WEP Warwick Evans Publishing

Transactions on Materials, Biotechnology and Life Sciences
ISSN: 2960-1932, eISSN: 2960-2319 | Volume 1
FBB 2023

Single-Cell RNA Sequencing: A Deep Dive into The Cellular Landscape and Interactions in Osteosarcoma

Sihan Zheng

Frankfurt International School (FIS), Frankfurt, 61440, Germany

Abstract. OSTEOSARCOMA is a leading cause of cancer-related deaths worldwide, with its progression highly influenced by the cellular interplay within the tumor microenvironment that is underexplored. Aiming to bridge this gap, our study utilizes single-cell RNA sequencing (scRNA-seq) to examine the cellular heterogeneity of OSTEOSARCOMA and investigate the roles of distinct cell populations. scRNA-seq was performed on eight DEN mice OSTEOSARCOMA samples, followed by bioinformatic analysis with Seurat package. 12 distinct cell populations were identified, with three unique macrophage populations suggestive of their role as tumor-associated macrophages (TAMs). The detected endothelial cells and pericytes hint at ongoing neoangiogenesis, with implications that endothelial cells might function as tumor-associated endothelial cells (TECs) and pericytes as carcinoma-associated fibroblasts (CAFs). Our findings provide insights into the potential roles of various cell populations in the OSTEOSARCOMA tumor microenvironment, which paves the way for developing novel therapies. These postulations, while offering a deeper understanding of OSTEOSARCOMA's cellular landscape, necessitate experimental validation for confirmation.

Keywords: Osteosarcoma, Single cell, microfluid, transcriptome.

1. Introduction

Cancer represents a diverse collection of diseases marked by unregulated cellular proliferation and dissemination stemming from accumulated oncogenic mutations and epigenetic alterations. Globally, cancer is a major health challenge, affecting millions of people. It is estimated that, in 2020, there were around 18.1 million new cancer cases and 10 million cancer-related deaths worldwide (Worldwide Cancer Statistics, 2015). Among all types of cancer, hepatocellular carcinoma (HCC) ranks fifth and remains the second leading contributor to cancer-related mortality (Wen et al., 2022). Even with modern technology, therapeutic response and prognosis of HCC are still poor due to therapeutic resistance, variable drug sensitivity in individual patients, and metastasis, all of which arise from the heterogeneity of cancer.

Tumor heterogeneity can manifest at multiple levels - intertumor, intratumor, and intermetastatic. In previous years, heterogeneity between tumors in different patients (intertumor heterogeneity) has been widely studied, and the variable responses to the same therapeutic regimens among distinct patients are widely attributed to intertumor heterogeneity. However, recent studies have revealed that neoplastic subclones of considerable genotypic and phenotypic variations are found within the same tumor (intratumor heterogeneity), which accounts for treatment failure, metastasis, and cancer evolution (Ramón Y Cajal et al., 2020). Intratumor heterogeneity originates from a number of factors, including genetic instability, epigenetic alterations, selection pressure, stochastic events in gene expression and protein stability, and interactions among tumor cells and the tumor microenvironment (Burrell et al., 2013).The tumor microenvironment, a significant contributing factor to intratumor heterogeneity and tumorigenesis, encompasses primarily the proliferating tumor cells, along with the tumor stroma, vascular networks facilitating nutrient supply to neoplastic cells, infiltrating immune inflammatory cells, microbiota, and a broad spectrum of associated tissue cells (Whiteside, 2008). The wide variety of cells interacts within the tumor microenvironment, generating an incredibly complex signaling network that shapes tumor pathogenesis (Hanahan & Weinberg, 2011).

Therefore, understanding intratumor heterogeneity and the cellular composition of the tumor microenvironment plays a crucial role in formulating novel therapies and realizing the ambition of precision medicine. The latest advances in single-cell transcriptome analysis by RNA sequencing offer great platforms to explore heterogeneity within a tumor and the subpopulations of cells in a tumor microenvironment. Unlike its traditional counterpart, bulk RNA sequencing, which averages the gene expression among the group of cells under study, single-cell RNA sequencing (scRNA-seq) provides a high-resolution view of distinct gene expression profiles of individual cells, which enables the revelation of diverse cell populations within a tumor, elucidating intratumoral heterogeneity. For solid tumors, tissue decomposition scRNA-seq can be performed to identify various cell types in the intratumoral immune microenvironment and genes with significant transcriptional upregulation or downregulation, which enhances insights into oncogenic cellular interactions and signaling pathways that contribute to drug resistance and metastasis, fueling the discoveries of cancer-associated biomarkers and druggable subclone targets, mapping of the complex intratumoral microenvironment, and the comprehension of the courses of neoplastic progression. scRNA-seq is also applicable to studying circulating tumor cells (CTCs), which facilitates the understanding of the underlying mechanisms of anoikis resistance, cluster-induced metastasis, epithelial-to-mesenchymal transition (EMT) transformation, and stemness (Zhu et al., 2017).

As the intricate cellular interplay within the HCC tumor microenvironment remains a challenge to decipher, this study aims to bridge this knowledge gap by uncovering distinct cell populations and shedding light on their potential roles and interactions in the tumor landscape, paving the way for the development of novel therapies.

2. Methods

2.1 Single-Cell RNA Sequencing

2.1.1 Gel Bead Synthesis

The gel beads in this study were generated with droplet microfluidic devices (Pan et al., 2022). Each bead was functionalized with oligonucleotides that comprise (i) a 16 nucleotide 10x Barcode for tagging the origin cell of the RNA molecules, (ii) a 12 nucleotide Unique Molecular Identifier (UMI) for preventing amplification bias, (iii) an Illumina TruSeq Read 1, (iv) and a 30 nucleotide poly-dT primer sequence for capturing the poly-A tail of the mRNA (*Chromium Single Cell 3' Reagent Kits User Guide (v3.1 Chemistry) - Official 10x Genomics Support, n.d.*). The overall reaction of bead synthesis was a radical chain-growth reaction that polymerized acrylamide and its cross linker bis-acrylamide (N, N'-methylene-bis-acrylamide) with the presence of redox initiator APS (Ammonium persulfate) and TEMED (N,N,N′,N′-Tetramethylethylenediamine).

In this reaction, APS, a strong oxidizing agent, decomposed to form sulfate radicals which initiated the polymerization, while TEMED catalyzed the decomposition of APS, generating more sulfate radicals. The acrylamide: bis-acrylamide solution, along with APS, TEMED, and oligonucleotides served as the aqueous phase for droplet generation, which was then introduced into a microfluidic device together with an immiscible oil phase that contained a surfactant stabilizing the droplet. As the two phases flowed through the microfluidic channels, aqueous droplets surrounded and separated by oil were formed, in which the sulfate radicals triggered the free radical-dependent polymerization of vinyl groups between acrylamide monomers and bis-acrylamide monomers. The products were gel beads with DNA primers trapped within and between the acrylamide chains cross-linked by bis-acrylamide (Kandow et al., 2007). The soft and deformable nature of the polyacrylamide gel beads allows them to pack closely in the microfluidic channel, thereby achieving a super-Poissonian distribution with a ~80% single gel bead encapsulation rate (Abate et al., 2009; Cao et al., 2023).

2.1.2 Microfluidics for GEM Generation

Single-cell RNA sequencing was performed using the Chromium Next GEM Single Cell 3' Reagent Kits v3.1 (10x Genomics). In the droplet-based microfluidic platform used for scRNA-seq, two

immiscible fluids were introduced at a junction between microfluidic channels, with the aqueous phase broken off by the oil phase, resulting in droplets in an aqueous dispersed phase contained within the continuous phase of the oil that prevented the droplets from merging and moved them around in the microfluidic channels. These water-in-oil (W/O) emulsions acted as monodispersed reaction chambers of limited volumes, which decreased the cost of reagents and increased the sensitivity and specificity of reactions due to the isolation of reactions and reduction of reaction competition (Salomon et al., 2019; Zhang et al., 2019). As the roles of the dispersed phase and the continuous phase were dictated by relative surface energies between the channel walls and the fluid (need rephrase), with the continuous phase having higher affinity for the surface than the dispersed phase, the microfluidic channels used in this study had hydrophobic surfaces to generate water-in-oil droplets(Chae et al., 2009). Specifically, in the 10x Genomics Chromium system microfluidics, the cells and gel beads were in the aqueous phase, with the partitioning oil serving as the immiscible oil phase that separated the aqueous phase into droplets of consistent dimension. As the cell occupancy rate followed Poisson distribution, the cells were delivered at a limiting dilution to minimize the chances of producing doublets. The size and frequency of droplet formation could be sophisticatedly manipulated by electro-gating or gas pressure. The droplet encapsulating the cell and the barcoded gel bead was the core technology of the 10X Genomics Chromium system, Gel bead in EMulsion (GEM). GEMs were generated in an 8-channel microfluidic chip that achieved a ~50% cell capture rate (Zheng et al., 2017).

2.1.3 cDNA Synthesis and Library Preparation

Once the cell and the gel bead got co-encapsulated, the cell lysed upon lysis mix in droplets, and the gel bead immediately dissolved due to the reduction of SS-bond with DTT, releasing the oligonucleotides with the poly-dT primer sequence that hybridized with the poly-A tail of the mRNA, capturing the mRNA in the cell lysate (Salomon et al., 2019). Subsequently, a Master Mix containing Moloney murine leukemia virus (MMLV) reverse transcriptase, template switching oligo (TSO) sequences, and other reverse transcription reagents were mixed with the GEMs. The MMLV reverse transcriptase reversely transcribed the mRNA by synthesizing the cDNA in a 5’ to 3’ direction, introducing C nucleotide overhangs only when the reverse transcriptase reached the end of the template mRNA without a premature termination (Matuła et al., 2020). The TSO sequence contained a poly(dG) tail that hybridized with the C nucleotide overhangs, which allowed the reverse transcriptase to switch the template from the mRNA to the cDNA, thereby ensuring the production of full-length cDNA sequences with common 5’ sequences for universal primers in following PCR amplification step (Zheng et al., 2017). The first cDNA strands with TSO were synthesized within the droplets. After demulsification, the first-strand cDNA was purified with DynaBeads MyOne Silane Beads (Thermo Fisher Scientific) from the post-GEM-RT reaction mixture. The cleaned-up cDNA underwent amplification, enzymatic fragmentation, end repair, A-tailing, SPRIselect, adaptor ligation, and sample index PCR according to the manufacturer's instructions (*Chromium Single Cell 3’ Reagent Kits User Guide (v3.1 Chemistry Dual Index) - Official 10x Genomics Support*, n.d.).

2.1.4 RNA Sequencing and Bioinformatics Analysis

Sequencing was performed on an Illumina Nextseq 2000 according to the sequencing protocol provided by 10x Genomics. The resulting data were processed using the Cell Ranger Single Cell Software Suite (10x Genomics) to perform sample demultiplexing, barcode processing, and single-cell 3' gene counting. Further downstream analysis was performed using the Seurat R package. With the help of a histogram showing the number of nFeature_RNA and its frequency, quality control parameters were set to between 200 and 7000 nFeature_RNA and less than 5% mitochondrial gene (Fig. S1). Principle component analysis (PCA) was performed after normalization and scaling of the top 2000 highly variable genes. Based on the principle component (PC) elbow plot (Fig. S2), the first 13 PCs were considered for the subsequent analysis. Two-dimensional data visualization was achieved by performing a nonlinear dimensionality reduction with a uniform manifold approximation and projection (UMAP) algorithm and a resolution of 0.2. A total of 9 clusters were generated, with the spatial relationship between single-cell clusters denoting the similarities/differences between their

transcriptomic landscape. Up-regulated marker genes were identified by the FindAllMarkers function, with a log fold change threshold of 0.25, minimum percentage of 0.2, and a logistic regression test only finding the positive markers. A heatmap showcasing the top 10 biomarker genes across each cluster was plotted. Cluster identities were labeled manually by using reference datasets such as Mouse Cell Atlas (MCA), PangloaDB, National Center for Biotechnology Information, and other published research (*Choose Gene Expression Markers | PanglaoDB*, n.d.; *Genes & Expression - Site Guide - NCBI*, n.d.; *MCA | Mouse Cell Atlas*, n.d.). Specifically, the top ten genes from each cluster were identified and served as key markers. The primary identification work is carried out with MCA, with the marker genes input into the MCA adult liver gallery and searched for the corresponding cell types. The marker genes not presented in MCA were researched otherwise using PangloaDB, etc. Certain markers that were poorly recognized by the FindAllMarkers function were manually identified and validated through dot plots and feature plots. Cell type-specific marker genes were identified for each of the nine cell types based on literature review and validation by dot plots and dim plots (Fig. S3-20). All feature plots in this study were plotted using a minimum cutoff value of 0.2, unless otherwise stated.

2.1.5 Results

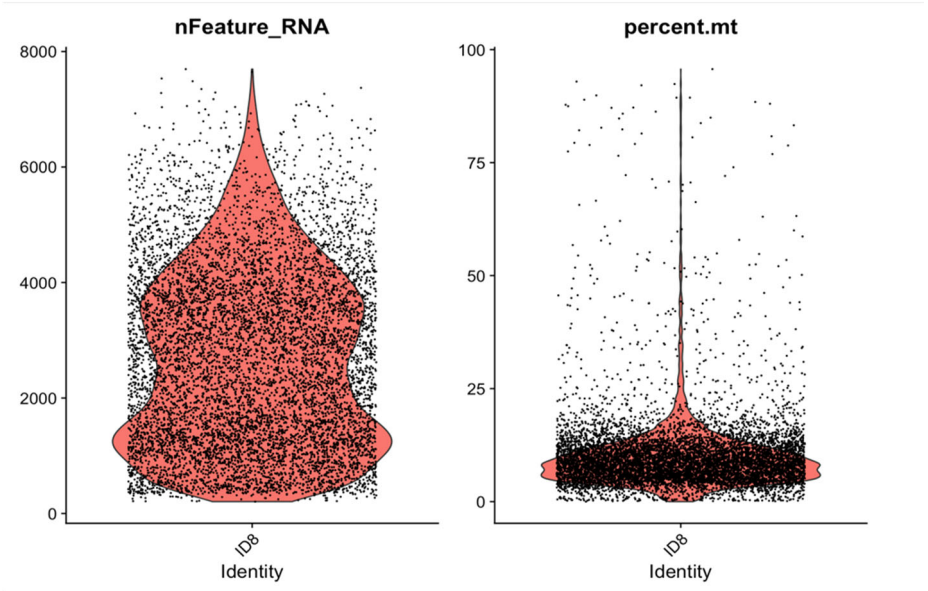

Figure 1. Violin plot of number of genes detected per cell and the percentage of transcripts derived from the mitochondria

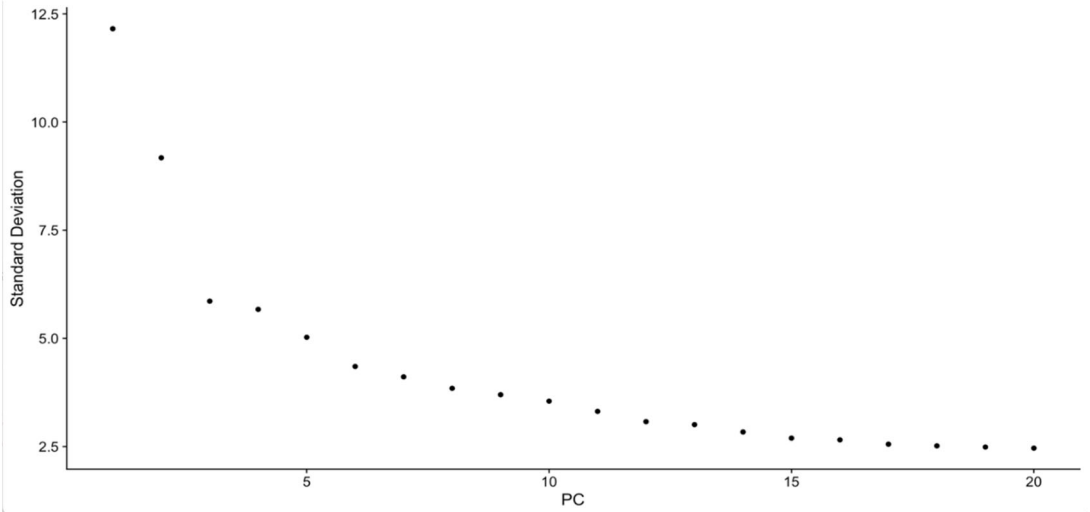

Figure 2. Standard deviation

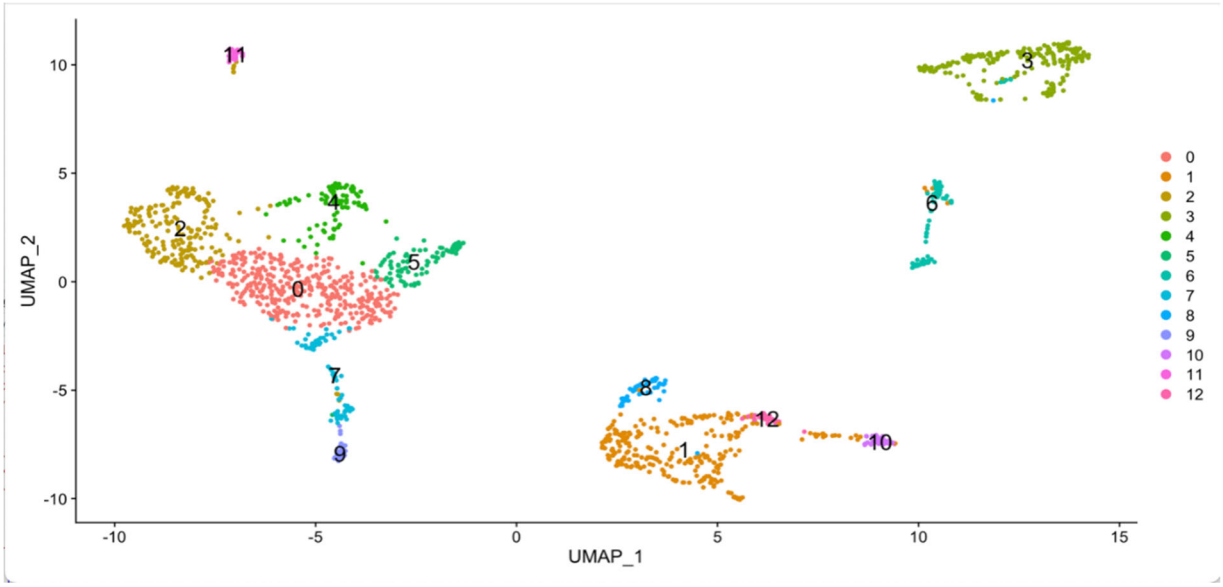

Figure 3. single cell cluster based on the similarity of gene expressed

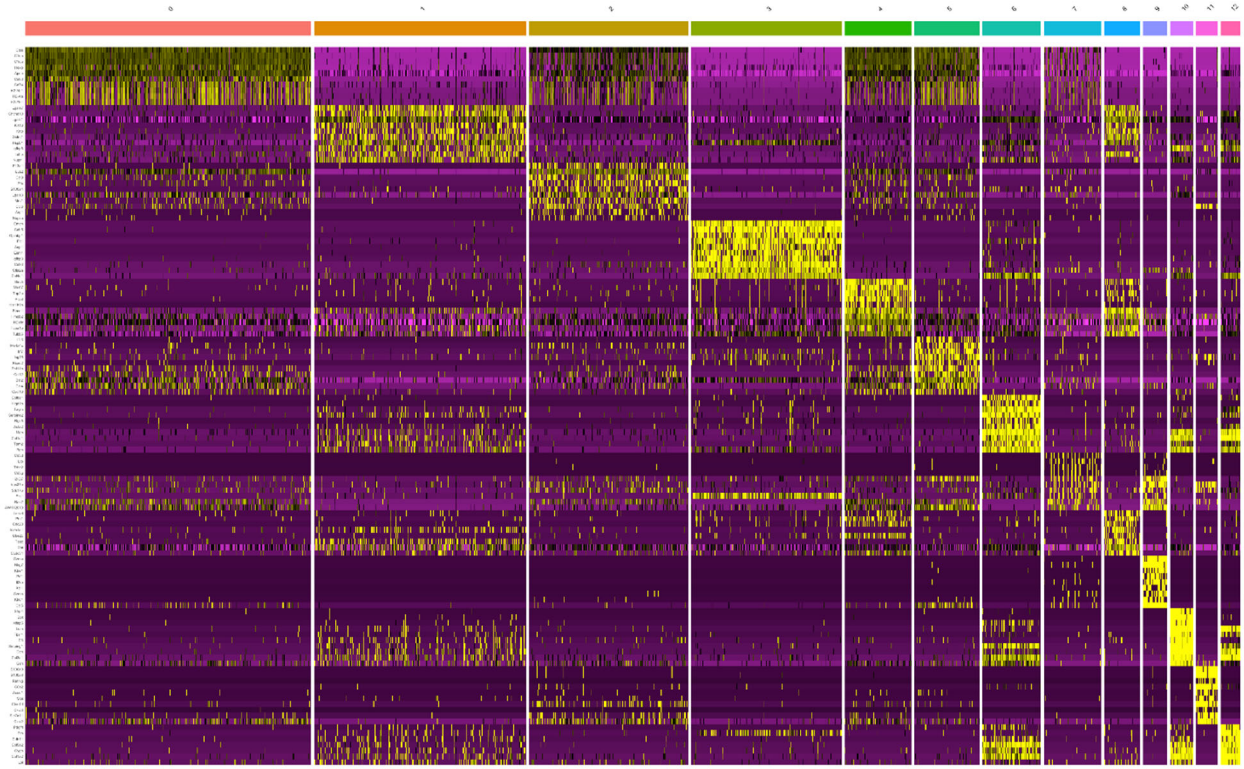

Figure 4. Heatmap of unusual gene expression

In this study, single-cell RNA sequencing (scRNA-seq) was performed on mouse Osteosarcoma tumor tissue to investigate the cellular composition and potential intercellular interactions within the tumor microenvironment. Nine distinct clusters of cells (Fig. 3) were identified, and the top 10 highly expressed genes of each clusters were determined using the FindAllMarker function with a logistic regression test and plotted onto a heatmap (Fig. 4).

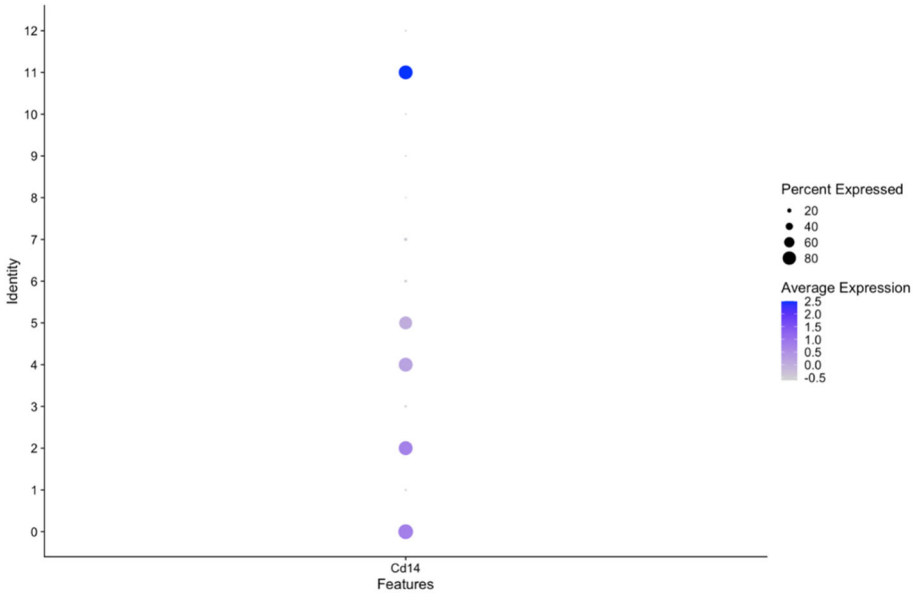

Figure 5. percentage and rate of expression in cloud 14

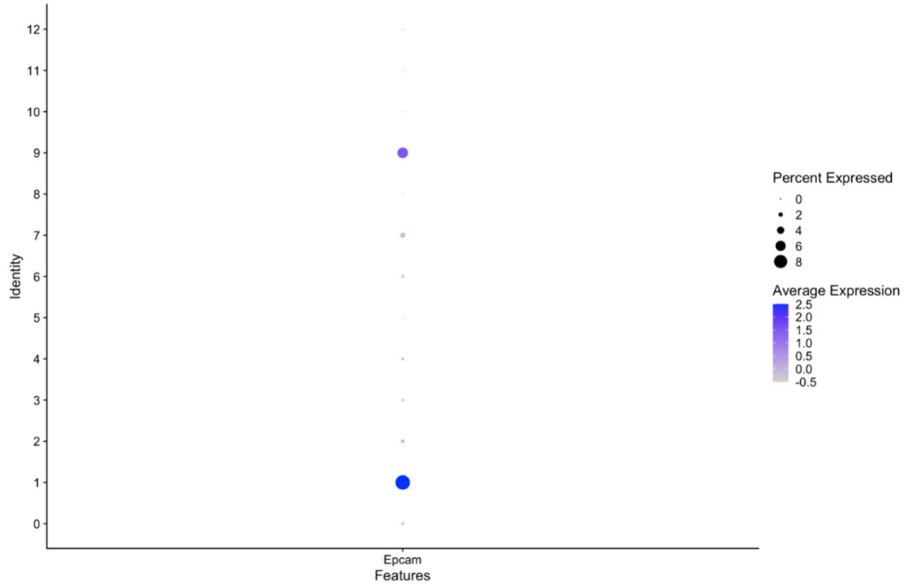

Figure 6. percentage and rate of expression in Epcam

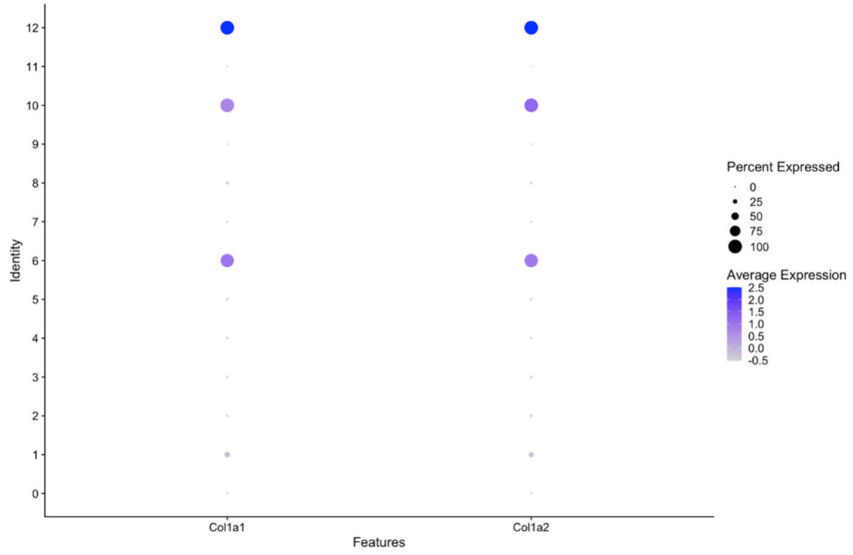

Figure 7. comparison of

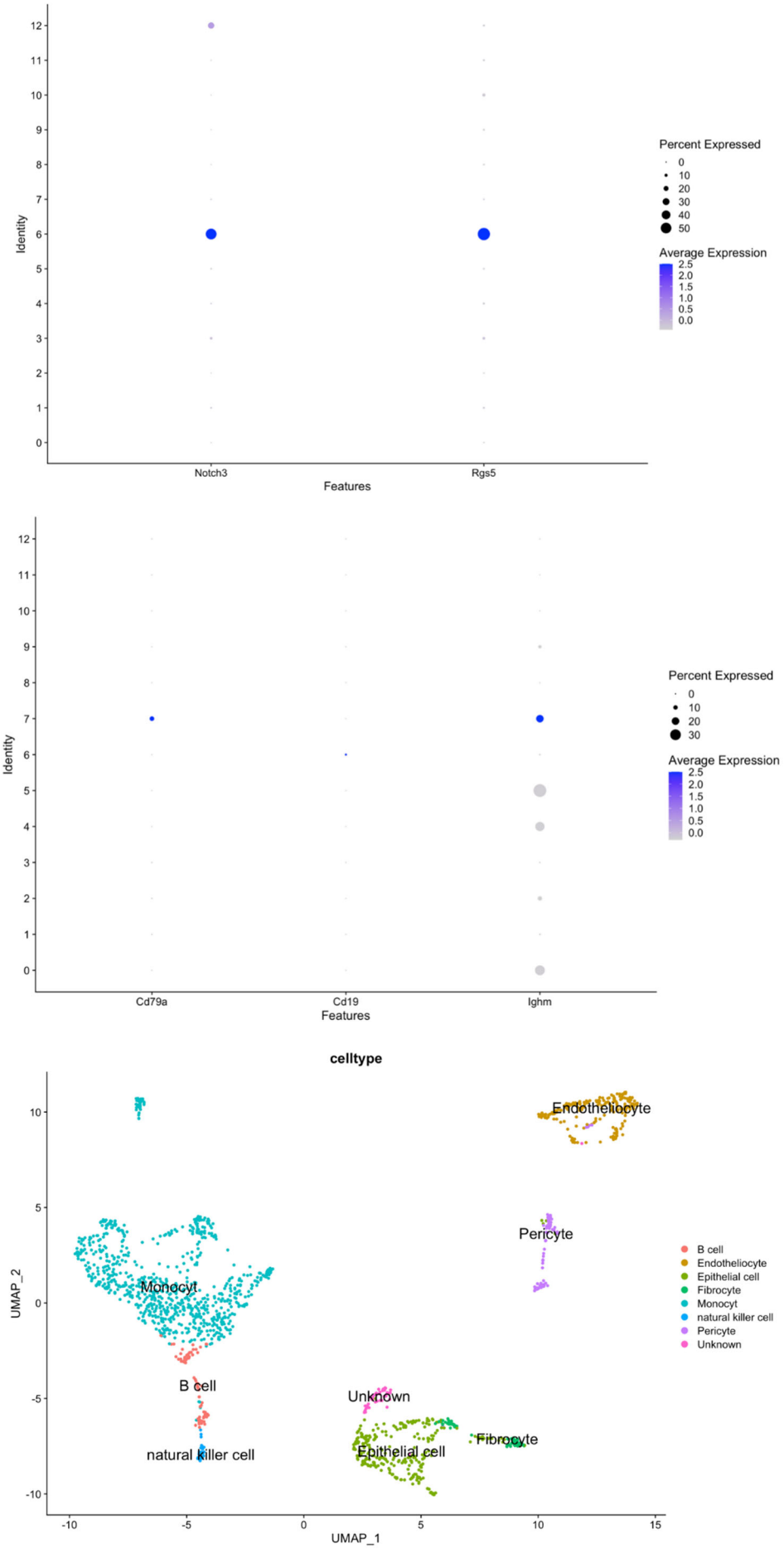

Figure 8. Diverse cell types in the tumor immune microevironment.

3. Conclusions

The dimensionality reduction plot shows the immune cell types identified in our Osteosarcoma sample, including Monocyte cells, Epithelial cell, B cell, endothelial cells, epithelial cells, NK cells, pericytes, fibrocyte. The differences and similarities in their gene expression are shown in the spatial distribution of these clusters, with closer proximity suggesting more similar expression profile, and vice versa. The relative percentages of each cluster are labeled, with the Monocyte cells being the most abundant and the NK cell being the least.

References

[1] Chen, X., & Xu, C. (2015). Transcription Profiles of Marker Genes Predict The Transdifferentiation Relationship between Eight Types of Liver Cell during Rat Liver Regeneration. 17(2), 339–354. https://doi.org/10.22074/cellj.2016.3756.

[2] Hielke van Splunder, Villacampa, P., Martínez-Romero, A., & Mariona Graupera. (2023). Pericytes in the disease spotlight. https://doi.org/10.1016/j.tcb.2023.06.001.

[3] Hosein, A. N., Huang, H., Wang, Z., Parmar, K., Du, W., Huang, J., Maitra, A., Olson, E., Verma, U., & Brekken, R. A. (2019). Cellular heterogeneity during mouse pancreatic ductal adenocarcinoma progression at single-cell resolution. JCI Insight, 4(16). https://doi.org/10.1172/jci.insight.129212.

[4] MCA | Gallery. (n.d.). Bis.zju.edu.cn. Retrieved August 2, 2023, from https://bis.zju.edu.cn/MCA/gallery.html?tissue=Adult-Liver.

[5] Rammal, H., Chaza Harmouch, Maerten, C., Gaucher, C., Fouzia Boulmedais, Schaaf, P., Voegel, J. C., Laurent-Maquin, D., Menu, P., & Halima Kerdjoudj. (2016). Upregulation of endothelial gene markers in Wharton's jelly mesenchymal stem cells cultured on polyelectrolyte multilayers. 105(1), 292–300. https://doi.org/10.1002/jbm.a.35868.