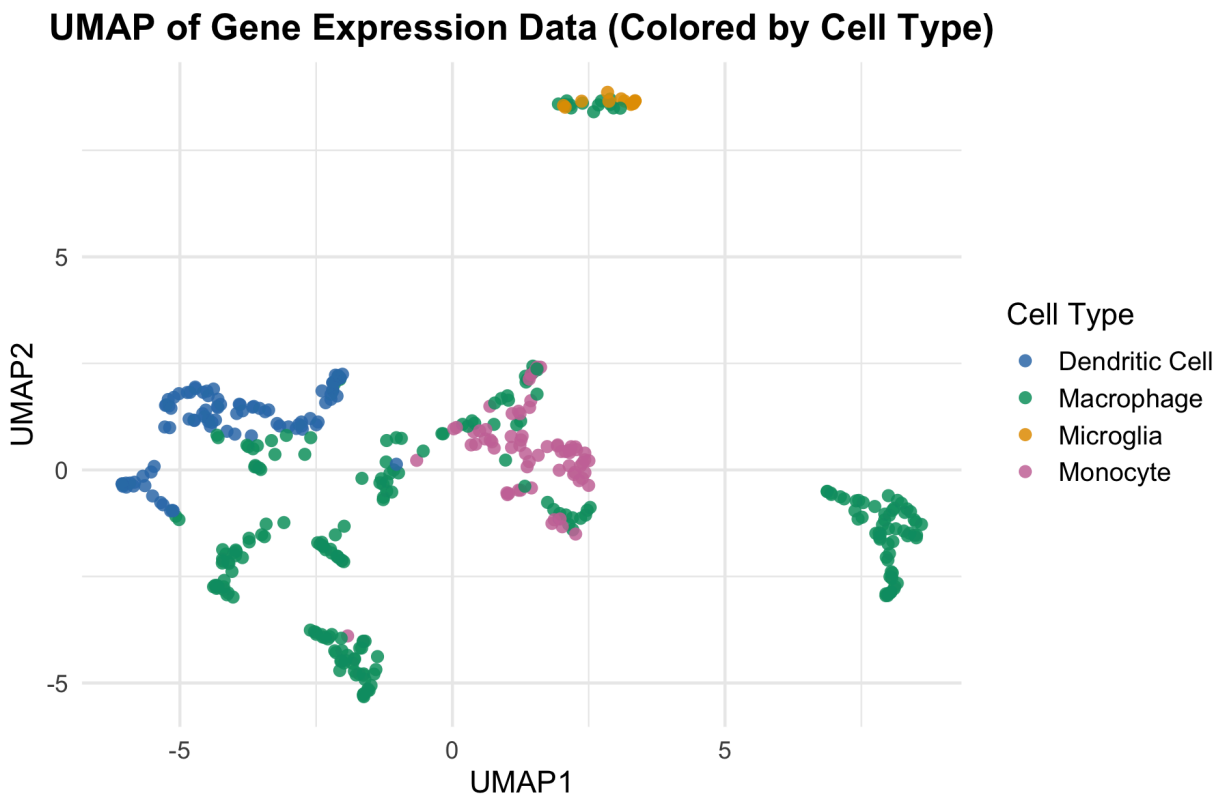


Fig. 4 shows the UMAP of Gene Expression Data Across MNP cell Types. This UMAP visualization represents the global structure of gene expression data, with each point corresponding to a sample and colored by cell type. Dendritic cells (blue) form a distinct, well-separated cluster, suggesting a highly specialized transcriptional program. Macrophages (green) exhibit more dispersed clustering, reflecting their metabolic and functional plasticity across tissues. Monocytes (pink) cluster separately but maintain some overlap with macrophages, indicating shared metabolic pathways and potential transitional states. Microglia (orange) form a tightly grouped cluster, supporting their unique gene expression profile as brain-resident macrophages. Compared to PCA, UMAP enhances the resolution of local relationships, revealing finer distinctions between cell types while also highlighting their functional similarities.

Whereas, UMAP clustering by cell type reveals that while macrophages, dendritic cells, and monocytes retain metabolic similarities within their respective groups, their spatial distribution suggests varying degrees of metabolic divergence. Microglia form a distinct, tightly clustered group, reinforcing their unique metabolic adaptations to the CNS, likely due to their reliance on oxidative metabolism and neuroimmune signaling. Dendritic cells also cluster separately, reflecting their specialized role in antigen presentation and immune surveillance. In contrast,

macrophages and monocytes exhibit more overlap, supporting the idea that they share core metabolic pathways, such as glycolysis and lipid metabolism, yet retain functional flexibility depending on their tissue environment. Notably, within each cell type, finer tissue-specific sub-clusters persist, suggesting that while cell-intrinsic programs contribute to metabolic identity, the tissue microenvironment remains a dominant factor in shaping metabolic specialization. Compared to PCA, UMAP provides a clearer resolution of these local structures, allowing for the identification of subtle metabolic shifts within and across cell types. These findings further support the conclusion that metabolic regulation in MNPs arises from a balance between conserved cell-type-specific programs and dynamic tissue-driven influences.

Therefore, by combining PCA and UMAP, my paper demonstrates that metabolic programs in MNPs are primarily dictated by tissue environment, with additional contributions from cell-type-specific pathways. PCA confirms that broad transcriptional variance is structured by tissue type, while UMAP reveals finer metabolic substructures. These findings showcase how metabolism supports functional specialization in MNPs and show the role of tissue microenvironments in shaping immune cell identity.

AI Policy

I have utilized AI in the following ways:

1. Improving coherence, spelling, and grammar in my paper.
2. Debug my code.

HC and LO Applications

HC/LO	Where and how is this concept applied in your assignment?	What does a strong application look like in the context of your assignment type?	How do you evaluate your work (including rubric score)? Please justify your reasoning.
#il181.007-molecularimmunology	I applied this LO in the background section of my paper by integrating molecular-level immune system concepts into the analysis of MNPs. I explained how pattern recognition receptors (PRRs), such as Toll-like receptors (TLRs) and C-type lectin receptors, initiate immune responses by recognizing pathogen-associated molecular patterns (PAMPs) and activating transcription factors like NF-κB, IRFs, and AP-1, which regulate cytokine production, antigen presentation, and immune activation. I also explored cytotoxicity and immune surveillance by describing how dendritic cells upregulate MHC-II and costimulatory molecules (CD80/CD86) upon activation, enabling them to prime T cells and drive adaptive immunity, while macrophages exhibit plasticity, transitioning between inflammatory (M1) and reparative (M2) states in response to cytokine signaling and tissue-specific cues.	A strong application is hard to define given it is a bioinformatic analysis but a good strategy could be to create an introduction or background section explaining the immunology aspects.	3/4 (I tried my best as I explained in the 2nd column and covered things breadth-wise but also in depth about different cells but as there is not enough information in the rubrics I am confused between a 3 and a 4).

	Additionally, I discussed how different metabolic pathways, such as glycolysis, oxidative phosphorylation, and lipid metabolism, influence immune cell specialization, emphasizing examples like cholesterol metabolism in dendritic cells supporting migration and glutathione metabolism in macrophages regulating inflammatory leukotriene production. By incorporating these molecular immunology concepts, I demonstrated how metabolic programs shape MNP function and highlighted the role of tissue microenvironments in directing immune cell metabolism and immune responses.		
#il181.007-systemsimmunology	I successfully applied this LO in my paper in the Figures and Analysis section by using omics-based computational approaches to study immune metabolism at a systems level. I employed principal component analysis (PCA) and uniform manifold approximation and projection (UMAP) to analyze transcriptional profiles from the ImmGen dataset (GSE122108), revealing how metabolic programs in mononuclear phagocytes (MNPs) are shaped by tissue environments.	A strong application is surely tackling a big-mid-size omics dataset like I did with RNA-seq with about 50,000 genes and 400 samples. It would have been nice to see a volcano plot or heat maps of differential gene expression analysis but the amount of data made it hard. Additionally, the metadata file was not clean and made it hard to even make	4 (I processed the raw dataset instead of using the data given in the paper and created my own plots with different normalization methods as explained in the data processing section. I wrote code from scratch while making PCA and UMAP visualizations and analyzed them across tissue and cell types. Lastly, I got a combined 23%

	My PCA plots demonstrated that tissue-driven metabolic constraints strongly influence immune cell clustering, with distinct metabolic signatures in liver, brain, and lung MNPs, while UMAP further refined these distinctions, uncovering finer metabolic substructures within MNP populations. Additionally, by comparing clustering patterns across tissues and cell types, I highlighted intercellular metabolic similarities, such as the overlap between macrophages and monocytes, indicating shared metabolic pathways, and the distinct clustering of dendritic cells, reflecting their specialized metabolic adaptations for antigen presentation. Through these visualizations, I provided a systems-level perspective on immune metabolism by integrating high-dimensional transcriptomic data with cellular specialization, successfully applying systems immunology concepts to elucidate the interplay between tissue microenvironments and immune function.	current plots. So for a strong application if someone can show they can handle such a large dataset, use their own code, make visualizations, comment, and analyze them should be enough.	variance between PCA1 and PCA2 which is considered acceptable for publishing papers. My combined variance is different than the paper as I used different normalization methods and the sample and gene count of my data was different from the paper (I am not sure if it was updated between when the authors published the paper and now) even some cell types are missing in the metadata due to which there is a difference between the published plots and my plots.)
#dataviz:	I successfully applied this HC in my figures and analysis by designing two PCA and two	Following the checklist and doing something new that you have never done	4 (The evaluation is similar to what I have written for systemimmunology

	UMAP visualizations to explore metabolic variation in mononuclear phagocytes (MNPs) across tissues and cell types. The PCA plots captured broad transcriptional variance, showing that tissue identity strongly influences metabolic clustering, with dendritic cells forming distinct groups while macrophages and monocytes exhibited overlap, suggesting shared metabolic pathways. The UMAP visualizations refined these patterns, preserving local structures and revealing finer metabolic subclusters within MNP populations, such as the unique metabolic signature of microglia and the metabolic plasticity of peritoneal cavity macrophages. I ensured clarity by using color-coded categories, labeled axes, and concise figure captions that describe major trends. My choice of PCA and UMAP was justified by the need to analyze high-dimensional transcriptomic data (ImmGen GSE122108), effectively distinguishing cell-intrinsic metabolic programs from tissue-driven adaptations. My analysis highlighted key findings, such as the metabolic convergence of immune-related tissues (spleen, lymph nodes) and	before.	and in the 2nd column here but I haven't made PCA and UMAP on omics data before and I accepted the challenge and delivered informative figures.)

	the role of cholesterol metabolism in dendritic cell migration. By structuring my visualizations to clearly communicate these insights, integrating figure captions with trend descriptions, and interpreting specific patterns such as clustering strength, dispersion, and overlap, I effectively used systems biology approaches to demonstrate how immune cell metabolism is shaped.		
#professionalism:	I have included at minimum 2 figures and mentioned the use of AI appropriately. I have provided my code in the GitHub repository and commented on all my code for each figure. I have cited relevant papers and included all the data. I tried replicating the Cell Report paper of Gainullina et al. (2023) but noticed that their data is not in a form that can easily be used so I utilized the data from the open-source platform. I made UMAPs of the data with PCA that the original paper didn't do to further explore the data, especially because the data I used had more samples and genes.	HC checklist	4 (I have explained that in the 2nd column)

References

Chen, S., Saeed, A. F. U. H., Liu, Q., Jiang, Q., Xu, H., Xiao, G. G., Rao, L., & Duo, Y. (2023). Macrophages in immunoregulation and therapeutics. *Signal Transduction and Targeted Therapy*, *8*(1), 1–35. https://doi.org/10.1038/s41392-023-01452-1

Gainullina, A., Mogilenko, D., Huang, L.-H., Sergushichev, A., & Artyomov, M. (2023). Network analysis of large-scale ImmGen and Tabula Muris datasets highlights metabolic diversity of tissue mononuclear phagocytes. *Cell Reports*, *42*. https://doi.org/10.1016/j.celrep.2023.112046

Guilliams, M., Mildner, A., & Yona, S. (2018). Developmental and Functional Heterogeneity of Monocytes. *Immunity*, *49*(4), 595–613. https://doi.org/10.1016/j.immuni.2018.10.005

Jang, J.-H., Shin, H. W., Lee, J. M., Lee, H.-W., Kim, E.-C., & Park, S. H. (2015). An Overview of Pathogen Recognition Receptors for Innate Immunity in Dental Pulp. *Mediators of Inflammation*, *2015*, 1–12. https://doi.org/10.1155/2015/794143

NCBI. (2024). *GEO Accession viewer*. Nih.gov. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE122108

Soto‐Heredero, G., Gómez de las Heras, M. M., Gabandé‐Rodríguez, E., Oller, J., & Mittelbrunn, M. (2020). Glycolysis – a key player in the inflammatory response. *The FEBS Journal*, *287*(16), 3350–3369. https://doi.org/10.1111/febs.15327

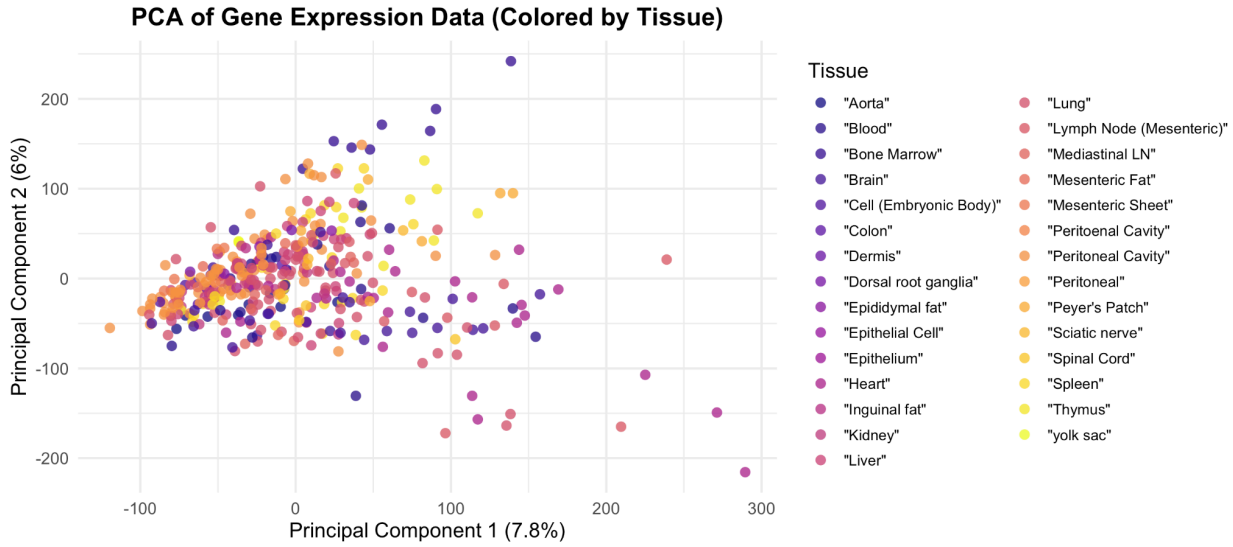
Tze, L. E., Horikawa, K., Domaschenz, H., Howard, D. R., Roots, C. M., Rigby, R. J., Way, D. A., Ohmura-Hoshino, M., Ishido, S., Andoniou, C. E., Degli-Esposti, M. A., & Goodnow, C. C. (2011). CD83 increases MHC II and CD86 on dendritic cells by opposing IL-10–driven MARCH1-mediated ubiquitination and degradation. *Journal of Experimental Medicine*, *208*(1), 149–165. https://doi.org/10.1084/jem.20092203

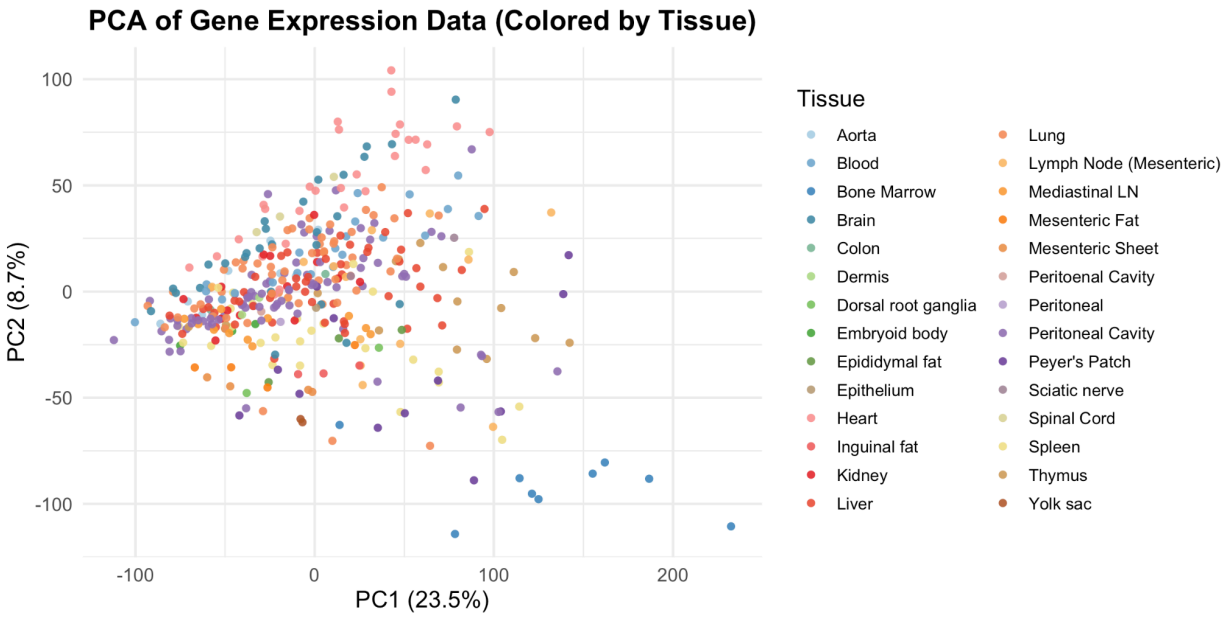
Wculek, S. K., Heras-Murillo, I., Mastrangelo, A., Mañanes, D., Galán, M., Miguel, V., Curtabbi, A., Barbas, C., Chandel, N. S., Enríquez, J. A., Lamas, S., & Sancho, D. (2023). Oxidative phosphorylation selectively orchestrates tissue macrophage homeostasis. *Immunity*, *56*(3), 516-530.e9. https://doi.org/10.1016/j.immuni.2023.01.011

You, Z., & Chi, H. (2023). Lipid metabolism in dendritic cell biology. *Immunological Reviews*, *317*(1), 137–151. https://doi.org/10.1111/imr.13215

Appendix

I am including the following figures to show what PCA1 and PCA2 variance looked like without reducing it to 8000 genes and implementing log2 stabilization. The initial intent was to utilize #designthinking.




PCA of Gene Expression Data (Colored by Tissue)
PC2 (8.7%)
PC1 (23.5%)
100
50
0
-50
-100
-100
0
100
200
Tissue
Aorta
Blood
Bone Marrow
Brain
Colon
Dermis
Dorsal root ganglia
Embryoid body
Epididymal fat
Epithelium
Heart
Inguinal fat
Kidney
Liver
Lung
Lymph Node (Mesenteric)
Mediastinal LN
Mesenteric Fat
Mesenteric Sheet
Peritoenal Cavity
Peritoneal
Peritoneal Cavity
Peyer's Patch
Sciatic nerve
Spinal Cord
Spleen
Thymus
Yolk sac