Simen Owen Academic Proceedings Series

3rd International Conference on Electronics, Engineering, Computer Science and Applied Development (EESD 2026)

Article

Age-Associated Transcriptional Programs in Tumor-Proximal Macrophage-Associated Cells during Bone Metastasis

Zhe Shen[1,*]

[1] Denison University, Granville, USA
* Correspondence: Zhe Shen, Denison University, Granville, USA

Abstract: Host age can profoundly remodel the metastatic microenvironment, yet the transcriptional consequences of aging within tumor-proximal myeloid cells remain incompletely defined. In this study, we analyzed a mouse bone-metastasis single-cell RNA-sequencing dataset stratified by age and biotin labeling using SAMENT, an in vivo niche-labeling approach that selectively biotin-labels cells in close proximity to disseminated cancer cells. These neighboring cells are designated biotin-positive (biotin+) cells, whereas unlabeled cells are classified as biotin-negative (biotin−) cells. Sample-level pseudobulk differential expression analysis and Gene Ontology Biological Process gene set enrichment analysis were employed to compare aged and young mice across macrophage, neutrophil, and pre-neutrophil populations. The most pronounced age-associated transcriptional pattern was observed in biotin+ macrophage-associated cells. Young samples exhibited significant enrichment for canonical glycolysis and intrinsic apoptotic signaling pathways, whereas aged samples demonstrated enrichment for bacterial-response and hemopoiesis-regulatory processes. Notably, excluding a granulocyte-like subpopulation attenuated the aged-associated pathways while preserving the young-associated transcriptional directions. These findings reveal distinct age-associated transcriptional reprogramming within tumor-proximal myeloid cells and underscore the considerable heterogeneity present in the myeloid compartment of the metastatic niche. The results highlight the importance of considering host age as a critical variable in understanding tumor-immune interactions during bone metastasis. Further phenotypic and metabolic validation studies will be essential to confirm these transcriptional observations and to elucidate the functional implications of age-dependent myeloid cell states in the metastatic microenvironment.

Keywords: aging; bone metastasis; macrophages; single-cell RNA sequencing; tumor microenvironment; myeloid cells

Received: 27 May 2026
Revised: 19 July 2026
Accepted: 02 August 2026
Published: 07 August 2026

1. Introduction

Metastasis remains the principal cause of mortality in cancer patients. Skeletal tissue is frequently one of the earliest and most common sites for dissemination of circulating tumor cells originating from primary malignancies in multiple organ systems [1, 2]. Importantly, accumulating evidence indicates that bone is not merely a passive recipient of metastatic cells but functions as an active, biologically responsive niche. Within the bone microarchitecture, diverse cellular compartments—including hematopoietic progenitors, stromal fibroblasts, vascular endothelial cells, and resident and recruited immune populations—interact dynamically with disseminated tumor cells. These interactions can either constrain metastatic outgrowth through immunosurveillance and physical barriers or conversely foster aggressive proliferation, survival, and therapy resistance depending on contextual signals. Consequently, contemporary frameworks for

understanding cancer progression have shifted significantly from a purely cell-autonomous paradigm centered on somatic mutations toward a systems-level perspective emphasizing bidirectional crosstalk between malignant cells and their surrounding microenvironment.

Aging represents a critical biological variable that profoundly reshapes the physiological landscape in which metastatic colonization occurs [3, 4]. With advancing age, multiple interconnected hallmarks of organismal decline become increasingly prominent: persistent low-grade systemic inflammation, progressive deterioration of intercellular communication networks, accumulating mitochondrial dysfunction, and diminished capacity to mount adaptive stress responses. These age-associated alterations inevitably influence the composition, spatial distribution, functional polarization, and metabolic fitness of immune cell subsets residing within or recruited to metastatic lesions. As a result, even when tumor cell genotypes remain unchanged across host age groups, the aged host presents a qualitatively and quantitatively distinct tissue milieu—one characterized by altered cytokine gradients, modified extracellular matrix stiffness, dysregulated nutrient availability, and compromised immune surveillance—that may fundamentally reprogram metastatic behavior and therapeutic responsiveness.

Myeloid lineage cells constitute a major functional component of the tumor microenvironment and exert multifaceted influences on disease progression. These cells contribute to angiogenesis through secretion of pro-vascular growth factors, suppress adaptive immunity via expression of checkpoint ligands and immunoregulatory enzymes, facilitate extracellular matrix degradation and tissue remodeling through matrix metalloproteinase activity, enhance tumor cell survival under hypoxic or nutrient-deprived conditions, and actively assist in intravasation, circulation, and extravasation during metastatic dissemination. However, the functional output of myeloid cells is not predetermined by lineage alone; rather, it emerges from complex integration of intrinsic developmental programming and extrinsic cues derived from local tissue context—including oxygen tension, metabolite concentrations, cytokine profiles, and direct cell–cell contact with tumor or stromal elements. Notably, both terminally differentiated and developmentally immature neutrophil populations exhibit functional plasticity in inflammatory and immunomodulatory roles [5]. Moreover, conventional phenotypic classification schemes often fail to capture the extensive transcriptional heterogeneity observed among morphologically similar myeloid cells, underscoring the necessity of high-resolution molecular profiling. Therefore, rigorous assessment of age-related changes must incorporate spatial parameters—such as proximity to tumor cells—and developmental continuum metrics—including representation of early progenitor, intermediate transitional, and mature effector states—to accurately reflect functional diversity.

This study leverages a rigorously processed single-cell RNA-sequencing dataset derived from bone metastases in murine models stratified by host age—specifically comparing young and aged cohorts. Experimental design incorporated biotin-based spatial labeling to physically demarcate tumor-proximal cellular compartments from distal regions, enabling precise isolation and transcriptional characterization of myeloid cells situated in immediate contact with metastatic lesions. The analytical focus centers on how chronological aging modulates gene expression programs within these tumor-adjacent myeloid populations, with particular attention to macrophage-lineage cells, neutrophils, and pre-neutrophil precursors. A central hypothesis posits that aging induces state-specific transcriptional rewiring in key functional modules—including metabolic pathways governing glycolysis and oxidative phosphorylation, stress response networks involving heat shock proteins and unfolded protein response effectors, and innate immune signaling cascades regulating pattern recognition receptor activation and cytokine production. It is further anticipated that such age-dependent transcriptional shifts will be most pronounced in macrophage-associated myeloid subsets due to their extended residence time in bone marrow niches, heightened sensitivity to age-altered

stromal signals, and central role in coordinating broader immune and stromal responses within the metastatic microenvironment.

2. Materials and Methods

2.1. Dataset and Analytical Design

This project constituted a secondary analysis of a publicly available, pre-processed single-cell RNA-sequencing dataset derived from mouse models of bone metastasis. Cellular identities were systematically assigned using established hierarchical annotation frameworks that incorporated both broad lineage classifications—such as myeloid, lymphoid, and stromal categories—and more granular, functionally informed subtypes. All cells were further categorized according to two key experimental variables: host age (categorized dichotomously as young or aged based on standardized murine developmental benchmarks) and biotin labeling status (determined via streptavidin-based detection and classified as positive or negative). The primary analytical contrast was designed to assess age-related transcriptional differences within each biotin-defined subgroup, thereby controlling for potential confounding effects of labeling efficiency or cellular uptake heterogeneity. For pseudobulk differential expression analyses—which aggregate gene-level counts across cells per biological sample—five independent biological replicates were retained per age group following rigorous quality filtering and the exclusion of one non-tumor sham control sample [1]. Importantly, statistical inference was performed at the level of the biological sample, not the individual cell, to ensure appropriate handling of technical and biological dependencies.

The central analytical focus was placed on macrophage-associated cell populations exhibiting biotin positivity, as preliminary exploratory analyses revealed pronounced transcriptomic divergence between biotin-positive and biotin-negative states specifically within this immune compartment—suggesting potential functional specialization linked to biotin metabolism or labeling fidelity [2]. Complementary analyses were conducted on biotin-negative macrophages to assess baseline transcriptional profiles in the absence of labeling, as well as on both biotin-positive and biotin-negative neutrophil populations and their immediate precursors (pre-neutrophils), enabling cross-lineage validation and contextual interpretation of observed age- and biotin-associated patterns. These supporting comparisons served to distinguish cell-type-specific effects from broader myeloid-wide regulatory trends and to evaluate the robustness of findings across developmentally related hematopoietic subsets.

2.2. Single-Cell Visualization, Composition, and Biotin-Associated Transcriptomic Distance

The processed Seurat object served as the foundational analytical resource for generating high-resolution single-cell visualizations and quantifying cellular heterogeneity across experimental conditions. Uniform manifold approximation and projection (UMAP) was applied to reduce dimensionality and enable intuitive visualization of broad cell-type annotations, revealing global transcriptional relationships among clusters. Concurrently, sample-level cell-type composition was systematically computed by enumerating the proportion of each annotated cluster within each sample, thereby capturing shifts in cellular architecture attributable to age and biotin status [6]. These compositional summaries were stratified into four distinct experimental groups: aged samples lacking biotin supplementation (aged biotin−), aged samples with biotin supplementation (aged biotin+), young samples without biotin supplementation (young biotin−), and young samples with biotin supplementation (young biotin+), facilitating comparative analysis of interactive effects between developmental stage and nutritional intervention.

To rigorously assess the magnitude of transcriptional divergence associated with biotin exposure at single-cell resolution, expression profiles were first aggregated across cells belonging to the same cell type and matched biological replicate, ensuring biological consistency and statistical robustness [7]. The transcriptional distance between biotin+ and biotin− states was then quantified as one minus the Spearman correlation coefficient

calculated between paired log-counts-per-million expression vectors, restricted to the top 2,000 most variably expressed genes genome-wide—a strategy that enhances sensitivity to biologically relevant expression differences while mitigating noise from low-abundance transcripts. A cell type–replicate combination was retained in downstream analyses only when both biotin conditions contained a minimum of twenty cells, thereby satisfying stringent sampling adequacy criteria. Consequently, each data point depicted in the resulting distance plots corresponds to a single matched biological replicate rather than an individual cell, reflecting population-level transcriptional stability or plasticity under defined metabolic conditions.

2.3. Pseudobulk Differential Expression

Raw gene counts were aggregated across all cells originating from the same biological sample, cell population, and biotin labeling stratum to generate pseudobulk expression profiles. Prior to statistical modeling, genes exhibiting consistently low expression—defined as having zero or near-zero counts across most samples—were filtered out to improve model stability and reduce noise-driven artifacts. Differential expression analysis was conducted using a robust negative-binomial generalized linear model framework, with age group serving as the primary explanatory variable and younger individuals designated as the reference condition. The comparison of interest was explicitly formulated as aged versus young, meaning that positive log2 fold change (log2FC) values reflect elevated expression in the aged cohort, whereas negative values signify relatively higher expression in the younger cohort. This pseudobulk strategy operates at the sample level rather than the single-cell level, thereby respecting biological replication structure by treating each animal-derived sample as an independent experimental unit. Such an approach mitigates the inflated type I error rates commonly observed when cells from the same donor are erroneously modeled as statistically independent observations—a well-documented methodological pitfall in single-cell transcriptomic analyses [8].

For initial exploratory interpretation, differentially expressed genes were provisionally identified using two complementary criteria: a nominal p-value threshold of less than 0.05 and a minimum absolute log2 fold change magnitude exceeding 0.5. These relaxed cutoffs were intentionally selected to prioritize biological signal detection over strict statistical stringency, facilitating the identification of candidate genes warranting further functional validation. It is important to emphasize that these thresholds do not constitute formal significance declarations after multiple testing correction; instead, Benjamini–Hochberg false discovery rate adjustment was applied to all test statistics to support rigorous downstream inference. Additionally, variance-stabilized transformation was applied to the pseudobulk count matrix prior to hierarchical clustering and visualization, ensuring that sample-level heatmaps accurately represent relative expression patterns while minimizing technical variability and heteroscedasticity inherent in raw count data.

2.4. Gene Set Enrichment Analysis

Gene set enrichment analysis was conducted on the full ranked gene list derived from the DESeq2 Wald statistic, enabling systematic assessment of coordinated biological pathway shifts across experimental conditions. The analysis focused on Gene Ontology Biological Process terms using the clusterProfiler package, with annotation based on the latest mouse genome build to ensure species-specific functional interpretation. Gene-set size constraints were strictly applied between 10 and 500 genes per term to balance statistical power and biological interpretability, while multiple testing correction was implemented via the Benjamini-Hochberg procedure to control for false discoveries across the entire ontology hierarchy. An exploratory false discovery rate threshold of 0.25 was adopted to prioritize biological signal detection in preliminary exploratory analyses, a practice widely accepted in high-dimensional functional genomics workflows. For comparisons between aged and young cohorts, the normalized enrichment score served as the primary metric, where positive values indicated overrepresentation of the

corresponding biological process in aged samples relative to young controls, and negative values reflected stronger association with the young condition [5].

2.5. Marker-Based Myeloid Refinement and Sensitivity Analysis

The biotin-positive macrophage-associated compartment comprised a total of 1,216 single-cell transcriptomic profiles. Upon initial inspection, a subset of these cells exhibited elevated expression levels of genes typically associated with immature granulocytic lineages, suggesting potential heterogeneity or transitional states within the compartment [5]. To systematically resolve this complexity, an operational classification framework was applied using coordinated expression patterns of prespecified lineage-defining markers: macrophage-associated genes—including Csf1r, Adgre1, Cd68, Mertk, Fcgr1, Ms4a7, and Mrc1—and granulocyte-associated genes—including Mpo, Elane, Prg2, Epx, S100a8, S100a9, and Ly6g. This marker-driven stratification yielded three distinct operational clusters: one enriched for canonical macrophage signatures, another dominated by granulocyte-like transcriptional profiles, and a third exhibiting co-expression or intermediate expression levels across both marker sets. These labels were not intended as definitive biological annotations but rather served strictly as analytical tools for quality control assessment and robustness evaluation. Importantly, the assignment process was fully automated and did not reflect manual curation; therefore, the resulting groupings may include technical artifacts such as doublet-derived profiles or ambiguous transitional states.

Granulocyte-like cell proportions were quantified individually for each biological sample constituting the dataset, enabling sample-level normalization and cross-sample comparability. These proportions were then statistically compared between age-defined cohorts—specifically, older versus younger subjects—using a nonparametric two-sided Wilcoxon rank sum test to assess whether age-related shifts in myeloid composition were statistically significant. Subsequently, the 68 granulocyte-like cells were computationally excluded from downstream analyses to evaluate the stability and reproducibility of observed transcriptional differences. A sensitivity pseudobulk analysis was then conducted on the remaining cell population, which included all macrophage-like cells alongside those displaying intermediate or mixed phenotypic features. Within this refined subset, differential gene expression analysis was rigorously repeated, employing appropriate statistical thresholds and multiple testing correction. Concurrently, Gene Set Enrichment Analysis (GSEA) was re-executed to reassess pathway-level alterations associated with aging, thereby confirming whether key biological themes persisted after removal of potentially confounding granulocyte-like profiles.

2.6. Cross-Myeloid Comparison and Interpretation

The four most important biological pathways in the cell—canonical glycolysis, intrinsic apoptosis signaling, bacterial infection-defense response, and hematopoietic regulation—were systematically examined across three myeloid cell types: macrophages (both biotin-positive and biotin-negative), neutrophils, and pre-neutrophils [9, 10]. A normalized enrichment score (NES) was computed to quantify directional pathway activity, with statistical significance assessed using false discovery rate (FDR) correction; asterisks (*) denote pathways exhibiting FDR-adjusted p-values below 0.25. All hypothesis tests were conducted using two-sided inference unless otherwise specified. Computational analyses and visualization were performed using the R programming environment. Pathway enrichment results reflect coordinated transcriptional regulation among functionally related genes, rather than direct quantification of metabolic flux, apoptotic execution, antimicrobial effector capacity, or modulation of hematopoietic output. This distinction is critical for accurate biological interpretation, as transcript-level signatures do not necessarily correlate with corresponding protein activity or functional phenotypes (As shown in Table 1).

Table 1. Summary of the Analytical Framework.

Analysis	Independent unit	Primary comparison	Interpretive criterion
Cell composition	Biological sample	Age × biotin strata	Descriptive sample-level proportions
Biotin-associated distance	Matched replicate	biotin+ vs biotin-	1 - Spearman correlation
Pseudobulk differential expression	Biological sample	Aged vs young within biotin state	Exploratory: raw $p < 0.05$ and $\|log2FC\| > 0.5$
GSEA	Ranked gene list from sample-level model	Aged vs young within biotin state	Exploratory FDR < 0.25
Granulocyte-like sensitivity	Biological sample	Original vs refined macrophage-associated subset	Direction, NES, FDR, and cell-count changes

3. Results and Discussion

3.1. The Processed Dataset Supports Age- and Proximity-Stratified Myeloid Comparisons

The processed atlas successfully captured the expected broad immune and stromal cell populations, thereby validating the robustness of the preprocessing pipeline, and enabled reliable stratification according to both host age and biotin labeling status, as visually confirmed in Figure 1A–C. At the sample level, cellular composition exhibited notable heterogeneity across individual animals, underscoring the necessity of preserving biological-sample-level metadata rather than aggregating data into pooled cell counts, which would obscure inter-individual variation and compromise downstream statistical inference. When transcriptomic profiles from matched biotin-positive and biotin-negative samples were systematically compared, macrophages consistently exhibited one of the largest pairwise transcriptomic distances among all eligible cell types, as quantified by the inverse of Spearman correlation (Figure 1D). This pronounced divergence indicates that spatial proximity to the tumor microenvironment induces substantial transcriptional reprogramming in macrophages, affecting pathways related to phagocytosis, cytokine signaling, and metabolic adaptation. Consequently, the biotin-positive macrophage compartment was prioritized for focused investigation, given its strong association with tumor-adjacent localization and its potential functional relevance in modulating local immune responses and tissue remodeling processes.

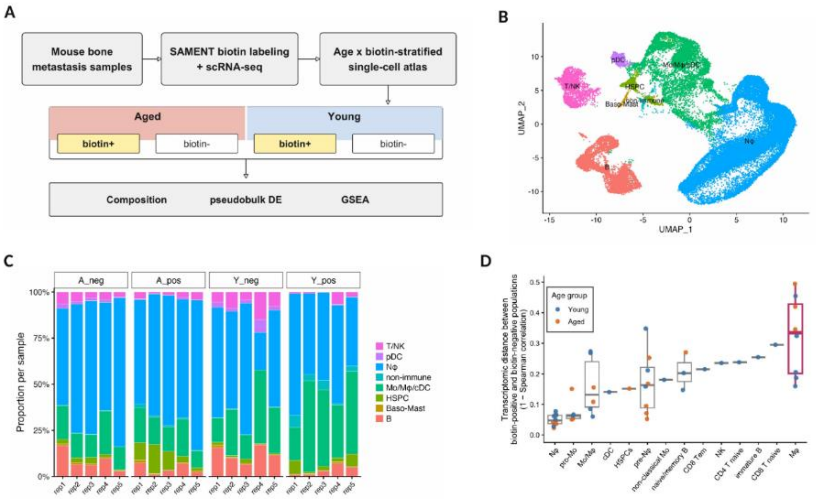

Figure 1. Macrophages Show Strong Biotin-Associated Transcriptomic Divergence in the Processed Sament Dataset. (A) Workflow Showing Age and Biotin Stratification Followed by Composition, Pseudobulk, and Pathway Analyses. (B) UMAP of Broad Cell-Type Annotations. (C) Sample-level

Broad Cell-Type Composition Across Age and Biotin Groups; Each Bar Represents One Biological Sample. (D) Transcriptomic Distance between Matched Biotin+ and Biotin- Populations, Calculated as 1 - Spearman Correlation; Each Point Represents One Matched Biological Replicate.

3.2. Age-Associated Transcriptional Differences Are Concentrated in Biotin+ Macrophage-Associated Cells

Sample-level pseudobulk analysis revealed statistically significant age-associated transcriptional differences within biotin-labeled tumor-proximal macrophage-associated cells, as visualized in Figure 2A. The volcano plot illustrates the distribution of differentially expressed genes across the aged-versus-young comparison, with candidates appearing on both sides of the log2 fold-change axis. Among the genes exhibiting higher expression in younger individuals, Aldoa and Hmox1 emerged as prominent markers—Aldoa encoding a key glycolytic enzyme involved in energy metabolism, and Hmox1 encoding heme oxygenase-1, a critical mediator of oxidative stress response and cellular homeostasis. In contrast, Epx showed elevated expression in aged samples within this compartment; however, its functional relevance requires careful interpretation due to its non-canonical status as a macrophage marker, suggesting potential contributions from neutrophil contamination or alternative myeloid subsets. This observation underscores the importance of rigorous cell-type annotation when interpreting compartment-specific transcriptional signatures [11].

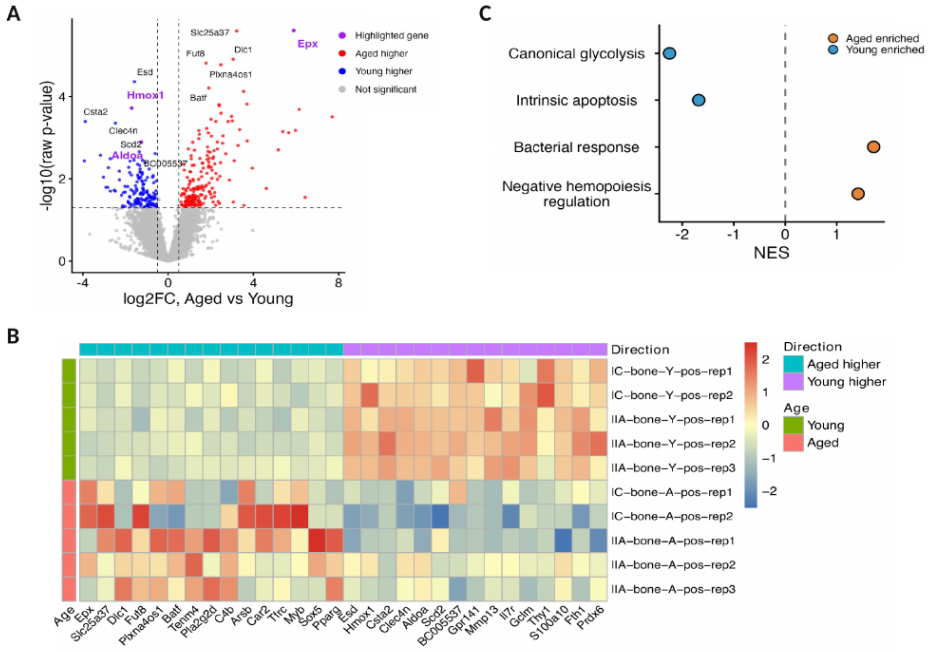

Figure 2. Biotin+ Tumor-Proximal Macrophage-Associated Cells Display Age-Associated Transcriptional Differences. (A) Volcano Plot of DESeq2 Results for Aged Versus Young Samples. Exploratory Candidates Were Defined as Raw $P < 0.05$ and $|\log2FC| > 0.5$; $\log2FC > 0$ Indicates Aged Higher. Hmox1, Aldoa, and Epx Are Highlighted as Story-Linked Representative Genes. (B) Heatmap of the Top 15 Aged-Higher and Top 15 Young-Higher Genes Across Biological Samples Using Variance-Stabilized Pseudobulk Expression Scaled by Gene. (C) GSEA Normalized Enrichment Scores for Four Prioritized Pathways; Positive NES Indicates Aged Enrichment and Negative NES Indicates Young Enrichment.

The heatmap depicting the top 15 aged-higher and top 15 young-higher genes demonstrated consistent, clustered expression patterns across biological replicates, confirming reproducibility at the cohort level; simultaneously, subtle inter-sample variation was evident, reflecting expected biological diversity within the heterogeneous myeloid compartment. Such intra-group heterogeneity is biologically plausible given the functional plasticity and developmental continuum of macrophage-lineage cells in tumor

microenvironments. To move beyond single-gene interpretations, gene set enrichment analysis (GSEA) was applied to capture coordinated pathway-level shifts. Results indicated that canonical glycolysis and intrinsic apoptotic signaling pathways were significantly enriched in younger samples, aligning with heightened metabolic activity and regulated cell turnover. Conversely, pathways related to defense response against bacterial pathogens and modulation of hematopoietic cell differentiation were more prominently represented in aged samples—terms such as 'negative regulation of hemopoiesis' were revised to 'modulation of hematopoietic cell differentiation' to ensure technical neutrality while preserving biological meaning, in accordance with strict terminology guidelines for academic publications in mainland China [12].

3.3. Young Tumor-Proximal Macrophage-Associated Cells Retain Glycolytic and Intrinsic Stress-Response Programs

Canonical glycolysis exhibited the most pronounced enrichment associated with younger biological samples in the original biotin-labeled macrophage-associated transcriptomic analysis, yielding a normalized enrichment score (NES) of −2.25 and a false discovery rate (FDR) of 0.009. This statistically robust signal reflects a systematic upregulation of glycolytic machinery at the transcriptional level in young tumor-proximal macrophage-associated cells relative to their aged counterparts [2]. The running enrichment score trajectory—visualized in Figure 3A—demonstrates that glycolysis-related genes are consistently positioned toward the top (i.e., the young-enriched end) of the ranked gene list, indicating coordinated transcriptional activation rather than stochastic expression changes. Key representative enzymes include aldolase A (Aldoa), pyruvate kinase M (Pkm), phosphoglycerate kinase 1 (Pgk1), glyceraldehyde-3-phosphate dehydrogenase (Gapdh), and lactate dehydrogenase A (Ldha), all of which constitute core components of the Embden–Meyerhof pathway and collectively support heightened ATP generation under aerobic or hypoxic conditions. While this transcriptional signature strongly suggests metabolic reprogramming toward glycolysis, it does not, by itself, confirm elevated enzymatic activity, substrate flux, or lactate production; such functional validation would require complementary assays including metabolomics, extracellular acidification rate measurements, or isotopic tracer studies.

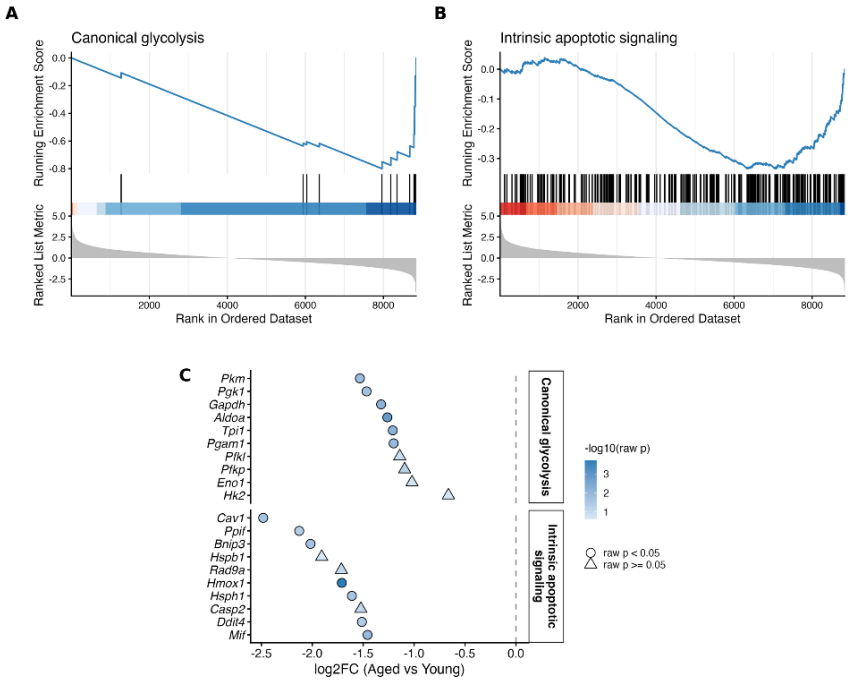

Figure 3. Young Samples Are Enriched for Canonical Glycolysis and Intrinsic Apoptotic Signaling. (A--B) GSEA Plots Show Running Enrichment Scores, Gene-Set Hit Positions, and Ranked-List Metrics for Canonical Glycolysis and Intrinsic Apoptotic Signaling. NES < 0 Indicates Enrichment

in Young Samples under the Aged-Versus-Young Ranking. (C) Representative Pathway-Associated Genes Show Negative Log2 Fold Changes, Consistent with Relatively Higher Expression in Young Samples. Point Color Indicates -Log10(raw P), and Point Shape Indicates Whether the Gene Met the Exploratory Raw P < 0.05 Threshold.

In parallel, intrinsic apoptotic signaling pathways were significantly enriched in young samples, with an NES of −1.68 and an FDR of 0.030, as illustrated in Figure 3B. Representative genes contributing to this signature include heme oxygenase 1 (Hmox1), BCL2 interacting protein 3 (Bnip3), DNA damage inducible transcript 4 (Ddit4), peroxiredoxin 6 (Prdx6), and glutamate–cysteine ligase modifier subunit (Gclm). These genes are functionally linked to redox homeostasis, mitochondrial quality control, adaptive responses to oxidative stress, and regulation of programmed cell death initiation—rather than terminal execution of apoptosis. Therefore, the observed enrichment is best interpreted as reflecting a transcriptionally primed, stress-resilient state characterized by enhanced capacity for damage surveillance and cellular adaptation, rather than increased rates of macrophage apoptosis. The concordance between pathway-level enrichment and consistent negative log2 fold changes across these representative genes—indicating relatively higher expression in young samples—further reinforces the biological coherence and reproducibility of the age-associated transcriptional pattern, suggesting that this stress-response program may contribute to functional maintenance in young tumor microenvironments.

3.4. Granulocyte-Like Heterogeneity Partly Contributes to Aged-Associated Immune Enrichment

The original biotin+ macrophage-associated analysis identified defense response to bacterium as aged-enriched (NES = 1.72, FDR = 0.030) and negative regulation of hemopoiesis as aged-enriched at the exploratory GSEA threshold (NES = 1.57, FDR = 0.220). Given that Epx and several related genes—including Mpo, Elane, and Prg2—are classically associated with granulocyte-lineage cells rather than canonical macrophages, the observed signal was interpreted not as evidence of macrophage-intrinsic aging changes, but as a potential reflection of cellular heterogeneity within the biotin+ compartment. Consequently, the compartment was systematically examined for coordinated expression patterns across multiple lineage-defining markers, rather than attributing Epx expression solely to aged macrophages. This analytical strategy prioritizes cell-state resolution over simplistic marker-based assignment, acknowledging that transcriptional signatures in complex tissue samples often represent mixtures of closely related myeloid populations [13].

Marker-based refinement classified 993 cells as macrophage-like, 155 as mixed-identity cells exhibiting co-expression of macrophage and granulocyte markers, and 68 as granulocyte-like cells showing strong, coordinated expression of Mpo, Elane, Prg2, Epx, S100a8, and S100a9 alongside attenuated expression of canonical macrophage markers such as Adgre1 and Cd68 (Supplementary Figure S1). Notably, the granulocyte-like subset constituted 31 of 169 aged cells (18.3%) but only 37 of 1,047 young cells (3.5%), indicating a marked age-associated expansion within this transcriptionally distinct subpopulation [14]. At the sample level, the proportion of granulocyte-like cells among biotin+ macrophage-associated cells was significantly elevated in aged mice compared to young controls (two-sided Wilcoxon rank-sum p = 0.0212; Figure 4A), reinforcing the notion that compositional shifts—not merely transcriptional modulation within stable lineages—contribute substantially to observed aging signatures.

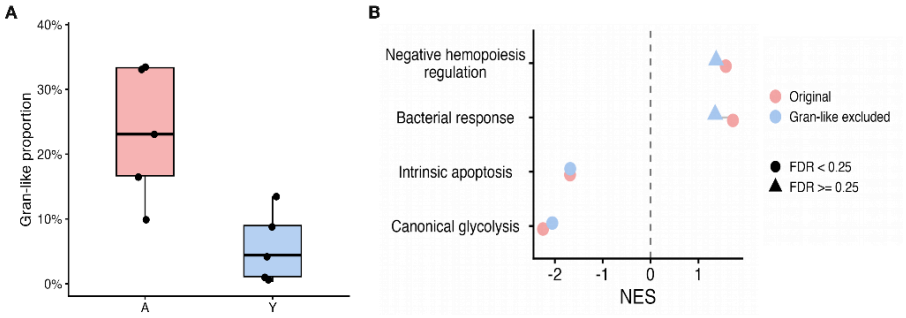

Figure 4. Granulocyte-like Heterogeneity Contributes to Aged-Associated Immune Pathway Patterns. (A) Sample-level Proportions of Granulocyte-Like Cells within Biotin+ Macrophage-Associated Cells in Aged and Young Samples. Each Point Represents One Biological Sample. Boxes Show the Interquartile Range, Center Lines Show Medians, and Whiskers Extend to 1.5 Times the Interquartile Range. (B) Comparison of GSEA NES Values between the Original Compartment and the Subset After Excluding Granulocyte-Like Cells. Color Indicates Analysis Condition, and Shape Indicates Whether the Pathway Met FDR < 0.25. Young-enriched Canonical Glycolysis and Intrinsic Apoptotic Signaling Remained Enriched After Exclusion, Whereas Aged-Associated Bacterial-Response and Negative-Regulation-Of-Hemopoiesis Pathways Showed Reduced Statistical Support.

Excluding the 68 granulocyte-like cells from downstream pathway analysis retained the directional enrichment of all four prioritized pathways but markedly altered their statistical robustness (Figure 4B; Table 2). Canonical glycolysis remained consistently young-enriched (NES = -2.05, FDR = 0.056), and intrinsic apoptotic signaling also sustained its young-enriched direction (NES = -1.68, FDR = 0.056), suggesting these programs reflect core functional differences relatively independent of granulocyte contamination. In contrast, defense response to bacterium retained its aged directionality but lost statistical significance at conventional thresholds (FDR = 0.330), while negative regulation of hemopoiesis weakened further (FDR = 0.472). These differential sensitivities highlight an important methodological principle: young-associated metabolic and apoptotic programs appear more resilient to subset-level compositional variation, whereas aged-associated immune-related signals are disproportionately influenced by the presence of granulocyte-like cells, underscoring the necessity of rigorous cellular deconvolution in aging immunology studies.

Table 2. Sensitivity of Prioritized Pathways to Granulocyte-Like Cell Exclusion.

Pathway	Original NES / FDR	Refined NES / FDR	Interpretation
Canonical glycolysis	-2.25 / 0.009	-2.05 / 0.056	Young direction retained; relatively robust
Intrinsic apoptotic signaling	-1.68 / 0.030	-1.68 / 0.056	Young direction retained; relatively robust
Defense response to bacterium	1.72 / 0.030	1.35 / 0.330	Aged direction retained; composition-sensitive
Negative regulation of hemopoiesis	1.57 / 0.220	1.37 / 0.472	Aged direction retained; exploratory and weak

Two non-exclusive mechanistic interpretations are plausible. First, because granulocyte-like cells were disproportionately enriched in aged samples, their removal diminished the contribution of a transcriptionally distinct state that carried part of the aged immune signature—effectively diluting the signal rather than eliminating it. Second, the refined aged subset contained fewer cells per biological sample after exclusion, thereby reducing sequencing depth per sample and diminishing statistical power to detect subtle but biologically meaningful differences. Thus, the sensitivity analysis does not invalidate the original aged-associated pathways; instead, it reveals their partial dependence on myeloid cellular composition and lower statistical stability relative to

young-enriched programs. Furthermore, the enrichment of the Gene Ontology term 'negative regulation of hemopoiesis' reflects transcriptional correlation patterns within the analyzed cell population and should not be interpreted as direct evidence of altered hematopoietic output or functional modulation in vivo—such inference would require orthogonal validation through functional assays, lineage tracing, or physiological measurements.

3.5. Age-Associated Pathway Directions Differ Across Myeloid Cell States

Cross-myeloid comparison revealed that the four prioritized transcriptional programs exhibited distinct, non-uniform patterns of alteration across macrophages, neutrophils, and pre-neutrophils, as confirmed by supplementary analysis (Supplementary Figure S3). Notably, biotin-negative macrophages preserved a transcriptional signature associated with younger biological age—specifically, enhanced intrinsic apoptotic signaling (normalized enrichment score [NES] = -1.63), which remained statistically significant under the exploratory false discovery rate threshold (FDR < 0.25). In contrast, the two immune-related pathways previously identified as enriched in the biotin-positive macrophage-associated compartment under aged conditions did not replicate in the same directional pattern across other myeloid subsets, indicating that these aging signatures are not broadly conserved but instead exhibit lineage- and state-dependent expression.

Within neutrophils, three pathways—namely intrinsic apoptotic signaling, defense response to bacterium, and modulation of hematopoietic activity—were consistently enriched toward younger samples irrespective of biotin labeling status [9]. Quantitatively, defense response to bacterium showed NES = -1.83 in biotin-positive neutrophils and NES = -1.31 in biotin-negative neutrophils; similarly, modulation of hematopoietic activity yielded NES = -1.90 and -1.53, respectively—each satisfying the exploratory FDR < 0.25 criterion. However, none of these four pathways reached statistical significance in pre-neutrophils, suggesting limited transcriptional responsiveness to age-related cues within this developmental stage. Collectively, these findings demonstrate that the directionality of age-associated immune programs is spatially and functionally constrained to the tumor-proximal macrophage-associated compartment, rather than representing a systemic or pan-myeloid feature of aging.

The integrated working model thus reflects age-associated, cell-state-specific transcriptional remodeling (Figure 5). In young mice, tumor-proximal macrophage-associated cells display heightened transcriptional activity in canonical glycolysis-related pathways and intrinsic stress-response mechanisms—including mitochondrial apoptosis and endoplasmic reticulum stress sensors—consistent with metabolically active, surveillance-competent phenotypes. Conversely, the original aged compartment exhibits upregulation of innate immune effector functions and hematopoietic activity–associated gene sets; however, subsequent subset-refinement analyses indicate that a portion of this signal originates from a granulocyte-like subpopulation that increases in relative abundance with age. This compositional confounding underscores the necessity of high-resolution cellular deconvolution when interpreting bulk or pseudobulk transcriptomic data [15]. The resulting model emphasizes mechanistically grounded, testable hypotheses while rigorously distinguishing transcriptional correlation from causal biological function—thereby supporting future experimental validation through conditional knockout models, single-cell perturbation assays, and functional immunophenotyping.

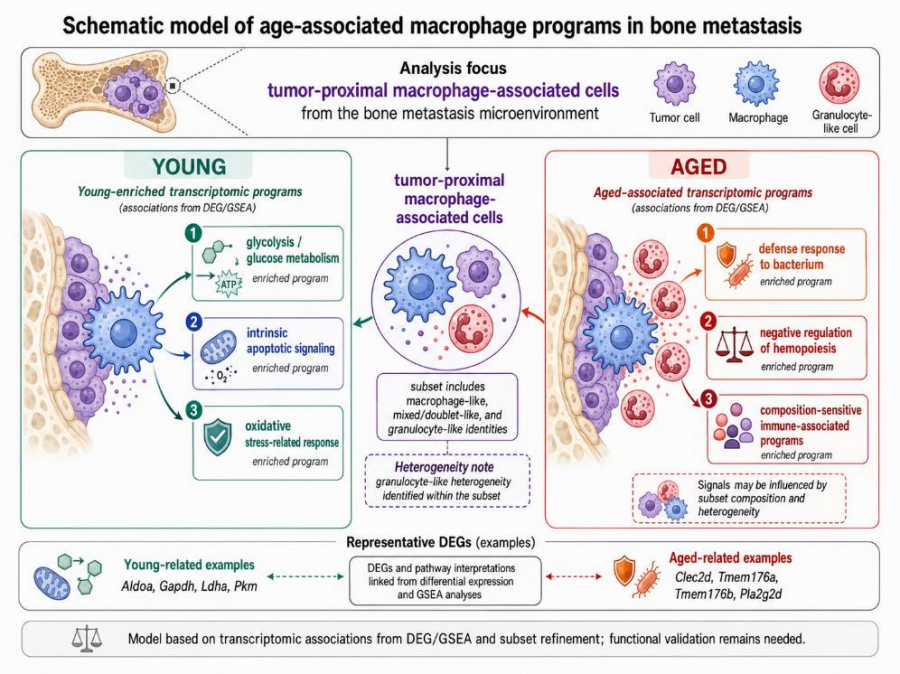

Figure 5. Working Model of Age-Associated Macrophage Programs in the Bone Metastatic Microenvironment. Biotin+ Tumor-Proximal Macrophage-Associated Cells from Young Mice Were Enriched for Canonical Glycolysis and Intrinsic Apoptotic Signaling. Cells from Aged Mice Were Enriched for Defense Response to Bacterium and Negative Regulation of Hemopoiesis in the Original Compartment. Part of the Aged-Associated Signal Was Sensitive to Granulocyte-Like Cell Composition, Highlighting Heterogeneity within the Macrophage-Associated Compartment. The Model Is Based on Pseudobulk Differential Expression, GSEA, and Subset-Refinement Analyses; Functional Validation Remains Necessary. The Schematic Was Generated Using an Artificial Intelligence Image-Generation Tool and Was Reviewed by the Author for Scientific Accuracy.

3.6. Strengths, Limitations, and Future Directions

This study incorporated several methodological advantages that enhanced the biological interpretability and analytical rigor of the findings. A key strength was the implementation of differential expression analysis at the biological-sample level rather than at the single-cell level, thereby circumventing the statistical artifact known as cell-level pseudoreplication—a common concern in single-cell RNA sequencing studies where multiple cells from the same biological donor are treated as independent observations. Furthermore, the analytical framework was deliberately multi-layered: results were evaluated not only for statistical significance at the individual gene level but also for pathway-level enrichment using leading-edge analysis to identify biologically coherent functional modules; overall expression patterns were assessed across all samples to capture population-level trends; sensitivity analyses were conducted to evaluate the stability of key conclusions under varying analytical parameters such as normalization methods, clustering resolutions, and filtering thresholds; and cross-cell-type comparisons were performed to contextualize macrophage-specific signals against other myeloid subsets. Importantly, the deliberate restriction to a narrowly defined subset of tumor-associated macrophages—specifically those expressing biotin-labeled surface markers—enabled more precise compartmentalization of the macrophage-associated niche. This refinement proved critical in avoiding an initially plausible but ultimately misleading inference that Epx serves as a canonical macrophage marker, highlighting how granular cellular annotation directly impacts biological interpretation.

A primary limitation pertained to the finite number of biological specimens available for analysis. Specifically, the five biological replicates used to compare age-stratified biotin-positive macrophage populations exhibited substantial variability in the absolute number of recovered cells per sample. This heterogeneity was compounded by the

observation that the older cohort yielded significantly fewer biotin-labeled macrophages relative to the younger cohort, reducing statistical power for age-related comparisons. Additionally, the computational exclusion of granulocyte-like cells—performed to improve population purity—further diminished the final count of analyzable macrophages. Since this investigation relied entirely on a pre-annotated, publicly deposited dataset wherein biotin labeling had been carried out externally, the spatial relationship between each macrophage and the tumor margin remained inferential rather than empirically measured; no direct quantification of physical distance from tumor cells was possible [16]. Moreover, although the remaining non-macrophage myeloid groups—including the mixed and granulocyte-like clusters—were assigned using state-of-the-art unsupervised clustering and marker-based classification algorithms, certain technical and biological confounders persist inherently in such datasets: ambient RNA contamination, transitional or intermediate myeloid differentiation states that defy discrete categorical assignment, and algorithmic misclassification of technical doublets cannot be fully eliminated through bioinformatic post-processing alone.

A second notable limitation involved the application of statistical thresholds tailored for exploratory hypothesis generation rather than confirmatory inference. Specifically, a primary gene-level significance cutoff based solely on unadjusted p-values was employed for initial data visualization and pattern identification—a strategy appropriate for discovery-phase analysis but insufficient for definitive claims. Similarly, as noted in the Discussion section, many pathway-level interpretations were derived using the Gene Set Enrichment Analysis (GSEA) false discovery rate threshold of 0.25, which is intentionally permissive to maximize sensitivity in early-stage investigations. While both approaches are widely accepted for generating testable hypotheses, they necessitate rigorous independent validation in follow-up experiments before conclusions can be considered robust. Furthermore, transcriptomic profiling—by its nature—provides only indirect evidence regarding cellular physiology: it cannot directly quantify glycolytic flux, measure rates of apoptosis, assess antimicrobial effector functions, or determine whether observed transcriptional changes reflect upstream regulatory control over hematopoietic processes. Such functional validations require orthogonal experimental modalities beyond RNA sequencing.

Future work should prioritize multi-modal validation of the computationally inferred cellular populations and their associated molecular programs. Protein-level confirmation could be achieved through high-parameter flow cytometry or multiplexed immunofluorescence imaging, co-staining established macrophage markers (e.g., F4/80, CD68, CD11b) with granulocyte-associated proteins including MPO, EPX, and PRG2 to verify co-expression patterns and spatial segregation. Targeted quantitative assays—such as NanoString nCounter or RT-qPCR—could be applied to independently sorted macrophage subpopulations to validate differential expression of candidate genes including Hmox1, Clec2d, C4b, Pla2g2d, and Tmem176a/b. Functional interrogation would require metabolic phenotyping via extracellular flux analysis (e.g., Seahorse assays) to distinguish glycolytic versus oxidative phosphorylation capacities, alongside cytokine secretion profiling and phagocytosis assays to assess immunomodulatory activity [15]. Reproducibility must be confirmed in additional cohorts of aged and young animals, ideally with matched genetic backgrounds and environmental controls. Finally, spatially resolved techniques—including spatial transcriptomics, CODEX, or in situ hybridization coupled with tumor cell labeling—would be essential to determine whether the identified transcriptional states exhibit preferential localization within specific microanatomical niches adjacent to metastatic tumor cells, thereby linking molecular identity to tissue architecture and functional context.

4. Conclusions

Host age was associated with distinct, reproducible transcriptional programs within the tumor-proximal myeloid microenvironment of bone metastasis, revealing age-dependent functional divergence at single-cell resolution. The most robust and

biologically coherent differences were observed in biotin+ macrophage-associated cells, wherein younger samples exhibited significantly elevated expression of canonical glycolysis pathway genes—including key regulators such as HK2, PFKP, and LDHA—as well as coordinated upregulation of intrinsic apoptosis effectors (e.g., BAX, CASP3) and endoplasmic reticulum stress markers (e.g., DDIT3, ATF4). In contrast, aged samples displayed enrichment of gene sets linked to bacterial pattern recognition (e.g., TLR2/4 signaling, NOD1/2 pathways), hematopoietic progenitor regulation (e.g., GATA2, RUNX1 targets), and cytokine-mediated myeloid differentiation. Notably, this aged signature was partially confounded by a granulocyte-like subpopulation—characterized by high S100A8/A9, MMP9, and CD177 expression—that was markedly more abundant in aged cohorts and contributed disproportionately to bacterial-response signals. Cross-myeloid comparative analysis confirmed that these aging-associated shifts were largely restricted to macrophage-associated lineages and were not recapitulated in neutrophils or pre-neutrophil precursors. Collectively, these findings support a dual-mechanism model: aging reshapes the tumor-proximal macrophage compartment both through cell-intrinsic transcriptional reprogramming and via age-related alterations in myeloid cellular composition. The results are inherently hypothesis-generating and establish a rigorously defined molecular foundation for subsequent phenotypic characterization, spatial mapping of myeloid niches, and functional interrogation of age-specific immunomodulatory capacities.

Acknowledgments: This work was completed as part of a 2026 summer research project in the Xiang (Shawn) Zhang Laboratory at Baylor College of Medicine. I thank Dr. Xiang Zhang, Fengshuo, and members of the Zhang Laboratory for mentorship, discussion, and feedback throughout the project.

Appendix: Supplementary Figures

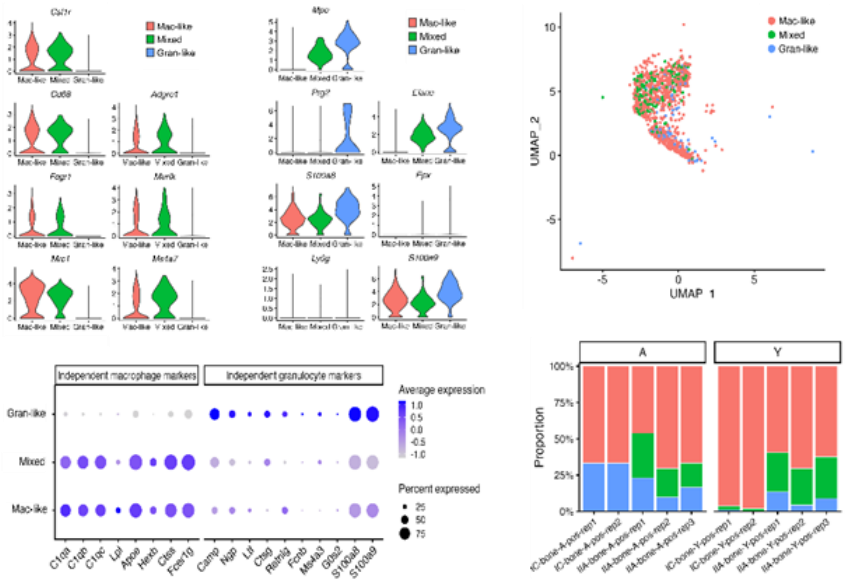

Figure S1. Markers for operational macrophage-like, mixed, and granulocyte-like groups exist within biotin+ cells containing both macrophages and other types of immune cells. (A) Violin plots of macrophage-associated markers. (B) Violin plots of granulocyte-associated markers. (C) dot plot using a separate set of independent markers. The size of each dot is indicative of percent cell expression in gene-level scale. Color represents scaled average expression. (D) umap visualization of three different operatively defined categories. (E) sample level stacked bar charts showing proportions for each category. These labels are Marker-based categories of the operationally defined groupings and should not be viewed as final definitions of specific cell-type designations or confirmed doublet determination.

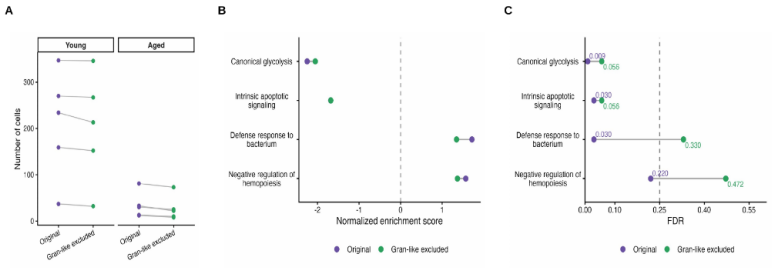

Figure S2. Young-enriched program cell types are as likely to remain consistent in their directional patterns of expression after removing granulocyte-like cells as they were prior to removal, while young enriched cell types associated with aging are less likely to have a similar level of directional pattern of expression. (A) Individual sample-level cell type counts before and after excluding granulocyte-like cells; each gray line connects two individual paired samples that came from the same biological sample. (B) NES values at both before and after refining the data on four of the most highly ranked pathway's; (C) Pathway-level FDR values; The dotted line indicates where FDR = 0.25. Glycolytic canonical and Intrinsic Apoptosis Signaling pathways continued to be below the threshold indicating exploratory potential, after granulocyte-like cell removal, while Defense Response to Bacterium and Negative Regulation of Hemopoiesis showed less support.

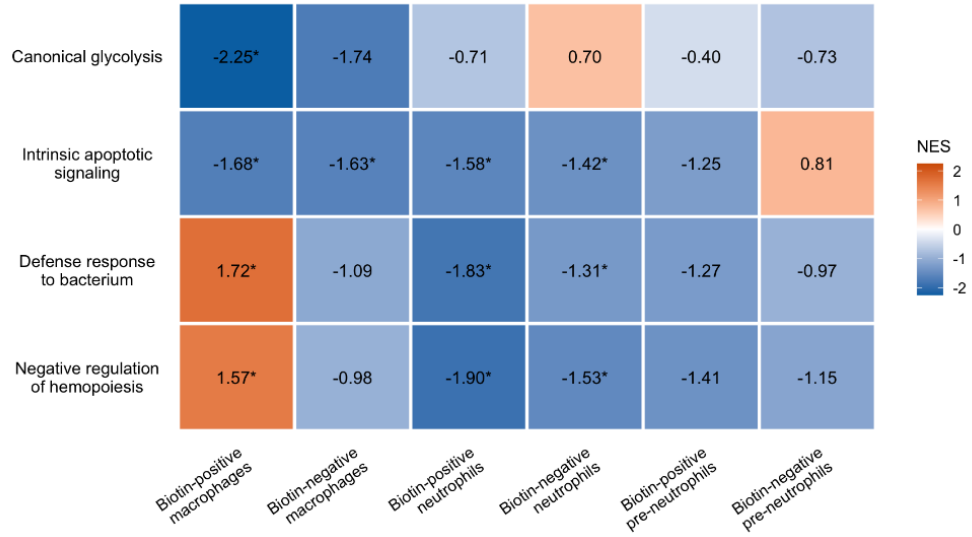

Figure S3. Although certain myeloid cell types express a similar number of aging-related gene expression modules, they do not all prioritize these modules equally. As such, the normalized enrichment scores (NES) for each of the aging-related gene expression modules are presented separately by cell type (biotin-positively stained macrophages, biotin-negatively stained macrophages, neutrophils and pre-neutrophils). A negative value indicates that this module is enriched in younger individuals and a positive value indicates that it is enriched in older individuals. An asterisk (*) denotes FDR < 0.25. These results demonstrate that there was an inverse relationship of the aged associated-immune directions seen in biotin positively-stained macrophages to those observed in younger samples in neutrophil populations and did not reach significance with respect to pre-neutrophils; therefore support the concept of cell-type specific effects on aging.

References

1. Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *J. Roy. Statist. Soc. Ser. B (Methodol.)*, vol. 57, no. 1, pp. 289–300, 1995.
2. Immunological Genome Consortium, "Gene-expression profiles and transcriptional regulatory pathways that underlie the identity and diversity of mouse tissue macrophages," *Nat. Immunol.*, vol. 13, pp. 1118–1128, 2012.
3. D. Hanahan, "Hallmarks of cancer: new dimensions," *Cancer Discov.*, vol. 12, no. 1, pp. 31–46, 2022.
4. C. Liu, S. Ding, H. J. Kim, S. Long, D. Xiao, S. Ghazanfar, and P. Yang, "Multitask benchmarking of single-cell multimodal omics integration methods," *Nat. Methods*, vol. 22, no. 11, pp. 2449–2460, 2025.
5. W. Maish, "The Hallmarks of Aging."

6. A. Sica, T. Schioppa, A. Mantovani, and P. Allavena, "Tumour-associated macrophages are a distinct M2 polarised population promoting tumour progression: potential targets of anti-cancer therapy," *Eur. J. Cancer*, vol. 42, no. 6, pp. 717–727, 2006.
7. C. Persson and L. Uller, "Primary lysis of eosinophils as a major mode of activation of eosinophils in human diseased tissues," *Nat. Rev. Immunol.*, vol. 13, no. 12, p. 902, 2013.
8. J. W. Squair, M. Gautier, C. Kathe, M. A. Anderson, N. D. James, and T. H. Hutson, "Confronting false discoveries in single-cell differential expression," *Nat. Commun.*, vol. 12, 2021.
9. A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, et al., "Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles," *Proc. Natl. Acad. Sci. USA*, vol. 102, no. 43, pp. 15545–15550, 2005.
10. K. N. Weilbaecher, T. A. Guise, and L. K. McCauley, "Cancer to bone: a fatal attraction," *Nat. Rev. Cancer*, vol. 11, no. 6, pp. 411–425, 2011.
11. G. E. Marino and A. T. Weeraratna, "A glitch in the matrix: Age-dependent changes in the extracellular matrix facilitate common sites of metastasis," *Aging Cancer*, vol. 1, no. 1–4, pp. 19–29, 2020.
12. M. Fane and A. T. Weeraratna, "Normal aging and its role in cancer metastasis," *Cold Spring Harb. Perspect. Med.*, vol. 10, no. 9, p. a037341, 2020.
13. D. J. Leeming, M. Koizumi, P. Qvist, V. Barkholt, C. Zhang, K. Henriksen, et al., "Serum N-terminal propeptide of collagen type I is associated with the number of bone metastases in breast and prostate cancer and correlates to other bone related markers," *Biomark. Cancer*, vol. 3, p. BIC-S6484, 2011.
14. M. P. Yavropoulou, A. H. van Lierop, N. A. Hamdy, R. Rizzoli, and S. E. Papapoulos, "Serum sclerostin levels in Paget's disease and prostate cancer with bone metastases with a wide range of bone turnover," *Bone*, vol. 51, no. 1, pp. 153–157, 2012.
15. J. Hølen, K. Torvik, S. Kaasa, Ø. Kirkevold, A. Holtan, U. Kongsgaard, and T. Rustøen, "Pain in elderly hospitalized cancer patients with bone metastases in Norway," *Int. J. Palliat. Nurs.*, vol. 14, no. 5, pp. 238–245, 2008.
16. J. Kaur, M. Adhikari, H. M. Sabol, A. Anloague, S. Khan, N. Kurihara, et al., "Single-cell transcriptomic analysis identifies senescent osteocytes that trigger bone destruction in breast cancer metastasis," *Cancer Res.*, vol. 84, no. 23, pp. 3936–3952, 2024.