





Micro Cells, Macro Variability: An Exploratory Single-Cell Analysis of Myeloid Ontogeny and T Cell Functional Heterogeneity in Pediatric High-Grade Glioma

Sara Althaqafi
sara.althaqafi@kaust.edu.sa

Supervised by:
Dr. Hugo Tavares
hm533@cam.ac.uk

August 2026

Code available at: https://github.com/saraalthaqafi/Pediatric-Glioma_SCA

Abstract

Pediatric high-grade glioma (pHGG) is an aggressive childhood malignancy with poor prognosis and limited response to immunotherapy. With recent literature affirming that pediatric gliomas are molecularly distinct from adult glioblastoma, immune mechanisms established in adults may not translate directly to this setting. The pHGG immune microenvironment is dominated by myeloid cells spanning a developmental continuum from resident microglia to recruited macrophages, but whether this ontogeny axis governs immunosuppressive function remains contested. We analyzed single-cell RNA-sequencing data from the ScPCA pediatric glioma cohort (SCPCP000001, n = 23), using literature-curated marker panels for myeloid lineage, immunosuppressive/scavenger state, CD8+ exhaustion, and cytotoxicity. CD8+ T cells were identified by CD8A expression, and scores were computed using AUCell and cross-validated against UCell. Myeloid ontogeny was not significantly associated with immunosuppressive state at the patient level (Spearman's rho = -0.26, p = 0.34, n = 15), consistent with a non-binary model of pHGG myeloid identity, although the modest sample size limits power. Within the CD8+ compartment, the dispersion of cytotoxicity scores differed significantly across patients (Levene's F = 5.95, p = $5.1 \times 10^{-}8$), whereas exhaustion-score dispersion did not (F = 0.72, p = 0.69), indicating distinct patterns of patient-level heterogeneity across these two transcriptional programs. Myeloid immunosuppressive score was also significantly higher in chemotherapy-exposed patients (Wilcoxon W = 4, p = 0.007). Together, these exploratory findings suggest that functional cytotoxicity and checkpoint-associated exhaustion may capture different dimensions of the pHGG CD8+ immune state and support further investigation of both axes in larger pediatric cohorts.

Contents

1 Introduction 4

1.1 Clinical Background . . . 4
1.2 Tumor Immune Microenvironment . . . 4
1.3 Problem Statement . . . 5
1.4 Objectives . . . 7

2 Related Work 7

2.1 Myeloid Ontogeny and Function in Adult Glioma . . . 7
2.2 Pediatric-specific Gap . . . 8
2.3 T Cell Exhaustion and Functional Dysfunction in Glioma . . . 9
2.4 Project Framing . . . 10

3 Methodology 12

3.1 Software and tools . . . 12
3.2 Dataset . . . 12
3.3 Preprocessing and QC . . . 12
3.4 Clustering and Batch Correction . . . 12
3.5 Compartment Classification . . . 13
3.6 Scoring Methodology . . . 13
3.7 Patient-level Aggregation . . . 15
3.8 Identifying Marker Genes . . . 15
3.9 Gene Ontology Enrichment . . . 15
3.10 Statistical Methods . . . 16

4 Results 17

4.1 Cohort Overview . . . 17
4.2 Batch Correction and Compartment Classification . . . 17
4.3 Myeloid Lineage Classification and Threshold Validation . . . 17
4.4 Marker Panel Expression . . . 19
4.5 Functional Score Distributions and Cluster-Level Expression . . . 21
4.6 Myeloid Ontogeny Index and Immunosuppressive Score Distribution . . . 21
4.7 Primary Result — Myeloid Ontogeny vs. Immunosuppressive State . . . 21
4.8 Clinical and Technical Covariate Analyses . . . 21
4.9 Secondary Results — Myeloid State vs. CD8+ Function . . . 23
4.10 CD8+ T Cell Functional Heterogeneity . . . 23
4.11 Differential Expression Analysis . . . 23
4.12 Gene Ontology Enrichment . . . 26

5 Discussion 26

5.1 Myeloid Identity Is a Continuum, Not Discrete . . . 27
5.2 Are pHGGs Truly Microglia-Dominated? . . . 28
5.3 The 191-vs-2 DEG Asymmetry . . . 29
5.4 Null GO is Informative . . . 30

5.5 Chemotherapy and Immunosuppression 30
5.6 Phenotype vs. Function — The CD8 Dissociation 31

6 Limitations and Future Work **34**

7 Conclusion **36**

References **37**

1 Introduction

1.1 Clinical Background

Gliomas are the most common type of central nervous system cancers in children and adolescents [1]. These tumors, which are derived from a variety of glial cells, are broadly classified into pediatric-type and adult-type, then sub-classified based on the cell of origin and aggressiveness. The World Health Organization (WHO) created this broader classification based on both the age group most commonly affected and the molecular behavior of the tumor cells [2].

Pediatric-type gliomas have inherently different molecular behavior from their adult counterparts, with distinct driver mutations, common anatomical sites, and treatment compatibilities [3]. The formal distinction between these two disease categories was only recognized by the WHO in the fifth edition of its Classification of Tumors of the Central Nervous System (WHO CNS5), published in 2021, which for the first time separated pediatric-type diffuse gliomas from adult-type diffuse gliomas as a direct response to accumulating molecular evidence that the two are biologically distinct entities rather than the same disease occurring at different ages [2]. Prior to this update, pediatric gliomas were frequently diagnosed and, by extension, treated using classification frameworks and clinical assumptions derived largely from adult glioma biology [1]. This relatively young distinction exposes a persistent gap in the field: much of the foundational molecular and immunological characterization that clinical decision-making for adult gliomas relies on either does not exist or remains substantially incomplete for pediatric disease, leaving open the question of how far adult-derived treatment paradigms can be safely or effectively extended to pediatric patients.

Pediatric gliomas are further subdivided into pediatric-type low-grade glioma, pediatric-type high-grade glioma, and diffuse midline glioma. Pediatric high-grade gliomas (pHGG) are considered a leading cause of cancer-related death in children, with five-year survival estimates reported as low as 10% for pHGG overall and close to 1% for diffuse midline glioma specifically [4]. This dismal prognosis, combined with the comparatively recent recognition of pHGG as a molecularly distinct disease entity, underscores the urgent need for research characterizing this tumor type at the molecular and cellular level, in order to pave the way for rationally designed, pediatric-specific molecular studies and clinical trials rather than continued reliance on adult-derived treatment protocols.

1.2 Tumor Immune Microenvironment

The tumor immune microenvironment plays a central role in determining how a tumor evades immune surveillance, and increasingly, how it responds or fails to respond to immunotherapy [5]. In gliomas, the tumor immune microenvironment is dominated by myeloid cells, which can arise from two distinct developmental origins: resident microglia, seeded in the brain during embryonic development from the yolk sac, and monocyte-derived macrophages, recruited from the peripheral circulation and infiltrating the tumor after it forms [6].

Whether this developmental distinction actually matters functionally, however, remains an active and unresolved debate in the field. One view holds that microglia and macrophages retain meaningfully distinct transcriptional and functional programs within the tumor, such that a tumor's balance of the two populations is informative about its overall immune permissiveness, with macrophage-skewed tumors generally assumed to be more immunosuppressive [6, 7]. An opposing and increasingly influential view is that both populations undergo substantial tumor-driven reprogramming — commonly termed re-education in the tumor-associated macrophage literature — that overrides their developmental origin, converging both lineages toward a

shared, tumor-imposed functional state regardless of where the cell originally came from [8]. Compounding this uncertainty, the marker genes conventionally used to distinguish the two populations by single-cell transcriptomics have themselves been reported to lose specificity within the glioma microenvironment specifically: peripheral macrophages have been shown to upregulate canonical microglia-associated markers, including TMEM119 and P2RY12, upon exposure to the brain tumor microenvironment [9], and comparable loss of specificity has been reported for SALL1 and HEXB under similar tumor-conditioned exposure [10]. This raises the possibility that at least part of the ontogeny distinction reported in some studies is a byproduct of classification methodology rather than a reflection of two cleanly separable underlying biological states.

Rather than existing as fixed, discrete cell types, resting microglia and infiltrating macrophages instead occupy a continuous activation spectrum, polarizing toward either a pro-inflammatory (M1-like) or anti-inflammatory (M2-like) phenotype in direct response to signals released by the tumor and surrounding tissue (Figure 1). Tumor-derived factors such as CSF-1, TGF-β, IL-4, and IL-10 skew myeloid cells toward an M2-like, anti-inflammatory state characterized by expression of SPP1, GPNMB, TREM2, APOE, IL10, and HMOX1 — the same gene set used to construct the immunosuppressive scoring panel in the present study — while damage-associated molecular patterns (DAMPs), interferon-γ, and GM-CSF instead promote an M1-like, pro-inflammatory phenotype capable of antigen presentation and direct engagement of cytotoxic CD8+ T cells via MHC-I, CD80/CD86, and chemokine signaling (CXCL9/10/11–CXCR3) [11, 12]. This M1-engaged, cytotoxic CD8+ state is itself not fixed: under chronic antigen stimulation and inhibitory signaling such as Galectin-9–TIM-3 and IDO1, cytotoxic T cells progressively transition along a cytotoxicity–exhaustion spectrum toward a terminally exhausted phenotype marked by PDCD1, HAVCR2, LAG3, TIGIT, and TOX — again, the same panel used for exhaustion scoring in this study — mediated in part through the PD-1/PD-L1 axis [13, 14]. This shared myeloid-polarization and T-cell-exhaustion axis is precisely the biology this study's marker panels were designed to capture, and is illustrated schematically in Figure 1.

Alongside this myeloid debate, tumor-infiltrating CD8+ T cells represent the primary cytotoxic arm of the anti-tumor immune response, and their functional state — whether cytotoxic and actively engaged, or exhausted and functionally impaired — is a major determinant of whether checkpoint immunotherapies targeting T cell exhaustion are likely to succeed [13]. Glioblastoma is further frequently considered "immunologically cold," or relatively T cell deficient compared to more immunotherapy-responsive tumor types [15], and much of the current understanding of both myeloid ontogeny and CD8+ T cell exhaustion in glioma has been established specifically in this adult, immunologically cold context [16]. Given the molecular distinction between pediatric and adult glioma established in Section 1.1, it cannot be assumed without direct investigation that these same relationships — between myeloid developmental origin and immunosuppressive function, or between T cell phenotype and cytotoxic capacity — hold in the same way in the pediatric disease context. Indeed, recent single-cell and spatial profiling of the pHGG immune landscape specifically has reported a striking absence of canonical CD8+ T cell exhaustion or activation signatures, in stark contrast to melanoma, a tumor type that responds well to checkpoint immunotherapy [17], suggesting that the adult glioma exhaustion paradigm may not transfer directly to the pediatric disease.

1.3 Problem Statement

Despite the established biological and clinical distinction between pediatric and adult high-grade glioma, the immune microenvironment of pHGG remains comparatively understudied at

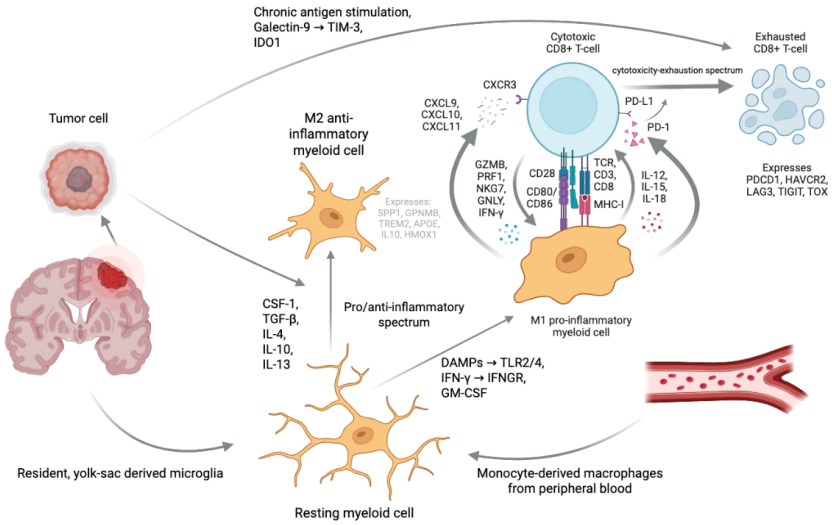


Figure 1 Schematic overview of myeloid polarization and CD8+ T cell exhaustion in the glioma tumor microenvironment. Tumor-derived cytokines (CSF-1, TGF-β, IL-4, IL-10) polarize resting microglia toward an M2-like, anti-inflammatory phenotype expressing SPP1, GPNMB, TREM2, APOE, IL10, and HMOX1, while DAMPs, IFN-γ, and GM-CSF instead promote an M1-like, pro-inflammatory phenotype capable of antigen presentation (MHC-I, CD80/CD86) and chemokine-mediated recruitment of cytotoxic CD8+ T cells (CXCL9/10/11–CXCR3). Engaged CD8+ T cells exert cytotoxic function via granzyme B (GZMB), perforin (PRF1), NKG7, granulysin (GNLY), and IFN-γ, but under chronic antigen stimulation and inhibitory signaling (Galectin-9–TIM-3, IDO1, PD-1/PD-L1) progressively transition toward a terminally exhausted state expressing PDCD1, HAVCR2, LAG3, TIGIT, and TOX. The gene sets shown for M2 microglia, cytotoxic CD8+ T cells, and exhausted CD8+ T cells correspond directly to the immunosuppressive, cytotoxicity, and exhaustion scoring panels used throughout this study. Created in BioRender.

single-cell resolution, and key relationships well-characterized in adult glioma have not been systematically tested in pediatric cohorts. In particular, it remains unclear whether the developmental origin of tumor-associated myeloid cells — resident microglia versus infiltrating monocyte-derived macrophages — predicts the degree to which these cells adopt an immunosuppressive phenotype in pHGG, and whether this myeloid state relates to the functional status of tumor-infiltrating CD8+ T cells, specifically their degree of exhaustion versus retained cytotoxic capacity. This gap carries direct clinical weight: immune checkpoint inhibition, a therapeutic strategy predicated on reversing T cell exhaustion, has shown limited efficacy in pediatric high-grade glioma to date [4], and resolving which axes of this pediatric-specific immune microenvironment — myeloid ontogeny, myeloid function, or T cell functional state — actually drive this limited response is a necessary step toward designing pediatric-specific rather than adult-derived immunotherapy strategies.

1.4 Objectives

This study aims to characterize the relationship between myeloid ontogeny and immunosuppressive function, and between CD8+ T cell phenotype and cytotoxic function, in pediatric high-grade glioma through bioinformatics analysis of publicly available single-cell RNA sequencing data. Specifically, this study seeks to:

1. Classify tumor-associated myeloid cells in pHGG according to their likely developmental origin (microglia-like versus macrophage-like) using a curated, literature-validated marker gene panel, and characterize the compositional structure of this classification across the cohort.
2. Test whether a continuous myeloid ontogeny index is associated with an independently scored immunosuppressive/scavenger functional signature at the patient level.
3. Characterize the exhaustion and cytotoxicity profiles of tumor-infiltrating CD8+ T cells across the cohort, and test whether either functional axis is associated with myeloid immunosuppressive state.
4. Examine whether cell-level heterogeneity in CD8+ exhaustion and cytotoxicity differs meaningfully between patients, independent of any relationship to myeloid state.
5. Interpret all findings within the broader context of known clinical and technical confounds, including chemotherapy exposure, patient age, sex, and diagnosis, in order to distinguish robust biological signal from artifacts of this specific cohort.

2 Related Work

2.1 Myeloid Ontogeny and Function in Adult Glioma

The relationship between tumor-associated myeloid ontogeny and function has been most extensively characterized in adult glioblastoma. Foundational work by Müller et al. used single-cell RNA sequencing to demonstrate that microglia-derived and monocyte-derived tumor-associated macrophages occupy distinct transcriptional states within the same tumor, and that their relative proportions vary regionally, with microglia-derived cells enriched at the invasive tumor margin and monocyte-derived cells enriched in the hypoxic tumor core [6]. Klemm et al. extended this work with a large cross-species single-cell atlas, further validating a marker-based approach to distinguishing these two lineages and linking macrophage ontogeny to broader patterns of immune activation across glioma grade [7]. Complementary characterization work has further reinforced this ontogeny-linked spatial patterning, describing glioma-associated

microglia and macrophages as occupying distinct microenvironmental niches and functional states across malignant progression [18, 19], consistent with the broader view that the glioma myeloid compartment is a major, ontogeny-structured driver of tumor pathobiology [20]. Together, this body of work forms the primary literature basis for the microglial and macrophage marker panels used throughout the present analysis, and represents the field's dominant paradigm: that ontogeny is a meaningful axis along which tumor-associated myeloid function can be predicted and stratified.

This paradigm has, however, been increasingly complicated by subsequent work emphasizing myeloid plasticity over fixed ontogeny-based identity. Bowman et al. showed directly that macrophage ontogeny underlies differences in tumor-specific "education" in brain malignancies, demonstrating that the tumor microenvironment actively reprograms myeloid cells toward a tumor-supportive phenotype in an ontogeny-dependent manner [8]. Reviews focused specifically on the glioma myeloid compartment echo this plasticity framing, describing glioma-associated microglia and macrophages as subject to continuous recruitment, polarization, and functional reprogramming by tumor-intrinsic and environmental cues [18]. Within glioma specifically, this plasticity has been shown to directly compromise the reliability of the ontogeny markers themselves: peripheral macrophages infiltrating glioma tissue have been reported to upregulate canonical microglia-restricted markers, including TMEM119 and P2RY12, upon exposure to the tumor microenvironment [9], with comparable specificity loss reported for SALL1 and HEXB under tumor-conditioned exposure [10]. This body of work collectively suggests that ontogeny-based classification, while methodologically well-established, may capture a less clean or less functionally meaningful distinction in situ than the foundational single-cell atlases initially suggested, motivating direct empirical testing of the ontogeny–function relationship in any new cohort rather than assuming it will replicate.

2.2 Pediatric-specific Gap

Pediatric high-grade glioma has been the subject of a growing, though still comparatively young, body of single-cell immune profiling work. Foundational pediatric single-cell glioma work, most notably Filbin et al.'s single-cell dissection of H3K27M-mutant diffuse midline glioma, focused primarily on malignant cell developmental hierarchies and oncogenic programs rather than on myeloid ontogeny or immune cell function [21], an emphasis broadly consistent with parallel single-cell efforts in adult glioblastoma that similarly prioritized malignant cell state heterogeneity over immune microenvironment characterization in their initial analyses [22]. DeSisto et al. subsequently demonstrated that tumor and immune cell types interact to produce highly heterogeneous phenotypes across pHGG samples [23], and the most direct pediatric counterpart to this study's aims is the recent, considerably larger-scale profiling effort by LaBelle et al., which combined single-cell RNA sequencing and spatial transcriptomics across 27 pediatric high-grade glioma and high-grade ependymoma patients to examine the immune microenvironment before and after chemoradiotherapy and checkpoint inhibition [17]. This work reported that pediatric myeloid states diverge substantially from adult glioma myeloid states, and separately reported a striking absence of canonical CD8+ T cell exhaustion or activation signatures relative to melanoma [17].

Both findings directly motivate the present study: if pediatric myeloid biology and pediatric T cell exhaustion biology both diverge from the adult paradigm, then the specific ontogeny–immunosuppression and exhaustion–cytotoxicity relationships established in adult glioma cannot be assumed to hold in pediatric disease without direct testing. Two other efforts have specifically reanalyzed the same public ScPCA Portal cohort (SCPCP000001) used in this study: one applying

antibody-based targeting analysis to the SPP1-CD44 axis in pHGG myeloid cells [24], and a separate spatial transcriptomic study identifying a novel SPP1+/GPNMB+ immunosuppressive myeloid population in an independent pHGG cohort [25] — both of which directly motivated the inclusion of SPP1 and GPNMB in the immunosuppressive scoring panel used here (Table 2). Neither of these efforts, however, examined myeloid ontogeny or CD8+ T cell functional state, and neither tested these axes against one another.

Beyond these specific adjacent studies, Messiaen et al. frame the pediatric glioma tumor immune microenvironment as an open question — "friend or foe" — noting substantially less mechanistic characterization than exists for adult glioblastoma [26], and du Chatinier et al. argue directly that this gap in myeloid characterization is a primary barrier to successful immunotherapy in pediatric brain tumors specifically [27]. Broader reviews of emerging biological concepts in pediatric high-grade glioma reinforce this same picture: a real and rapidly growing, but still comparatively thin, evidence base relative to adult glioma [28].

2.3 T Cell Exhaustion and Functional Dysfunction in Glioma

T cell exhaustion in adult glioblastoma has been characterized as unusually severe relative to other solid tumors. Woroniecka et al. demonstrated that glioblastoma-infiltrating T cells show a particularly pronounced exhaustion signature, marked by prominent upregulation of multiple immune checkpoints and stereotyped transcriptional programs matching classical virus-induced exhaustion, distinguishing glioblastoma from other murine tumor models tested in the same study [16], a severity consistent with broader reviews describing T cell exhaustion as a central mechanism of immune evasion and immunotherapy resistance across human cancers [29, 30]. Mechanistically, this exhaustion phenotype has been linked directly back to the myeloid compartment: Ravi et al. identified interleukin-10-releasing myeloid cells as a direct driver of T cell dysfunction in the glioblastoma microenvironment via a pseudotime-resolved transcriptional trajectory, providing a concrete mechanistic bridge between myeloid state and T cell dysfunction of exactly the kind this study tests for, in adult disease [31].

More recent work has also reframed exhaustion itself as a structured continuum rather than a single terminal state. Wherry and Kaech's foundational framework, substantially updated in recent reviews by Baessler and Vignali and by Sun and Dong, distinguishes progenitor or precursor-exhausted T cells, which retain proliferative and effector capacity and remain responsive to checkpoint blockade, from terminally exhausted T cells, which show stable and largely irreversible functional decline [13, 14, 32]. This distinction is directly relevant to any transcriptional exhaustion score of the kind used in the present study, since such scores capture phenotypic checkpoint receptor expression without necessarily distinguishing where a given cell sits along this continuum, or whether phenotypic exhaustion has actually translated into functional loss of cytotoxic capacity. Zebley and Youngblood similarly argue that resolving where a given T cell population sits along this continuum, rather than treating exhaustion as binary, is essential for designing next-generation immunotherapies [33].

Evidence that exhaustion severity varies meaningfully by glioma subtype, rather than being a fixed property of glioma as a disease category, further supports treating exhaustion as an empirical question rather than an assumed constant. Work comparing pediatric low-grade and high-grade glioma has reported greater CD8+ exhausted T cell content in low-grade relative to both adult and pediatric high-grade glioma, with a distinct PD1+/TIGIT+ exhausted population shown to be functionally important for maintaining tumor growth in that specific disease context [34]. As noted above, the pediatric high-grade-specific dataset examining this axis directly found yet another pattern again: a comparative absence of exhaustion and activation signatures relative

to melanoma [17]. Taken together, these findings indicate that exhaustion phenotype in glioma is subtype- and age-dependent rather than uniform, directly raising the question this study investigates from a different statistical angle: rather than testing central tendency of exhaustion, whether patient-level heterogeneity exists in exhaustion, cytotoxicity, or both, and whether these two functional axes vary independently of one another within a pediatric high-grade cohort specifically.

2.4 Project Framing

Considered together, the literature reviewed above establishes three separate empirical claims in adult glioma — that myeloid ontogeny predicts immune function [6, 7], that myeloid state drives T cell dysfunction [31], and that glioblastoma T cell exhaustion is severe and stereotyped [16, 29, 30] — each of which is complicated or directly contradicted by more recent pediatric-specific and plasticity-focused work [8–10, 17, 34], summarized in Table 1. Pediatric high-grade glioma single-cell immunology is not an empty field — several substantial efforts, including two that have independently reanalyzed the same public cohort used here [24, 25], have emerged in just the past two years — but to the author's knowledge, no prior study has directly tested the ontogeny–immunosuppression relationship and the exhaustion–cytotoxicity relationship together within the same pediatric high-grade glioma cohort, using a hard, marker-panel-based ontogeny classification alongside continuous functional scoring of both myeloid and CD8+ compartments.

This study addresses that gap using an approach that differs from prior work along two dimensions, not just cohort selection or question choice. First, rather than assuming a discrete ontogeny classification a priori, myeloid lineage identity is treated as an open empirical question: model-based clustering is used explicitly to test whether the data support a discrete microglia/macrophage split before any hard threshold is applied, directly addressing the plasticity concerns raised throughout this review [8–10]. Second, and more centrally, this study tests for patient-level heterogeneity in CD8+ functional state using tests of variance (Levene's test) rather than relying solely on tests of central tendency (correlation and group comparison). This distinction matters directly in light of the exhaustion-continuum literature reviewed above [13, 14, 32]: if phenotypic exhaustion and functional cytotoxicity are genuinely dissociable processes, a relationship between them may be invisible to a correlation test on patient means while still being detectable as a difference in how consistently each axis behaves within and across patients. This framing — prioritizing the internal consistency of a functional signature across patients, not only its average level — is, to the author's knowledge, not present elsewhere in the glioma myeloid or T cell exhaustion literature reviewed here, and constitutes the primary methodological contribution of this study alongside its empirical, pediatric-specific findings.

This framing carries both advantages and limitations relative to prior work. Its principal advantage is methodological transparency and direct testability: by using an independent, publicly available cohort (ScPCA Portal, project SCPCP000001) and a marker panel drawn explicitly from established adult glioma literature [6, 7], this study provides a direct, out-of-sample test of whether adult-derived relationships generalize to pediatric disease, rather than assuming generalization by default. Its principal limitation, addressed at length in the Discussion and Limitations sections of this report, is scale: at 15–23 patients depending on the specific analysis, this study is considerably smaller than the 27-patient LaBelle cohort [17] and far smaller than the atlas-scale efforts underlying the foundational adult ontogeny literature [6, 7], limiting this work to a hypothesis-generating pilot role rather than a definitively powered confirmatory study.

Table 1 Comparison of the present study against the most directly relevant prior single-cell glioma immune profiling efforts.

Study	Cohort	Age group	Ontogeny/function finding	T cell exhaustion finding
Müller et al. [6]	Adult GBM	Adult	Microglia- and monocyte-derived TAMs occupy distinct spatial niches & transcriptional states	Not examined
Klemm et al. [7]	Cross-species atlas	Adult (+mouse)	Ontogeny classification links macrophage origin to immune activation across grade	Not examined
Filbin et al. [21]	pHGG (H3K27M DMG)	Pediatric	Not examined (focus: malignant cell hierarchy)	Not examined
van den Broek et al. [25]	pHGG (spatial)	Pediatric	Novel SPP1+/GPNMB+ immunosuppressive myeloid population	Not examined
LaBelle et al. [17]	pHGG, n=27	Pediatric	Pediatric myeloid states diverge from adult glioma	Striking absence of exhaustion/activation vs. melanoma
This study	pHGG, n=15–23	Pediatric	Ontogeny does not predict immunosuppressive score; continuum, not discrete	Exhaustion-score dispersion similar; cytotoxicity-score dispersion heterogeneous across patients

3 Methodology

3.1 Software and tools

All of the bioinformatics analyses performed on this study are done via R and RStudio. We convert the data into a Seurat Object [35] and consequently use the Seurat package for many downstream analyses. An overview of the full pipeline is shown in Figure 2.

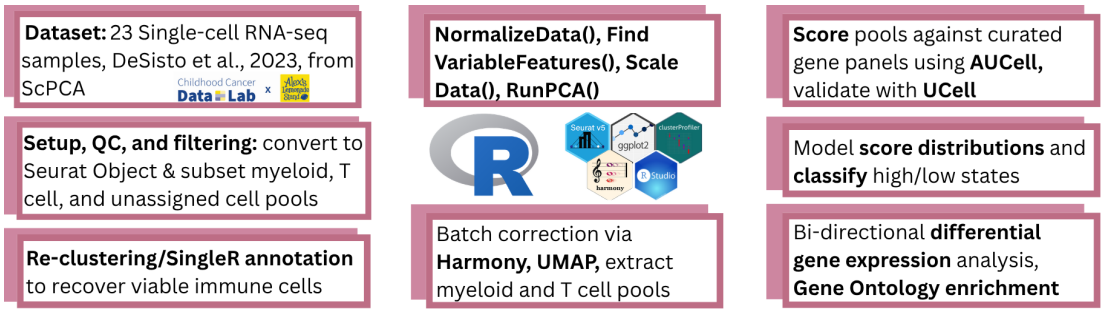


Figure 2 Overview of the analysis pipeline, from data acquisition through statistical testing.

3.2 Dataset

The dataset was obtained from the Single-cell Pediatric Cancer Atlas [36] (SCPCP000001), which was created by Alex's Lemonade Stand Foundation for Childhood Cancer and the Childhood Cancer Data Lab. The data is publicly available at `https://scpca.alexslemonade.org/projects/SCPCP000001`. It consists of 22 samples of patients clinically diagnosed with pediatric-type high grade glioma, and one control. The patients' ages range from 4-24, and some samples were obtained in vivo, while others were obtained at autopsy. The data is annotated with other metadata such as sex, diagnosis, sub-diagnosis, treatment type, and tissue location. The data was available to download as unfiltered, filtered, and processed, which included manual annotation of cell types.

3.3 Preprocessing and QC

The processed data were downloaded as SingleCellExperiment objects and converted to Seurat objects, with the available ScPCA quality-control metadata retained. Cells failing the study's quality-control criteria for elevated mitochondrial transcript content and/or doublet status were removed prior to downstream analysis. In total, 2,009 cells were excluded, leaving 25,301 cells for subsequent analysis.

3.4 Clustering and Batch Correction

Despite the data being manually annotated by the authors, examination revealed many redundant labels, and cells with an "Unknown" label reaching up to 97% in a single sample. For this reason, we decided to re-cluster in an attempt to salvage any relevant immune cells which may be in the "Unknown" group. We first created a combined pool with all myeloid, unknown, and T cells. Cell-type re-annotation was performed using SingleR [37]. AnnotationDBI [38] was used for gene-identifier annotation and mapping where required, rather than for cell-type classification. SingleR re-annotation recovered 10,285 cells from the previously Unknown-labelled group, which were subsequently incorporated into the appropriate broad compartments for downstream analysis. For downstream analyses, the cells were hereon split into the myeloid pool and the T cell pool.

Subsequently, normalisation, feature selection, data scaling, and principal component analysis (PCA) were performed before batch correction to test if there was a batch effect. `nfeatures` was set to 2000, and `npcs` was set to 30. We synthesized a UMAP in order to visualise the clusters, selecting the first 15 PCs (`dims = 1:15`). The pre-integration UMAP showed pronounced sample-associated structure. Because sample identity can reflect both technical variation and genuine inter-patient biology, Harmony [39] was used to reduce sample-associated structure specifically for integrated clustering and visualization. Harmony-corrected dimensions were used to construct the neighbor graph, perform clustering, and generate the integrated UMAP. Importantly, the downstream AUCell functional scores and patient-level statistical analyses were calculated from gene-expression data rather than from the Harmony embedding, so Harmony was not used to directly define the functional differences tested between patients. We ran `FindNeighbors` (`k.param = 15`) and `FindClusters` then re-ran the UMAP, which resolved the batch effect.

3.5 Compartment Classification

Following SingleR re-annotation, cells were separated into broad candidate T cell and myeloid pools. Cluster-level marker expression was then examined as an additional validation and refinement step rather than as the primary source of cell identity. T cell identity was assessed using the CD3 complex (CD3D, CD3E, CD3G) together with the T-cell receptor constant-region genes TRAC, TRBC1, and TRBC2. Myeloid identity was supported by expression of the combined microglial and macrophage marker panels described below (P2RY12, SALL1, TMEM119, MERTK, P2RY13, CD68, ADGRE1, CCR2, CD200R1, CD5L, CHIT1). Clusters showing neither T cell nor myeloid signatures were assigned to the residual Other/Tumor compartment. The lineage-specific microglia and macrophage scores used in subsequent analyses were calculated only after this broad compartment assignment.

3.6 Scoring Methodology

The basis of the downstream analyses, and indeed the entire study, depends on the scoring of the cells, as it determines which cells corresponded to which subtype or state. Two key decisions arose: what are the cells being scored against, and how will they be scored?

Table 2 Curated gene panels used for AUCell scoring of myeloid lineage and functional states. All genes were individually supported by established literature and restricted to five to six per panel to preserve biological defensibility.

Panel	Genes
Microglia	P2RY12, SALL1, TMEM119, MERTK, P2RY13
Macrophage	CD68, ADGRE1, CCR2, CD200R1, CD5L, CHIT1
Immunosuppressive	SPP1, GPNMB, TREM2, APOE, IL10, HMOX1
Exhaustion	PDCD1, HAVCR2, LAG3, TIGIT, TOX, ENTPD1
Cytotoxicity	GZMB, PRF1, NKG7, GNLY, IFNG, CX3CR1

3.6.1 Gene Panels

The first question is answered by selecting the genes used to score each category. Gene panels were curated from established literature [6, 7] for microglia, macrophage, immunosuppressive, exhaustion, and cytotoxicity states. Panels were restricted to five to six genes each, all individually supported in the literature, to avoid diluting biological signal with weakly-justified markers. Final panels are shown in Table 2.

3.6.2 Comparing Packages

Scoring was run using the two most widely-used R packages for this purpose: AUCell [40] and UCell [41]. UCell is a more recent version based on AUCell architecture, swapping the latter's Area-Under-Curve (AUC) enrichment score for the Mann-Whitney-U test. As such, the authors "expect similar results." [41]. During single-cell raw score plotting, however, we discovered that UCell's ranking-based scoring produced a floor-collapse artefact with small gene panels: genes ranked below the cutoff for a given cell are effectively treated as equivalent, compressing scores toward a shared minimum rather than reflecting graded expression differences. As our panels were intentionally restricted to five to six genes, this artefact was pronounced enough to warrant pivoting toward an alternative method. We also evaluated Seurat's AddModuleScore() as an alternative expression-based module-scoring approach. Unlike AUCell and UCell, which are rank-based methods, AddModuleScore calculates a per-cell module score relative to matched control gene sets. Because the primary analysis was designed around rank-based single-cell scoring, and AUCell avoided the floor-compression observed with UCell while retaining straightforward cell-level interpretation, AUCell was selected as the primary scoring method. UCell was retained as an independent rank-based validation method.

The decision process behind both the scoring method (this section) and the myeloid lineage classification method (Section 3.6.3 below) is summarized in Figure 3.

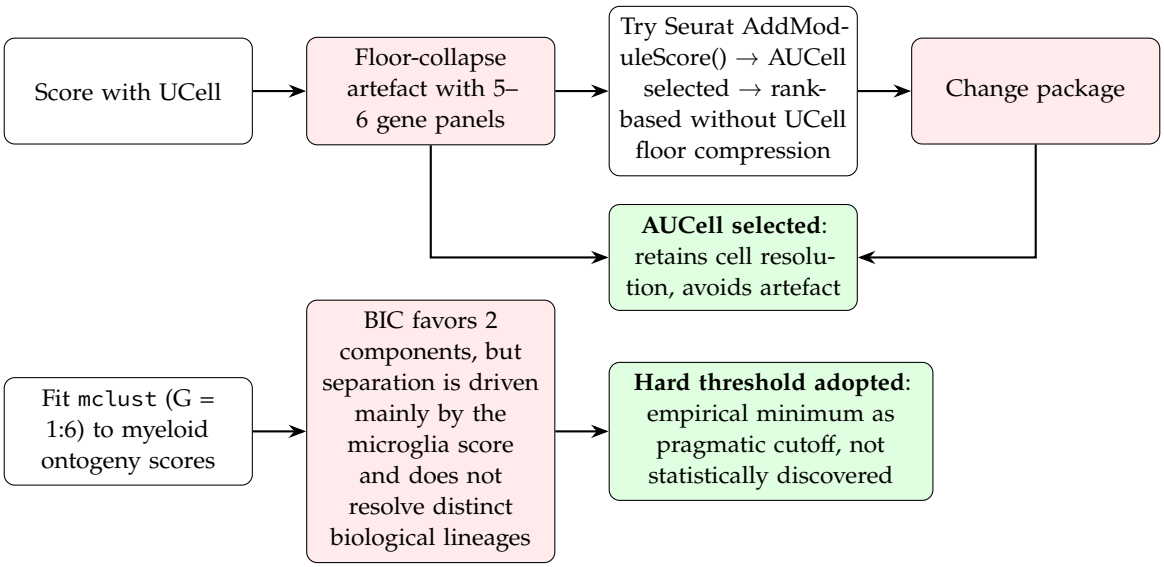


Figure 3 Decision flow for (top) the choice of scoring method, and (bottom) the choice of myeloid lineage classification method. Both decisions followed an initial attempt that failed on a specific, identifiable criterion before a more defensible alternative was adopted.

3.6.3 Myeloid Lineage Classification & Ontogeny Index

Scoring the myeloid cells against the microglial and macrophage gene panels helped determine whether myeloid cells were more similar to microglia, macrophages, or neither. In order to confirm as precisely as possible which cells belonged to which group, we viewed the distributions of the cells for each score. Due to the variability of the data, the distributions were roughly bi- or tri-modal, with the largest peaks around zero, central peaks, and occasionally large peaks at higher scores. We attempted to discretize the data using mclust [42], a Gaussian finite mixture model, allowing the number of components to be selected automatically (G = 1:6) rather than assumed in advance. BIC favored a two-component solution. However, inspection of the fitted

components showed that this separation was driven predominantly by variation along the microglia-score axis, while the macrophage score showed no comparably distinct structure. The two statistical components therefore did not correspond to the biologically expected Microglia-like and Macrophage-like populations. We consequently interpreted the myeloid compartment as lacking a clean, biologically meaningful discrete ontogeny separation and adopted pragmatic hard thresholds for descriptive classification.

In an effort to adhere to biology, we also subtracted the myeloid scores to create the continuous ontogeny index, since transcriptional state cannot definitively determine lineage.

3.6.4 Immunosuppressive Scoring

All myeloid cells, regardless of lineage, were scored against the immunosuppressive gene panel using AUCell and UCell.

3.6.5 CD8+ Subsetting and T Cell Scoring

Exhaustion and cytotoxicity are only applicable in the context of "cytotoxic T cells"; T cells that naturally have an anti-tumor function and express the marker CD8, but not CD4. Therefore, the pool of T cells was subsetted to CD8+ cells. To ensure that the CD8+ pool was not contaminated, we removed all cells which expressed CD4. This subset was then scored against the exhaustion and cytotoxicity panels respectively.

3.7 Patient-level Aggregation

As part of a typical single-cell analysis pipeline, we performed patient-level aggregation or pseudobulking in order to draw relevant clinical conclusions. After this, examination of the data showed some sparsity in cell counts across patients. In order to ensure that conclusions are as biologically meaningful as possible, we created thresholds of at least 20 myeloid and 20 T cells for samples to be included. Myeloid reliability filtering retained 15 for the primary myeloid analysis, while the CD8+ requirement reduced the secondary joint analyses to 10. No samples were excluded on the basis of diagnosis; reductions in patient number resulted solely from the compartment-specific cell-count reliability criteria.

3.8 Identifying Marker Genes

Within the myeloid compartment, we investigated whether there were genes differentially expressed between the Microglia-like and Macrophage-like lineage groups specifically; this analysis was not performed on CD8+ T cell subgroups, as exhaustion and cytotoxicity were treated as continuous scores throughout rather than discretized. We performed an exploratory, bi-directional cell-level differential expression analysis between the Microglia-like and Macrophage-like groups using Seurat's `FindMarkers()`. Because individual cells are nested within patients and the two groups were themselves defined using lineage-associated marker expression, this analysis was treated as descriptive and confirmatory rather than as an independent patient-level differential-expression test.

3.9 Gene Ontology Enrichment

Finally, in order to better understand the biological context in which differentially expressed genes exist (DEGs), we performed gene ontology enrichment using `clusterProfiler::enrichGO` [43] after mapping the Ensembl IDs to the Entrez IDs.

3.10 Statistical Methods

All correlation analyses used Spearman's rank correlation coefficient (ρ), selected over Pearson correlation due to the small sample sizes (n=15 for the primary analysis; n=8–10 for secondary analyses) and non-normal, right-skewed distributions typical of single-cell-derived pseudobulk scores. Group comparisons were assessed using non-parametric rank-based tests, with Wilcoxon rank-sum or Kruskal–Wallis tests applied as indicated in Tables 4 and 5. Kruskal–Wallis testing was used for multi-category diagnosis comparisons.

To assess whether CD8+ T cell functional heterogeneity differed by patient, Levene's test was applied at the single-cell level, testing for equality of variance in the distributions of exhaustion and cytotoxicity scores between patients. Unlike the correlation and group-comparison tests above, which test differences in central tendency, Levene's test specifically evaluates differences in within-group dispersion — appropriate for testing whether patients differ in the internal consistency of their T cell populations, rather than in average score level.

Throughout, a significance threshold of $\alpha = 0.05$ was applied. For Gene Ontology enrichment analysis, Benjamini-Hochberg correction was applied to control the false discovery rate across all tested terms within each ontology (Biological Process, Molecular Function, Cellular Component); results not surviving correction are reported separately and explicitly labeled as exploratory.

An overview of all statistical methods used are outlined in Table 3.

Table 3 Statistical tests used in this study

Test	Purpose	Applied to
Spearman correlation	Association between continuous variables	Myeloid ontogeny vs. immunosuppressive score; secondary myeloid - CD8 correlations; confound checks (age, n_cells)
Two-group rank sum test	Two-group comparison	Chemotherapy status, sex
Rank comparison across groups	Comparison across 3+ groups	Diagnosis
Levene's test	Equality of variance	CD8+ exhaustion/cytotoxicity by patient

4 Results

4.1 Cohort Overview

Of the 23 samples in SCPCP000001, all were initially retained and no diagnosis-based exclusion criterion was applied. Diagnoses included glioblastoma, high-grade glioma (NOS), anaplastic astrocytoma, anaplastic glioma, diffuse midline glioma, pleomorphic xanthoastrocytoma, and one non-cancerous reference sample. Downstream sample numbers were determined by cell-count reliability thresholds rather than diagnosis. Because several diagnostic categories were represented by only one or two patients, some categories were no longer represented after reliability filtering when the corresponding samples did not meet the required cell-count threshold. (Figure 4). Age at diagnosis skewed toward adolescence, and sex distribution was approximately balanced. Quality control removed 2,009 cells for doublets and mitochondrial content, retaining 25,301 cells. Following compartment classification, cells were assigned to the myeloid compartment, the T cell compartment, or a residual Other/Tumor compartment.

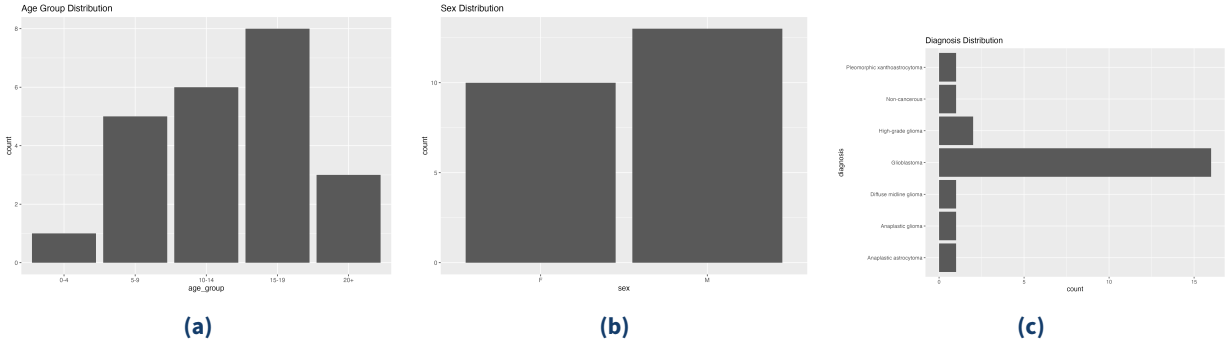


Figure 4 Cohort composition. (a) Age group distribution. (b) Sex distribution. (c) Diagnosis distribution across the 23 retained samples.

4.2 Batch Correction and Compartment Classification

Principal component analysis of the combined immune-cell pool showed the expected elbow, with variance leveling off around 10 components (Figure 5). UMAP embeddings computed prior to Harmony integration showed pronounced sample-driven clustering, with cells from individual patients forming largely non-overlapping islands (Figure 6a). Following Harmony correction, cells mixed substantially across samples throughout the embedding (Figure 6b). Graph-based clustering of the batch-corrected embedding (Figure 7a) was collapsed into three broad compartments — Myeloid, T cell, and Other/Tumor — based on marker expression (Figure 7b).

4.3 Myeloid Lineage Classification and Threshold Validation

Microglia and macrophage AUCell scores were both strongly right-skewed with a large zero-inflated mass and no clear bimodal separation (Figure 8). Hard thresholds (Microglia = 0.075; Macrophage = 0.03) were applied to classify each myeloid cell into one of four lineage groups (Figure 9), collapsed to three groups for visualization (Figure 10). As a validation check, AUCell and UCell scores agreed closely across all three functional axes (Spearman: ontogeny index $\rho = 0.75$, $p < 0.001$, $n = 5{,}965$; exhaustion $\rho = 0.66$, $p = 1.4 \times 10^{-85}$, $n = 672$; cytotoxicity $\rho = 0.83$, $p = 9.3 \times 10^{-174}$, $n = 672$), supporting AUCell as the primary scoring method. The Microglia score consistently exceeded the Macrophage score across nearly all myeloid clusters, consistent

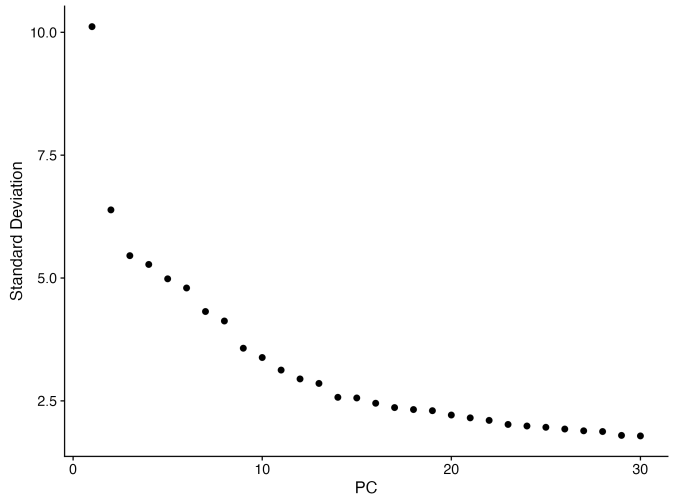


Figure 5 Elbow plot of principal component standard deviations for the combined immune-cell pool.

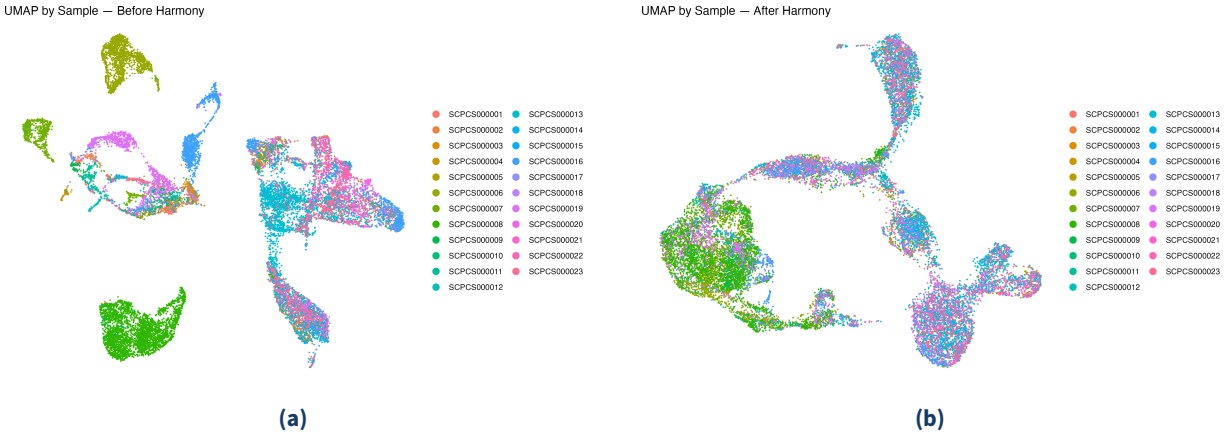


Figure 6 Batch correction assessment. UMAP embeddings colored by sample identity (a) before and (b) after Harmony integration.

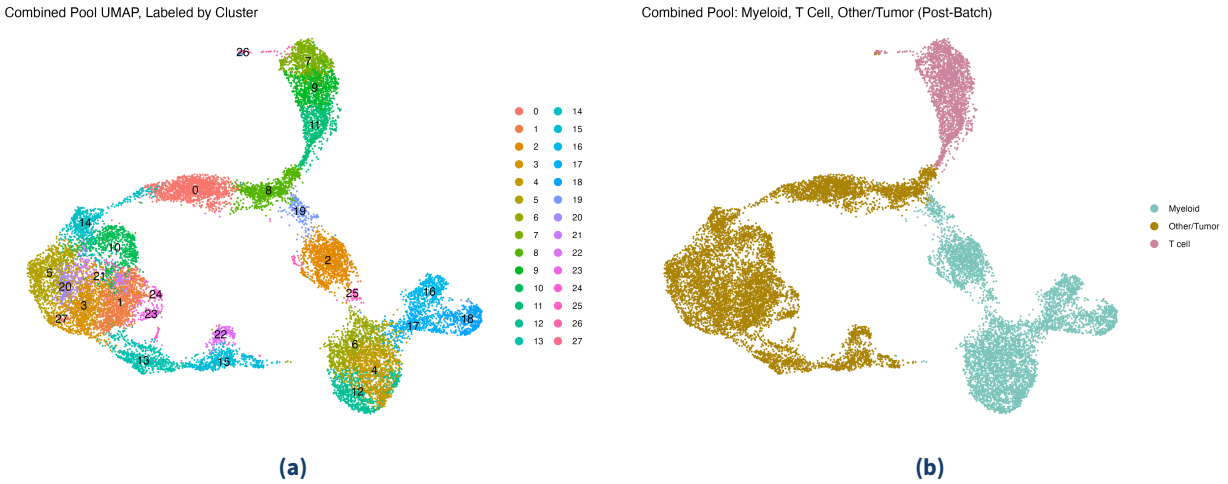


Figure 7 Compartment classification of the combined immune-cell pool. (a) Batch-corrected UMAP labeled by Louvain cluster. (b) The same embedding labeled by assigned compartment.

with a Microglia-dominant compartment mirrored in the continuous ontogeny index, which was negative (Microglia-skewed) for all 15 evaluable patients (Figure 11).

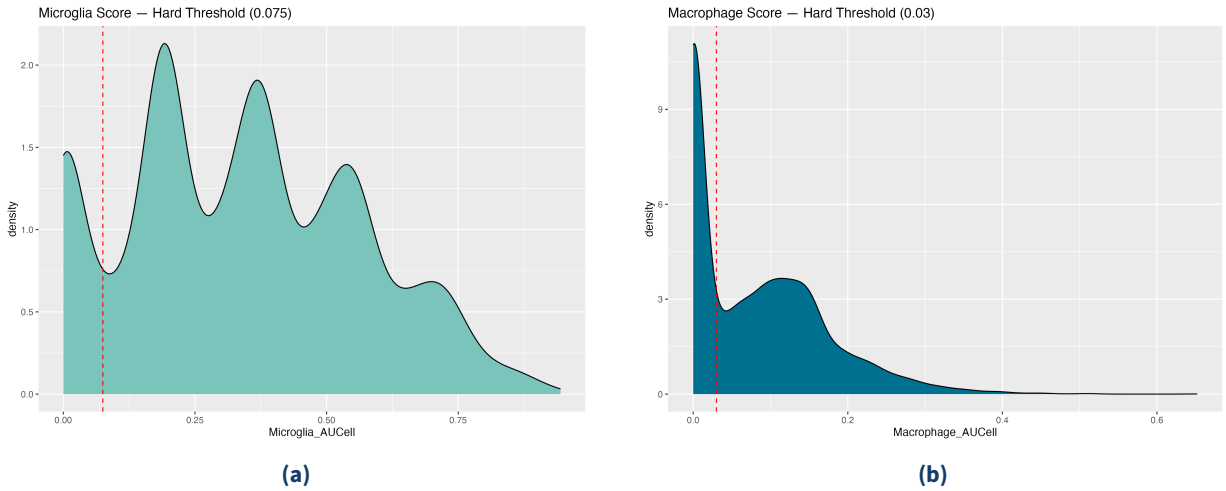


Figure 8 Density distributions of raw AUCell scores with hard thresholds overlaid for (a) Microglia and (b) Macrophage panels.

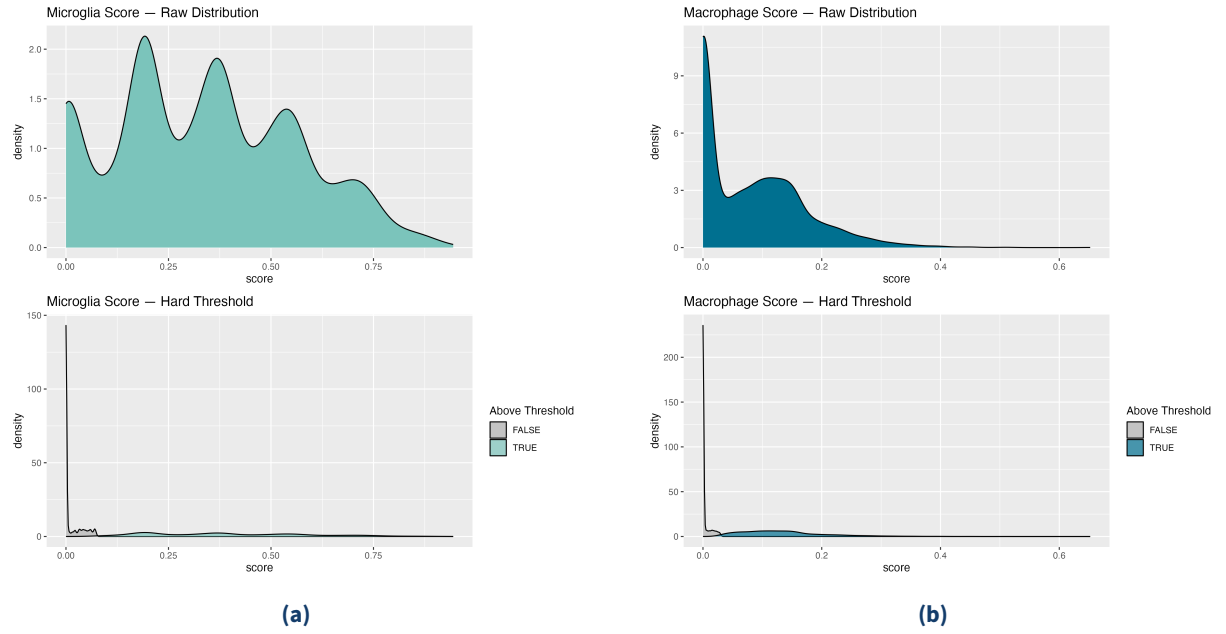


Figure 9 Raw score distributions (top) versus hard-threshold classification (bottom) for (a) Microglia and (b) Macrophage scores.

4.4 Marker Panel Expression

Dot plot visualization (Figure 12) confirmed strong, consistent expression of microglial markers (P2RY12, SALL1, TMEM119, MERTK, P2RY13) specifically in Microglia-like cells, with near-absent expression in Macrophage-like and Ambiguous cells. Macrophage markers (CD68, ADGRE1, CCR2, CD200R1, CD5L, CHIT1) showed comparatively weaker, less specific expression across all three lineage classes.

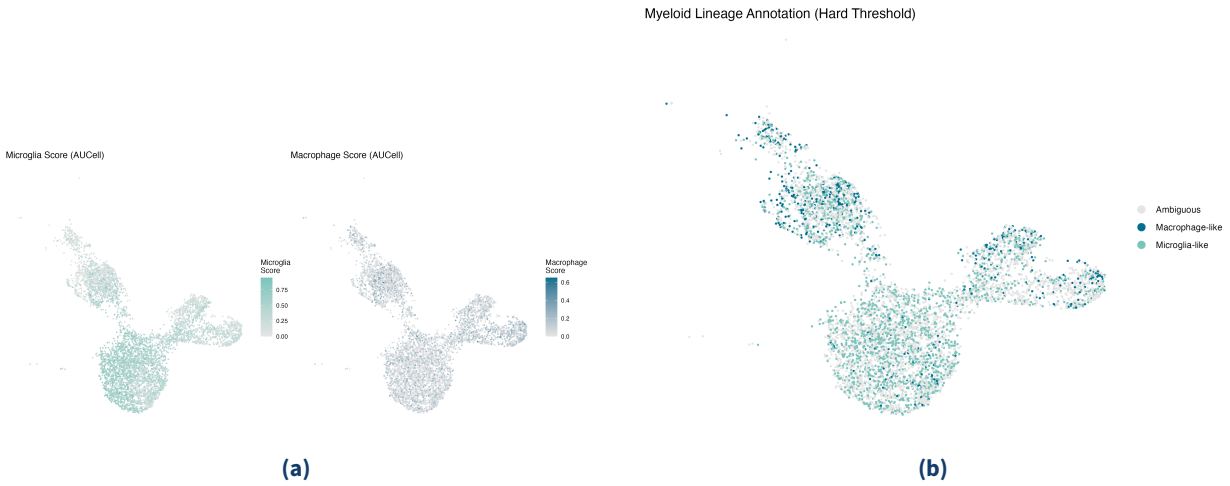


Figure 10 Myeloid lineage identity across the batch-corrected UMAP. (a) Continuous scores. (b) Discrete lineage annotation.

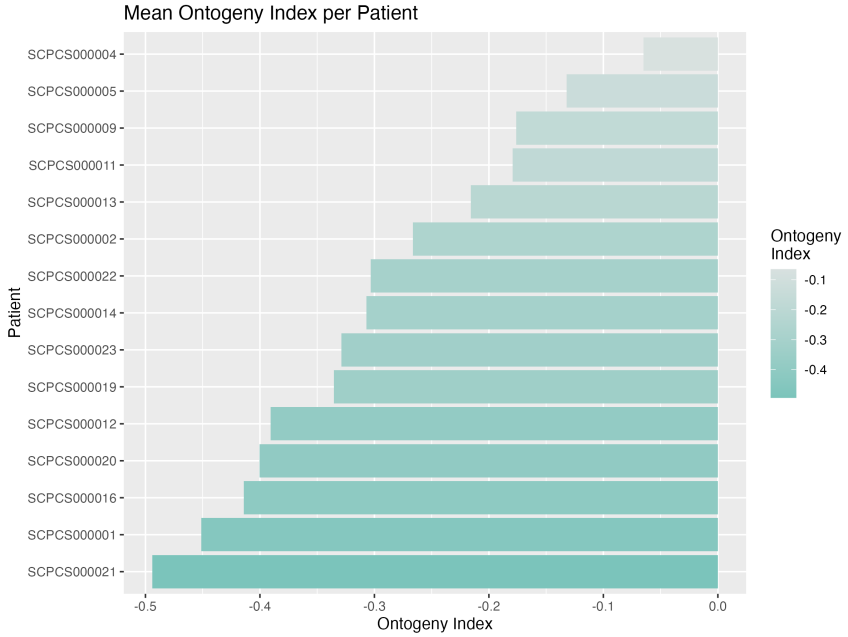


Figure 11 Mean myeloid ontogeny index per patient. All 15 evaluable patients fell on the Microglia-skewed side.

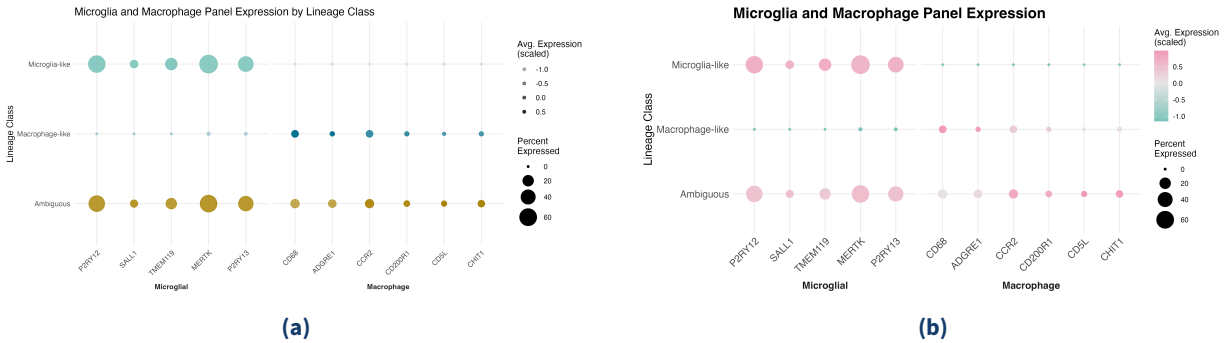
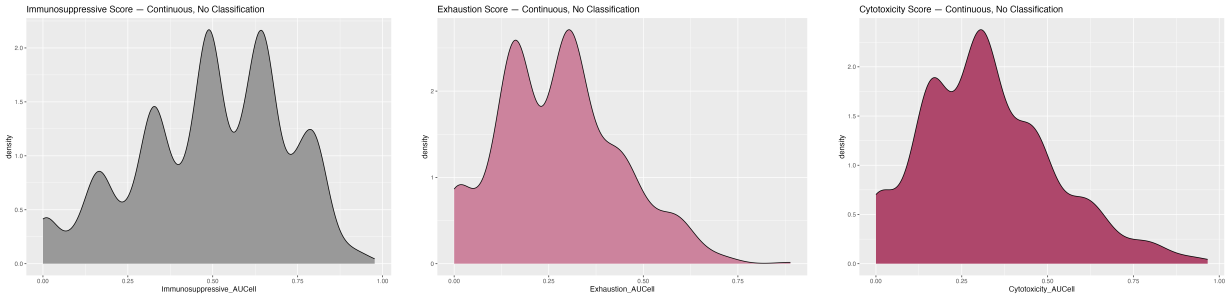


Figure 12 Microglia and Macrophage panel gene expression by lineage class.

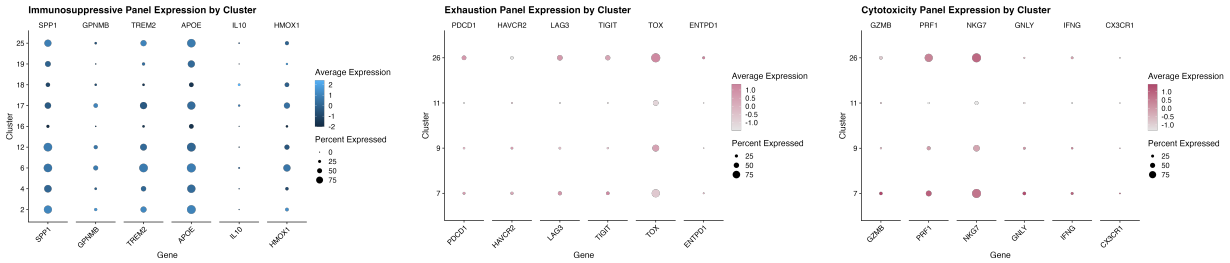
4.5 Functional Score Distributions and Cluster-Level Expression

Immunosuppressive, exhaustion, and cytotoxicity scores were treated as continuous throughout (Figure 13), all showing multimodal, right-skewed distributions consistent with the non-normal, small-n statistical approach adopted. Cluster-level dot plots (Figure 14) showed APOE, SPP1, and TREM2 as the most consistently expressed immunosuppressive markers; TOX as the dominant exhaustion marker; and NKG7 and PRF1 as the most broadly expressed cytotoxicity markers.



(a) (b) (c)

Figure 13 Cell-level AUCell score distributions, continuous, no classification applied.



(a) (b) (c)

Figure 14 Panel gene expression by cluster.

4.6 Myeloid Ontogeny Index and Immunosuppressive Score Distribution

Both the ontogeny index and immunosuppressive score varied continuously across the myeloid compartment with no discrete spatial substructure evident (Figure 15).

4.7 Primary Result — Myeloid Ontogeny vs. Immunosuppressive State

The myeloid ontogeny index was not significantly associated with the mean immunosuppressive AUCell score at the patient level (Spearman's $\rho = -0.26$, $p = 0.34$, $n = 15$; Figure 16).

4.8 Clinical and Technical Covariate Analyses

Exploratory clinical and technical covariate analyses are summarized in Table 4. The ontogeny index showed no significant association with sex, diagnosis, chemotherapy status, age, or myeloid cell count. Immunosuppressive score showed no significant association with sex or diagnosis, but was significantly higher among chemotherapy-exposed patients (Wilcoxon W = 4, p = 0.007; Figure 16). For the CD8+ outcomes (Table 5) neither exhaustion nor cytotoxicity differed significantly by sex or diagnosis, and exhaustion was not significantly associated with age.

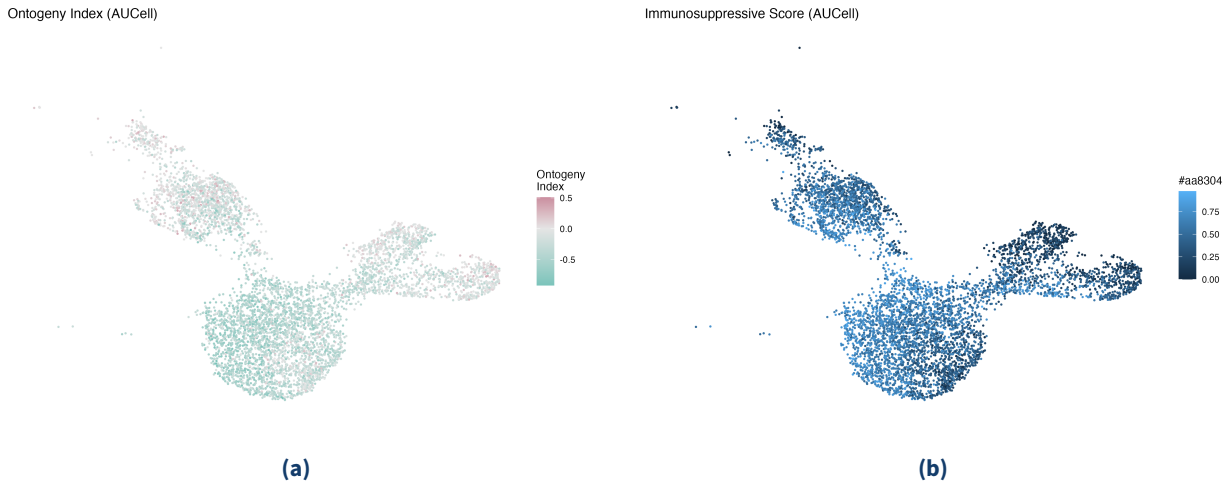


Figure 15 (a) Continuous myeloid ontogeny index. (b) Immunosuppressive AUCell score. Both across the batch-corrected UMAP.

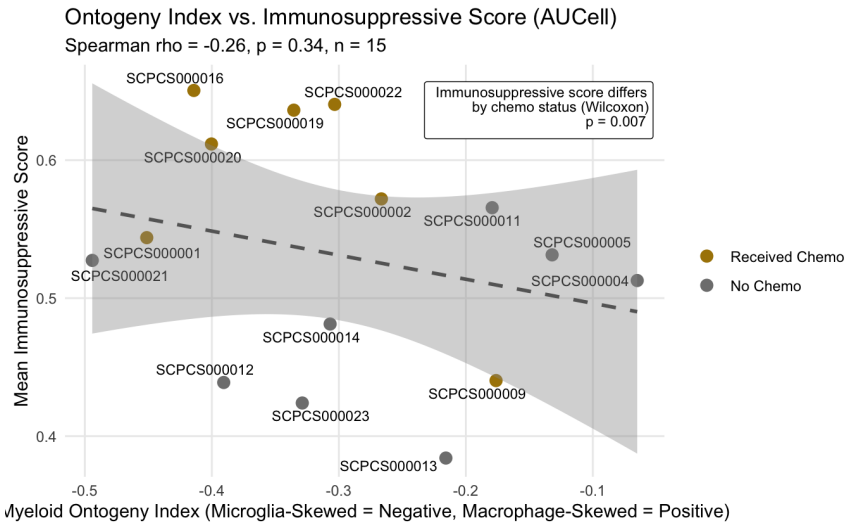


Figure 16 Primary result: myeloid ontogeny index versus mean immunosuppressive AUCell score. Spearman's $\rho = -0.26$, $p = 0.34$, $n = 15$.

Table 4 Analysis for the primary (myeloid ontogeny, immunosuppressive score) outcomes, $n = 15$.

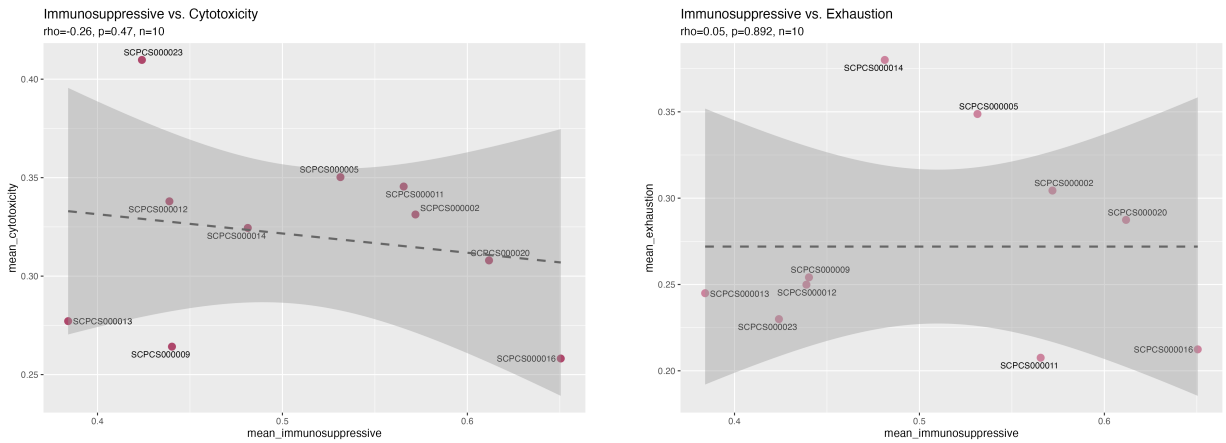
Outcome	Confound	Test	Statistic	p-value
Ontogeny index	Sex	Kruskal–Wallis	0.35	0.56
Immunosuppressive score	Sex	Kruskal–Wallis	0.01	0.91
Ontogeny index	Diagnosis	Kruskal–Wallis	6.20	0.40
Immunosuppressive score	Diagnosis	Kruskal–Wallis	4.20	0.65
Ontogeny index	Chemotherapy	Wilcoxon	31	0.46
Immunosuppressive score	Chemotherapy	Wilcoxon	4	**0.007**
Ontogeny index	Age	Spearman	$\rho = 0.26$	0.34
Ontogeny index	Myeloid cell count	Spearman	$\rho = -0.47$	0.076

Table 5 Analysis for the secondary (CD8+ exhaustion, cytotoxicity) outcomes, $n = 10$.

Outcome	Confound	Test	Statistic	p-value
Exhaustion	Sex	Wilcoxon	12	1.00
Cytotoxicity	Sex	Wilcoxon	13	0.91
Exhaustion	Diagnosis	Kruskal–Wallis	5.36	0.25
Cytotoxicity	Diagnosis	Kruskal–Wallis	4.64	0.33
Exhaustion	Age	Spearman	$\rho = -0.29$	0.42

4.9 Secondary Results — Myeloid State vs. CD8+ Function

Mean immunosuppressive score was not significantly correlated with mean CD8+ cytotoxicity ($\rho = -0.26$, $p = 0.47$, $n = 10$; Figure 17a) or mean CD8+ exhaustion ($\rho = 0.05$, $p = 0.89$, $n = 10$; Figure 17b). The ontogeny index itself was also not significantly correlated with exhaustion ($\rho = 0.22$, $p = 0.54$, $n = 10$). Both exhaustion and cytotoxicity scores showed broadly overlapping spatial distributions across the CD8+ T cell UMAP (Figure 18), and cell-level exhaustion and cytotoxicity scores showed no discernible relationship (Figure 19).



(a) (b)

Figure 17 Secondary correlations. (a) Immunosuppressive vs. cytotoxicity ($\rho = -0.26$, $p = 0.47$). (b) Immunosuppressive vs. exhaustion ($\rho = 0.05$, $p = 0.89$).

4.10 CD8+ T Cell Functional Heterogeneity

Levene's test on cell-level score distributions, grouped by patient, showed that the spread of CD8+ cytotoxicity scores differed significantly across patients ($F = 5.95$, $p = 5.1 \times 10^{-8}$, $n = 672$), while the spread of exhaustion scores did not ($F = 0.72$, $p = 0.69$, $n = 672$; Figure 20). This asymmetry indicates that within-patient cytotoxicity-score dispersion differed across patients, whereas no corresponding difference in exhaustion-score dispersion was detected.

4.11 Differential Expression Analysis

Exploratory cell-level differential expression analysis between Microglia-like and Macrophage-like cells identified 191 genes significantly upregulated in Microglia-like cells and 2 genes upregulated in Macrophage-like cells at the cell-level testing threshold ($p_{\text{adj}} < 0.05$; Wilcoxon rank-sum test, `FindMarkers`). Given the nested cell structure and classification-related circularity described

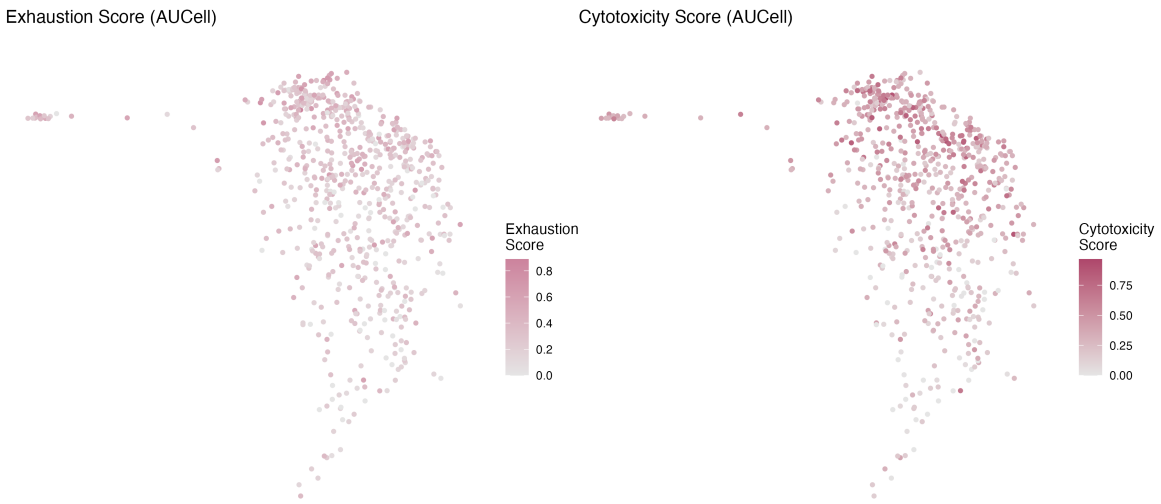


Figure 18 Exhaustion and cytotoxicity AUCell scores across the batch-corrected CD8+ T cell UMAP.

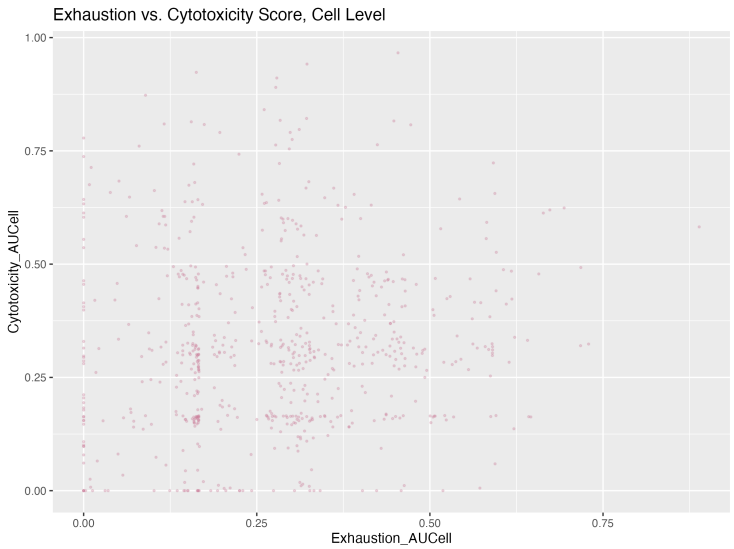


Figure 19 Cell-level exhaustion versus cytotoxicity AUCell scores across all reliable CD8+ T cells.

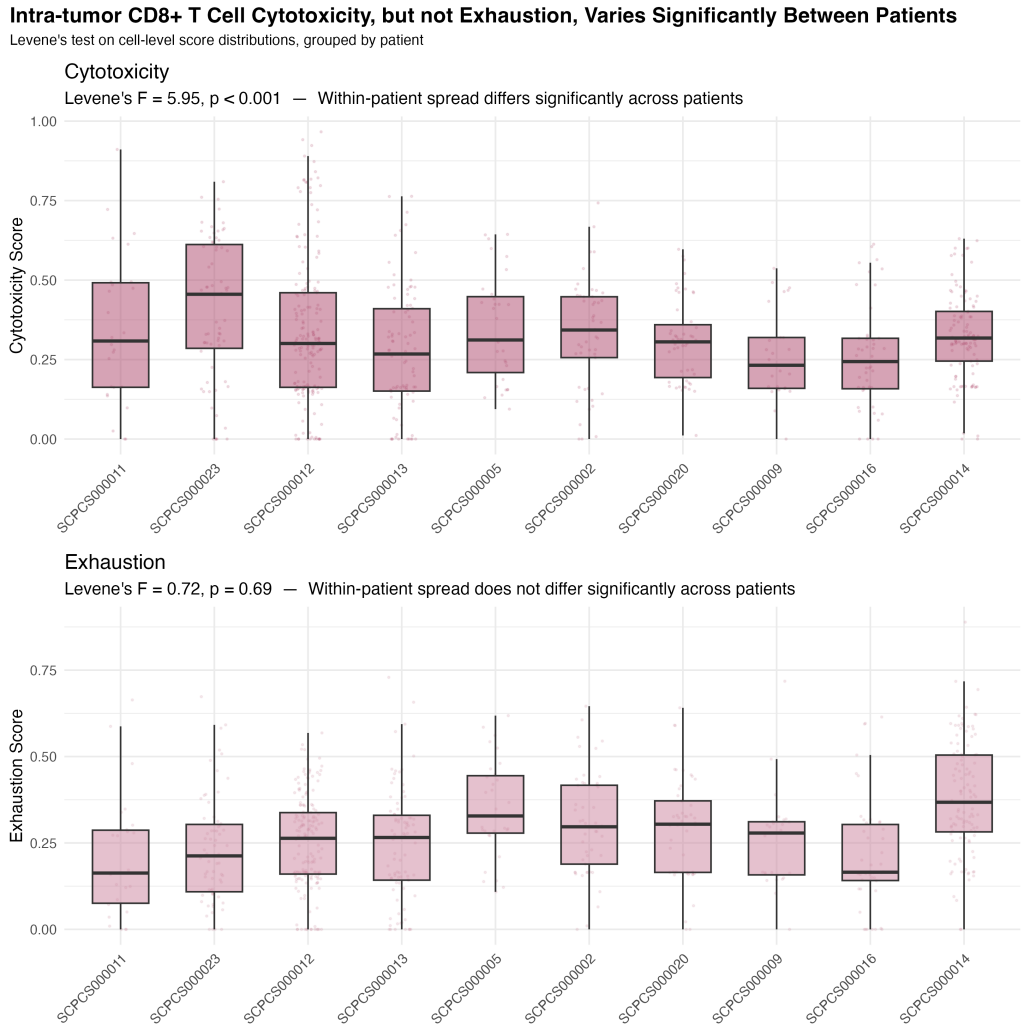


Figure 20 Within-patient CD8+ T cell cytotoxicity-score dispersion differed significantly across patients, whereas exhaustion-score dispersion did not. Levene's test: cytotoxicity, F = 5.95, p = 5.110⁻8; exhaustion, F = 0.72, p = 0.69.

in Section 3.8, these results were interpreted descriptively rather than as patient-replicated differential expression. This asymmetry mirrors the compositional skew described in Section 4.3.

4.12 Gene Ontology Enrichment

No GO terms reached significance after Benjamini–Hochberg correction in any ontology tested. Exploratory analysis at unadjusted $p < 0.05$ identified 85 Biological Process terms — comparable to the ~65 expected under the null given 1,295 terms tested, and largely circular given that top terms were driven by the same panel genes used for lineage classification. The same pattern held for Cellular Component (Figure 21) and Molecular Function (Figure 22) ontologies.

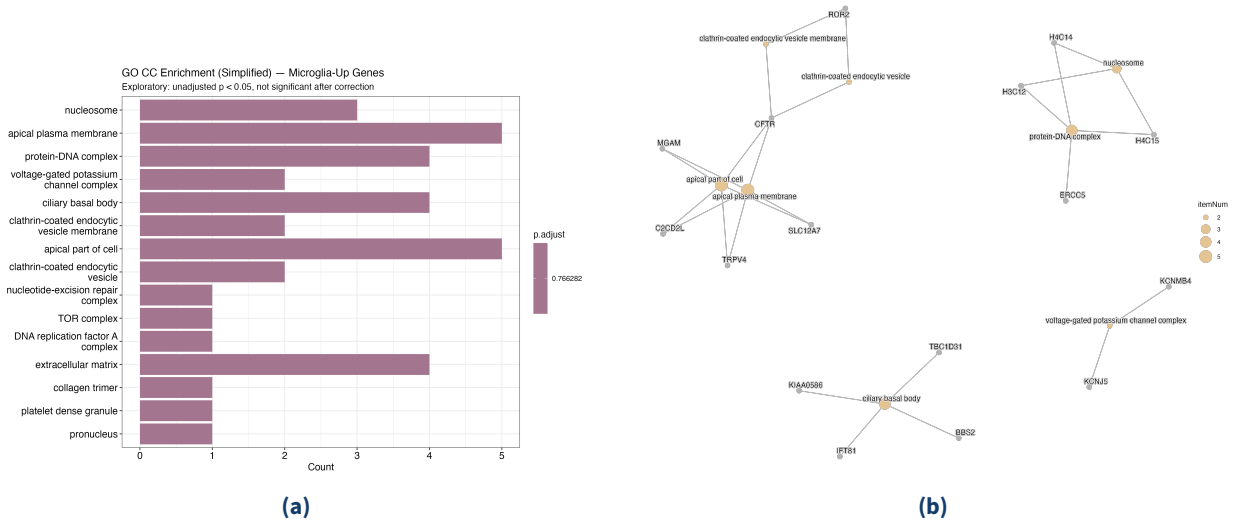


Figure 21 Exploratory GO enrichment (Cellular Component), unadjusted $p < 0.05$, not significant after correction.

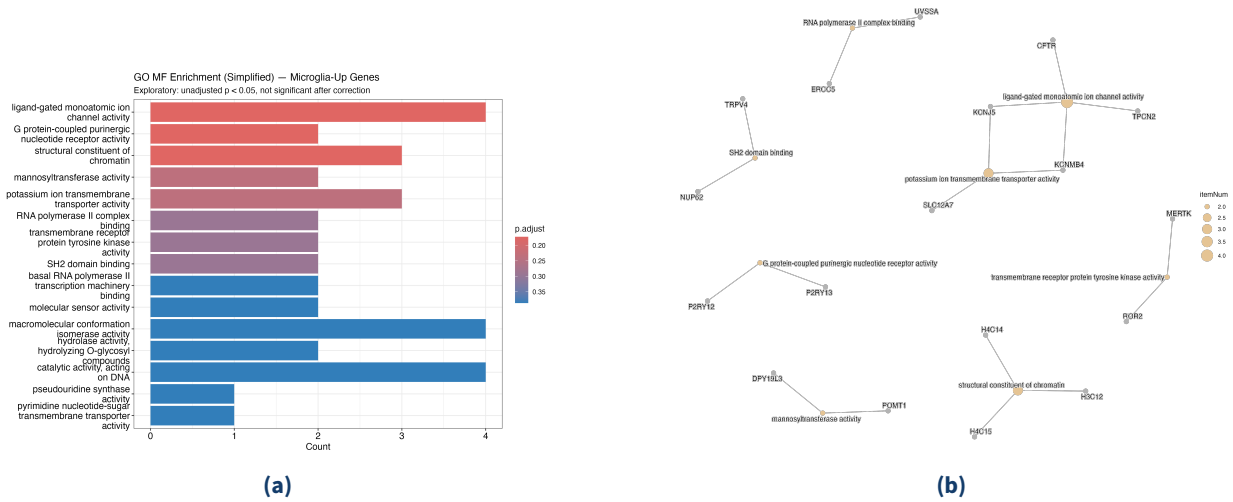


Figure 22 Exploratory GO enrichment (Molecular Function), unadjusted $p < 0.05$, not significant after correction.

5 Discussion

5.1 Myeloid Identity Is a Continuum, Not Discrete

The central aim of this study was to test whether myeloid ontogeny predicts the immunosuppressive state of tumor-associated myeloid cells in pHGG. This analysis found no significant correlation between the two ($\rho = -0.26$, $p = 0.34$, $n = 15$). Part of this null result is attributable to the small number of patients passing the reliability threshold, which limited statistical power (see Section 6 for a fuller discussion of the sample-size limitation affecting this result). However, an observation made earlier in the analysis pipeline offers a more substantive explanation: the myeloid compartment in this dataset did not resolve into two discrete populations, but instead formed a single, continuous distribution of ontogeny scores.

This conclusion was not immediate. Substantial effort went into attempting to recover a genuine two- or four-population structure from the myeloid ontogeny scores before the continuum interpretation was accepted. Model-based clustering via `mclust` was applied to the AUCell score distributions specifically to test whether a discrete mixture model — ideally recovering the four expected lineage groups (Microglia-like, Macrophage-like, Double-high, Double-low), or at minimum a clean two-component split — fit the data better than a single homogeneous population. Model-based clustering via mclust favored a two-component statistical solution. However, the separation was driven primarily by variation in the microglia score and did not resolve corresponding Microglia-like and Macrophage-like populations across both lineage axes. The fitted components were therefore not interpreted as biologically distinct ontogeny groups, indicating the data did not support a clean, biologically interpretable microglia–macrophage population structure. Only after this result was reproduced across parameterizations was the decision made to abandon a data-driven threshold in favor of the pragmatic hard cutoffs (0.075 for Microglia, 0.03 for Macrophage) used throughout the remainder of this analysis, a decision made explicitly, and confirmed as reasonable, rather than treated as a statistically discovered boundary (see Section 6 for further discussion of the sensitivity of downstream results to this threshold choice).

Several factors may account for this.

Biologically, resident microglia and monocyte-derived macrophages are known to arise from distinct embryological lineages, yet in pHGG — or at least in this cohort — they appear to converge toward a shared phenotypic signature.

Part of this convergence likely reflects a broader limitation of transcriptomic classification itself. Gene expression does not directly encode developmental origin; it reflects the transcriptional state a cell occupies at the moment of sampling. Because that state is shaped by the surrounding tissue, and tumor cells are known to actively reprogram the immune cells around them, confidently assigning lineage identity from expression data alone is inherently difficult. This problem is compounded by the limited literature establishing validated transcriptional markers for tumor-associated macrophages in pHGG specifically, and by the still-incomplete understanding of the glioma tumor microenvironment more broadly — both of which make it difficult to disentangle genuine biological ambiguity from a methodological artifact.

The restriction to 5–6 marker genes per panel may also have contributed to this ambiguity. This restriction was a deliberate choice: rather than relying on a broader but less-validated gene set, we prioritized genes with consistent support across the literature to preserve the biological defensibility of the classification. We acknowledge, however, that this conservative approach may have reduced our sensitivity to detect a distinct macrophage signal.

Despite these considerations, we consider active reprogramming of myeloid cells by the tumor microenvironment — sometimes described in the glioma immunology literature as tumor-driven "re-education" of TAMs [8] — to be the most likely contributor to the phenotypic

ambiguity observed here. Under this framework, myeloid cells may lose or blur their ontogeny-associated transcriptional identity as they adopt a tumor-imposed functional state, which would independently explain both the continuum observed here and the absence of a clean ontogeny–immunosuppression relationship.

Taken together, this null primary result does not rule out a relationship between myeloid ontogeny and immunosuppressive function in pHGG; rather, it points toward tumor-driven re-education as a candidate mechanism obscuring that relationship, and highlights the need for a more robust, pHGG-specific gene panel to resolve microglia and macrophage identity with greater confidence in future work.

5.2 Are pHGGs Truly Microglia-Dominated?

A separate but related observation was the strong compositional skew of the myeloid compartment toward a microglia-like phenotype, present consistently across every patient in this cohort. At least three, non-mutually-exclusive explanations are worth considering, summarized in Figure 23.

The first is genuine biology. Adjacent work profiling the pediatric high-grade glioma immune landscape has reported that pediatric myeloid states diverge substantially from those seen in adult glioma [17], where the tumor-associated myeloid compartment is typically described as more heavily infiltrated by monocyte-derived macrophages [6, 7]. If pediatric tumors are less permissive to peripheral monocyte recruitment, or if the developing brain environment favors expansion of resident microglia over infiltrating macrophages, a microglia-dominant compartment in pHGG would represent a genuine and biologically meaningful pediatric-versus-adult contrast rather than an artifact of this analysis.

The second is technical or panel-driven bias, and here the concern is sharper than a simple asymmetry in marker sensitivity. Beyond the macrophage panel showing weaker and less specific expression overall, the specificity of the microglial panel itself cannot be assumed to hold inside tumor tissue. Work examining marker specificity directly in glioblastoma has shown that canonical microglia-restricted genes — including SALL1, TMEM119, P2RY12, and HEXB, three of which are used in the present microglial panel — are not reliably microglia-exclusive in the glioma context: infiltrating macrophages have been shown to express these same markers after entering brain and tumor tissue, with cross-expression varying spatially and the specificity of these genes breaking down most severely toward the intratumoral compartment [10]. Complementary in vitro work in the same study found that tumor-conditioned medium alone was sufficient to downregulate microglial marker expression, independent of any change in true cell identity. Taken together, this raises the possibility that some fraction of the cells classified here as Microglia-like are in fact infiltrating macrophages that have acquired microglial marker expression under tumor influence, rather than true yolk-sac-derived resident microglia. Under this explanation, the observed skew would be partly an artifact of marker specificity collapsing inside the tumor microenvironment, rather than a faithful readout of true lineage composition.

The third is a sampling or dissociation bias inherent to the single-cell method itself. Enzymatic tissue dissociation protocols used to generate single-cell suspensions from solid tumors are known to preferentially recover certain cell populations over others, and resident, tissue-adherent cells such as microglia may survive dissociation more readily than more fragile or loosely adherent infiltrating monocyte-derived cells. Under this explanation, the observed skew would partly reflect a technical limitation of the ScPCA processing pipeline rather than the true in vivo composition of the tumor microenvironment.

Distinguishing between these three explanations is not possible with this dataset alone, since all three would produce the same observed pattern of a microglia-dominant myeloid

compartment (Figure 23). Resolving this question would require either an independent pHGG cohort processed with a dissociation-minimizing protocol (e.g., single-nucleus rather than single-cell sequencing), spatial transcriptomic validation that does not require tissue dissociation, or a validated, pHGG-specific macrophage marker panel developed independently of the one used here.

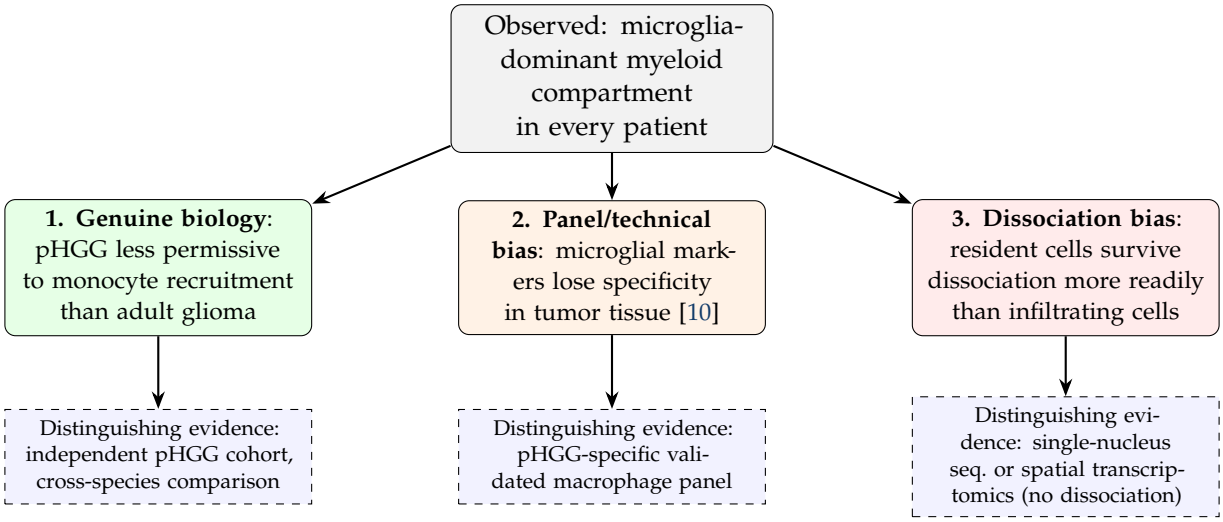


Figure 23 Three non-mutually-exclusive explanations for the observed microglia-dominant myeloid composition, each producing the identical observed pattern and therefore indistinguishable using this dataset alone. Dashed boxes indicate the type of additional evidence required to distinguish between explanations.

5.3 The 191-vs-2 DEG Asymmetry

A secondary, confirmatory analysis was performed on the classification described above: bidirectional differential gene expression analysis between the Microglia-like and Macrophage-like sub-populations identified 191 genes significantly upregulated in Microglia-like cells, compared to only 2 genes significantly upregulated in Macrophage-like cells.

Many of the top genes driving this asymmetry were the same genes used to construct the microglial and macrophage AUCell panels in the first place. This is an expected, and in one sense reassuring, result: it confirms that the cells assigned the Microglia-like label are indeed transcriptionally consistent with the panel used to define that label, functioning as a basic sanity check on the classification pipeline rather than a novel biological discovery. Because this circularity is built into the design of the confirmatory DGE step, the specific magnitude of the asymmetry should not be over-interpreted as evidence that Macrophage-like cells are transcriptionally quiescent or uninteresting; rather, it reflects that the classification is doing what it was designed to do.

The size of the asymmetry is also partly a direct consequence of the compositional skew discussed above. With far fewer cells assigned to the Macrophage-like class, the differential expression test has correspondingly less statistical power to detect genes upregulated in that direction, independent of any true underlying difference in transcriptional activity between the two populations. The 191-vs-2 split should therefore be read as a combination of expected circularity and a sample-size effect, rather than as an independent line of evidence for the biological distinctiveness of the two classes. A natural next question is whether this same gene set reveals anything further at the pathway level once the obvious panel-gene signal is accounted

for — a question addressed directly in the following section.

5.4 Null GO is Informative

Gene Ontology enrichment analysis was performed on the 191 genes upregulated in Microglia-like cells to test whether this signal reflected a coherent underlying biological program beyond the panel genes used for classification. No terms reached significance in any ontology (Biological Process, Molecular Function, or Cellular Component) after Benjamini–Hochberg correction. At an unadjusted threshold of $p < 0.05$, 85 Biological Process terms were nominally significant — a number that is not meaningfully different from the approximately 65 terms expected to appear by chance alone given 1,295 terms tested at this threshold.

The content of these unadjusted terms reinforces this conclusion rather than complicating it. Rather than surfacing categories with plausible relevance to microglial identity, tumor immunology, or myeloid function, the top nominally significant terms were dominated by generic, thematically unrelated processes such as spermatid differentiation, sperm flagellum assembly, motile cilium assembly, and general chromosome organization. These categories share no obvious biological connection to the classification task at hand, and their appearance is far more consistent with the pattern expected from a gene list enriched by chance — or by a handful of large, non-specific gene sets happening to overlap with the DEG list — than with a genuine, coherent microglia-associated transcriptional program. Had the top unadjusted terms instead clustered around immune signaling, neuroinflammation, or myeloid activation pathways, that would have offered at least suggestive support for a real biological signal beneath the correction threshold; the fact that they did not further supports treating this enrichment analysis as null in substance as well as in adjusted significance.

This null result should be read as informative rather than as a failure of the analysis. Given the circularity established in the previous section — where the DEG list is substantially composed of the same genes used to define the Microglia-like label — a strongly significant, coherent pathway enrichment would actually have been the more suspicious outcome, since it could simply reflect the panel genes' own annotated functions being reported back as an "enriched" signal rather than genuine independent biology. The absence of any robust enrichment beyond chance level, combined with the biologically incoherent content of the nominal hits, is therefore consistent with, rather than contradictory to, the broader picture developed in this Discussion: that the transcriptional distinction between Microglia-like and Macrophage-like cells in this cohort is real enough to be constructed by a marker panel, but does not extend into a distinct, independently discoverable functional program at the pathway level. This adds a second, independent line of evidence — alongside the continuum result in Section 5.1 — that the microglia/macrophage distinction in this cohort is weaker and more diffuse than a hard classification pipeline assumes.

5.5 Chemotherapy and Immunosuppression

Although myeloid ontogeny did not significantly predict immunosuppressive score, a confound analysis conducted to rule out clinical and technical drivers of the primary result surfaced a striking secondary finding: immunosuppressive score differed significantly by chemotherapy status (Wilcoxon, $p = 0.007$), with patients who had received chemotherapy showing consistently higher immunosuppressive scores than those who had not. Given that this comparison emerged from a confound check rather than the primary hypothesis, its statistical strength is notable, and it arguably represents the single clearest signal in this entire dataset.

A plausible mechanism is that chemotherapy itself drives myeloid cells toward an immunosuppressive phenotype, as part of a broader tissue damage or stress response, independent of

the developmental origin of the cell. Chemotherapy-induced myeloid reprogramming toward an immunosuppressive state has precedent in other cancer contexts, and would fit a model in which tumor-associated myeloid function in pHGG is shaped more strongly by treatment history and the local injury environment than by whether a cell descends from resident microglia or infiltrating monocytes. If this relationship holds, it has real practical value: unlike ontogeny, which is fixed and cannot be intervened on, chemotherapy timing, sequencing, and regimen are modifiable variables. A confirmed link between chemotherapy exposure and myeloid immunosuppression would motivate directly testable questions with clinical relevance — for instance, whether adjusting chemotherapy scheduling relative to planned immunotherapy could avoid compounding an already immunosuppressive myeloid state, or whether specific agents differ in how strongly they drive this effect, information that could inform the design of combination treatment protocols for a disease with very few effective options.

At the same time, this result cannot be reduced to "patients who received chemotherapy end up with more suppressive myeloid programs" without care regarding causal direction, because the causality could plausibly run the other way. Patients selected for or steered toward chemotherapy may have differed systematically at baseline — in tumor burden, aggressiveness, or an already-immunosuppressive microenvironment that influenced treatment decisions in the first place — meaning a pre-existing immunosuppressive state could have shaped treatment choice rather than resulted from it. The cross-sectional, single-timepoint design of this dataset cannot distinguish between chemotherapy driving immunosuppression, immunosuppression driving chemotherapy allocation, or some bidirectional combination of the two, and this comparison was not pre-specified, so it should be treated as hypothesis-generating rather than confirmatory (see Section 6 for further discussion of the cross-sectional design limitation).

Within this cohort, chemotherapy status showed a stronger association with myeloid immunosuppressive state than ontogeny did. However, the present univariate analyses cannot establish independence between these variables or determine causal direction. Resolving this relationship through a larger, longitudinally sampled cohort capable of comparing myeloid state before and after treatment, and capturing treatment timing and specific agent rather than a binary status, is therefore a natural next step arising from this pilot analysis.

5.6 Phenotype vs. Function — The CD8 Dissociation

Because both scores were computed on the same CD8+ T cell subset, this asymmetry indicates different patterns of within-patient dispersion across the two transcriptional programs: cytotoxicity heterogeneity varied significantly between patients, whereas no corresponding difference was detected for exhaustion. The most substantive positive result of this study emerged not from a comparison of central tendency, but from a comparison of within-tumor variance. Levene's test on cell-level score distributions, grouped by patient, showed that the spread of CD8+ cytotoxicity scores differed significantly across patients ($F = 5.95$, $p < 0.001$), while the spread of exhaustion scores did not ($F = 0.72$, $p = 0.69$). Both scores were computed on the identical CD8+ T cell subset per patient, so this asymmetry reflects a genuine difference in how these two functional axes vary across the cohort, rather than an artifact of differing sample composition between the two tests.

This result is best interpreted in light of a broader shift in how T cell exhaustion itself is understood. Rather than a single, uniform dysfunctional endpoint, exhaustion is now widely described as a hierarchically organized, epigenetically constrained differentiation continuum, spanning progenitor or precursor-exhausted cells (Tpex), through intermediate or transitory states, to terminally exhausted cells [13, 14, 32], illustrated in Figure 24. This distinction matters directly for the present analysis, because it separates *phenotypic* exhaustion — surface expression of

inhibitory checkpoint receptors such as PD-1, TIM-3, and LAG-3, which is what a transcriptional exhaustion score of the kind used here actually measures — from *functional* exhaustion, the genuine loss of proliferative and cytotoxic capacity. Precursor-exhausted cells are known to co-express several of these same checkpoint markers while retaining substantial effector potential and responsiveness to checkpoint blockade, whereas only terminally exhausted cells show the stable, largely irreversible functional decline the term "exhaustion" originally described [13, 14, 32]. A relatively uniform phenotypic exhaustion score across patients is therefore consistent with a population of CD8+ T cells sitting at broadly similar points along this checkpoint-expression axis, without that uniformity implying anything about how much genuine cytotoxic function each patient's population has actually retained — which is precisely the axis on which this cohort showed significant heterogeneity instead.

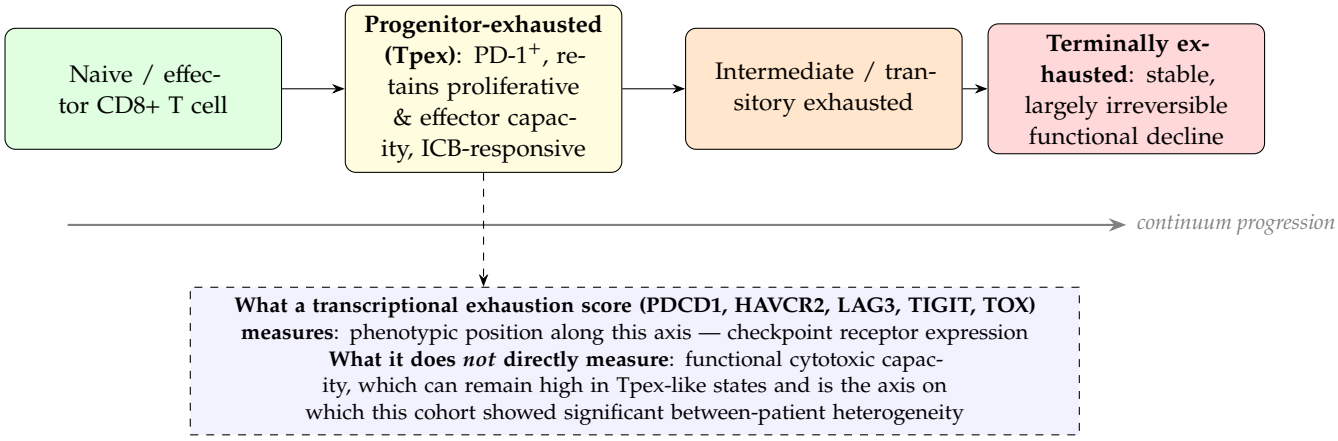


Figure 24 The T cell exhaustion continuum, from naive/effector through progenitor-exhausted (Tpex) and intermediate states to terminal exhaustion [13, 14, 32]. A transcriptional exhaustion score measures phenotypic position along this axis but does not directly capture functional cytotoxic capacity, which can remain substantial in progenitor-exhausted cells — the distinction underlying the CD8 dissociation reported in this study.

This interpretation is further supported by direct evidence from the pediatric high-grade glioma literature. In the largest single-cell and spatial transcriptomic profiling of the pHGG immune landscape to date, LaBelle et al. reported a striking absence of canonical CD8+ T cell exhaustion or activation signatures, in stark contrast to melanoma, a tumor type that responds well to checkpoint immunotherapy [17]. If low, relatively uniform phenotypic exhaustion is a general feature of the pHGG CD8+ compartment rather than a peculiarity of this cohort, it would directly explain why exhaustion scores showed no significant between-patient heterogeneity here, while leaving open — as the pattern observed in this study suggests — that meaningful, clinically relevant variability can still exist in the cytotoxic function of these cells even when their phenotypic exhaustion profile looks broadly similar across tumors. Taken together, these findings are consistent with phenotypic exhaustion and cytotoxicity capturing partially distinct dimensions of CD8+ T cell state. The present analysis does not establish that these programs are statistically independent, but it shows that their patterns of patient-level heterogeneity are not identical.

The biological drivers of this patient-specific cytotoxic ceiling cannot be determined from this dataset and remain speculative. Candidate explanations include differences in tumor antigen burden or immunogenicity, TCR clonal diversity within the infiltrating CD8+ population, time elapsed since diagnosis or treatment, and prior therapeutic exposure — none of which were directly measured here and each of which represents a distinct, testable hypothesis for future work.

If this dissociation holds in a larger cohort, it carries a specific translational implication worth stating carefully. Therapeutic strategies built around reversing T cell exhaustion, such as checkpoint blockade, implicitly assume that phenotypic exhaustion tracks functional exhaustion closely enough that relieving the former will restore the latter. The present result, read alongside the broader literature on the exhaustion continuum and the specific absence of a strong exhaustion signature already reported in pHGG, suggests that phenotypic exhaustion may instead be a relatively fixed, low-variance feature of this tumor's immune microenvironment, while the ceiling on cytotoxic capacity — the feature that actually varies patient to patient in this cohort — may be the more clinically informative axis to measure and target. This reframing is offered here as a hypothesis for direct testing in future, adequately powered cohorts, not as a claim established by the present pilot analysis (see Section 6 for a full discussion of the functional-validation limitation affecting this interpretation). It is this dissociation, rather than the null ontogeny–immunosuppression result, that constitutes the principal empirical contribution of this study.

6 Limitations and Future Work

Several limitations qualify the results presented in this study, most of which have already been raised in context within the Discussion but are consolidated here for clarity.

The most fundamental limitation is sample size. The primary analysis was based on only 15 patients passing the reliability threshold for the ontogeny–immunosuppression correlation, and secondary CD8+ analyses were restricted to 10 patients meeting the ≥20-cell reliability criterion. These are small samples for correlation and variance-based statistical testing, yielding wide confidence intervals and limited power to detect real effects of moderate size. The null results reported throughout this study — most notably the primary ontogeny–immunosuppression correlation — cannot be interpreted as strong evidence of a true absence of effect; they are equally consistent with a real but modest effect that this cohort was underpowered to detect.

Related to sample size is the reliability filtering itself. The ≥20-cell threshold applied to CD8+ T cells, while necessary to avoid drawing conclusions from noise-dominated per-patient scores, excluded the majority of the original cohort from the secondary analyses. If the patients excluded by this filter differ systematically from those retained — for instance, in tumor size, immune infiltration, or diagnosis — the resulting secondary analyses may not generalize to the full pHGG population, and this possibility cannot be ruled out with the present data.

The hard AUCell thresholds used to classify myeloid cells as Microglia-like or Macrophage-like (0.075 and 0.03, respectively) were selected as a pragmatic, visually and distributionally justified cutoff rather than a threshold discovered through an independent statistical procedure. Model-based clustering via `mclust` favored a two-component statistical solution, but this separation was driven predominantly by the microglia-score axis and did not correspond to clearly separable Microglia-like and Macrophage-like populations. While this pragmatic threshold was adopted deliberately after this alternative failed to resolve a clean structure, it means the resulting cell counts and downstream compositional statistics are sensitive to a threshold choice that is not itself statistically derived, and a different reasonable choice of cutoff could shift the precise numbers reported here.

The confirmatory differential expression and Gene Ontology analyses are subject to a circularity limitation discussed at length above: because the DEG list is substantially composed of the same genes used to construct the classification panels, both the magnitude of the DEG asymmetry and the content of the GO enrichment results should be interpreted as validating the classification pipeline rather than as independent evidence of novel microglia-macrophage biology in this cohort.

This study is also cross-sectional, drawing on a single timepoint per patient. No dynamic or longitudinal information was available, meaning the analysis cannot speak to how myeloid ontogeny, immunosuppressive state, or CD8+ functional heterogeneity change over the course of disease progression or in response to treatment. This limitation is particularly relevant to the chemotherapy-immunosuppression association reported above, where the direction of causality cannot be established from single-timepoint data.

Finally, all conclusions in this study rest on transcriptomic inference alone. No functional assays (e.g., cytotoxicity assays, ex vivo killing assays) or spatial validation were performed to confirm that the transcriptional signatures used here correspond to genuine functional differences in cytotoxic capacity, immunosuppressive activity, or myeloid lineage identity in situ. The interpretation of the central CD8 dissociation finding, in particular, would benefit substantially from orthogonal functional validation — for instance, staining for progenitor-exhaustion markers such as TCF7 or SLAMF6 (Figure 24) would allow direct testing of whether the uniform phenotypic exhaustion signature observed here reflects a genuinely progenitor-like, functionally reversible

state, or a more heterogeneous mixture that a bulk transcriptional score cannot resolve.

These limitations motivate several concrete directions for future work, summarized on a near- to long-term timeline in Figure 25. A larger, ideally multi-cohort replication — potentially incorporating the GSE227983 and GSE271936 accessions identified during the initial scoping of this project as candidate pooling datasets — would directly address the power limitations affecting both the primary and secondary analyses, and would allow the reliability thresholds used here to be relaxed without sacrificing patient-level confidence. Longitudinal or paired pre-/post-treatment sampling would allow the chemotherapy–immunosuppression relationship to be tested for directionality rather than mere association, addressing the chicken-or-egg ambiguity raised in Section 5.5 directly. Distinguishing between the three candidate explanations proposed for the observed microglia-dominant composition (Section 5.2, Figure 23) would require either single-nucleus sequencing to minimize dissociation bias, spatial transcriptomic profiling that avoids tissue dissociation altogether, or an independently validated, pHGG-specific macrophage marker panel. Most directly, the central finding of this study — significant differences in within-patient CD8+ cytotoxicity-score dispersion across patients, with no corresponding difference detected for exhaustion-score dispersion — warrants dedicated functional follow-up, ideally combining single-cell transcriptomics with ex vivo functional assays or spatial proteomics, to establish whether this transcriptional dissociation corresponds to a genuine, clinically actionable difference in anti-tumor cytotoxic capacity between patients.

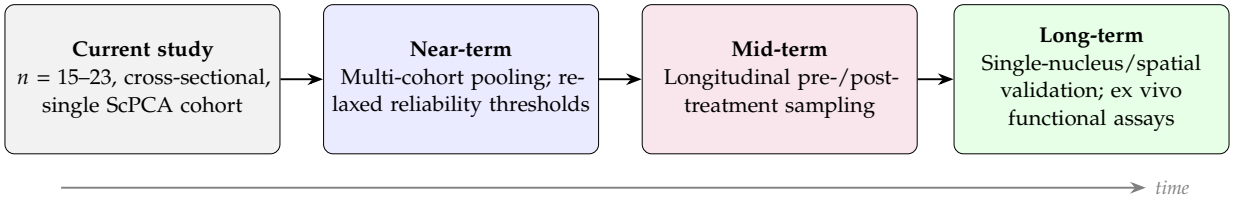


Figure 25 Proposed roadmap for future work, addressing the power (near-term), causal-direction (mid-term), and functional-validation (long-term) limitations of the present study.

7 Conclusion

This study explored whether myeloid ontogeny predicts immunosuppressive functional state in pediatric high-grade glioma, and, secondarily, how myeloid state relates to CD8+ T cell exhaustion and cytotoxicity. Myeloid ontogeny did not significantly predict immunosuppressive score, a null result consistent with a continuum, rather than discrete, model of microglia–macrophage identity in this tumor context. In its place, this pilot analysis surfaced a more substantive finding: within-patient CD8+ T cell cytotoxicity-score dispersion differed significantly across patients, whereas no corresponding difference was detected for exhaustion-score dispersion, suggesting that checkpoint-associated exhaustion and cytotoxicity exhibit different patterns of patient-level heterogeneity in pHGG and may capture distinct dimensions of CD8+ T cell state. Taken together with the observed association between chemotherapy exposure and myeloid immunosuppressive state, these results reposition cytotoxic reserve, rather than myeloid developmental origin, as the more promising axis for future investigation into immunotherapy responsiveness in this disease.

References

[1] Chris Jones, Lara Perryman, and Darren Hargrave. "Pediatric high-grade glioma: biologically and clinically in need of new thinking". In: *Neuro-Oncology* 19.2 (2017), pp. 153–161.

[2] David N. Louis et al. "The 2021 WHO Classification of Tumors of the Central Nervous System: A Summary". In: *Neuro-Oncology* 23.8 (2021), pp. 1231–1251. DOI: 10.1093/neuonc/noab106.

[3] Amelia Foss and Manav Pathania. "Pediatric Glioma Models Provide Insights into Tumor Development and Future Therapeutic Strategies". In: *Developmental Neuroscience* 46.1 (2024), pp. 22–43. DOI: 10.1159/000531040.

[4] Dalia Haydar, Jorge Ibañez-Vega, and Giedre Krenciute. "T-Cell Immunotherapy for Pediatric High-Grade Gliomas: New Insights to Overcoming Therapeutic Challenges". In: *Frontiers in Oncology* 11 (2021), p. 718030. DOI: 10.3389/fonc.2021.718030.

[5] Nicole M. Anderson and M. Celeste Simon. "The Tumor Microenvironment". In: *Current Biology* 30.16 (2020), R921–R925.

[6] Sören Müller et al. "Single-cell profiling of human gliomas reveals macrophage ontogeny as a basis for regional differences in macrophage activation in the tumor microenvironment". In: *Genome Biology* 18 (2017), p. 234. DOI: 10.1186/s13059-017-1362-4.

[7] Florian Klemm et al. "Interrogation of the Microenvironmental Landscape in Brain Tumors Reveals Disease-Specific Alterations of Immune Cells". In: *Cell* 181.7 (2020), pp. 1643–1660.

[8] Robert L. Bowman et al. "Macrophage Ontogeny Underlies Differences in Tumor-Specific Education in Brain Malignancies". In: *Cell Reports* 17.9 (2016), pp. 2445–2459.

[9] Georgios Solomou, Adam M. H. Young, and Harry J. C. J. Bulstrode. "Microglia and macrophages in glioblastoma: landscapes and treatment directions". In: *Molecular Oncology* 18.12 (2024), pp. 2906–2926. DOI: 10.1002/1878-0261.13657.

[10] Alexander D. Bungert et al. "Microglia-specificity of different markers is overridden in glioblastoma specimens". In: *Scientific Reports* 16 (2026), p. 14687. DOI: 10.1038/s41598-026-52315-y.

[11] Ariel J. Boutilier and Sherine F. Elsawa. "Macrophage Polarization States in the Tumor Microenvironment". In: *International Journal of Molecular Sciences* 22.13 (2021), p. 6995. DOI: 10.3390/ijms22136995.

[12] Michael Kilian et al. "MHC Class II-Restricted Antigen Presentation Is Required to Prevent Dysfunction of Cytotoxic T Cells by Blood-Borne Myeloids in Brain Tumors". In: *Cancer Cell* 41.2 (2023), 235–251.e9. DOI: 10.1016/j.ccell.2022.12.007.

[13] E. John Wherry and Susan M. Kaech. "Molecular and cellular insights into T cell exhaustion". In: *Nature Reviews Immunology* 15.8 (2015), pp. 486–499.

[14] Angela Baessler and Dario A. A. Vignali. "T Cell Exhaustion". In: *Annual Review of Immunology* 42.1 (2024), pp. 179–206.

[15] Pakawat Chongsathidkiet et al. "Sequestration of T Cells in Bone Marrow in the Setting of Glioblastoma and Other Intracranial Tumors". In: *Nature Medicine* 24.9 (2018), pp. 1459–1468. DOI: 10.1038/s41591-018-0135-2.

[16] Karolina Woroniecka et al. "T-Cell Exhaustion Signatures Vary with Tumor Type and Are Severe in Glioblastoma". In: *Clinical Cancer Research* 24.17 (2018), pp. 4175–4186. DOI: 10.1158/1078-0432.CCR-17-1846.

[17] James J. LaBelle et al. “Dissecting the immune landscape in pediatric high-grade glioma reveals cell state changes under therapeutic pressure”. In: *Cell Reports Medicine* 6.5 (2025), p. 102095.

[18] Yi Zheng et al. “Consensus statement on microglial and macrophage functions in gliomas”. In: *Acta Neuropathologica* 151.1 (2026), p. 39.

[19] Ruixue Ma, Xiaoxin Chen, and Xiaokun Li. “Macrophages/Microglia in the Glioblastoma Tumor Microenvironment”. In: *International Journal of Molecular Sciences* 22.11 (2021), p. 5775.

[20] David H. Gutmann and Helmut Kettenmann. “Microglia/Brain Macrophages as Central Drivers of Brain Tumor Pathobiology”. In: *Neuron* 104.3 (2019), pp. 442–449.

[21] Mariella G. Filbin et al. “Developmental and oncogenic programs in H3K27M gliomas dissected by single-cell RNA-seq”. In: *Science* 360.6386 (2018), pp. 331–335.

[22] Cyril Neftel et al. “An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma”. In: *Cell* 178.4 (2019), pp. 835–849.

[23] John DeSisto et al. “Tumor and immune cell types interact to produce heterogeneous phenotypes of pediatric high-grade glioma”. In: *Neuro-Oncology* 26.3 (2024), pp. 538–552.

[24] J. Cvitkovic et al. “Antibody-based targeting of the SPP1-CD44 axis in pediatric high-grade glioma through single-cell and structural bioinformatics”. In: *bioRxiv* (2025). DOI: 10.1101/2025.05.01.651763.

[25] T. J. M. van den Broek et al. “Single-cell spatial analysis of pediatric high-grade glioma reveals a novel population of SPP1+/GPNMB+ myeloid cells with immunosuppressive and tumor-promoting capabilities”. In: *bioRxiv* (2025). DOI: 10.1101/2025.03.18.643953.

[26] Jarne Messiaen, Steven A. Jacobs, and Frederik De Smet. “The tumor micro-environment in pediatric glioma: Friend or foe?” In: *Frontiers in Immunology* 14 (2023), p. 1227126.

[27] Antoinette du Chatinier et al. “Microglia in pediatric brain tumors: The missing link to successful immunotherapy”. In: *Cell Reports Medicine* 4.11 (2023), p. 101246. DOI: 10.1016/j.xcrm.2023.101246.

[28] Amir Yoel et al. “Emerging and Biological Concepts in Pediatric High-Grade Gliomas”. In: *Cells* 13.17 (2024), p. 1492.

[29] Kai Kang et al. “T cell exhaustion in human cancers”. In: *Biochimica et Biophysica Acta - Reviews on Cancer* 1879.5 (2024), p. 189162.

[30] Qiang Wang, Yuan Qin, and Bo Li. “CD8+ T cell exhaustion and cancer immunotherapy”. In: *Cancer Letters* 559 (2023), p. 216043.

[31] Vidhya M. Ravi et al. “T-cell dysfunction in the glioblastoma microenvironment is mediated by myeloid cells releasing interleukin-10”. In: *Nature Communications* 13.1 (2022), p. 925.

[32] Qian Sun and Chen Dong. “Regulators of CD8+ T cell exhaustion”. In: *Nature Reviews Immunology* 26.2 (2026), pp. 129–151.

[33] Caitlin C. Zebley and Ben Youngblood. “Mechanisms of T cell exhaustion guiding next-generation immunotherapy”. In: *Trends in Cancer* (2022).

[34] Rasha Barakat et al. “Human single cell RNA-sequencing reveals a targetable CD8+ exhausted T cell population that maintains mouse low-grade glioma growth”. In: *Nature Communications* 15.1 (2024), p. 10312. DOI: 10.1038/s41467-024-54569-4.

[35] Yuhan Hao et al. “Integrated analysis of multimodal single-cell data”. In: *Cell* 184.13 (2021), pp. 3573–3587.

[36] A. G. Hawkins et al. “The Single-Cell Pediatric Cancer Atlas: Data portal and open-source tools for single-cell transcriptomics of pediatric tumors”. In: *Cell Genomics* (2026).

[37] Dvir Aran et al. “Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage”. In: *Nature Immunology* 20.2 (2019), pp. 163–172. DOI: 10.1038/s41590-018-0276-y.

[38] H. Pagès et al. *AnnotationDbi: Manipulation of SQLite-based annotations in Bioconductor*. R package version. Bioconductor. https://bioconductor.org/packages/AnnotationDbi.

[39] Ilya Korsunsky et al. “Fast, sensitive and accurate integration of single-cell data with Harmony”. In: *Nature Methods* 16.12 (2019), pp. 1289–1296.

[40] Sara Aibar et al. “SCENIC: single-cell regulatory network inference and clustering”. In: *Nature Methods* 14.11 (2017), pp. 1083–1086.

[41] Massimo Andreatta and Santiago J. Carmona. “UCell: Robust and scalable single-cell gene signature scoring”. In: *Computational and Structural Biotechnology Journal* 19 (2021), pp. 3796–3798.

[42] Luca Scrucca et al. “mclust 5: Clustering, Classification and Density Estimation Using Gaussian Finite Mixture Models”. In: *The R Journal* 8.1 (2016), pp. 289–317.

[43] Tianzhi Wu et al. “clusterProfiler 4.0: A universal enrichment tool for interpreting omics data”. In: *The Innovation* 2.3 (2021), p. 100141.