DISSERTATION

USE OF SINGLE-CELL RNA SEQUENCING AND COMPARATIVE IMMUNO-ONCOLOGY TO GAIN INSIGHTS INTO SPONTANEOUS CANINE CANCERS

Submitted by

Dylan T. Ammons

Department of Microbiology, Immunology, and Pathology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2023

Doctoral Committee:

Advisor: Steven Dow

Douglas Thamm
Anne Avery
Randall Basaraba

ABSTRACT

USE OF SINGLE-CELL RNA SEQUENCING AND COMPARATIVE IMMUNO-ONCOLOGY TO GAIN INSIGHTS INTO SPONTANEOUS CANINE CANCERS

Advances in human clinical medicine stem from discoveries and reports in model systems, therefore the use of biologically relevant models in essential for developing effective human therapeutics. Traditionally, small mammals, such as mice and rats, have been used to address basic science questions and they have contributed substantially to our understanding of biology. Despite widespread use and accessibility of rodent models, there is a growing awareness that findings in rodents frequently fail to translate to human medicine. In recent years, pet dogs have been proposed as an ideal model system to facilitate translational research. As such, the overarching themes of this dissertation are to (1) build upon the dog as a model by providing novel cell type transcriptomic references for immuno-oncology research and (2) investigate immunological correlates with treatment responses in clinical trials using dogs with spontaneously arising tumors.

First, the introductory chapter discusses the dog as a model for human disease with a focus on the application in glioma and osteosarcoma (OS). The biological and molecular features of each tumor type are described, then current therapeutic approaches in dogs and human are discussed. After introducing the tumor types, two cell types, myeloid-derived suppressor cells (MDSCs) and tumor-associated macrophages (TAMs), are discussed in detail as they are key cell types throughout the dissertation. In the final section of the introduction, single-cell RNA (scRNA) sequencing, the technology foundational to the work presented here, is discussed in detail.

In chapters 2 through 5 we focus on OS, a malignant tumor of the bone with minimal therapeutic options. In chapter 2 we generated a reference scRNA dataset of canine circulating

leukocytes, then applied the dataset to investigate how the presence of a primary OS tumor impacts systemic immune cell transcriptomes. Through evaluation of 74,067 cells from 17 dogs (7 healthy, 10 OS) we identified relative increases in the abundances of polymorphonuclear (PMN-) and monocytic (M-) MDSCs and provided their transcriptomic signatures for further study. The reference aspect of the work constituents a comprehensive database with gene signatures for each of the 36 cell types identified in canine blood. This work provides key insights into OS induced changes to circulating immune cells while also providing a broadly applicable reference that can be applied to many different areas of canine research.

In chapter 3 we generate another comprehensive database, this time focusing on characterizing the heterogeneity within canine OS tumors. Through analysis of 35,310 cells we identified exhausted T cells, mature regulatory dendritic cells (mregDCs), and 8 transcriptomically distinct macrophage/monocyte populations and provide their transcriptomic signatures. We used cell-cell interaction inference approaches to investigate active immune suppressive pathways in OS and found TAMs and mregDCs to be major contributors to T cell suppression. Lastly, we obtained an external human OS scRNA dataset to evaluate cell type homologies between dogs and human which suggested a high degree of similarities between the species. We hope the data generated in this chapter can be applied to enhance canine OS research and shed light on conserved immune suppressive pathways in OS.

In chapter 4 we apply the datasets generated in chapters 2 and 3 to investigate how the tumor microenvironment (TME) impacts the transcriptional programs of infiltrating immune cells. To complete the analysis, we used data from circulating leukocytes of the 10 OS dogs in chapter 2 and the OS tumor-infiltrating immune cells identified in chapter 3. Through direct comparison of infiltrating and circulating immune cells we were able to confirm several tumor-induced changes reported in humans are also apparent in the dog. Key confirmatory findings in infiltrating immune cells included the upregulation of activation markers on T cells, increased relative abundance in exhausted T cells, and increased expression of immune suppressive molecules on myeloid cells.

Overall, the analysis suggests overarching tumor-induced immunological changes are conserved between human and dogs.

In chapter 5 we apply scRNA sequencing to investigate how a myeloid targeted combination therapeutic (losartan, ladarixin, and toceranib) impacts intratumoral and systemic immune responses. Analysis revealed broad immune cell depletion in the tumor and increases in circulating M-MDSCs in dogs receiving treatment. We identified modulation to multiple chemokine signaling axes which shed light on mechanisms associated with treatment-induced immune cell depletion. Finally, the analysis revealed profound impacts to tumor cells and fibroblasts, with treatment skewing transcriptomic profiles toward a hypoxic phenotype and increased insulin-like growth factor associated gene expression. Ultimately, this study represents the first insights into how any therapeutic modulates the OS tumor microenvironment at the single-cell level.

Finally, in chapter 6 we conducted a canine glioma clinical trial to investigate the utility of another myeloid targeted therapy (vaccination, losartan, and propranolol). We observed treatment to induce partial tumor regression in 2 and stable disease in 6 of 10 dogs, for an overall clinical benefit rate of 80%. Through evaluation of antibody responses to vaccination, we identified a subset of patients to be immunological responders, which we found exhibited enhanced overall survival times relative to dogs that did not generate antibody responses. The findings from the clinical study suggest that myeloid targeted therapy for treatment of glioma may be a valuable approach that warrants further investigation in canine and human glioma patients.

In conclusion, our work applying single-cell RNA sequencing resulted in the generation of valuable canine-specific cell type reference datasets and revealed key insights in osteosarcoma immunobiology. The work evaluating myeloid therapeutics in the setting of osteosarcoma and glioma provide mechanistic and clinical insight that can be applied to further study of the therapeutic approach. Overall, we hope the body of work presented here strengthens the foundation of the dog as a model for translational biomedical research.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my PhD advisor, Dr. Steven Dow. Steve has been instrumental in my training and the work presented herein would not have been possible without his support and contributions. I came into the Dow laboratory with an interest in cancer immunology and the desire to conduct clinically relevant basic research. Dr. Dow identified projects that matched my interests and enable me to develop my skillset as a scientist. I greatly appreciate all the opportunities that were presented during the completion of my dissertation work and have been honored with ability to complete cutting-edge veterinary medical research. My experiences in the Dow laboratory have shaped the way I approach science and will continue to influence me as I progress though my career as a veterinary biomedical researcher.

Thank you to my committee members, Drs. Douglas Thamm, Anne Avery, and Randall Basaraba for their time and contributions to the shaping of my dissertation. Collectively, your insights have enhanced my dissertation work and enabled me to think about the data from a different perspective. I would also like to thank all the members of the Dow laboratory, but especially to Jade Kurihara and Lyndah Chow who were always there to help when I needed it. I also would like to thank Adam Haris for taking the time to learn about single-cell RNA sequencing together, and for always being there to bounce ideas around.

I would like to express my gratitude to the combined DVM-PhD program at Colorado State University. The program created an environment that fostered my development as a scientist and presented opportunities to learn about the many roles of veterinary medical researchers within the greater realm of science. The program directors secured funding for a NIH T32 biomedical research training program, which I was privileged to have been supported by.

Thank you to the clinical trials team at the Flint Animal Cancer Center, especially Lynelle Lopez, Kara Hall, and Allister Aradi who were instrumental in obtaining the clinical samples used throughout the projects.

I would like to express my thanks to the University of Colorado Boulder Research Computing staff for their help and efforts to maintain the Alpine supercomputer. The high-performance computing resources were instrumental in the completion of this work.

Lastly, I would like to thank my family members for their love and support. In particular my parents who provided undying support throughout all of my studies, and I would not be where I am today if it was not for all they have done for me. And to my partner, Sophie, who has been there to support me scientifically, personally, and professionally. I appreciate your willingness to listen to me ramble on about science while also providing perspective to help me develop my thoughts.

DEDICATION

This dissertation is dedicated to Riot, Mir-cat, Row-row, Grant the beagle, and Echo boi.

TABLE OF CONTENTS

ABSTRACT ii
ACKNOWLEDGEMENTS v
DEDICATION vii
CHAPTER 1: Introduction of the dog as a translational model for human disease and the power of single-cell RNA sequencing 1
1.1 The dog as a translational model for human disease 1
1.1.1 An overview of human and canine glioma 2
1.1.2 Canine osteosarcoma: an ideal clinical model for the human disease 4
1.1.3 Cancer therapeutic advancements from canine to human 5
1.2 Cancer immunity: a focus on immune suppressive cellular mediators 8
1.2.1 The myeloid-derived suppressor cell 9
1.2.2 Immune suppressive macrophages and their impact on tumor immunology 12
1.3 Molecular dissection of complex tissues 15
1.3.1 The evolution of single-cell RNA sequencing 15
1.3.2 From basic to advanced analysis approaches 17
1.3.3 New discoveries in the era of single-cell 'omics 28
1.3.4 A summary of single-cell RNA sequencing studies in dogs to date 30
References 36
CHAPTER 2: A single-cell RNA sequencing atlas of circulating leukocytes from healthy and osteosarcoma affected dogs 46
Summary 46
Introduction 47
Methods 48
Results 55
Discussion 75
References 80
CHAPTER 3: Molecular dissection of primary canine osteosarcoma using single-cell RNA sequencing 84
Summary 84
Introduction 84
Methods 86
Results 92

Discussion 116
References 119
CHAPTER 4: Molecular characterization of tumor microenvironment induced changes to infiltrating immune cells 123
Summary 123
Introduction 123
Methods 125
Results 129
Discussion 144
References 146
CHAPTER 5: Single-cell RNA sequencing reveals the impacts of losartan, ladarxin, and toceranib combination therapy on the canine osteosarcoma microenvironment 148
Summary 148
Introduction 148
Methods 150
Results 155
Discussion 165
References 168
CHAPTER 6: Reprogramming the canine glioma microenvironment with tumor vaccination plus oral losartan and propranolol induces objective responses 170
Summary 170
Introduction 171
Methods 173
Results 180
Discussion 189
References 193
CONCLUSIONS 197

CHAPTER 1: Introduction of the dog as a translational model for human disease and the power of single-cell RNA sequencing

1.1 The dog as a translational model for human disease

Therapeutic advances in human medicine stem from discoveries and experimentation in model systems that recapitulate aspects of human disease processes. For decades rodent models have been used to address basic scientific questions with the idea that findings can be translated to human medicine. However, in recent years it has become apparent that findings from traditional animal models frequently fail to translate to the human clinic, with over 90% of therapeutic findings failing to translate (1). The failure to translate findings stems in part from animal models inaccurately recapitulating human biology. For example, a novel immunotherapeutic target remained unidentified for years due to the absence of expression in murine models, but the molecule is proving to have clinical relevance in human medicine (2). Thus, there is a need to enhance the translational potential of research findings and use model systems that accurately model human disease. The dog is a large animal model that develops disease spontaneously and, in many settings, may provide an avenue for highly translatable research. Two settings in which the dog is regarded to have potential are glioma and osteosarcoma (3,4). In this section the biological characteristics of these cancers will be discussed in the context of human and canine cancers. Although this section will focus on glioma and osteosarcoma, it should be noted that dogs are also considered valuable models for melanoma, lymphoma, leukemia, soft tissue sarcoma, head and neck, and bladder carcinoma (5,6).

1.1.1 An overview of human and canine glioma

In humans and dogs, glioma is the most common primary brain tumor, with high grade glioblastoma being among the deadliest brain tumors in humans. Glioma arises from the malignant transformation of stromal cells (astrocytes, oligodendrocytes, and ependymal cells) and the term glioma encompasses three major subtypes: astrocytoma, oligodendroglioma, and ependymoma. In 2016, the World Health Organization (WHO) proposed the use of histological features and a four-point grading scale (I–IV) to classify gliomas (7). In response to the classification scheme, the canine diagnostic parameters were modified to parallel the 2016 WHO recommendations (8). The canine glioma classification scheme included subclassifications of glioma into three categories (astrocytoma, oligodendroglioma, or undefined glioma), then further recommended classification as high- or low-grade based on histological properties. Despite attempts to synchronize human and canine glioma subtypes, recent advances in the molecular characterization of human glioma, and the inability of histological characteristics to provide prognostic insight, led the WHO (2021) to recommend coupling histological data with mutation status in key genes (7). The two key mutations identified through the molecular characterization of human glioma are isocitrate dehydrogenase (IDH1/2) mutational status and the presence or absence of a codeletion of the 1p and 19q chromosomal arms. Both mutational statuses have been determined to have prognostic implications in humans (9).

The recent advances in molecular characterization of glioma have led to three major classifications being proposed by the WHO: (1) IDH-mutant, non-codeleted astrocytoma, (2) IDH-mutant, 1p/19q codeleted oligodendroglioma, and (3) IDH-wildtype glioblastoma. IDH enzymes play key roles in cellular metabolism, with the catalytic sites of IDH1/2/3 converting isocitrate to alpha-ketoglutarate (10). The conversion of isocitrate to alpha-ketoglutarate makes up an essential step in redox regulation, lipogenesis, glutamine metabolism, and the Krebs cycle. Tumor cells with IDH mutations frequently harbor the mutation on one copy of the gene and form catalytic

heterodimers which alters the function of the enzyme to drive the conversion of α-KG into D-2-hydroxyglutarate (D-2-HG). The modified enzyme function results in a reduction of glycolytic capacity and promotes metabolic reprogramming of the tumor cells.

The metabolic changes in IDH-mutated tumor cells have been reported to have reduced proliferative capacity, reduced the likelihood of disease recurrence, and enhanced survival outcomes relative to IDH wild-type tumors (IDH wild type: 15 months; IDH mutant: 31 months) (11). A study using the 2016 WHO classification scheme identified IDH1 mutations in greater than 80% of low-grade gliomas, whereas the rate of IDH mutation in glioblastoma was only 3.7% (12). Thus, in humans, the mutation is more common in low grade glioma compared to glioblastoma. Dogs are reported to not harbor IDH mutations and therefore would be predicted to resemble glioblastoma more closely than low grade glioma (13).

Difficulty differentiating oligodendroglioma from other glioma tumor types based solely on histological features led to the use of molecular screening for a 1p/19q codeletion, with deletion being indicative of an oligodendroglioma. The molecular diagnosis of oligodendroglioma involves the presence of both a 1p/19q codeletion and an IDH mutation. Overall, the prognosis of patients that harbor a 1p/19q codeletion have a more favorable outcome relative to patients that do not (9).

The major classification scheme discussed above was tailored to adult glioma, but it is important to note that there are also key distinctions with pediatric glioma that require additional molecular characterization. Analysis of DNA copy number variations and gene-expression profiles in pediatric gliomas revealed 44% of tumors exhibited somatic mutations in the H3.3 chromatin remodeling pathway (14). This finding led to a dichotomization of the disease based on H3.3 mutational status, as tumors with H3.3 mutations carry a dismal prognosis due to the inhibition of the methylation pathways (15). An investigation of the molecular landscape of human and canine glioma revealed canine glioma most closely resembles pediatric glioma (16). This finding was supported by highly conserved methylation pattens and similar rate of aneuploidy between canine

and pediatric gliomas, with all canine glioma subtypes closely resembling pediatric gliomas. Overall, canine glioma most closely resembles that of pediatric glioma which makes canine glioma particularly relevant for translational research in pediatric glioma.

1.1.2 Canine osteosarcoma: an ideal clinical model for the human disease

Osteosarcoma (OS) is a malignant tumor of the bone which carries a dismal prognosis in humans and dogs. The tumor has a propensity to affect children and adolescence resulting in a substantial impact on children and their families. Therapeutic advancements for OS have been stagnant for decades, which is in part due to the rare occurrence of the cancer. While there are only approximately 400 humans diagnosed with OS each year in the United States, the rate of disease occurrence in dogs is closer to 8,000 cases a year. When factoring in the size of human and canine populations, it is estimated that the disease prevalence in dogs is up to 75 times the rate of disease in humans (17).

The increased incidence of OS in dogs allows for more clinical data to be collected regarding OS biology in a naturally occurring tumor and overall, the disease pathobiology appears to be conserved between species (18). In both species most tumors develop on appendicular bones and arise in the intramedullary cavity of metaphyseal bones. The tumor microenvironment is characterized by a dense osteoid matrix with minimal immune cell infiltration. Histologically, OS across species bears remarkable similarities and are characterized into three major subtypes: chondroblastic, fibroblastic, and osteoblastic. The ability to identify distinct histological subtypes led to work attempting to correlate histology with clinical outcomes, and the findings are suggestive that fibroblastic tumors have favorable outcomes relative to other subtypes, although the prognostic relevance of histological subtypes is still debated (3).

OS is a locally invasive tumor that results in the lysis of neighboring normal bone. The tumors do not extend across joints without entering circulation and reseeding the adjacent bone.

The frequency of metastasis is high in both species with the primary tissue tropism being the lung and the less frequent site of metastasis being distant bone. Despite overwhelming similarities, there are subtle differences in age of onset, with dogs typically being diagnosis in middle aged years, although a bimodal distribution in which younger dogs develop disease has been described (19). Differences in the rate of metastasis also exist, however in the context of using the dog as model, the earlier rates of metastasis act to provide an avenue for more timely evaluation of therapeutics.

The molecular characterization of both human and canine osteosarcoma has further highlighted similarities in the tumor type between species. For instance, there is a low mutational rate in OS (~2 mutations per Mb in humans, ~1.98 mutations per Mb in dogs) and copy number aberrations involving CDKN2A, CDKN2B, TP53, MYC, and PTEN have been reported in both species (20). Furthermore, both ERK and PI3K–mTOR pathway hyper-activation has been reported in human and canine OS at a rate of 17% and 37%, respectively (20). Together the molecular landscape, histological properties, and clinical similarities between human and dogs makes canine OS an ideal model to evaluate novel therapeutics with translational relevance to human medicine.

1.1.3 Cancer therapeutic advancements from canine to human

To date most of the advancement in anticancer therapeutics have been translated from humans to dogs, however there is ever-growing interest to reverse the discovery pipeline and use dogs to inform human medicine. Standard of care for osteosarcoma consists of radical surgical debulking of the primary tumor coupled with adjuvant chemotherapy. This long-standing approach arose from clinical trials in humans and dogs which suggested cisplatin or doxorubicin administered shortly after surgical debulking of the tumor resulted in enhanced survival outcomes relative to surgery alone (21). Despite the persistence of the standard of-care approach, the

survival rate for metastatic osteosarcoma remains dismal. For humans and dogs with metastatic disease the survival rate is estimated to be approximately 20%, highlighting the need for more therapeutic options (22,23).

The advent of immunotherapeutics, such as PD-1/PD-L1 and CTLA4 checkpoint blockade, led to substantial improvements in survival outcomes of patients with non-small cell lung, melanoma, and multiple other tumor types (24). Despite high expression of PD-L1 on tumor cells in OS and glioma, the application of immunotherapeutics in these cancer types have yielded marginal clinical results (25). Immunotherapeutic approaches that rely on blockade of immune suppressive pathways are dependent on antitumoral T cells being present which upon treatment will enable the reinvigoration of antitumor responses. Because both glioma and osteosarcoma lack adequate basal antitumor T cell infiltration, checkpoint blockade alone may not be sufficient to reinvigorate T cell responses and enhance clinical outcomes (26). Ongoing studies are attempting to combine immune checkpoint blockade with radiation therapy or vaccination to introduce neoantigens and enable T cell mediated immunity to benefit from checkpoint blockade (27).

Cellular immunotherapeutics, such as chimeric antigen receptor T cells (CAR T cells), are gaining interest as a therapeutic in glioma and OS, and pre-clinical studies in murine models have shown promise. Due to broad expression on tumor cells by immunohistochemistry, CD276 (B7-H3), HER2, and GD2 are some of the most widely explored targets in osteosarcoma and glioma. As such, CAR T cells targeting CD276 have shown promise in pre-clinical murine models and recent advances in canine therapeutics led to the infusion of CD276 CAR T cells into healthy dogs to evaluate safety (28,29). Despite unprecedented outcomes of CAR T cells used to treat leukemias, the therapeutic approach applied to solid tumors has been less promising. Efforts in CAR T cell approaches are now attempting to enhance homing to the tumor with dual valent CAR T cells or through modification of the tumor microenvironment (30). Major strategies of tumor microenvironment modification consist of depleting stromal cells or immune suppressive myeloid

cells and will be discussed in more detail below. Overall, the similarities between human and dog, coupled with the recent introduction of CD276 CAR T cells into dogs, makes canine cancer a potentially valuable model to expedite exploration of combinatorial therapeutic approaches aimed at enhancing CAR T cell efficacy in solid tumors.

Macrophages (and microglia in glioma) are the most abundant immune cells in OS and glioma. Tumor-associated macrophages (TAMs) are traditionally thought to play key roles in maintenance of an immune suppressive microenvironment and have become a therapeutic target. Across several tumor types, increased TAM abundances have been correlated with poorer survival outcomes (31–33). Despite this, correlation of TAM abundances in the setting of OS have been contrary to other tumors. Such unexpected findings include reports suggesting increased TAM abundance is correlated with a lower rate of lung metastasis and prolonged survival (34,35). Interestingly, these correlations have been observed in both human and canine OS (34,35). It should be noted that some human OS researchers identified conflicting results in which increased TAM infiltrates was found to be correlated with poorer clinical outcomes (36). Thus, the role of TAMs in OS is still debated and there is a need to further investigate the underlying biology.

Despite uncertainty regarding the role of macrophages in OS, TAM targeted therapies are being actively investigated. The major therapeutic approaches consist of TAM depletion, blocking monocyte recruitment, or reprogramming TAMs toward antitumoral phenotypes. Depletion strategies applied to OS have consisted of broad depletion with liposomal clodronate or targeted CD163 TAM depletion. Both strategies have been reported to increase T cell activation, enhance survival, and reduce metastatic rate in preclinical murine models (37–39). We have identified blockade of monocyte recruitment to the tumor using losartan to block CCL2-CCR2 mediated recruitment enhanced survival outcomes in dogs with metastatic osteosarcoma (40). Efforts to reprogram TAMs have taken multiple approaches, but the use of mifamurtide and trans retinoic acid have each been demonstrated to skew TAMs away from a pro-tumoral phenotype toward an antitumoral phenotype and may represent useful therapeutics (41,42). The mTOR/PI3K pathway

represents a widely studied axis which in preclinical models revealed treatment can reprogram TAMs toward tumor suppressing characteristics. Despite promising preclinical findings, a canine osteosarcoma clinical trial identified the use of adjuvant sirolimus (an mTOR inhibitor) to have no clinical impact relative to amputation plus adjuvant chemotherapy (43).

Overall, advancements in osteosarcoma and glioma therapeutics will require greater understanding of tumor biology and thoughtful combination therapies to bolster antitumor immune responses. The dog represents a valuable spontaneous model that recapitulates human disease and can act to provide an avenue for expedited understanding of tumor biology and evaluation of clinical efficacy of novel therapeutics.

1.2 Cancer immunity: a focus on immune suppressive cellular mediators

Cancer encompasses a heterogenous group of diseases which are all interrelated based on the presence of key hallmarks (44). Although cancer typically arises from a focal lesion which induces the malignant transformation of a cell, the disease can exhibit profound systemic effects in an affected individual. The number of reported cancer-induced changes is forever growing, but it has been consistently reported that the expansion of circulating immune suppressive immune cell populations, reprograming of tumor-infiltrating immune cells to promote tumor growth, and dysfunction of cells in peripheral tissues represent key immune-related cancer-induced changes. Ultimately the progression of cancer is multifaceted, in which the many subtle changes work together to impact major biological systems. Thus, without downplaying the significance of other contributors to cancer immunity, this section will focus on the impacts of circulating myeloid-derived suppressor cells and immune suppressive tumor-associated macrophages.

1.2.1 The myeloid-derived suppressor cell[1]

Myeloid-derived suppressor cells (MDSCs) are a key population of cells known to play a prominent role in cancer progression (45,46). MDSCs were first described in human cancer patients, yet they have since been defined as a collection of myeloid cells reported to accumulate in mammalian tissues in response to prolonged inflammation (47,48). Under physiological conditions, this cell population is believed to quell excessive inflammation, promote tissue rebuilding, and assist in the restoration of tissues to the pre-inflammatory state (49). However, it is now appreciated that cancer is akin to a chronically inflamed wound, and that the inflammatory signals emanating from the tumor microenvironment (TME) recruit MDSCs (50). Once present in the TME, these highly suppressive cells restrict the activity of antitumor immune cells and allow tumors to evade the immune system (51). As a result, strategies for preventing or reducing the negative impact of MDSCs on antitumor immunity are of great interest in cancer research.

Currently there are two known subdivisions of MDSCs, polymorphonuclear (PMN) and monocytic (M)-MDSCs, and these cells are typically identified by the expression of stereotypic surface markers via flow cytometry. In humans, $CD11b^{+}CD14^{-}CD15^{+}$ or $CD11b^{+}CD14^{-}CD66b^{+}$ cells are considered PMN-MDSCs and $CD11b^{+}CD14^{+}HLA\text{-}DR^{lo}CD15^{-}$ cells are considered M-MDSCs (47). While in mice, PMN-MDSCs are defined as $CD11b^{+}Ly6G^{+}Ly6C^{lo}$ and M-MDSCs as $CD11b^{+}Ly6G^{-}Ly6C^{+}$ cells. Although research into canine MDSCs is less robust, early reports indicate that immunosuppressive MDSC-like populations can be identified as cells that are $CD11b^{low}$/$CADO48A^{low}$ or $CD11b^{+}$/$CD14^{-}$/$MHCII^{-}$ (52). Subsequent characterization and RNA sequencing data have further defined MDSCs generally as $MHCII^{-}CD5^{-}CD21^{-}CD11b^{+}$ with the ability to subclassify MDSCs as polymorphonuclear or monocytic based on the expression of CD14 and a pan-neutrophil marker, CADO48a (53). Using this classification scheme, PMN-MDSCs are represented in the $CADO48A^{+}CD14^{-}$ population and M-MDSCs are $CADO48A^{-}CD14^{+}$ cells.

Despite definitive markers to differentiate between PMN- and M-MDSCs in dogs, there is a limitation in methods to reliably distinguish between “normal” neutrophils and PMN-MDSCs. The only method to differentiate neutrophils from PMN-MDSCs is based on changes in their cellular density (47). It has been reported that PMN-MDSCs become hypodense and will separate out from “normal” neutrophils when completing density centrifugation. Although there are many potential confounders, such as activated neutrophils having an altered density and time from blood collection to density centrifugation, this is currently the only method to differentiate these cells. Further confounders include species differences in neutrophil densities, which increases the potential for findings in one species to fail to translate to another. An important consideration in dogs is that density centrifugation of whole blood has long been reported to allow varying amounts of neutrophils to be isolated (54). Thus, there is a great need for the identification of additional surface markers that allow for consistent and accurate identification of PMN-MDSCs. Until such a marker is found, the easily confounded approach of density centrifugation remains the only method to identify PMN-MDSCs.

Although phenotypic markers are important for efficient identification of these populations, the defining characteristic of MDSCs is their ability to functionally suppress immune populations, particularly T cells. T cell suppression is often measured using co-culture experiments in which T cells are activated in vitro, and their proliferation is measured (55). Cells suspected of being MDSCs are added to the culture to determine if they are capable of reducing T cell activation. Considering that most cancer treatments are at least partially dependent on effective T cell mediated antitumor immune responses, it is evident that MDSC mediated T cell suppression is a major hurdle that antitumor immune responses must overcome. As a result, recent research has begun to delineate the mechanisms through which MDSCs suppress antitumor immunity.

One such mechanism that MDSCs use is the release of immunosuppressive cytokines, such as transforming growth factor β (TGF-β) and interleukin-10 (IL-10), that act to directly suppress T cell function (56–59). In addition, these cytokines can also facilitate the accumulation

of regulatory T cells, which can further promote an immune suppressive environment through mechanisms of their own. Furthermore, MDSCs can express a high amount of program cell death ligand 1 (PD-L1) while in circulation and within the TME (60). The expression of PD-L1 is essential in regulating immune responses, but its expression can have detrimental effects on antitumor responses if overexpressed. Through binding of its cognate receptor, programmed cell death protein 1 (PD-1), T cell signaling is induced which leads to reduced T cell function and promotes exhaustion (61). In addition, MDSCs express elevated levels of the catabolic protein arginase 1 (ARG-1), which leads to the depletion of L-arginine in the surrounding area. T cells require arginine for their function, thus metabolic competition is another mechanism that can profoundly suppress T cell function (62,63). MDSCs can also produce numerous types of reactive oxygen species (ROS). These molecules are toxic to many cell types, but T cells are particularly vulnerable. When exposed to high levels of ROS, T cells exhibit reduced antitumor function and increased levels of apoptosis, further reducing the efficacy of the immune response (64). Another mechanism used by MDSCs to inhibit antitumor immunity is through the conversion of adenosine triphosphate (ATP) into adenosine. ATP is a highly immunogenic molecule that can promote the initiation of robust antitumor immune responses, while adenosine is an immunosuppressive molecule that subdues immune responses. MDSCs can express the ectonucleotidase enzymes CD39 and CD73, which catabolize ATP into adenosine, and therefore, shift the balance towards adenosine (65). Taken together, the many different MDSC mediated immunosuppressive pathways are often devastating to anticancer immunity, and there is a need to understand the processes that contribute to exacerbation of MDSC accumulation and function.

[1]Text modified from Ammons, D.T., MacDonald, C.R., Chow, L., Repasky, E.A. and Dow, S., 2023. Chronic adrenergic stress and generation of myeloid-derived suppressor cells: Implications for cancer immunotherapy in dogs. Veterinary and Comparative Oncology, 21(2), pp.159-165. **https://doi.org/10.1111/vco.12891**

1.2.2 Immune suppressive macrophages and their impact on tumor immunology

Recent technological advances have enabled detailed characterizations of the myeloid compartment of tumors and have revealed the presence of range of macrophage cell types. Early studies of macrophage heterogeneity relied heavily on the dichotomization of macrophages as pro- or antitumoral which mirrored the pro-inflammatory (M1) and anti-inflammatory (M2) macrophage states that have been widely studied using in vitro systems. Subsequently, the roles of macrophages have been identified as a spectrum of functions, with dichotomization largely failing to accurately capture the role of macrophage in the tumor (66). The advent of single-cell RNA (scRNA) sequencing has generated an overwhelming amount of data that must be closely evaluated to understand the biologically relevant subdivisions of macrophage populations. As such a recent review article by Ma et al. summarized findings across many tumor types to identify 7 conserved TAM populations (67). The 7 proposed TAM subtypes are: interferon-primed TAMs (IFN-TAMs), immune regulatory TAMs (RegTAMs), inflammatory cytokine-enriched TAMs (Inflam-TAMs), lipid-associated TAMs (LA-TAMs), proangiogenic TAMs (Angio-TAMs), resident-tissue macrophage-like TAMs (RTM-TAMs), and proliferating TAMs (Prolif-TAMs). In addition to the TAM populations, tumor-infiltrating monocytes (TIMs) exhibit distinct transcriptomic signatures and make up a key TAM precursor cell type. Of the 8 TAM/TIM subtypes, LA-TAMs, RegTAMs, Angio-TAMs, and TIMs are reported to exhibit some of the most profound immune suppressive properties and will be discussed in detail.

As the name suggests, LA-TAMs favor lipid metabolism and are thought to actively suppress antitumor immune responses via an ARG1 dependent mechanism. Transcriptionally, this cell type is defined by SPP1, C1QA/B/C, FABP5, and TREM2 expression, while flow cytometry of $TREM2^{+}$ (Triggering Receptor Expressed on Myeloid cells 2) macrophages isolated from tumors were defined by high expression of immune suppressive molecules ARG-1, CD206, and IL-10 (68). Indirect functional evidence of immune suppressive potential was presented in the

context of breast cancer in which increased abundances of LA-TAMs were found in patients who failed to respond to checkpoint blockade (69). Furthermore, it was reported that genetic depletion of LA-TAMs using a TREM2$^{-/-}$ mouse model reduced the rate tumor growth (69). Additional functional and correlative analysis is needed to define the role of LA-TAMs in tumor biology, but overall, the current evidence suggests LA-TAMs are a pro-tumoral macrophage subtype with immune suppressive properties.

Immune regulatory TAMs are closely related to anti-inflammatory macrophages with ARG1, MRC1, and CX3CR1 being key transcriptomic markers of the cell type. The group of Reg-TAMs overlaps with other cell types, such as LA-TAMs and RTM-TAMs based on TREM2 expression, so TREM2 positivity likely corresponds to a heterogenous group of cells. Therefore, studies investigating the role of TREM2^{+} macrophages should take the heterogeneity into consideration and apply additional markers to separate functional entities. Nonetheless there is growing evidence that TREM2+ macrophages broadly have of a protumor role which makes the TREM2 signaling axis an appealing therapeutic target (70). As such, a recent report identified blockade of TREM2 using a monoclonal antibody in combination with NK cell activating therapeutics resulted in reduced tumor burden in a murine metastatic lung cancer model (71). As therapeutics targeting TREM2 are developed consideration of heterogeneity within the target population may be an important consideration when evaluating clinical outcomes and predicting responsiveness to intervention.

Angio-TAMs are defined by an angiogenic gene signatures which includes expression of VEGFA, SPP1, VCAN, FCN1, and THBS1. Evaluation of gene regulatory networks in angio-TAMs further revealed CEBPB, FOSL2, and HIFA transcription factors to be key factors in angio-TAM differentiation. Spatially, this TAM subset is predicted to be enriched near tumor vasculature and in hypoxic regions of tumors, as evidenced by a regulatory role of HIFA in angio-TAMs. Given that VEGF expression is a defining feature of this cell type, angio-TAMs may contribute to cancer progression through promotion of epithelial-to-mesenchymal transition (EMT) in malignant cells

(72). To this end, a study which used angio-TAM gene signatures to estimate cell type abundances and found increased angio-TAM signatures to be associated with poorer prognosis (73). Currently, angio-TAMs represent an attractive therapeutic target which warrants further investigation to identify mechanisms to best target them.

Tumor-infiltrating monocytes have been identified in multiple tumor types and their presence highlights the complexity of macrophage biology, as traditionally monocytes have been discussed as transitioning to macrophages upon extravasation. Despite a distinct monocyte gene signature, tumor-infiltrating monocytes differ from their circulating counterparts based on the upregulation of inflammatory molecules and reduced expression of adhesion molecules (73). Furthermore, studies have been able to consistently differentiate between $CD14^+$ (classical) and $CD16^+$ (non-classical) monocytes in human tumor tissues, which suggests a degree of heterogeneity in TIMs. Further study of TIM heterogeneity may reveal novel targets to block monocytes recruitment to the tumor microenvironment.

Only the TAM populations with reported immune suppressive properties were discussed, but further detailed descriptions of the other TAM populations are available in Ma et al. Although efforts have been made to segment macrophage populations into distinct entities, it is important the keep in mind macrophages are a highly plastic cell type that may not fall cleanly into segmentations. To this end, single-cell transcriptomics is able to capture unprecedented heterogeneity within TAM populations, but the approaches used to identify distinct cell types are susceptible to over clustering of cells into divisions that do not accurately recapitulate biological divisions of cell types. In the case of Reg-TAMs and LA-TAMs it is possible that TREM2 expression is the biologically relevant distinction and that functionally Reg-TAMs and LA-TAMs are more similar than dissimilar. Ultimately, the classifications made using scRNA sequencing must be functionally validated to ensure they represent a biologically relevant subdivision of macrophages.

1.3 Molecular dissection of complex tissues

In recent years there have been substantial technological advances that are enabling the dissection of complex tissues to reveal previously unknown aspects of biology. The key behind these advances is the ability to complete in-depth characterization of cell types at a single cell level. Several different platforms have been developed to breakdown tissues into their cellular constituents, but a leading method is the use of microfluidic partitioning of single cell suspensions and the study of cellular transcriptomics using molecular barcoding. These powerful approaches allow researchers to move past traditional molecular approaches, such as bulk RNA sequencing, to begin to understand how individual cells contribute to the complex tissue systems. In this section the evolution, analytical approaches, and key discoveries resulting from the use of single-cell transcriptomics will be discussed.

1.3.1 The evolution of single-cell RNA sequencing

Ribonucleic acid (RNA) is a ubiquitous molecular component that has many functions, but the key function discussed here is the role in the conversion of deoxyribonucleic acid (DNA) into functional protein units that enable cells to carry out functions. Through the study RNA, specifically messenger RNA (mRNA), researchers can obtain insights into how cells are responding to their environments. The beauty of transcriptomics is that mRNA is evolutionarily conserved which makes it so the technologies can be applied to nearly any biological system (74).

The study of individual transcripts began many years ago when the first reverse transcriptase (RT) enzyme was isolated from a retrovirus and adopted for use in polymerase chain reaction (PCR) (75). RT-PCR enabled the study of gene expression within cells, but the approach was limited to evaluating the expression of one or few genes of interest. Around 35 years after the discovery of the reverse transcriptase enzyme, next-generation sequencing approaches were developed which built up the conceptual framework of RT-PCR, but enabled a more complete analytic approach in which an entire transcriptome could be studied in one assay (76). The most

widely used approach to study transcriptomic changes between conditions and tissues became known as bulk RNA sequencing. These approached rely on the extraction of RNA from tissues, enrichment for mRNA, cDNA generation and fragmentation, followed by ligation of sequencing adaptors to generate a library compatible with next-generation sequencers (77). Bulk RNA sequencing enabled a rapid expansion in our understanding of molecular characteristics of cells, but the technology was limited to averages within in the sample and could not resolve out how individual cells are responding.

In 2009, Tang et al. completed the first experiment which laid the foundation for the study of individual cell transcriptomes (78). The approach used manual isolation of single oocyte or blastomere cells from mice and subsequent library preparation to provide a proof of concept that individual cell transcriptomes could be captured with a sensitivity similar to bulk RNA sequencing. The first assay was low throughput in which less than 10 cells were profiled with deep sequencing resulting in a high cost per cell. From there, a rapid expansion of approaches followed, with increases in the throughput jumping from 1 cell to 100's of cells, to 1000's of cells, with the most recent approaches enabling the profiling of 1,000,000 individual cells (79). The technological advances progressed from manual isolation to FACS-based to microfluidic-based to combinatorial barcoding (80). Each change in approach increased the throughput of the technology, but the increase was achieved in part by shifting from full-length sequencing approaches to approaches that are biased toward the 3' (or sometimes 5') portion of the mRNA molecule.

Today, the most widely used approach used to complete scRNA sequencing is 10x Genomics microfluidic-based assay which uses specialized hardware and microfluidic channels to pair cells with molecular barcodes. Briefly, the approach uses microfluidics to partition cells with a gel bead containing a unique oligonucleotide barcode sequence. Individual cells interact with one gel bead then enter partitioning oil to form a Gel Bead-in-Emulsion (GEM). Once in the GEM, the cell is lysed, all transcripts are labeled with a unique barcode, and reverse transcribed to create a cDNA for each cell. The cDNA is then pooled, PCR amplified, cleaned up, and Illumina

sequencing adapters are added. From there, the library is sequenced to gain insight into individual cell transcriptomes.

1.3.2 From basic to advanced analysis approaches

With the rapid and widespread adoption of scRNA sequencing there has been an equally rapid development of analytic tools to aid researchers in addressing pertinent scientific questions. The sheer number of tools (an up-to-date list of scRNA sequencing tools can be found at https://www.scrna-tools.org/) can be intimidating and the lack of current gold standards in analysis approaches further complicates establishment of analysis pipelines. Therefore, the aim of this section is to describe a general scRNA sequencing pipeline and discuss the application of different approaches.

Alignment and data normalization

Data analysis begins when raw FASTQ files are generated on sequencer. From the raw data, reads are aligned to a reference genome and a cell by gene count matrix is generated. Spliced Transcripts Alignment to a Reference (STAR) is the most widely used aligner with 10x Genomics Cell Ranger platform and STARsolo using a STAR aligner (81). The count matrices generated from scRNA sequencing are remarkably different from bulk RNA sequencing matrices due to the sparsity of the count matrices. All scRNA datasets are zero inflated with greater than 90% of elements in the matrix being zeros (82). The zeros can arise from true biological phenomena (i.e., CD20 expression on B cells not T cells) or technical biases such as drop out or inefficient capture of a cell's transcriptome. Nonetheless the zero inflated datasets have profound impacts on analytical approaches that can be applied and should be acknowledged by experimenters using the technology.

Data normalization is the most fundamental step in analysis, yet there is no single-cell specific approach that has been widely accepted (83). The most basic normalization approach consists of normalizing within each cell by dividing the counts for a given feature by the total number of counts for that cell. The fraction for each feature is then scaled (typically multiplied by 10,000) and natural log transformed. The resulting counts can be used for differential gene expression (DGE) analysis and other downstream analytic approaches. Although this “simple normalization” approach is the most widely used (default method using the NormalizeData() function of the Seurat package) there are assumptions that may not be met in certain datasets and should be considered when normalizing data (84). The key assumption that should be met before using a simple normalization approach is that each cell’s transcriptome should be approximately equal in size across cell types and treatments. Given this assumption, simple normalization may perform poorly when there is a mix of low diversity (i.e. granulocytes) and high diversity (i.e., epithelial/tumor cells) cells within a dataset. Scran, an alternative normalization approach that uses pooled variances (uses clustering and average number of features in a cluster to create pools), was developed to enable normalization without relying on the assumption that all cells have equal sized transcriptomes (85). While this normalization approach improves downstream analysis in certain datasets, the most effective normalization method for a given dataset should be determined through evaluation of multiple different normalization approaches and the cellular composition of the dataset. Outside of scran and simple normalization, there are many other proposed methods (83), but the applicability of each approach is variable and currently simple normalization appears to represent the most versatile method. Depending on experimental setup, subsetting and re-normalizing data using only cell types with similar transcriptomic sizes can overcome some of the limitations of simple normalization and increase the accuracy of downstream analysis. Ultimately, normalization is the most fundamental data processing step and directly impacts all downstream analysis; therefore, it is important to consider

biases introduced through normalization and determine which normalization approach is best suited for a given dataset.

Quality control

Once data are normalized the next step in analysis is to remove low quality cells (and complete background correction if desired). In the context of microfluidic-based scRNA sequencing approaches, low quality cells will typically present as three different types: dead cells, cell doublets, and empty droplets. Each low quality will be discussed in the subsequent text.

Dead cells are most commonly filtered out by evaluating the percentage of reads mapping to mitochondrial features within a given cell (86). This approach relies on the notion that as cells die their plasma membranes breakdown and become leaky. Once the membrane is compromised cytosolic contents, including mRNA, will be released. Due to the size of mitochondria, mitochondrial mRNA persists in a dying cell longer which results in an enrichment of mitochondrial-associated reads.

Cell doublets within datasets can impact clustering and DGE analysis so it important to take steps to reduce their impact on downstream analysis. Microfluidic-based approaches rely on the partitioning of individual cells into droplets, but there is an expected doublet rate which varies based on the number of cells loaded into the instrument. Outside of the technical doublet rate, experimenters must also consider biologically relevant doublets in their sample (i.e., incomplete digestion, cell clumps) and should modify the expected doublet rate to reflect this. Basic doublet filtering consists of removing cells that present as high-end outliers with elevated numbers of UMIs and features. Machine learning doublet prediction methods should also be used to identify cell doublet candidates. Multiple algorithms are available, with each using the cells within a sample to train an in silico generated dataset to predict which datapoints represent putative cell doublets (87). All current approaches (with one exception) artificially create cell doublets then

score cells in the dataset to estimate the likelihood a cell represents a doublet. These algorithms are effective at identifying heterotypic cell doublets (two distinct cells; i.e., T cell-Monocyte doublet), but typically perform poorly in the identification of homotypic cell doublets (two like cells; i.e., T cell-T cell doublet).

The final contributor to low quality cells is empty droplets. These “cells” occur when there is background mRNA in solution during capture and enough mRNA is present that an oligonucleotide coated bead captures the free-floating mRNA. The empty droplets can be filtered out by removing datapoints with low numbers of UMIs and features. However, if a study is interested in low diversity cell types (i.e., granulocytes), careful setting of the QC cutoffs or avoiding applying the low-end filters of granulocytes clusters may be required. This background RNA has been described as “soup” and the mRNA associated with soup typically resembles a tissue signature (88). Due to the potential confounding of background signal when completing DGE analysis, tools have been developed to estimate and remove the mRNA soup (88). These approaches may be valuable when comparing datasets between tissue types (i.e., tumor and normal tissue).

Dimension reduction, visualization, and unsupervised clustering

Once low-quality cells are removed from a dataset, the next steps are to complete dimension reduction, visualization, and unsupervised clustering to transform the high dimensional data into a simplify plotting scheme. The first steps in dimension reduction are to identify features that exhibit highly variable expression within the dataset (89). The approach to identify highly variable features differs based on the tool being used and there is evidence of inconsistencies between approaches, so evaluation of consistency between experiments and approaches can help determine biological relevance of feature selection (90). Regardless of approach used, between 1,500 and 5,000 of the most highly variable features are typically used to construct

principal components for further dimension reduction. Once principal components are established the data can be visualized on a principal component analysis plot or more commonly used to complete non-linear dimension reduction. The non-linear dimension reduction approaches widely applied to scRNA datasets are uniform manifold and approximation (UMAP) and t-distributed stochastic neighbor embedding (t-SNE) (91,92). Both approaches attempt to reduce multivariate data into a two-dimensional plot that indicates the relative relatedness between cells. The choice of which dimension reduction approach to use largely comes down to experimenter preference, but there are some subtle distinctions between the two approaches. t-SNE was developed prior to UMAP and tends to have a longer processing time with greater resource requirements. Thus, if working with large datasets, UMAP may be more amiable as the faster computing time can facilitate evaluation of multiple parameters. An important distinction between the two approaches is that UMAP is generally better at conserving the global data structure, while both approaches are able to retain local data structure.

Data harmonization

Harmonization of samples within a dataset is necessary to remove batch effects and enable accurate interpretation of the data. Similar to a previous comment on normalization tools, the utility of integration approaches will perform better under certain circumstances than others. Therefore, evaluation of multiple approaches is recommended to ensure the data are analyzed in a biologically relevant manner while minimizing artifact introduced from batch effect. As highlighted in several integration benchmark reports, skipping integration will lead to misleading clustering and potentially contribute to inappropriate interpretation of the data (93,94). There are two major approaches to integrating samples to remove batch effects: (1) cell-level similarity search and (2) cluster-level similarity search. The first approach identifies similarities between individual cells across samples whereas the second approach takes a broader approach and uses

cell clusters to evaluate similarities across samples. One downside of cell-level similarity searches is that they can require large computation resources and in turn reduce the scalability of the approach. The downside of using a cluster-level search is that small clusters may be placed into other similar cell types following integration and in turn reducing the resolution of downstream clustering. Both approaches can perform well, and multiple approaches should be investigated to identify the method that works best for each dataset.

Cell classification

Once data are integrated, cell classification is the next step which is another fundamental step, as all downstream approaches rely on accurate classification of cell types. Although annotation of datasets may seem trivial, in practice the process is not clear cut and should be approached using an iterative process and multiple difference approaches to assign cell identities. Algorithmic approaches which use machine learning and anchor identification between a reference and query dataset are becoming widely used (95). In humans and mice, multiple references exist for the training of datasets to increase the likelihood of accurate cell classification, but in the context of non-traditional animal models, references are lacking and the use of reference-based classifications is limited (96,97). Although human or murine datasets can be used to identify cell types in non-traditional animal models, the approach may result in cell types being forced into categories that do not accurately capture subtle species-specific differences in cell types. Therefore, until adequate reference datasets are available for algorithmic-based classification for each non-traditional animal model, manual classification will be a leading approach to assign cell type identities. Ideal human and murine references consist of multimodal Cellular Indexing of Transcriptomes and Epitopes by Sequencing (CITE-seq) datasets which allow for researchers to assign identities based on well-studied cell types, and further use transcriptomic data to describe the heterogeneity within antibody-based nomenclature (98). In non-traditional

animal models where CITE-seq might not be available, the use of paired flow cytometry analysis to correlate cell type percentages with traditional gating schemes can assist in validating cell types for reference generation.

Single-cell statistics

Statistical inference is an essential part of analysis that enables researchers to identify changes between conditions as well as generate cell type specific gene signatures. There are two major approaches that can be applied to identify differentially expressed features: (1) cell replicate based and (2) sample replicate based.

Cell replicate based approaches are the most widely used and will have the greatest sensitivity. These approaches use individual cells as replicates then use statistical tests (Wilcoxon Rank Sum test the most widely used). Due to the fact that each cell is used as a replicate the approach is using a unit (cells) that has limited reproducibility between experiment. Furthermore, the use of cells as replicates inflates sample size which acts to increase statistical significance. Overall, the application of cell replicate based approaches are best when applied to define cell type specific gene signatures, but the approaches should be used cautiously when contrasting between conditions (99,100).

Sample replicate-based approaches collapse data into biological replicates then complete DGE analysis. This approach has been termed pseudobulk conversion, as the resulting count matrix closely resembles that of a count matrix observed in bulk RNA sequencing. Thus, the approach captures biological replicates rather than cell replicates which reduces the impacts of sample size inflation, better captures data variability, and results in more stringent (and potentially more biologically relevant) DGE analysis. The approach has the benefit of overcoming impacts of dropout because counts within like cells are pooled, but relative to bulk RNA sequencing the counts are still low and filtering steps should be completed to remove lowly expressed features

prior to DGE analysis. Additionally, the use of sample replicate based approaches may not be possible when studying rare cell types, at which point cell replicate based approaches may have to be used. In summary, cell replicate based are most useful for identify cell type gene signatures, while replicate based approaches are between for completing contracts across conditions.

Specialized analysis approaches

When completing scRNA sequencing on samples containing tumor cells it is important to distinguish healthy cells from malignant cell. In certain cancers copy number aberrations are frequently observed and can act as an avenue to identify genomic changes that can be suggestive of malignancy. Although copy number variation (CNV) analysis has traditionally been used with whole exome or whole genome sequencing, the use of scRNA sequencing datasets to infer CNV status is becoming widely used (101–103). The basic concept of CNV analysis in scRNA datasets is that if a deletion or amplification of a chromosome occurs in a given cell, the counts of features within the genomic block will be reduced or increased, respectively. As such the available approaches bin (segment) the genome then assesses each cell for changes in expression within each bin to infer genomic aberrations. The two most widely used tools are infercnv and CopyKAT, but algorithms were designed for use in murine and human datasets, so application in non-traditional animal models should be interpreted cautiously.

Trajectory analysis is a computational approach that relies on the assumption that a sample contains cells at various states along differentiation process, then uses cell transcriptomes to describe how cells are progressing. This approach is especially useful if working with serial timepoint datasets or tissues in which cells are actively differentiating (bone marrow, thymus, etc.). There are two main approaches used to study the trajectory of cell types: (1) RNA velocity and (2) unsupervised time ordering and clustering. RNA velocity is an unsupervised approach that captures transitional cell states using the concept that as a transcript matures it progresses from an unspliced to a spliced form before being translated to a protein (104). Therefore, by using the

output BAM file from alignment each transcript is classified as ‘spliced’, ‘unspliced’ or ‘ambiguous’. Using the classifications, progressive cell states can be established by evaluating the expression of spliced and unspliced transcripts. The second overarching method, unsupervised time ordering and clustering, takes a different approach and rather than using alignment files the algorithm instead assigns a pseudotimeline based on dimension reduced data. These algorithms computationally reconstruct the transcriptional transitions and should only be applied to datasets in which there is an expected biologically relevant progression of cell types. Several different algorithms exist to reconstruct pseudotime and each tool has a different capacity to handle branchpoints in the reconstructed trajectory (105). Currently there are over 45 different algorithms available which each tool’s performance highly dependent on the structure of the trajectory being reconstructed. Due to the variable efficacy of algorithms, current best practices should use multiple unsupervised time ordering tools coupled with RNA velocity to achieve the highest likelihood of discovering biologically accurate reconstructions of cell progression.

In an effort to extract all biologically relevant insight from scRNA datasets, several approaches to infer how cells are interacting within tissues have been developed. One of the greatest limitations of the analysis is that when using microfluidic approaches there is no method to evaluate the spatial relationship between cells to know if they are in a close enough proximity to interact. The advent of spatial transcriptomics enables analysis to take into consideration the proximity of cell types, but the application will not be discussed here. All available approaches require the use of a reference dataset consisting of a list of known receptor-ligand pairs. Therefore, the current approaches do not allow for the discovery of novel cell-cell interactions, but rather infer the likelihood that known R-L pairs are interacting. A majority of available approaches use the law of mass action to infer interaction potential based on the expression level of Rs and Ls between two cell types (106,107). Recently, a novel approach to infer the likelihood of cell-cell interactions which uses the downstream signaling molecules of well-studied R-L pairs rather than expression of the R-L pair was developed (108). While both approaches should yield

similar results, when evaluating cell-cell interactions, findings can be strengthened if predicted interactions are significant using multiple approaches. It is important to note that if these approaches are applied to non-traditional animal models, the results may be impacted by species-specific differences. To this end, the use downstream signaling molecules to infer activity may be inaccurate in non-traditional animal models and approaches using law of mass action may yield more biologically relevant results.

Biases to consider in an "unbiased" approach

Single-cell RNA sequencing is often discussed as an unbiased approach used to investigate transcriptomic heterogeneity in complex tissues. Despite widespread use of the term, unbiased is a misnomer. In a classical sense the technology overcomes obvious biases, such as using predefined markers to study cell populations, but bias is present throughout sample preparation and into data analysis. The introduction of bias begins during the capture of mRNA, in which there is an overrepresentation of long unspliced mRNA in a variety of datasets (109). The size bias coupled with most platforms only sampling roughly 30% of an individual cell's transcriptome leads to the underrepresentation of various features (110). Given only a fraction of a cell's transcriptome is captured, there is a high dropout rate in scRNA sequencing. Dropout occurs when a cell is expressing a transcript, but the capture platform fails to capture the presence or a low enough frequency that the sequencing depth does not pick up the unique transcript. Several approaches to combat drop out have been proposed, but to-date there is no widely accepted method to account for dropout (82,111). It is important that investigators are aware of the phenomenon and are not alarmed when every cell within a cluster does not express a canonical marker.

Despite the platform capturing full length transcripts, the fragmentation of the transcript results in a strong 3' bias which only captures the sequence of a short fragment of a transcript near the poly-A tail (112). Therefore, unlike bulk RNA sequencing in which an entire transcript can

be rebuilt, scRNA sequencing libraries are limited to evaluating a short sequence (typically 75-150 bp) starting approximately 50-500 nucleotides upstream of the poly-A tail. Due to the biases, longer sequencing approaches largely do not yield additional data because reads are generated in the 5' to 3' direction, so the further downstream the read goes the more likely poly-A sequences will be encountered. Therefore, the nucleotides which run into the poly-A tail will ultimately need to be trimmed during alignment or they can negatively impact the alignment rate (this is built into Cell Ranger pipelines). While the 3' bias does not have a profound impact in species with well-established genome annotations, the bias can lead to lower mapping rates in non-traditional animal models due to incomplete 3' annotations. If the mapping rate of sequencing data is less than 30% in any species, the use of additional methods to extend 3' annotations may enhance downstream analysis. Currently, there are a few methods to increase mapping rates to 3' regions which may be worth exploring if the mapping rate is impacting downstream analysis (113,114).

After alignment most of the sources of bias have already been encountered. However, during certain downstream analysis approaches additional sources of bias may be encountered. For instance, when working in non-traditional animal models there may be interest in completing a cross-species comparison to investigate if cell type gene signatures conserved. While this analysis can be useful, much of the nomenclature and gene signatures used to assign cell type identities come from human and murine research. Therefore, the comparison between a non-traditional animal model and human data has an inherent propensity to exaggerate the apparent similarities. As such, it is important to complete deeper transcriptomic comparisons such as a previously reported approach used to compare cell types between mouse and human, or to complete functional analysis of sorted cells to confirm similar functionality between species (115). Ultimately, comparisons between species are impacted by various confounding factors, including genome annotation quality, and the results of cross-species comparisons should be interpreted with caution.

1.3.3 New discoveries in the era of single-cell 'omics

The introduction of single cell 'omics has revolutionized our ability to understand complex tissues in health and disease. Several reference datasets across essentially all human and murine tissues have provided a plethora of data to identify and investigate previously undescribed cell types. A few key discoveries will be discussed next.

The myeloid cell compartment of tumors has historically proven difficult to dissect and many researchers were attempting to dichotomize macrophage into pro- and anti-tumoral despite the abundance of evidence suggesting macrophage function falls on a spectrum (66). Cheng et al. applied scRNA sequencing to describe the myeloid cell heterogeneity across 15 human tumors and identified 16 transcriptomically distinct monocyte/macrophage populations (73). The first key finding they made was that mast cells from different tumor types differentially express TNF and VEGFA in a mutually exclusive manner. Thus, their data suggest mast cells play different roles depending upon the cancer type they are found in. Using gene signatures for mast cells and the TCGA dataset they found that enrichment of mast cells resulted in poorer prognosis in all cancers except for nasopharyngeal carcinoma. They went on to identify nasopharyngeal carcinoma was the only tumor type to preferentially have TNF^+ mast cells, which implicates TNF^+ mast cells as an antitumoral entity, but $VEGFA^+$ mast cells are protumor. Furthermore, they were among the first to report the presence of $LAMP3^+$ dendritic cells across tumor types which additional reports have defined this cell population as mature regulatory DCs (mregDC) (116). This population has been proposed to exhibit immune suppressive properties, but their functional role is the tumor is still being understood. The authors also provided evidence of the broad presence of interferon associated TAMs which exhibit expression of ISG15, IFIT1, IFITM1, and CXCL10. Lastly, they identified $SPP1^+$ and $C1QC^+$ macrophage populations in which $SPP1^+$ were determined to resemble angiogenesis-associated macrophages and have been associated with a poorer prognosis.

Of particular importance to this dissertation are two recently released human scRNA atlases of treatment-naïve osteosarcoma tumors. Between the two studies, over 125,000 cells obtained from 13 primary, 2 recurrent, and 2 lung metastatic osteosarcoma lesions and were profiled in detail (117,118). Both studies report similar major cell types identified, with 11 reported in Zhou et al. and 9 reported in Liu et al. Differences in major cell type annotations arose from the groups deciding to use different nomenclature, and because the Zhou et al. dataset integrated 7 primary, 2 recurrent, and 2 lung metastatic osteosarcoma lesions whereas the Liu et al. dataset was limited to 6 primary lesions. Both report the presence of a distinct fibroblast cluster with an overlapping gene signature of COL1A1, FAP, ACTA2, DCN, LUM, and THY1 while also noting marked heterogeneity within TAMs and malignant osteoblasts. A notable difference between the two studies was that a population of "activated CCR7^{+} DCs" was observed in Zhou et al., but missing from Liu et al. The authors suggested the population may be involved in metastasis, but they did not investigate further. Fortunately, a different group, Hu et al., complete a reanalysis of the dataset release by Zhou et al., in which they document the "activated CCR7^{+} DCs" closely resembles mature regulatory dendritic cells (119). Zhou et al. then went on to bring in external human peripheral blood scRNA datasets to provide evidence of a tissue specificity of mregDCs. The collective contributions from both groups provide a valuable resource that can be used for further investigation by other groups. As such, in chapter 3 we apply the dataset presented by Liu et al. to make inferences about cell type similarities between human and canine primary tumors.

Continuing with a discussion of dendritic cells, a paradigm shifting discovery was made in 2017 when See et al. completed a detailed investigation of dendritic cells (120). In their analysis they unexpectedly identified a population of precursor DCs (preDCs) that have a transcriptomic signature that overlaps with plasmacytoid DCs. Their report went on to demonstrate the preDCs contributed a substantial amount of previously ascribed pDC functions and they were able to propose modified approaches to distinguish between the two cell types in future studies. Their revised approach to study preDCs also shifted the expected abundance of the cell type in

circulation which increased >10-fold from the previously reported 0.001% of cells in peripheral blood. Overall, the contribution revealed unrecognized heterogeneity within dendritic cells and devised an approach to distinguish between cell types to better study the role of the different cell types.

Many additional discoveries have been made using scRNA sequencing, but the final study discussed here is that of Zhang et al. in which mechanisms behind anti-MCSF depletion were investigated (121). Anti-MCSF therapies exhibited promising preclinical impacts on cancer progression, but when applied to the clinic they failed to enhance patient outcomes. Therefore, this group aimed to use scRNA sequencing to investigate mechanisms that reduce the efficacy of anti-MCSF therapeutics. They identified anti-MCSF depletion reduced the abundances of all macrophage populations except for $SPP1^{+}$ angio-TAMs expressing VEGFA, PD-L1, and ARG1. Ultimately, their findings revealed that although overall macrophage abundances are reduced following anti-MCSF treatment, the depletion favors $C1QC^{+}$ TAMs over $SPP1^{+}$ TAMs which is associated with worse prognostic outcomes.

The select scRNA sequencing studies described above highlight some of the ways the technology can be applied to shed light on complex biological process and lead to paradigm shifting discoveries that have profound impacts on our understanding of biology.

1.3.4 A summary of single-cell RNA sequencing studies in dogs to date

To date there have been 12 publications (including our publication based on chapter 2) and 1 pre-print using some form of single-cell RNA sequencing in dogs. Select published contributions will be summarized in this section.

Jackson et al. applied scRNA sequencing to investigate heterogeneity in $CD11b^{+}$ cells isolated from canine whole blood with the goal of further describing canine myeloid-derived suppressor cells (122). They profiled 4,427 cells from a healthy dog and 13,200 from a dog with

treatment-naïve oral melanoma. Their use of CD11b$^+$ sorted whole blood provides unique insights into the myeloid populations found in circulation. However, the downside of the approach is that whole blood has an overwhelming proportion of neutrophils. This is apparent in their dataset, as evidenced by the minuscule diversity in RNA within cells (mean of < 500 unique features per cell). The use of scRNA sequencing to study the granulocytes is notoriously challenging due to small transcriptomes, high RNase content, and sensitivity to sample processing (123). With that said, their analysis identified two clusters to be nearly unique to the dog with oral melanoma. The authors go on to propose that the two clusters may represent myeloid-derived suppressor cells (suspected polymorphonuclear MDSCs), due to the specificity to the cancer dog. However, the gene signatures were not consistent with previously reported canine MDSC gene signatures (which was noted by the authors). Furthermore, the clustering results suggest the suspected MDSC cells are markedly distinct from neutrophil clusters. As PMN-MDSCs are closely related to neutrophils it is unlikely that they would cluster in a distinct location. Collectively, it is more likely that one of the suspected MDSCs clusters, defined by GATA2/MS4A2 expression, is basophils or mast cells, not PMN-MDSCs (125). The authors note that there are discrepant findings and discuss the classification as a possibility rather than as an absolute classification. Overall, the data provide useful insight as the authors used whole blood sorted for CD11b$^+$ cells rather than density centrifugated samples.

Through release of two separate reports Fastrès et al. used single-cell RNA sequencing to characterize the cellular heterogeneity within canine bronchiolar lavage fluid (BALF) (126). In their first report they presented an atlas of 5,710 cells obtained from BALF of four healthy dogs. The authors presented high level cell annotations without the completion of independent reclustering to further resolve cell type heterogeneity. A concerning, but potentially harmless observation, was that the feature plots depicting expression of canonical markers extend into negative ranges. There is insufficient information to determined what expression data are represented by the plots, but the plots could be depicting scaled expression or SCT normalized

data (which can contain negative values). The concern comes from the possibility of the authors using SCT normalized data to complete DGE analysis between clusters, as the results may be inaccurate due to the approach being not recommended for use in DGE analysis (Seurat documentation). It is likely that log normalized counts were used to complete DGE analysis between clusters, but this is not explicitly stated. Importantly, the group released gene lists for each cluster identified through unsupervised clustering which increases the accessibility of the data by enabling researchers to use and evaluate the gene signatures for cell types.

After generating a reference dataset, they released a separate follow-up study where they completed scRNA sequencing on an additional 19,255 cells obtained from 3 healthy West Highland white terrier (WHWT) and 5 WHWTs with idiopathic pulmonary fibrosis (127). The authors focused on macrophages and completed independent reclustering to resolve out 5 macrophages populations. They observed an overrepresentation of one of the clusters (“M2”) in dogs with idiopathic pulmonary fibrosis, which they classified as monocytes. Interestingly, the cluster had the highest TREM2 and SPP1 expression (as determined in supplemental table 3 of the published manuscript) of all macrophage subtypes and may represent a lipid-associated-like macrophage cluster. The authors go on to discuss disease induced changes in another macrophage cluster (“M1”) which revealed FN1, SPP1, CXCL8, and PLAU to be overexpressed in dogs with idiopathic pulmonary fibrosis. However, closer inspection of cell distributions (as depicted in Figure 2 of the manuscript) indicates that clustering may not have been to a high enough resolution as there is a cluster (around coordinates x=-30, y=5) that is largely composed of cells from the healthy dogs. This change in cell distribution likely drove the DGE analysis and further evaluation of the dataset at a higher resolution is warranted to understand what the unresolved subcluster represents. The last point to highlight is that the authors detected a cluster they called “mature dendritic cells” (cluster 11), which after inspection of the gene list (as determined in supplemental table 2 of the published manuscript), likely corresponds to the first report of mature regulatory (aka migratory) dendritic cells in dogs. They report that the mature

dendritic cluster was overrepresented in dogs with idiopathic pulmonary fibrosis, which likely represents a biologically relevant mechanism of the disease process that was not fully explored in the report.

Gingrich et al. employed scRNA sequencing as part of a larger project aiming to describe canine NK cells and their relationship to human and murine cell types (128). For scRNA sequencing analysis they used a 3' RNA-seq protocol employing the QIAseq UPX 3' Transcriptome Kit to profile 96 cells sorted for $CD3^{-}NKp46^{+}$ or $CD5^{dim}$ immunolabeling properties. They used the technology to investigate heterogeneity within each population and confirmed suspicions that $CD5^{dim}$ NK cell population is more heterogenous than $CD3^{-}NKp46^{+}$ populations. The authors went on to use scRNA sequencing to describe how co-culture with K562 human feeder cells impacted the gene expression profile of $CD3^{-}NKp46^{+}$ and CD5dim cells. When completing dimension reduction and visualization on a t-SNE they found activated cells to cluster distinctly from resting. Lastly, their analysis identified KLRB1, a widely used NK cell marker in mice, is a defining feature of canine NK cells.

As part of larger project ("Pandora") two groups used single-cell RNA sequencing of 11 species to make inferences about cross-species similarities (129,130). The canine specific datasets generated through the Pandora project consisted of normal lung tissue and peripheral blood mononuclear cells (PBMCs). They used lung tissue to make inferences about susceptibility each species to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection (the causative agent of coronavirus disease 2019). The PBMCs were used to study the interactome between the species. The processed data were made accessible though the Pandora webpage and are available for browsing (http://120.79.46.200:81/Pandora/PBMC.html). Through exploration of the canine PBMC dataset it became apparent that cell type annotations were inaccurate (i.e. cells annotated as "monocytes" exhibited high expression of CD3E which suggests a classification of T cells is more appropriate). The misclassification impacted downstream analysis of cross-species interactome in which canine "monocytes" more closely

resembled T cells of other species and vice versa. Although the data presented through the Pandora project is a valuable resource, the annotations should be used with caution and reanalysis of the datasets is recommended if using the datasets as a reference.

The most recent canine scRNA sequencing report, released by Eschke et al., was the first published report applying a Chromium Next GEM Single Cell 5' Kit plus Chromium Single Cell V(D)J Enrichment Kit in dogs (131). It should be noted that a preprint by Skidmore et al. was release prior to Eschke et al. in which they reported the use of V(D)J 5' sequencing in dogs (132). The application of the technology, which was previously only applied to mouse and human, represents a substantial advancement in veterinary transcriptomics. The group completed scRNA sequencing on 31,909 αβ T cells sorted from the peripheral blood of 4 healthy beagles. They reported 22 distinct T cell clusters and discussed the importance of carefully evaluating CD4 expression in non-traditional animal models due to poor mapping rate (132). Overall, the group adequately supported their cell type classifications and provided cell type gene lists for evaluation/use by other research groups. One downside to their analysis approach was the lack of subsetting on CD4 and CD8 T cells and completing independent reclustering. Although their analysis appears to have accurately captured the heterogeneity within canine T cells, rare cell populations could have been missed due to the variable features used for dimension reduction not enabling their identification. The application of V(D)J sequencing enabled the authors to identify CD8 MAIT T cells (TCR α invariant cells) to disproportionately use TRAV9-11 and TRAJ28 which provides evidence that they are TCR α invariant. Thus, the corresponding gene signature from CD8 MAIT T cells can be applied to 3' prime sequencing datasets to identify the population without having to complete 5' TCR sequencing. Overall, the report represents a key advancement in our understanding of canine T cell biology and serves as a reference for further research.

In addition to the detailed discussions above, Ayers et al. applied scRNA sequencing to investigate heterogeneity within canine OS cell lines (134). Sparling et al. used scRNA sequencing to investigate impacts of atopic dermatitis on skin cells (135). Früh et al. used a

combination of scRNA sequencing and flow cytometry to study atopic dermatitis induced changes to circulating leukocytes (136). Lastly, Zhou et al. completed an elegant comparison of human and canine hippocampal cells, which suggested high degree of cell type transcriptome conservation between species (137).

In conclusion, veterinary medical researchers can apply single-cell RNA sequencing to make rapid advancements in our ability to study the canine immune system and evaluate similarities to humans. As discussed above, the foundation of scRNA sequencing analysis relies on the accurate classification of cell types which correspond to biologically/functionally distinct entities. Therefore, effort should be placed on generating comprehensive canine scRNA sequencing atlases that can be used for subsequent functional investigation of cell types and hypothesis testing research using transcriptomics. The importance of accurately annotating datasets is paramount, as each reference has the potential to be used broadly and a misclassification could lead to a downstream interpretation error. To facilitate transparency and ensure the veterinary medical field generates a solid cell annotation foundation, it is essential that researchers deposit not only raw data, but also processed data for other groups to critically evaluate so that the field can reach consensus annotations at a more rapid pace. Overall, the application of scRNA sequencing in veterinary medicine is rapidly growing and the datasets generated will have wide reaching impacts for canine translational research.

References

1. Mak IWY, Evaniew N, Ghert M. Lost in translation: animal models and clinical trials in cancer treatment. Am J Transl Res. 2014;6(2):114.

2. Bhatt RS, Berjis A, Konge JC, Mahoney KM, Klee AN, Freeman SS, et al. KIR3DL3 Is an Inhibitory Receptor for HHLA2 that Mediates an Alternative Immunoinhibitory Pathway to PD1KIR3DL3 Is an Immunoinhibitory Receptor for HHLA2. Cancer Immunol Res. 2021;9(2):156–69.

3. Gustafson DL, Duval DL, Regan DP, Thamm DH. Canine sarcomas as a surrogate for the human disease. Pharmacol Ther. 2018;188:80–96.

4. Chen L, Zhang Y, Yang J, Hagan JP, Li M. Vertebrate animal models of glioma: understanding the mechanisms and developing new therapies. Biochim Biophys Acta (BBA)-Reviews Cancer. 2013;1836(1):158–65.

5. Gardner HL, Fenger JM, London CA. Dogs as a model for cancer. Annu Rev Anim Biosci. 2016;4:199–222.

6. Liu D, Xiong H, Ellis AE, Northrup NC, Dobbin KK, Shin DM, et al. Canine spontaneous head and neck squamous cell carcinomas represent their human counterparts at the molecular level. PLoS Genet. 2015;11(6):e1005277.

7. Louis DN, Perry A, Reifenberger G, Von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 World Health Organization classification of tumors of the central nervous system: a summary. Acta Neuropathol. 2016;131:803–20.

8. Koehler JW, Miller AD, Miller CR, Porter B, Aldape K, Beck J, et al. A revised diagnostic classification of canine glioma: towards validation of the canine glioma patient as a naturally occurring preclinical model for human glioma. J Neuropathol Exp Neurol. 2018;77(11):1039–54.

9. Ceccarelli M, Barthel FP, Malta TM, Sabedot TS, Salama SR, Murray BA, et al. Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma. Cell. 2016;164(3):550–63.

10. Waitkus MS, Diplas BH, Yan H. Biological role and therapeutic potential of IDH mutations in cancer. Cancer Cell. 2018;34(2):186–95.

11. Han S, Liu Y, Cai SJ, Qian M, Ding J, Larion M, et al. IDH mutation in glioma: molecular mechanisms and potential therapeutic targets. Br J Cancer. 2020;122(11):1580–9.

12. Nobusawa S, Watanabe T, Kleihues P, Ohgaki H. IDH1 mutations as molecular signature and predictive factor of secondary glioblastomas. Clin Cancer Res. 2009;15(19):6002–7.

13. Reitman ZJ, Olby NJ, Mariani CL, Thomas R, Breen M, Bigner DD, et al. IDH1 and IDH2 hotspot mutations are not found in canine glioma. Int J Cancer. 2010;127(1):245.

14. Schwartzentruber J, Korshunov A, Liu X-Y, Jones DTW, Pfaff E, Jacob K, et al. Driver mutations in histone H3. 3 and chromatin remodelling genes in paediatric glioblastoma. Nature. 2012;482(7384):226–31.

15. Karremann M, Gielen GH, Hoffmann M, Wiese M, Colditz N, Warmuth-Metz M, et al. Diffuse high-grade gliomas with H3 K27M mutations carry a dismal prognosis independent of tumor location. Neuro Oncol. 2018;20(1):123–31.

16. Amin SB, Anderson KJ, Boudreau CE, Martinez-Ledesma E, Kocakavuk E, Johnson KC, et al. Comparative molecular life history of spontaneous canine and human gliomas. Cancer Cell. 2020;37(2):243–57.

17. Schiffman JD, Breen M. Comparative oncology: what dogs and other species can teach us about humans with cancer. Philos Trans R Soc B Biol Sci. 2015;370(1673):20140231.

18. Withrow SJ, Khanna C. Bridging the gap between experimental animals and humans in osteosarcoma. Pediatr Adolesc Osteosarcoma. 2010;439–46.

19. Klosowski M, Haines L, Alfino L, McMellen A, Leibowitz M, Regan D. Naturally occurring canine sarcomas: Bridging the gap from mouse models to human patients through cross-disciplinary research partnerships. Front Oncol. 2023;13.

20. Gardner HL, Sivaprakasam K, Briones N, Zismann V, Perdigones N, Drenner K, et al. Canine osteosarcoma genome sequencing identifies recurrent mutations in DMD and the histone methyltransferase gene SETD2. Commun Biol. 2019;2(1):266.

21. Link MP, Goorin AM, Horowitz M, Meyer WH, Belasco J, Baker A, et al. Adjuvant chemotherapy of high-grade osteosarcoma of the extremity: updated results of the Multi-Institutional Osteosarcoma Study. Clin Orthop Relat Res. 1991;270:8–14.

22. Szewczyk M, Lechowski R, Zabielska K. What do we know about canine osteosarcoma treatment?–review. Vet Res Commun. 2015;39:61–7.

23. Mialou V, Philip T, Kalifa C, Perol D, Gentet J, Marec-Berard P, et al. Metastatic osteosarcoma at diagnosis: prognostic factors and long-term outcome—the French pediatric experience. Cancer Interdiscip Int J Am Cancer Soc. 2005;104(5):1100–9.

24. Naimi A, Mohammed RN, Raji A, Chupradit S, Yumashev AV, Suksatan W, et al. Tumor immunotherapies by immune checkpoint inhibitors (ICIs); the pros and cons. Cell Commun Signal. 2022;20(1):1–31.

25. D'Angelo SP, Mahoney MR, Van Tine BA, Atkins J, Milhem MM, Jahagirdar BN, et al. Nivolumab with or without ipilimumab treatment for metastatic sarcoma (Alliance A091401): two open-label, non-comparative, randomised, phase 2 trials. Lancet Oncol. 2018;19(3):416–26.

26. Wu C-C, Beird HC, Andrew Livingston J, Advani S, Mitra A, Cao S, et al. Immuno-genomic landscape of osteosarcoma. Nat Commun. 2020;11(1):1008.

27. Takahashi Y, Yasui T, Tamari K, Minami K, Otani K, Isohashi F, et al. Radiation enhanced the local and distant antitumor efficacy in dual immune checkpoint blockade therapy in osteosarcoma. PLoS One. 2017;12(12):e0189697.

28. Zhang Q, Zhang Z, Liu G, Li D, Gu Z, Zhang L, et al. B7-H3 targeted CAR-T cells show highly efficient antitumor function against osteosarcoma both in vitro and in vivo. BMC Cancer. 2022;22(1):1–12.

29. Zhang S, Black RG, Kohli K, Hayes BJ, Miller C, Koehne A, et al. B7-H3 specific CAR T cells for the naturally occurring, spontaneous canine sarcoma model. Mol Cancer Ther. 2022;21(6):999–1009.

30. Foeng J, Comerford I, McColl SR. Harnessing the chemokine system to home CAR-T cells into solid tumors. Cell Reports Med. 2022;100543.

31. Nishie A, Ono M, Shono T, Fukushi J, Otsubo M, Onoue H, et al. Macrophage infiltration and heme oxygenase-1 expression correlate with angiogenesis in human gliomas. Clin Cancer Res. 1999;5(5):1107–13.

32. Kumar AT, Knops A, Swendseid B, Martinez-Outschoom U, Harshyne L, Philp N, et al. Prognostic significance of tumor-associated macrophage content in head and neck squamous cell carcinoma: a meta-analysis. Front Oncol. 2019;9:656.

33. Zhang Y, Cheng S, Zhang M, Zhen L, Pang D, Zhang Q, et al. High-infiltration of tumor-associated macrophages predicts unfavorable clinical outcome for node-negative breast cancer. PLoS One. 2013;8(9):e76147.

34. Gomez-Brouchet A, Illac C, Gilhodes J, Bouvier C, Aubert S, Guinebretiere JM, et al. CD163-positive tumor-associated macrophages and CD8-positive cytotoxic lymphocytes are powerful diagnostic markers for the therapeutic stratification of osteosarcoma patients: An immunohistochemical analysis of the biopsies fromthe French OS2006 phase 3 trial. https://doi.org/101080/2162402X20171331193 [Internet]. 2017 Sep 2 [cited 2023 Jun 7];6(9). Available from: https://www.tandfonline.com/doi/abs/10.1080/2162402X.2017.1331193

35. Withers SS, Skorupski KA, York D, Choi JW, Woolard KD, Laufer-Amorim R, et al. Association of macrophage and lymphocyte infiltration with outcome in canine osteosarcoma. Vet Comp Oncol. 2019;17(1):49–60.

36. Cascini C, Chiodoni C. The immune landscape of osteosarcoma: implications for prognosis and treatment response. Cells. 2021;10(7):1668.

37. Piaggio F, Kondylis V, Pastorino F, Di Paolo D, Perri P, Cossu I, et al. A novel liposomal Clodronate depletes tumor-associated macrophages in primary and metastatic melanoma: Anti-angiogenic and antitumor effects. J Control Release. 2016;223:165–77.

38. Han Q, Shi H, Liu F. CD163+ M2-type tumor-associated macrophage support the suppression of tumor-infiltrating T cells in osteosarcoma. Int Immunopharmacol. 2016;34:101–6.

39. Maloney C, Kallis MP, Edelman M, Tzanavaris C, Lesser M, Soffer SZ, et al. Gefitinib inhibits invasion and metastasis of osteosarcoma via inhibition of macrophage receptor interacting serine-threonine kinase 2. Mol Cancer Ther. 2020;19(6):1340–50.

40. Regan DP, Chow L, Das S, Haines L, Palmer E, Kurihara JN, et al. Losartan blocks osteosarcoma-elicited monocyte recruitment, and combined with the kinase inhibitor toceranib, exerts significant clinical benefit in canine metastatic osteosarcoma. Clin Cancer Res. 2022;28(4):662–76.

41. Punzo F, Bellini G, Tortora C, Di Pinto D, Argenziano M, Pota E, et al. Mifamurtide and TAM-like macrophages: effect on proliferation, migration and differentiation of osteosarcoma cells. Oncotarget. 2020;11(7):687.

42. Zhou Q, Xian M, Xiang S, Xiang D, Shao X, Wang J, et al. All-trans retinoic acid prevents osteosarcoma metastasis by inhibiting M2 polarization of tumor-associated macrophages. Cancer Immunol Res. 2017;5(7):547–59.

43. LeBlanc AK, Mazcko CN, Cherukuri A, Berger EP, Kisseberth WC, Brown ME, et al. Adjuvant Sirolimus Does Not Improve Outcome in Pet Dogs Receiving Standard-of-Care Therapy for Appendicular Osteosarcoma: A Prospective, Randomized Trial of 324 DogsAdjuvant Sirolimus in Canine Appendicular Osteosarcoma. Clin Cancer Res. 2021;27(11):3005–16.

44. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646–74.

45. Yang Y, Li C, Liu T, Dai X, Bazhin A V. Myeloid-derived suppressor cells in tumors: from mechanisms to antigen specificity and microenvironmental regulation. Front Immunol. 2020;11:1371.

46. Gabrilovich DI. Myeloid-derived suppressor cells. Cancer Immunol Res. 2017;5(1):3–8.

47. Bronte V, Brandau S, Chen S-H, Colombo MP, Frey AB, Greten TF, et al. Recommendations for myeloid-derived suppressor cell nomenclature and characterization standards. Nat Commun. 2016;7(1):12150.

48. Talmadge JE, Gabrilovich DI. History of myeloid-derived suppressor cells. Nat Rev Cancer. 2013;13(10):739–52.

49. Youn J, Gabrilovich DI. The biology of myeloid-derived suppressor cells: the blessing and the curse of morphological and functional heterogeneity. Eur J Immunol. 2010;40(11):2969–75.

50. Kumar V, Patel S, Tcyganov E, Gabrilovich DI. The nature of myeloid-derived suppressor cells in the tumor microenvironment. Trends Immunol. 2016;37(3):208–20.

51. Pyzer AR, Cole L, Rosenblatt J, Avigan DE. Myeloid-derived suppressor cells as effectors of immune suppression in cancer. Int J cancer. 2016;139(9):1915–26.

52. Goulart MR, Pluhar GE, Ohlfest JR. Identification of myeloid derived suppressor cells in dogs with naturally occurring cancer. PLoS One. 2012;7(3):e33274.

53. Goulart MR, Hlavaty SI, Chang Y-M, Polton G, Stell A, Perry J, et al. Phenotypic and transcriptomic characterization of canine myeloid-derived suppressor cells. Sci Rep. 2019;9(1):3574.

54. Buurman WA, Vegt PA, Groenewegen G, Linden C, Jeunhomme G. Analysis of buoyant density of canine peripheral blood leukocytes with PVP-Silica (Percoll) density gradients. Vet Immunol Immunopathol. 1982;3(6):547–56.

55. Bruger AM, Vanhaver C, Bruderek K, Amodio G, Tavukçuoğlu E, Esendagli G, et al. Protocol to assess the suppression of T-cell proliferation by human MDSC. In: Methods in enzymology. Elsevier; 2020. p. 155–92.

56. Chikamatsu K, Sakakura K, Toyoda M, Takahashi K, Yamamoto T, Masuyama K. Immunosuppressive activity of CD 14+ HLA-DR− cells in squamous cell carcinoma of the head and neck. Cancer Sci. 2012;103(6):976–83.

57. Hart KM, Byrne KT, Molloy MJ, Usherwood EM, Berwin B. IL-10 immunomodulation of myeloid cells regulates a murine model of ovarian cancer. Front Immunol. 2011;2:29.

58. Chen W, Jin W, Hardegen N, Lei K, Li L, Marinos N, et al. Conversion of peripheral CD4+ CD25− naive T cells to CD4+ CD25+ regulatory T cells by TGF-β induction of transcription factor Foxp3. J Exp Med. 2003;198(12):1875–86.

59. Grütz G. New insights into the molecular mechanism of interleukin-10-mediated immunosuppression. J Leukoc Biol. 2005;77(1):3–15.

60. Lu C, Redd PS, Lee JR, Savage N, Liu K. The expression profiles and regulation of PD-L1 in tumor-induced myeloid-derived suppressor cells. Oncoimmunology. 2016;5(12):e1247135.

61. Dosset M, Vargas TR, Lagrange A, Boidot R, Végran F, Roussey A, et al. PD-1/PD-L1 pathway: an adaptive immune resistance mechanism to immunogenic chemotherapy in colorectal cancer. Oncoimmunology. 2018;7(6):e1433981.

62. Munder M. Arginase: an emerging key player in the mammalian immune system. Br J Pharmacol. 2009;158(3):638–51.

63. Liu Y, Yu Y, Yang S, Zeng B, Zhang Z, Jiao G, et al. Regulation of arginase I activity and expression by both PD-1 and CTLA-4 on the myeloid-derived suppressor cells. Cancer Immunol Immunother. 2009;58:687–97.

64. Chen X, Song M, Zhang B, Zhang Y. Reactive oxygen species regulate T cell immune response in the tumor microenvironment. Oxid Med Cell Longev. 2016;2016.

65. Li J, Wang L, Chen X, Li L, Li Y, Ping Y, et al. CD39/CD73 upregulation on myeloid-derived suppressor cells via TGF-β-mTOR-HIF-1 signaling in patients with non-small cell lung cancer. Oncoimmunology. 2017;6(6):e1320011.

66. Xue J, Schmidt S V., Sander J, Draffehn A, Krebs W, Quester I, et al. Transcriptome-Based Network Analysis Reveals a Spectrum Model of Human Macrophage Activation. Immunity. 2014 Feb 20;40(2):274–88.

67. Ma R-Y, Black A, Qian B-Z. Macrophage diversity in cancer revisited in the era of single-cell omics. Trends Immunol. 2022;

68. Zhang H, Liu Z, Wen H, Guo Y, Xu F, Zhu Q, et al. Immunosuppressive TREM2 (+) macrophages are associated with undesirable prognosis and responses to anti-PD-1 immunotherapy in non-small cell lung cancer. Cancer Immunol Immunother. 2022;71(10):2511–22.

69. Timperi E, Gueguen P, Molgora M, Magagna I, Kieffer Y, Lopez-Lastra S, et al. Lipid-associated macrophages are induced by cancer-associated fibroblasts and mediate immune suppression in breast cancer. Cancer Res. 2022;82(18):3291–306.

70. Binnewies M, Pollack JL, Rudolph J, Dash S, Abushawish M, Lee T, et al. Targeting TREM2 on tumor-associated macrophages enhances immunotherapy. Cell Rep. 2021;37(3):109844.

71. Park MD, Reyes-Torres I, LeBerichel J, Hamon P, LaMarche NM, Hegde S, et al. TREM2 macrophages drive NK cell paucity and dysfunction in lung cancer. Nat Immunol. 2023;1–10.

72. Kong D, Zhou H, Neelakantan D, Hughes CJ, Hsu JY, Srinivasan RR, et al. VEGF-C mediates tumor growth and metastasis through promoting EMT-epithelial breast cancer cell crosstalk. Oncogene. 2021;40(5):964–79.

73. Cheng S, Li Z, Gao R, Xing B, Gao Y, Yang Y, et al. A pan-cancer single-cell transcriptional atlas of tumor-infiltrating myeloid cells. Cell. 2021;184(3):792–809.

74. Ekblom R, Galindo J. Applications of next generation sequencing in molecular ecology of non-model organisms. Heredity (Edinb). 2011;107(1):1–15.

75. Coffin JM, Fan H. The discovery of reverse transcriptase. Annu Rev Virol. 2016;3:29–51.

76. Hu T, Chitnis N, Monos D, Dinh A. Next-generation sequencing technologies: An overview. Hum Immunol. 2021;82(11):801–11.

77. Head SR, Komori HK, LaMere SA, Whisenant T, Van Nieuwerburgh F, Salomon DR, et al. Library construction for next-generation sequencing: overviews and challenges. Biotechniques. 2014;56(2):61–77.

78. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, et al. mRNA-Seq whole-transcriptome analysis of a single cell. Nat Methods. 2009;6(5):377–82.

79. Jovic D, Liang X, Zeng H, Lin L, Xu F, Luo Y. Single-cell RNA sequencing technologies and applications: A brief overview. Clin Transl Med. 2022;12(3):e694.

80. Lee J, Hyeon DY, Hwang D. Single-cell multiomics: technologies and data analysis methods. Exp Mol Med. 2020;52(9):1428–42.

81. Brüning RS, Tombor L, Schulz MH, Dimmeler S, John D. Comparative analysis of common alignment tools for single-cell RNA sequencing. Gigascience. 2022;11.

82. Jiang R, Sun T, Song D, Li JJ. Statistics or biology: the zero-inflation controversy about scRNA-seq data. Genome Biol. 2022;23(1):1–24.

83. Lytal N, Ran D, An L. Normalization methods on single-cell RNA-seq data: an empirical survey. Front Genet. 2020;11:41.

84. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol. 2018;36(5):411–20.

85. L Lun AT, Bach K, Marioni JC. Pooling across cells to normalize single-cell RNA sequencing data with many zero counts. Genome Biol. 2016;17(1):1–14.

86. Ilicic T, Kim JK, Kolodziejczyk AA, Bagger FO, McCarthy DJ, Marioni JC, et al. Classification of low quality cells from single-cell RNA-seq data. Genome Biol. 2016;17(1):1–15.

87. Xi NM, Li JJ. Benchmarking computational doublet-detection methods for single-cell RNA sequencing data. Cell Syst. 2021;12(2):176–94.

88. Young MD, Behjati S. SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. Gigascience. 2020;9(12):giaa151.

89. Yang P, Huang H, Liu C. Feature selection revisited in the single-cell era. Genome Biol. 2021;22:1–17.

90. Yip SH, Sham PC, Wang J. Evaluation of tools for highly variable gene discovery from single-cell RNA-seq data. Brief Bioinform. 2019;20(4):1583–9.

91. McInnes L, Healy J, Melville J. Umap: Uniform manifold approximation and projection for dimension reduction. arXiv Prepr arXiv180203426. 2018;

92. Van der Maaten L, Hinton G. Visualizing data using t-SNE. J Mach Learn Res. 2008;9(11).

93. Luecken MD, Büttner M, Chaichoompu K, Danese A, Interlandi M, Müller MF, et al. Benchmarking atlas-level data integration in single-cell genomics. Nat Methods. 2022;19(1):41–50.

94. Ryu Y, Han GH, Jung E, Hwang D. Integration of single-cell RNA-seq datasets: a review of computational methods. Mol Cells. 2023;46(2):106.

95. Pasquini G, Arias JER, Schäfer P, Busskamp V. Automated methods for cell type annotation on scRNA-seq data. Comput Struct Biotechnol J. 2021;19:961–9.

96. Franzén O, Gan L-M, Björkegren JLM. PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. Database. 2019;2019.

97. 9 C-UTMCCL lcai@ caltech. edu 21 b SJ 9 TC 9 LS shinlin@ uw. edu 2 e JD, 8 UTMCZK kzhang@ bioeng. ucsd. edu 15 b SX 15 JS 24 HJ 25 PG 26 KP, 35 CI of TTTDCL lcai@ caltech. edu 21 b YG-C 35 ZQ 35 DR, 37 HTTDYP peng_yin@ hms. harvard. edu 36 37 b SSK 36 37 KJY 36 37 WY 36 37 GI 36, 39 PTTDLJ jlaskin@ purdue. edu 10 b YD 10 38 B-JKE 39 PPD 39 AC 39 ZY, 40 STTDHP harbury@ stanford. edu 11 b DT 40 MJ 11 CP 11 NM. The human body at cellular resolution: the NIH Human Biomolecular Atlas Program. Nature. 2019;574(7777):187–92.

98. Stoeckius M, Hafemeister C, Stephenson W, Houck-Loomis B, Chattopadhyay PK, Swerdlow H, et al. Simultaneous epitope and transcriptome measurement in single cells. Nat Methods. 2017;14(9):865–8.

99. Squair JW, Gautier M, Kathe C, Anderson MA, James ND, Hutson TH, et al. Confronting false discoveries in single-cell differential expression. Nat Commun. 2021;12(1):1–15.

100. Zimmerman KD, Espeland MA, Langefeld CD. A practical solution to pseudoreplication bias in single-cell studies. Nat Commun. 2021;12(1):738.

101. Gao R, Bai S, Henderson YC, Lin Y, Schalck A, Yan Y, et al. Delineating copy number and clonal substructure in human tumors from single-cell transcriptomes. Nat Biotechnol. 2021;39(5):599–608.

102. Puram S V, Tirosh I, Parikh AS, Patel AP, Yizhak K, Gillespie S, et al. Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. Cell. 2017;171(7):1611–24.

103. Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science (80-). 2014;344(6190):1396–401.

104. La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, et al. RNA velocity of single cells. Nature. 2018;560(7719):494–8.

105. Saelens W, Cannoodt R, Todorov H, Saeys Y. A comparison of single-cell trajectory inference methods. Nat Biotechnol. 2019;37(5):547–54.

106. Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. Nat Protoc. 2020;15(4):1484–506.

107. Dimitrov D, Türei D, Garrido-Rodriguez M, Burmedi PL, Nagai JS, Boys C, et al. Comparison of methods and resources for cell-cell communication inference from single-cell RNA-Seq data. Nat Commun. 2022;13(1):3224.

108. Browaeys R, Saelens W, Saeys Y. NicheNet: modeling intercellular communication by linking ligands to target genes. Nat Methods. 2020;17(2):159–62.

109. Gorin G, Pachter L. Length biases in single-cell RNA sequencing of pre-mRNA. Biophys Reports. 2023;3(1).

110. Zheng GXY, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, et al. Massively parallel digital transcriptional profiling of single cells. Nat Commun. 2017;8(1):14049.

111. Qiu P. Embracing the dropouts in single-cell RNA-seq analysis. Nat Commun. 2020;11(1):1169.

112. Baran-Gale J, Chandra T, Kirschner K. Experimental design for single-cell RNA sequencing. Brief Funct Genomics. 2018;17(4):233–9.

113. Fu R, Riemondy KA, Sheridan RM, Hesselberth JR, Jordan CT, Gillen AE. scraps: an end-to-end pipeline for measuring alternative polyadenylation at high resolution using single-cell RNA-seq. bioRxiv. 2022;2008–22.

114. Derr A, Yang C, Zilionis R, Sergushichev A, Blodgett DM, Redick S, et al. End Sequence Analysis Toolkit (ESAT) expands the extractable information from single-cell RNA-seq data. Genome Res. 2016;26(10):1397–410.

115. Pombo Antunes AR, Scheyltjens I, Lodi F, Messiaen J, Antoranz A, Duerinck J, et al. Single-cell profiling of myeloid cells in glioblastoma across species and disease stage reveals macrophage competition and specialization. Nat Neurosci. 2021;24(4):595–610.

116. Li J, Zhou J, Huang H, Jiang J, Zhang T, Ni C. Mature dendritic cells enriched in immunoregulatory molecules (mregDCs): A novel population in the tumour microenvironment and immunotherapy target. Clin Transl Med. 2023;13(2):e1199.

117. Liu Y, Feng W, Dai Y, Bao M, Yuan Z, He M, et al. Single-cell transcriptomics reveals the complexity of the tumor microenvironment of treatment-naive osteosarcoma. Front Oncol. 2021;11:709210.

118. Zhou Y, Yang D, Yang Q, Lv X, Huang W, Zhou Z, et al. Single-cell RNA landscape of intratumoral heterogeneity and immunosuppressive microenvironment in advanced osteosarcoma. Nat Commun. 2020;11(1):6322.

119. Liu W, Hu H, Shao Z, Lv X, Zhang Z, Deng X, et al. Characterizing the tumor microenvironment at the single-cell level reveals a novel immune evasion mechanism in osteosarcoma. Bone Res. 2023;11(1):4.

120. See P, Dutertre C-A, Chen J, Günther P, McGovern N, Irac SE, et al. Mapping the human DC lineage through the integration of high-dimensional techniques. Science (80-). 2017;356(6342):eaag3009.

121. Zhang L, Li Z, Skrzypczynska KM, Fang Q, Zhang W, O'Brien SA, et al. Single-cell analyses inform mechanisms of myeloid-targeted therapies in colon cancer. Cell. 2020;181(2):442–59.

122. Jackson K, Milner RJ, Doty A, Hutchison S, Cortes-Hinojosa G, Riva A, et al. Analysis of canine myeloid-derived suppressor cells (MDSCs) utilizing fluorescence-activated cell sorting, RNA protection mediums to yield quality RNA for single-cell RNA sequencing. Vet Immunol Immunopathol. 2021;231:110144.

123. Wigerblad G, et al. Single-cell analysis reveals the range of transcriptional states of circulating human neutrophils. J Immunol (2022) 209:772–82. doi: 10.4049/jimmunol.2200154

124. Li Y, Gao J, Kamran M, Harmacek L, Danhorn T, Leach SM, et al. GATA2 regulates mast cell identity and responsiveness to antigenic stimulation by promoting chromatin remodeling at super-enhancers. Nat Commun. 2021;12(1):494.

125. Ohmori S, Ishijima Y, Numata S, Takahashi M, Sekita M, Sato T, et al. GATA2 and PU. 1 collaborate to activate the expression of the mouse Ms4a2 gene, encoding FcεRIβ, through distinct mechanisms. Mol Cell Biol. 2019;39(22):e00314-19.

126. Fastrès A, Pirottin D, Fievez L, Marichal T, Desmet CJ, Bureau F, et al. Characterization of the bronchoalveolar lavage fluid by single cell gene expression analysis in healthy dogs: a promising technique. Front Immunol. 2020;11:1707.

127. Fastrès A, Pirottin D, Fievez L, Tutunaru A-C, Bolen G, Merveille A-C, et al. Identification of pro-fibrotic macrophage populations by single-cell transcriptomic analysis in West

highland white terriers affected with canine idiopathic pulmonary fibrosis. Front Immunol. 2020;11:611749.

128. Gingrich AA, Reiter TE, Judge SJ, York D, Yanagisawa M, Razmara A, et al. Comparative immunogenomics of canine natural killer cells as immunotherapy target. Front Immunol. 2021;12:670309.

129. Chen D, Sun J, Zhu J, Ding X, Lan T, Wang X, et al. Single cell atlas for 11 non-model mammals, reptiles and birds. Nat Commun. 2021;12(1):7083.

130. Li Z, Sun C, Wang F, Wang X, Zhu J, Luo L, et al. Molecular mechanisms governing circulating immune cell heterogeneity across different species revealed by single-cell sequencing. Clin Transl Med. 2022;12(1):e689.

131. Eschke M, Moore PF, Chang H, Alber G, Keller SM. Canine peripheral blood TCRαβ T cell atlas: Identification of diverse subsets including CD8A+ MAIT-like cells by combined single-cell transcriptome and V (D) J repertoire analysis. Front Immunol. 2023;14.

132. Skidmore ZL, Rindt H, Chu S, Fisk B, Fronick C, Fulton R, et al. Single Cell T Cell Receptor Repertoire Profiling for Dogs. bioRxiv. 2021;2006–21.

133. Patel RS, Tomlinson JE, Divers TJ, Van de Walle GR, Rosenberg BR. Single-cell resolution landscape of equine peripheral blood mononuclear cells reveals diverse cell types including T-bet+ B cells. BMC Biol. 2021;19:1–18.

134. Ayers J, Milner RJ, Cortés-Hinojosa G, Riva A, Bechtel S, Sahay B, et al. Novel application of single-cell next-generation sequencing for determination of intratumoral heterogeneity of canine osteosarcoma cell lines. J Vet Diagnostic Investig. 2021;33(2):261–78.

135. Sparling BA, Moss N, Kaur G, Clark D, Hawkins RD, Drechsler Y. Unique Cell Subpopulations and Disease Progression Markers in Canines with Atopic Dermatitis. J Immunol. 2022;209(7):1379–88.

136. Früh SP, Saikia M, Eule J, Mazulis CA, Miller JE, Cowulich JM, et al. Elevated circulating Th2 but not group 2 innate lymphoid cell responses characterize canine atopic dermatitis. Vet Immunol Immunopathol. 2020;221:110015.

137. Zhou Q-J, Liu X, Zhang L, Wang R, Yin T, Li X, et al. A single-nucleus transcriptomic atlas of the dog hippocampus reveals the potential relationship between specific cell types and domestication. Natl Sci Rev. 2022;9(11):nwac147.

CHAPTER 2: A single-cell RNA sequencing atlas of circulating leukocytes from healthy and osteosarcoma affected dogs[2]

Summary

Translationally relevant animal models are essential for the successful bridging of basic science findings into clinical medicine. While rodent models are widely accessible, there are numerous limitations that prevent the extrapolation of findings to human medicine. Dogs represent an alternative model which spontaneously develop diseases, recapitulate the natural progression of human disease, and live in similar environments alongside humans. Despite the value of the canine model, species specific reagent limitations have hampered in depth characterization of canine immune cells, which constrains the conclusions that can be drawn from canine immunotherapy studies. To address this need, we used single-cell RNA sequencing to characterize the heterogeneity of circulating leukocytes in healthy dogs (n = 7) and osteosarcoma (OS) affected dogs (n = 10). We present a cellular atlas of leukocytes in healthy dogs, then employ the dataset to investigate the impact of primary OS tumors on the transcriptome of circulating leukocytes. We identified 36 unique cell populations amongst dog circulating leukocytes, and then generated gene signatures for each cell type. In our comparison of healthy dogs and dogs with OS, we identified relative increases in the abundances of polymorphonuclear (PMN-) and monocytic (M-) myeloid-derived suppressor cells (MDSCs), as well as aberrations in gene expression within myeloid cells. Furthermore, our data suggest LTF, CAMP, and S100A12 may represent valuable biomarkers for evaluating MDSC burden in dogs. Overall, this study provides a detailed atlas of canine leukocytes and investigates how the presence of osteosarcoma alters the transcriptional profiles of circulating immune cells.

[2]Ammons DT, Harris RA, Hopkins LS, Kurihara J, Weishaar K and Dow S (2023) A single-cell RNA sequencing atlas of circulating leukocytes from healthy and osteosarcoma affected dogs. Front. Immunol. 14:1162700. https://doi.org/10.3389/fimmu.2023.1162700

Introduction

Traditional animal models, such as rodents, have been used for decades as a steppingstone in understanding human disease. Unfortunately, due to environmental and genetic disparities between humans and model species, research findings frequently fail to translate to human medicine (1). While these animal models are valuable in preclinical research, there has been a push to incorporate the use of more biologically relevant models to facilitate the translation of key research findings to human clinical trials (2). Client owned dogs offer an excellent model to investigate novel therapeutics as they are outbred, spontaneously develop disease, and share living space with humans. Interestingly, certain cancers, such as osteosarcoma (OS), follow similar disease progression and occur at a rate roughly 75 times that of humans (3). Therefore, clinical trials in dogs allow for thorough investigation of novel therapeutics for the treatment of diseases considered to be rare and difficult to study in humans. With the increasing use of canine models to investigate novel therapeutics, there is a need to better describe canine immune cell populations.

Previous characterizations of canine immune cells have relied on antibody-based assays, such as flow cytometry, cell sorting, and immunohistochemistry, to describe cell populations (4–6). While these approaches have been fundamental in understanding canine immunology, the limited availability of markers to define cell populations has the tendency to introduce pre-selection bias. An alternative and novel approach to describe the heterogeneity of canine immune cells is to use single-cell RNA sequencing which enables the unsupervised characterization of individual immune cell transcriptomes. The flexibility of this platform allows researchers to overcome species-specific reagent barriers that have limited previous characterizations of canine immune cells.

Circulating immune cells play a key role in responding to disease and have been reported to be altered in cancer patients, with increased immune suppression being a predominate finding

(7). For example, two recently defined cell populations, polymorphonuclear (PMN-) and monocytic (M-) myeloid-derived suppressor cells (MDSCs), have been reported to be expanded in individuals with cancer, and increases in circulating MDSCs have been demonstrated to have prognostic implications (8). These two distinct MDSC populations exhibit immature phenotypes with potent immune suppressive activity. Recently, canine MDSCs were characterized using flow cytometry-based assays and were defined to be MHCII negative monocytes and neutrophils isolated through density centrifugation (9,10). Given the limited antibody selection for flow cytometry-based approaches there is a need to evaluate heterogeneity within MDSC populations, as well as investigate how all leukocyte transcriptomes are altered in dogs with cancer.

In the present study, we first complete single-cell transcriptomic analysis on seven middle-aged healthy dogs of varying breeds and genders to develop a comprehensive reference database for canine leukocytes. Then we apply the database to investigate how leukocyte abundances and transcriptomes are altered in ten dogs with spontaneously arising osteosarcoma. In depth analysis revealed key changes in immune cell abundances as well as OS induced transcriptomic changes. The data presented here highlight the differences in circulating leukocytes between healthy and osteosarcoma affected dogs, as well as provide an open-access annotated canine leukocyte database for future investigations.

Methods

Patient selection

Clinically healthy client-owned dogs without preexisting conditions were selected for inclusion in the healthy reference database. All dogs were followed for at least two months to ensure no disease was observed. Osteosarcoma patients were selected based on the presence of a primary tumor and all dogs were naïve to treatment. All dogs presented with radiographic evidence of OS and underwent amputation in the weeks following blood collection. All studies

were approved by the Colorado State University (CSU) Institutional Animal Care and Use Committee and the CSU Clinical Review Board. All dog owners provided informed consent prior to blood donation.

Sample preparation

Approximately 10 mL of whole blood was obtained in an EDTA collection tube then processed within 30 minutes using Ficoll Paque (Cytiva; Marlborough, MA) to complete density gradient centrifugation. Whole blood was mixed with 12 mL of phosphate buffered saline, pH = 7.40 (PBS), then layered onto 12 mL of Ficoll Paque. To isolate leukocytes, the layered sample was then centrifugated at room temperature for 40 minutes at 500 rcf with acceleration at maximum and brake off. Leukocytes were isolated through collection of the cell interface then washed one time with PBS, resuspended in 10 mL of Ammonium-Chloride-Potassium (ACK) lysis buffer for 3-7 minutes at room temperature, then washed an additional time with PBS. A final wash at 100 rcf x 15 minutes was completed to remove platelets and other small contaminates. Finally, cells were resuspended in 0.04% molecular grade BSA (Sigma-Aldrich; St. Louis, MO) in PBS and adjusted to obtain a cell count between 700-1200 cells/μL. Once in solution, leukocytes were transported to a chromium instrument (10x Genomics; Pleasanton, CA) and captured within 2 hours of preparation.

Library preparation and sequencing

Single cells were isolated and tagged with unique cell barcodes using a Chromium controller or Chromium iX instrument (10x Genomics) with a target of 5,000 cells per sample. Single cells were isolated and processed using a Chromium Next GEM Single Cell 3' Kit v3.1 following manufacture recommended protocols. Once cells were barcoded and unique molecular identifiers (UMIs) added, a standard Illumina library preparation was completed using a dual index library construction kit (10x Genomics). Sample quality was analyzed using a TapeStation bioanalyzer and/or LabChip (Agilent Technologies; Santa Clara, CA/PerkinElmer; Waltham, MA)

then submitted for sequencing on an Illumina NovaSeq 6000 sequencer (Novogene Corporation; Sacramento, CA) with a target of 50,000 150 bp paired-end reads per cell. Raw data was demultiplexed by the sequencing core then transferred for downstream analysis.

Read mapping and quantification

A Cell Ranger analysis pipeline (version 7.0.0, 10x Genomics) was utilized to process raw FASTQ sequencing data, align reads to the canine genome, and generate a count matrix. First, the CanFam3.1 Ensembl annotation (gtf) file was filtered for protein_coding, lincRNA, antisense, and immunoglobulin gene biotypes. Then a canine reference package using the filtered gtf file and CanFam3.1 genome (FASTA) was created using the cellranger mkref command. The reference package and raw FASTQ files were then used to complete read mapping and quantification of UMIs using the cellranger count command. Each sample was aligned once using include-introns mode set to true and once with include-introns mode set to false. We observed that include-introns mode increased the sensitivity of the assay, but also increase the abundance of low-quality cell clusters. By using both methods we were then better able to filter out artifactual clusters while retaining clusters with low transcript abundance. Due to incomplete annotation of the canine genome, we also aligned the healthy dog samples to an alternate canine genome (ROS_Cfam_1.0) using Cell Ranger version 7.0.0 with include-introns mode set to true. The output count matrices obtained under each alignment protocol in the format of cell barcode x feature (column x row) were then exported and used for downstream analysis.

Data filtering and integration

For each sample, the count matrix was imported into R using the Read10X() function then converted to a Seurat object using the CreateSeuratObject() function (11). To estimate the number of dead/poor quality cells, the percent mitochondrial reads per cell was calculated using PercentageFeatureSet() to count all reads mapped to features with the prefix "MT-". Each object was then filtered to only retain cells which met the following requirements: 200 < nFeature_RNA < 4500, percent.mt < 10, and 500 < nCount_RNA < 20000. An initial low-resolution unsupervised

clustering was completed to remove contaminating red blood cells and platelets. Next, DoubletFinder, was used to identify and remove putative cell doublets (12). After filtering each sample individually, all samples were integrated into one object using a SCTransform workflow (13). During this step, we regressed out the percent mitochondrial reads and percent platelet-associated reads to minimize the impacts on clustering and used 2000 features as integration anchors. Following data integration, ideal clustering parameters of each subset of data were determined using the R package clustree (14). Dimension reduction and visualization was then completed, and the data were projected using 2-dimensional, non-linear uniform manifold approximation and projection (UMAP) plots.

Cell classification

Cells were classified using the integrated dataset containing 7 healthy and 10 OS dogs. Annotations were then transferred to the healthy only dataset to create gene lists for each cell population. Major cell population identities were assigned using compiled data gathered from singleR reference mapping, Seurat reference mapping, gene set enrichment analysis, and manual annotation based on the literature (15,16). Further high-resolution cell identification was completed through use of independent reclustering on cells within each major immune cell population. Confirmation of T cell subsets was completed using the AddModuleScore() function and gene lists provided in a comprehensive human T cell dataset (17). The gene lists presented as supplemental data were generated using the FindAllMarkers() function (Wilcoxon Rank Sum test) on final cell type classification in healthy only (Supplemental data 1) and combined datasets (Supplemental data 2). The final short gene lists were generated using a selection of the top features that define a cell type (as defined in the healthy only dataset) with preference given to unique features in the top 50 genes (weighted by adjusted P value) identified to define each cell population.

Feature visualization

Feature expression was visualized using dot plots, feature plots, and violin plots. Selected features were chosen based on prior biological knowledge and features identified using the FindAllMarkers() function. Y-axis scales for violin plots within a figure are fixed, so they are all on the same scale. Feature plots show normalized expression for each feature and are on variable scales. Dot plots used scaled expression data which depicts deviation from the average value for a gene across the cells being sampled. Visualization of T Cell Receptor Delta Constant (TRDC) was completed by transferring normalized expression counts from the healthy samples aligned to ROS_Cfam_1.0 to the CanFam3.1 aligned healthy samples.

Cell abundance analysis

All cell abundance comparisons were made using percentage of parent. When visualizing cellular contribution on a UMAP, all samples were down sampled to the value equal to the sample which contributed the fewest number of cells. By doing this, we obtained equal number of cells from each sample and avoided visual biasing of data presentation. To make statistical inferences on changes in cell abundances, two-sided Wilcoxon Rank Sum tests were used, and exact P values were reported (18). P values less than 0.05 were considered to be statistically significant.

Pseudotime analysis

To complete pseudotime analysis and predict cell lineages, we used Slingshot (19). Only high-quality cell clusters were used for this analysis. A starting node was selected based on biological knowledge, then an unsupervised approach was used to infer branches between clusters. Following branch identification, the Slingshot getLineages() function was used to identify predicted cell lineage pathways. Custom functions were then used to extract the data and plot the branch patterns/lineages.

Differential gene expression analysis

Differential gene expression analysis was completed using two different approaches (1) Wilcoxon Rank Sum test (FindMarkers() function) and (2) pseudobulk conversion. When possible,

we used pseudobulk conversion followed by a DESeq2 pipeline to evaluate differential gene expression (20). We required a minimum of 25 cells in a sample to be included in pseudobulk conversion and only applied this approach to compare cells within clusters or between groups of cells that had limited heterogeneity. Prior to running DESeq2, low abundance features, defined as features with less than 10 raw counts across all cells sampled, were filtered out. Features that had an adjusted P value of less than 0.01 and a log2(fold change) greater than 0.58 were considered to be statistically significant. For differential gene expression analysis completed using FindMarkers(), we obtained the average normalized count of each feature grouped by classification (Y versus X), then plotted the values on a scatter plot. Values that fell below a line of y-intercept = 0; slope = 1 and were determined to be higher in the X category using the FindMarkers() function (log2(fold change) < -0.58 and adjusted P value < 0.01), were discussed as increased in X or decreased in Y. Alternatively, values that fell above a line of y-intercept = 0; slope = 1 and were determined to be higher in the Y category using the FindMarkers() function (log2(fold change) > 0.58 and adjusted P value < 0.01), were discussed as increased in Y or decreased in X. Any subsequent pathway analysis was completed using lists of upregulated or downregulated genes and the enricher() function from clusterProfiler was used with the hallmarks gene sets (21,22).

Human-canine homology analysis

An annotated human leukocyte dataset (blish_covid.seu.rds) consisting of 6 healthy adults was obtained from https://zenodo.org/record/4021967/ and integrated with the 7 healthy dog samples generated in this study (18). SCT normalization and integration of the merged canine and human datasets was completed using 2000 variable features, while also regressing the percentage of mitochondrial reads. Following integration, the canine and human cell type annotations were prepended with “can_” or “hu_” then SCT normalized data was used to complete hierarchical clustering of the prepended cell types, an approach adopted from Cheng et al. (23).

Hierarchical clustering was completed using the hclust() function with using complete Euclidean distance to complete the analysis.

Flow cytometric analysis

A subset of cells obtained from the samples used for single-cell RNA sequencing were used for paired flow cytometric analysis of immune cell types. Approximately 500,000 cells were plated and used for immunolabeling. Three antibody panels were used per sample (Supplemental table 2.1). Cells were blocked with 5 uL of normal dog serum (panel 1 & 2) or 5 uL of normal donkey serum (panel 3) for 15 minutes on ice (Jackson Laboratory; Bar Harbor, ME). Primary antibodies diluted in FACS buffer (5% FBS plus 0.1% sodium azide in PBS) were added for 30 minutes (only mouse anti-CADO48a for panel 3), then for panel 3 a donkey anti-mouse secondary antibody (30 minutes) followed by strepdavidin-Qdot800 (15 minutes) and directly conjugated antibodies were added (30 minutes) on ice. Cells were washed twice with FACS between each labeling step. After completion of immunolabeling, 5 uL of 7-aad were added to each sample then run on a Beckman Coulter Gallios 3-laser flow cytometer. A total of 150,000 events were targeted during data acquisition.

Data and software availability

Raw sequencing data (FASTQ format) is available on the NCBI Gene Expression Omnibus database under the accession number GSE225599. In addition to the raw data, processed Seurat objects are provided as supplemental files in the NCBI Gene Expression Omnibus database. The processed dataset is available for browsing at UCSC Cell Browser (https://canine-leukocyte-atlas.cells.ucsc.edu) (24). An associated GitHub page containing all the code and software versions used to analyze the data presented in this manuscript is available at https://github.com/dyammons/Canine_Leukocyte_scRNA (https://zenodo.org/record/7884518). Any additional data requests can be made by contacting the corresponding author or through the GitHub repository.

Results

Establishment of a healthy canine leukocyte reference database

The first objective of this study was to establish a comprehensive canine reference leukocyte database that can be used to further define immune cell transcriptomes and be available for use by other research groups. To establish the reference dataset, we obtained a total of 32,028 cells from 3 male and 4 female middle-aged, clinically healthy dogs (Table 2.1).

Table 2.1. Dog demographics.

Group	ID	Sex	Age	Breed
Healthy	H_1	FS	8.3	Mixed
	H_2	MC	8.4	Australian Shepherd
	H_3	FS	6.9	Standard Poodle
	H_4	MC	7.4	Beagle
	H_5	FS	8.8	Mixed
	H_6	FS	7.7	Labrador Retriever
	H_7	MC	7.7	Newfoundland
Osteosarcoma	OS_1	MC	10	Labrador Retriever
	OS_2	FS	9	Saint Bernard
	OS_3	FS	6.8	Great Dane
	OS_4	MC	11.5	Catahoula
	OS_5	MC	12.1	Rottweiler
	OS_6	MC	6.8	German Shepherd
	OS_7	FS	7.3	Great Dane
	OS_8	FS	5.7	Staffordshire Terrier mix
	OS_9	MI	7.8	Bernese Mountain Dog/Great Pyrenees
	OS_10	MI	10.4	Saint Bernard

The average number of cells collected per dog was 4,575 and on average each cell was sequenced to a depth of 60,686 reads per cell.

In total, 42 unique cell clusters were identified with major immune cell populations clustering in distinct regions of the uniform manifold approximation and projection (UMAP) plot (Figure 2.1a). Cell identities were assigned based on feature plots using stereotypic markers reported in the literature (Figure 2.1b/c) (16). To further support our classifications, we used reference mapping to human databases (Supplemental figure 2.1a-e). Unexpectedly, cell

classifications based on human references were highly variable and largely ineffective at assigning cell identities, especially to CD8 T cell, NK cell, and myeloid cell populations. This variability is likely the result of incomplete canine genome annotation and distinct cellular transcriptomes between species. Following cell classification, we evaluated the relative contribution of each sample to every cluster (Figure 2.1d/e; Supplemental figure 2.2a). This analysis revealed that the dataset was well integrated without overt batch effects, and most dogs contributed equally to each cluster. The exceptions were within two neutrophil clusters, Clusters 4 and 9, which were largely composed of cells from healthy dog 7. In mice and humans, neutrophils are reported to have a cellular density that is too dense to be isolated in the polymorphonuclear cell (PBMC) layer when completing density centrifugation (25). However, consistent with previous reports, we present evidence that a sizeable population of canine neutrophils have a density that allows them to be collected with PBMC isolation (26). Additionally, there is evidence of marked inter-sample variability in the number of neutrophils isolated through density centrifugation, which can alter overall cellular proportions in a given sample.

Although UMAP plots do not perfectly reconstruct the relative spatial relationship between cells, we noted that myeloid cells largely clustered together, while T cells and B cells clustered in different regions of the plot, indicating the data clustered in a biologically relevant manner. Interestingly, there was one rare myeloid cell cluster (Cluster 38) which clustered in a distant location from the other neutrophils and lacked CD4 expression. Additionally, one T cell population (Cluster 29) was plotted in a distant location relative to other T cells; this population was identified as a gd (γδ) T cell population based on GATA3 and TRDC expression (Figure 2.1b, Supplemental figure 2.1f). Another interesting observation was that the CD8 T cell populations clustered in two separate regions of the UMAP, with naïve CD8 T cells appearing to be more similar to naïve CD4 T cells than effector CD8 T cells. This clustering of T cells is likely a result of naïve CD8 T cells lacking cytotoxic properties that arise following interaction with their cognate antigen.

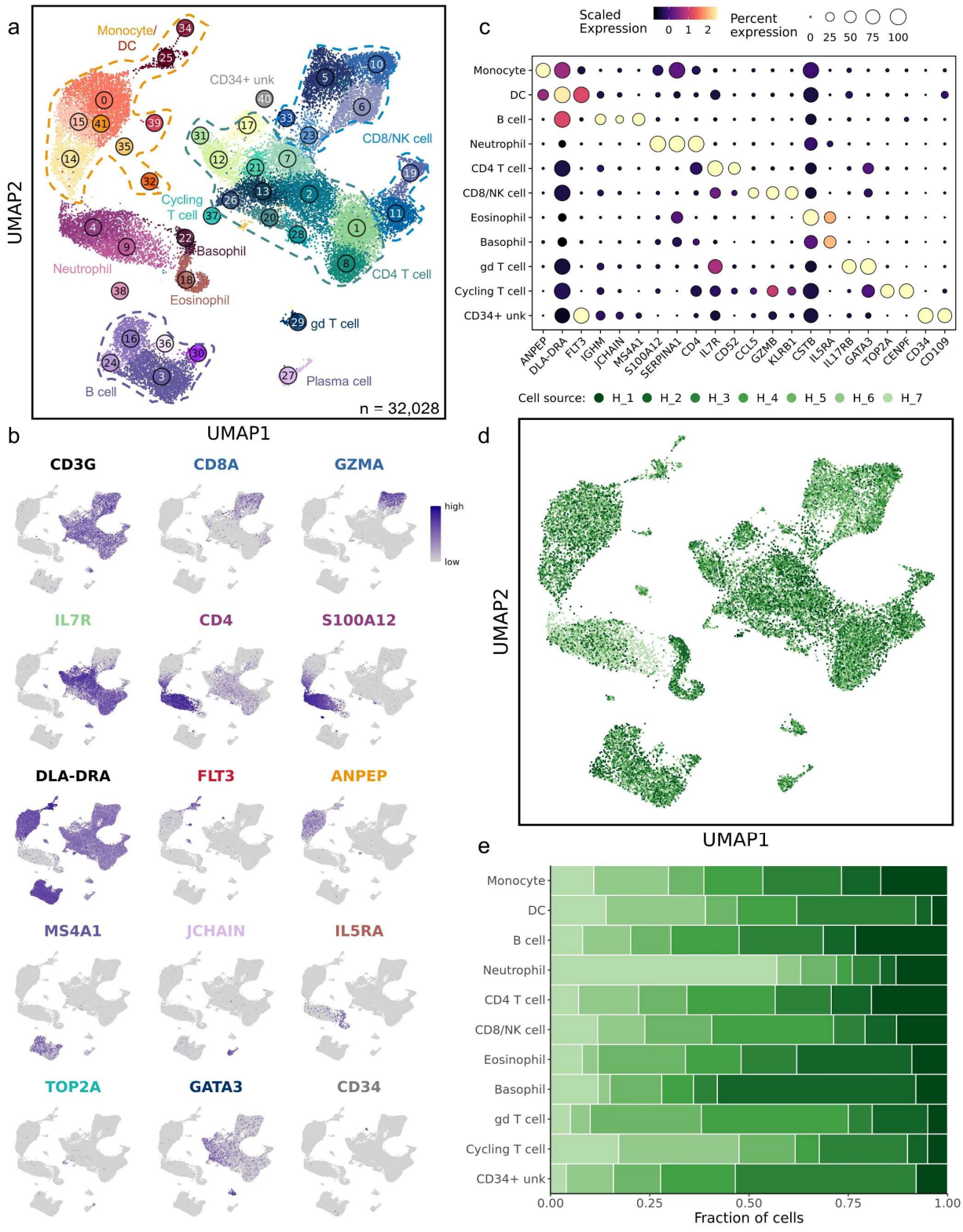

a
Monocyte/
DC
CD34+ unk
CD8/NK cell
Cycling
T cell
Basophil
Neutrophil
Eosinophil
CD4 T cell
gd T cell
B cell
Plasma cell
n = 32,028
UMAP2
UMAP1
c
Scaled
Expression
0 1 2
Percent
expression
0 25 50 75 100
Monocyte
DC
B cell
Neutrophil
CD4 T cell
CD8/NK cell
Eosinophil
Basophil
gd T cell
Cycling T cell
CD34+ unk
ANPEP
DLA-DRA
FLT3
IGHM
JCHAIN
MS4A1
S100A12
SERPINA1
CD4
IL7R
CD52
CCL5
GZMB
KLRB1
CSTB
IL5RA
IL17RB
GATA3
TOP2A
CENPF
CD34
CD109
Cell source:
H_1
H_2
H_3
H_4
H_5
H_6
H_7
b
CD3G
CD8A
GZMA
high
low
IL7R
CD4
S100A12
DLA-DRA
FLT3
ANPEP
MS4A1
JCHAIN
IL5RA
TOP2A
GATA3
CD34
d
UMAP2
UMAP1
e
Monocyte
DC
B cell
Neutrophil
CD4 T cell
CD8/NK cell
Eosinophil
Basophil
gd T cell
Cycling T cell
CD34+ unk
0.00
0.25
0.50
0.75
1.00
Fraction of cells

Figure 2.1. Unsupervised clustering reveals 42 unique cell populations in healthy dog leukocytes. (a) UMAP representation depicting the unsupervised clustering results of 32,028 leukocytes from 7 healthy dogs. Major immune cell populations are labeled on the plot (CD34+ unk = CD34+ unclassified, DC = Dendritic cell, gd = γδ) and each number corresponds to a unique cluster, with numbers increasing as cluster size becomes smaller. (b) Feature plots depicting expression of stereotypic genes used to define cell populations. (c) Dot plots illustrating scaled gene expression grouped by major immune cell populations. (d) UMAP representation of healthy canine leukocytes down sampled to obtain equal sampling of each dog and colorized by biological replicate (n = 7 dogs). (e) Stacked bar graph depicting contribution of each sample to every cell type.

Upon classification of all cell populations identified in the 7 healthy dogs, we generated gene lists that define each cell population and provide these lists in Supplemental data 1. In summary, we present a road map of healthy canine leukocytes and provide transcriptomic signatures for each distinct immune cell type that can be applied to study canine immunology as well used as a reference for deconvolution of bulk RNA sequencing data. Next, we applied the reference to investigate how the presence of osteosarcoma alters leukocyte transcriptional programs.

Comparison of healthy and osteosarcoma affected canine leukocytes

We isolated circulating leukocytes from 6 male and 4 female middle-aged tumor bearing dogs diagnosed with osteosarcoma (OS) (Table 2.1; Supplemental table 2.2). The cells obtained from the cancer burdened dogs were integrated with the 7 healthy samples to obtain a complete dataset of 74,067 cells. Similar to the unsupervised clustering of the healthy samples, we identified 46 unique clusters with the major immune cell types apparent (Figure 2.2a). Evaluation of data integration revealed uniform distribution of cells across all samples, except for clusters 9, 12, and 26 (Figure 2.2b/c; Supplemental figure 2.2b).

Clusters 9 and 12 (B cell populations) and Cluster 26 (a double negative [$CD3^+/CD4^-/CD8^-$] T cell population) were found to be underrepresented in the OS dogs. Interestingly, there has been one report of B cell and double negative (DN) T cell reductions in the peripheral blood of non-small cell lung cancer patients (27). While these findings agree with our data, other reports

have documented reductions in B cell and DN T cell abundances with age (28,29). Therefore, to investigate further, we compared the relative contribution of cell types in the middle-aged OS (aged 5-8 years old) and old OS dogs (aged 9-12 years old). We found no difference in B cell or DN T cell abundances between middle-aged and old OS dogs, while comparisons between healthy and middle-aged OS dogs also suggested no differences in abundance (Supplemental figure 2.3). Therefore, due to the limited sample size and previous reports of age-dependent B cell and DN T cell reductions, we believe the observed reductions in our dataset are likely in part due to age and not a direct effect of OS.

Aside from changes in cell abundances, we also observed that the $CD4^-$ neutrophil population identified in the healthy dataset (Cluster 38 in Figure 2.1a) shifted its position on the UMAP to be closer to other neutrophil populations (Cluster 27 in Figure 2.2a). This shift is attributed to the addition of cell numbers increasing the sample size and enabling the UMAP dimension reduction to better reconstruct the relative position on this cell population. While this population visually appeared to be expanded, statistically, there was no difference in abundance when evaluating the percentage out of total leukocytes using the clustering results from all immune cells.

Finally, we evaluated differentially expressed genes in all cells and each of the major immune cell populations (Figure 2.2d). This analysis indicated that several features were overrepresented in dogs with OS, with the most prominent differences arising from changes to myeloid cell gene expression. In particular, we observed IL1B and LTF upregulation in dogs with OS and a slight reduction of CD4 expression in dogs with OS. To further evaluate transcriptional differences in health and diseased states, we subset the database on each major cell population and completed independent unsupervised clustering to obtain greater resolution.

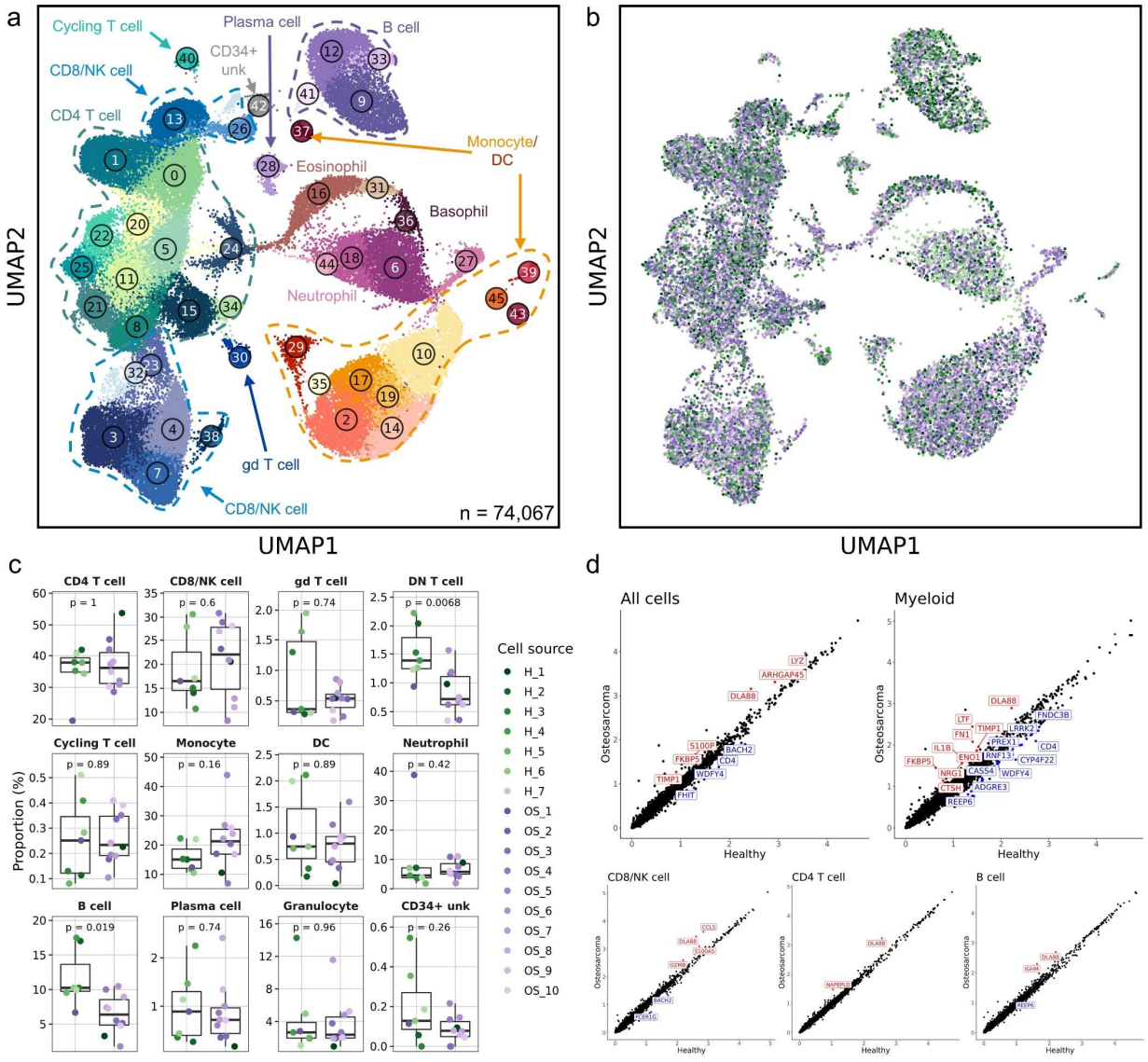


Figure 2.2. Canine osteosarcoma contributes to reductions in the relative abundances of double negative T cells and B cells, while myeloid cells contribute to most transcriptomic aberrations. (**a**) UMAP representation depicting the unsupervised clustering results of 74,067 leukocytes from 7 healthy and 10 osteosarcoma dogs (CD34+ unk = CD34+ unclassified). (**b**) UMAP representation of healthy (n = 7) and osteosarcoma (n = 10) canine leukocytes down sampled to obtain equal sampling of each dog and colorized by biological replicate. Legend is shared with (c). (**c**) Box plots quantifying changes in cell abundances between healthy and cancer dogs (Granulocyte = Eosinophil and Basophil clusters; DN T cell = Cluster 26 of CD8/NK cell group). P values were obtained using exact two-sided P values by the Wilcoxon Rank Sum test. (**d**) Scatter plots comparing average feature expression in osteosarcoma (y-axis) versus healthy cells (x-axis) with all (74,067) cells and 4 of the major immune cell populations, with labeled features significantly altered, as determined using a Wilcoxon Rank Sum test.

CD8 T cells, double negative T cells, and NK cells

In depth analysis of CD8 T cell, double negative T cell, and NK cell populations consisted of 15,846 cells with unsupervised clustering revealing 12 distinct clusters. Within this major immune cell population, we identified DN T cells, naïve CD8 T cells, effector CD8 T cells, memory CD8 T cells, NK cells, and NK T cells (Figure 2.3a/b, Supplemental figure 2.4a-c). Unexpectedly, we also identified a population of CD8+ gd T cells (Cluster 11) based on TRDC expression (Supplemental figure 2.4d). Of note, we found TRDC expression to extend outside of the CD8+ gd T cell cluster, with intermittent gene expression in effector and naïve CD8 T cells as well as uniform expression in NK cells. In addition to manual and algorithmic classification methods, we completed differential gene expression analysis to determine which features define key populations.

The results of differential gene expression using pseudobulk analysis from comparing one cluster to all other cells allowed for a complete description of the features that define each cell type. We found that the DN T cell cluster had reduced expression of cytotoxic genes (GZMB, NCR3, and KLRK1) and increased CTLA4 expression relative to all other cytotoxic cells (Figure 2.3c, Supplemental data 3). These gene patterns suggest the population may have reduced cytotoxic potential and may exhibit immune suppressive properties. When completing the same analysis on the NK cell population we found a reduction in T cell markers (CD3E, CD3G, CD8A) with an increase in certain cytotoxic features (GZMA, CD96 and KLRF1) relative to other cells (Figure 2.3d, Supplemental data 3). Finally, we applied this approach to further investigate Cluster 11, a $CD8^+$ gd T cell population (Figure 2.3e, Supplemental data 3). This analysis revealed that several cytotoxic features were downregulated (GZMA, GZMB, GZMK, KLRK1, and NCR3) while also indicating IKZF2, SOX5, and ELOVL5 were upregulated relative to other cytotoxic cells.

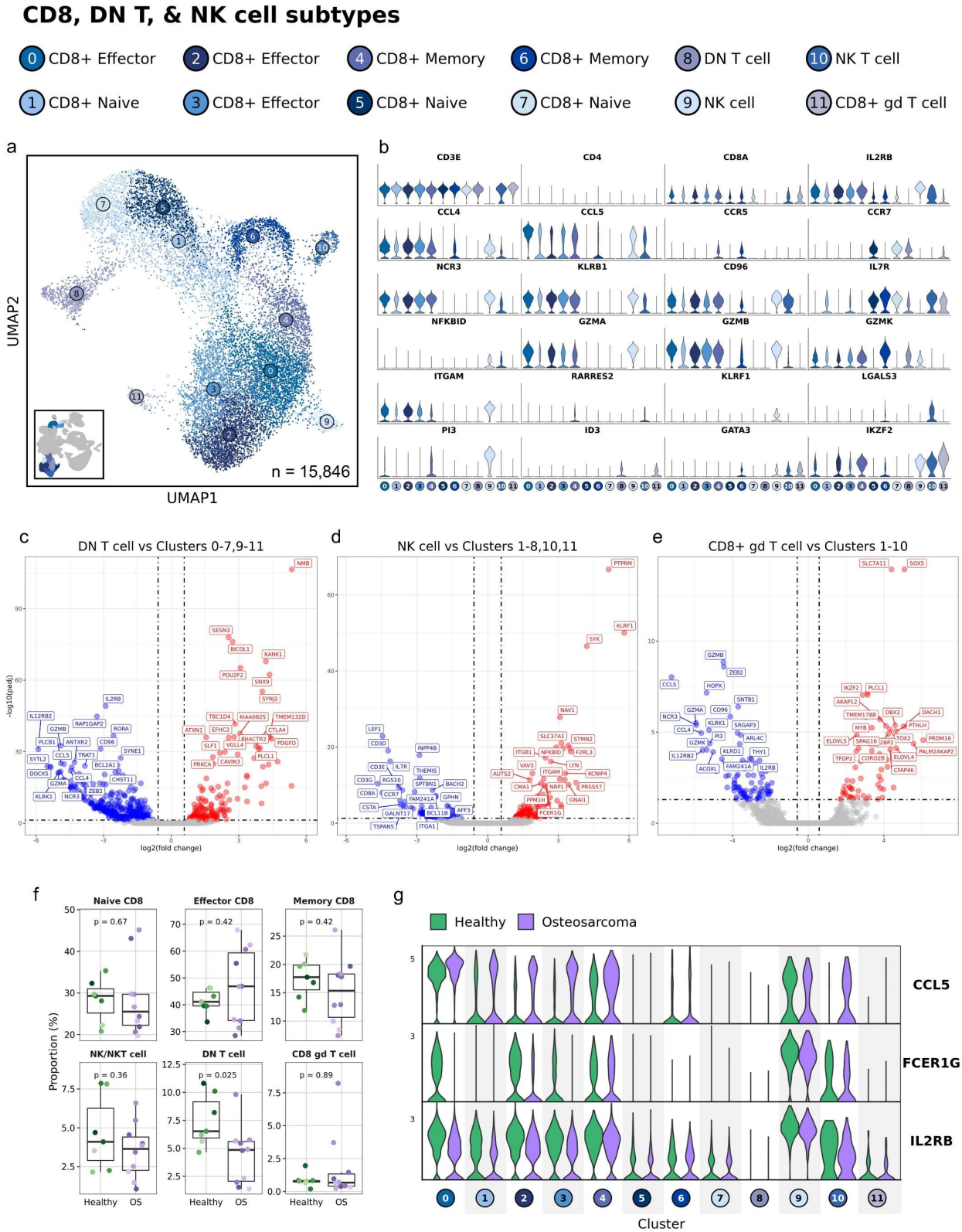

CD8, DN T, & NK cell subtypes
0 CD8+ Effector
2 CD8+ Effector
4 CD8+ Memory
6 CD8+ Memory
8 DN T cell
10 NK T cell
1 CD8+ Naive
3 CD8+ Effector
5 CD8+ Naive
7 CD8+ Naive
9 NK cell
11 CD8+ gd T cell
a
UMAP2
UMAP1
n = 15,846
b
CD3E
CD4
CD8A
IL2RB
CCL4
CCL5
CCR5
CCR7
NCR3
KLRB1
CD96
IL7R
NFKBID
GZMA
GZMB
GZMK
ITGAM
RARRES2
KLRF1
LGALS3
PI3
ID3
GATA3
IKZF2
c
DN T cell vs Clusters 0-7,9-11
-log10(padj)
log2(fold change)
d
NK cell vs Clusters 1-8,10,11
log2(fold change)
e
CD8+ gd T cell vs Clusters 1-10
log2(fold change)
f
Naive CD8
p = 0.67
Effector CD8
p = 0.42
Memory CD8
p = 0.42
NK/NKT cell
p = 0.36
DN T cell
p = 0.025
CD8 gd T cell
p = 0.89
Proportion (%)
Healthy
OS
g
Healthy
Osteosarcoma
CCL5
FCER1G
IL2RB
Cluster

Figure 2.3. Double negative T cell abundances are reduced in dogs with cancer, while CCL5 expression is increased in effector CD8 T cell clusters. (**a**) UMAP representation depicting the unsupervised clustering results of 15,846 leukocytes (n = 7 healthy and 10 osteosarcoma dogs). (**b**) Violin plots depicting key feature expression for each cytotoxic cell cluster. Axis scales are fixed across all features. (**c-e**) Volcano plots depicting differential gene expression when comparing DN T cell, NK cell, and $CD8^+$ gd T cell populations to all remaining cell clusters. (**f**) Box plots quantifying changes in cell cluster abundances (percentage of total cells, n = 15,846) between healthy and cancer dogs. P values were obtained using exact two-sided Wilcoxon Rank Sum test. (**g**) Violin plots of three features (CCL5, FCER1G aka Fc Epsilon Receptor Ig, and IL2RB aka CD122) identified as increased (CCL5, FCER1G) or decreased (IL2RB) in at least one cell cluster.

With cell classifications established, we next investigated the earlier observation that DN T cells were reduced in dogs with OS. To complete this analysis, we evaluated the cellular abundances as a percentage of the total cells (15,846 cells) in this immune cell subset and the results were consistent with the earlier approach (Figure 2.3f). Despite the marked reduction in DN T cells, we believe the decrease is in part due to age and should be interpreted cautiously. Unexpectedly, there were minimal transcriptomic differences when comparing healthy and OS CD8 T cell populations, but we consistently observed an increase in CCL5 expression on effector CD8 T cells in OS dogs (Figure 2.3g). Although less pronounced, we also identified decreases in FCER1G and IL2RB on effector CD8 T cell populations in dogs with OS. Together, these transcriptomic changes suggest there may be altered T cell recruitment signals in effector CD8 T cells of dogs with OS.

CD4 T cells

To continue a deeper investigation of major immune cell populations, we next focused our analysis on the most heterogenous group of cells: the CD4 T cells. Following independent reclustering, an additional 3 clusters were revealed which together indicated the presence of 15 transcriptionally unique CD4 T cell populations (Figure 2.4a, Supplemental figure 2.5a). One additional cluster was determined to be of poor quality and is depicted (grey) but was excluded from downstream analysis. Comparisons between healthy and OS CD4 T cells revealed no

apparent aberrations in cell abundances or transcriptomes (Supplementary figure 2.5b). Therefore, our analysis of CD4 T cells focused on describing the transcriptomic signature of each unique CD4 T cell population.

We used stereotypic markers and reference mapping to assign cell identities (30,31). This enabled the identification of naïve, central memory (TCM), effector memory (TEM), T regulatory (Tregs), and interferon (IFN) signature CD4 T cell populations (Figure 2.4b/c). The IFN signature cell type (Cluster 14) was defined by high expression of features associated with IFN response pathways when completing gene set enrichment analysis (Supplemental figure 2.5c) (32). The heterogeneity within the TEM clusters was further investigated and revealed the presence of Th1-like, Th2-like, and Th17-like clusters (Supplemental figure 2.5d) (17). Next, we used pseudotime trajectory analysis to investigate how the CD4 T cell populations were related. To complete this analysis, we established the naïve CD4 T cell Cluster 0 as the root node, then used Slingshot to infer how cell transcriptomes progressed (19). The analysis indicated several branchpoints, with 4 major lineages identified (Supplemental figure 2.5e/f). All lineages began at the assigned naïve CD4 T cell cluster (Cluster 0), progressed through a TCM cluster (Cluster 5), then through a TEM, and finally diverged toward an endpoint. Lineage 1 was determined to represent the progression of Th1-like cells, lineage 2 tracked the Th2-like progression, lineage 3 illustrated the Th17-like progression, and lineage 4 represented the progression of Tregs. Finally, for each of the four-lineage endpoints we completed differential gene expression analysis to further define the cell populations (Figure 2.4d, Supplemental data 3). Each population had upregulation of canonical features (Th1 = IL18R1/TBX21; Th2 = GATA3; Th17 = RORA/RORC, Treg = CTLA4/FOXP3) which acted to further validate cell classifications and provide novel markers to distinguish each cell type (30,33). In summary, we observed minimal aberrations in cell abundance or transcriptomes of CD4 T cell population in OS dogs but were able to identify 4 major lineages which provides insight into CD4 T cell biology in dogs.

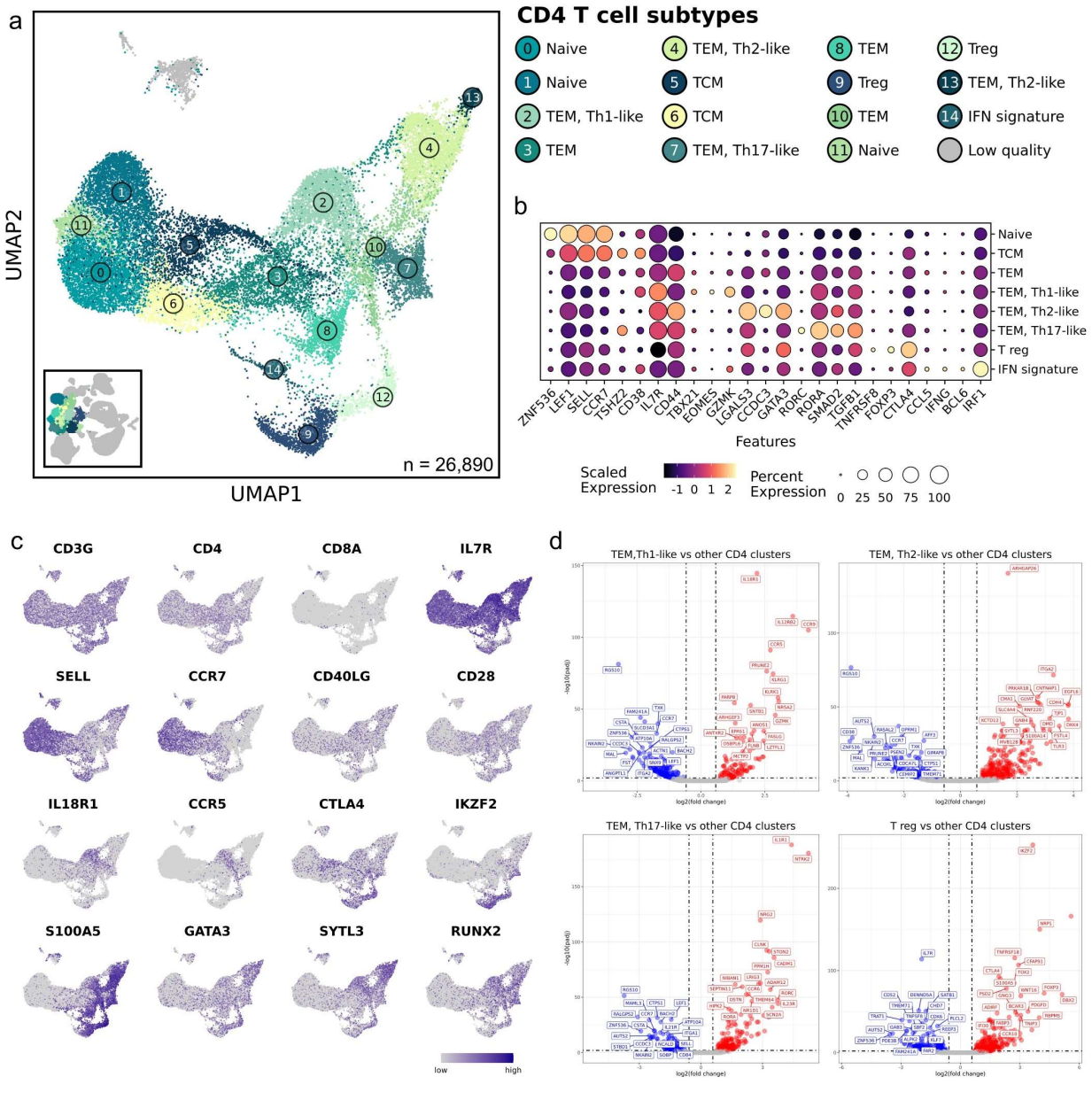


Figure 2.4. Analysis reveals 4 major CD4 T cell subtypes that were unaltered when comparing between healthy and osteosarcoma leukocytes. (**a**) UMAP representation depicting the unsupervised clustering results of 26,890 leukocytes (n = 7 healthy and 10 osteosarcoma dogs). (**b**) Dot plots illustrating gene expression by major CD4 T cell classification. Cell sub-classifications were collapsed to obtained 1 cell type group. (**c**) Feature plots depicting expression of stereotypic features used to define CD4 T cell populations. (**d**) Volcano plots depicting differentially expressed gene that were identified when comparing TEM Th1-like, Th2-like, Th17-like, and regulatory T cell (Treg) clusters to all other CD4 T cell clusters.

Myeloid cells

The next major immune cell population we analyzed consisted of monocytes, dendritic cells, neutrophils, basophils, and eosinophils. Through completion of independent reclustering, we identified 22 high quality myeloid cell populations and 2 low quality clusters that were excluded from downstream analysis (Figure 2.5a, Supplemental figure 2.6a). Monocytes were the most heterogenous cell type with 9 unique clusters, followed by 5 dendritic cell clusters, 5 neutrophil clusters, 2 eosinophil clusters, and 1 basophil cluster. Stereotypic features were then used to confirm cell identities (Figure 2.5b) (4,16,34). Although CD4 is reported to be expressed on canine neutrophils by flow cytometry, we identified a distinct neutrophil population that lacks CD4 expression (Cluster 12) (35). This unique neutrophil population appeared to be immature with immune suppressive gene expression patterns and was assigned an identity of polymorphonuclear myeloid-derived suppressor cells (PMN-MDSCs). When annotating the monocyte populations, we intended to define subpopulations by CD14/CD16 expression, as used for humans and mice. However, both CD14 and CD16 lack annotation in the primary canine reference genome (CanFam3.1) used in this study. As an alternative approach, we used CD64 (FCGR1A), MHCII (DLA-DRA), and CD86 to characterize to evaluate these cell populations in the context of human nomenclature (Supplemental figure 2.6b) (36). We determined that $CD4^+$ monocytes and M-MDSCs resembled classical monocytes ($CD14^+/CD16^-$), while $CD4^-$ monocytes resembled non-classical monocytes ($CD14^-/CD16^+$). Despite some overlap in nomenclature, the human classifications largely did not translate to our dataset, so we used CD4 expression to define canine monocyte populations.

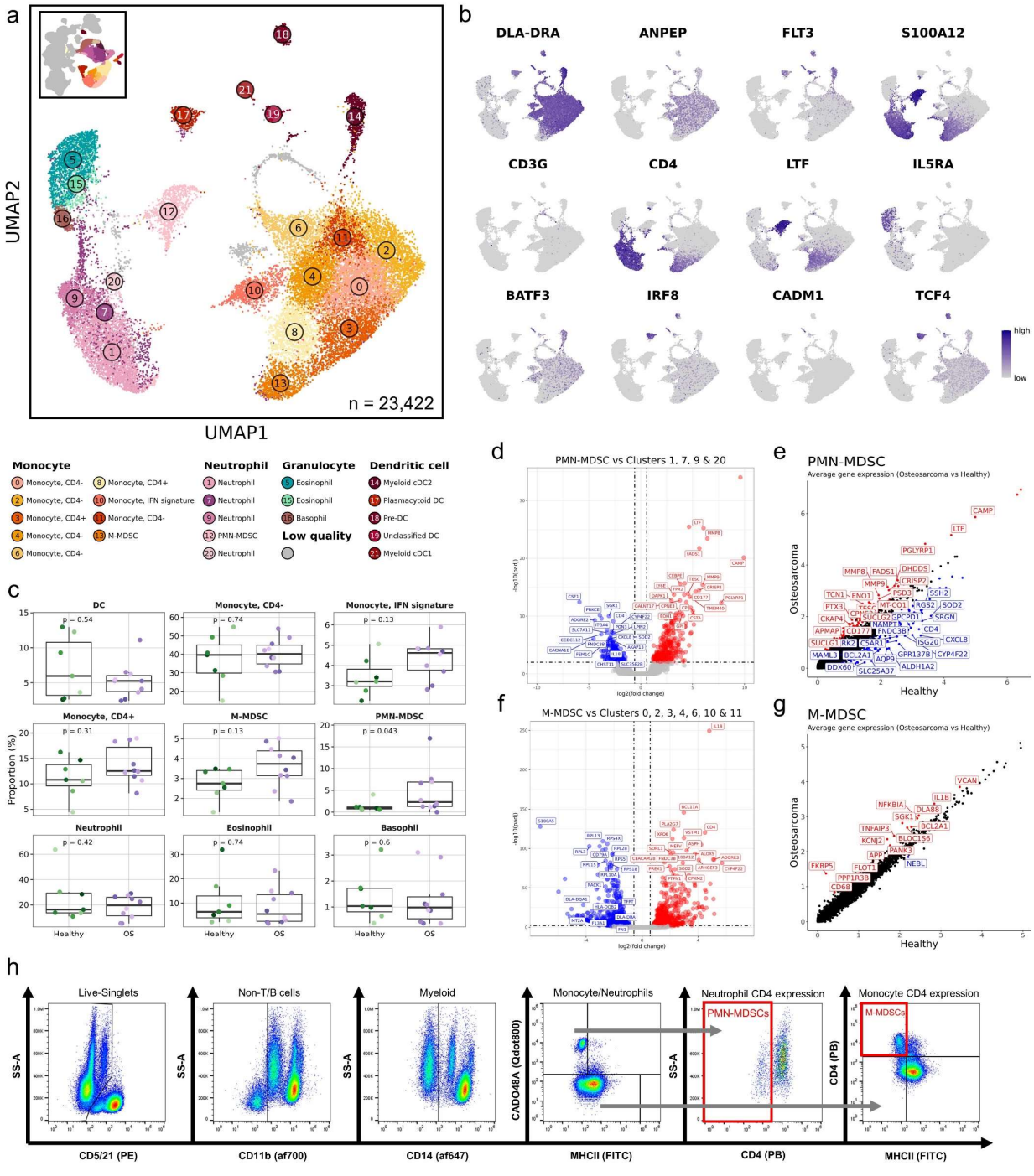

a
UMAP2
UMAP1
n = 23,422
b
DLA-DRA
ANPEP
FLT3
S100A12
CD3G
CD4
LTF
IL5RA
BATF3
IRF8
CADM1
TCF4
high
low
Monocyte
0 Monocyte, CD4-
2 Monocyte, CD4-
3 Monocyte, CD4+
4 Monocyte, CD4-
6 Monocyte, CD4-
8 Monocyte, CD4+
10 Monocyte, IFN signature
13 Monocyte, CD4-
13 M-MDSC
Neutrophil
1 Neutrophil
7 Neutrophil
9 Neutrophil
12 PMN-MDSC
20 Neutrophil
Granulocyte
5 Eosinophil
15 Eosinophil
18 Basophil
Low quality
Dendritic cell
14 Myeloid cDC2
17 Plasmacytoid DC
18 Pre-DC
19 Unclassified DC
21 Myeloid cDC1
c
DC
p = 0.54
Monocyte, CD4-
p = 0.74
Monocyte, IFN signature
p = 0.13
Monocyte, CD4+
p = 0.31
M-MDSC
p = 0.13
PMN-MDSC
p = 0.043
Neutrophil
p = 0.42
Eosinophil
p = 0.74
Basophil
p = 0.6
Proportion (%)
Healthy
OS
d
PMN-MDSC vs Clusters 1, 7, 9 & 20
log2(fold change)
e
PMN-MDSC
Average gene expression (Osteosarcoma vs Healthy)
Osteosarcoma
Healthy
f
M-MDSC vs Clusters 0, 2, 3, 4, 6, 10 & 11
log2(fold change)
g
M-MDSC
Average gene expression (Osteosarcoma vs Healthy)
Osteosarcoma
Healthy
h
Live-Singlets
Non-T/B cells
Myeloid
Monocyte/Neutrophils
Neutrophil CD4 expression
PMN-MDSCs
Monocyte CD4 expression
M-MDSCs
SS-A
CD5/21 (PE)
CD11b (af700)
CD14 (af647)
CADO48A (Qdot800)
MHCII (FITC)
CD4 (PB)
MHCII (FITC)

Figure 2.5. CD4 negative neutrophils are expanded in dogs with osteosarcoma and transcriptomically resemble polymorphonuclear myeloid-derived suppressor cells. (**a**) UMAP representation depicting the unsupervised clustering results of 23,422 myeloid cells (n = 7 healthy and 10 osteosarcoma dogs). (**b**) Feature plots depicting expression of key features used to define myeloid cell populations. (**c**) Box plots for each major myeloid cell population quantifying changes in cell abundances between healthy and cancer dogs. P values were obtained using a two-sided Wilcoxon Rank Sum test. (**d**) Volcano plots depicting differentially expressed genes identified when comparing the PMN-MDSC cluster (Cluster 12) to all other neutrophil clusters (Clusters 1,7,9,22). (**e**) Scatter plot comparing average feature expression in osteosarcoma (y-axis) versus healthy (x-axis) cells within the PMN-MDSC cluster. (**f**) Volcano plots highlighting DEGs identified when comparing the M-MDSC cluster (Cluster 13) to all other monocyte clusters (Clusters 0,2,3,4,6,8,10,11). (**g**) Scatter plot comparing average feature expression in osteosarcoma (y-axis) versus healthy (x-axis) cells within the M-MDSC cluster. (**h**) Representative flow cytometry gating strategy used to identify PMN- and M-MDSC populations.

To investigate how cell abundances were altered by OS, we compared the cellular frequencies of major myeloid cell sub-populations (Figure 2.5c). Consistent with the human literature, this analysis revealed PMN-MDSCs were expanded in dogs with OS (37,38). While not statistically significant as a percentage of myeloid cells, we observed a statistically significant increase in monocytic (M-) MDSCs (Cluster 13) when evaluating the cells as percentage of total leukocytes (Table 2.2). Differential gene expression analysis between OS and healthy dogs within monocytes exhibiting an IFN signature (Cluster 10) revealed a reduction in IFN-related gene expression in dogs with OS (Supplemental figure 2.6 c/d). This interesting change in gene expression indicates that dogs with OS may have a reduced IFN response potential in this subset of monocytes.

We next evaluated differentially expressed genes to determine the transcriptomic signatures for the suspected PMN-MDSC cluster (Figure 2.5d, Supplemental data 3). We found stereotypic PMN-MDSC features (CD177, LTF, CAMP, MMP9, and MMP8) to be upregulated in the PMN-MDSC cluster relative to the other neutrophil clusters (Clusters 1, 7, 9, & 22) (39). This analysis supported the classification of Cluster 12 as PMN-MDSCs. To determine if there were transcriptomic differences between healthy and OS PMN-MDSCs, we used a Wilcoxon Rank Sum

test to identify differentially expressed genes and found the expression of CD4, ISG20, and CXCL8 to be higher in cells from healthy dogs (Figure 2.5e). These findings suggest that either the healthy dog PMN-MDSCs are distinct from the OS derived PMN-MDSCs, or more likely, that normal neutrophils were misclassified during unsupervised clustering. Another immune suppressive cell population previously identified to be expanded in cancer patients are M-MDSCs (40). We suspected that the $CD4^+$ monocyte population, Cluster 13, represented M-MDSCs. When evaluating differentially expressed genes between Cluster 13 and the remaining monocytes (Clusters 0, 2, 3, 4, 6, 10, and 11), we found down regulation of ribosomal transcripts (RPS/RPL) and MHCII (DLA-DRA) features and upregulation of CD4, S100A12, and IL18 (Figure 2.5f, Supplemental data 3). While not definitive, this expression profile supports the classification of Cluster 13 as M-MDSCs, while also providing the gene signature of Cluster 13 (41,42). Investigation of transcriptomic differences in disease revealed subtle changes in gene expression, with higher IL1B and lower DLA-DRA in the dogs with OS within the M-MDSC cluster (Figure 2.5g).

Following the identification of M-MDSC and PMN-MDSC populations, we developed a clinically accessible flow cytometry-based assay to monitor MDSC populations in dogs with cancer (Figure 2.5h, Supplemental figure 2.6e, Supplemental table 2.1). The assay builds upon previously proposed canine MDSC immunolabeling protocols, but includes an anti-CD4 antibody to further distinguish MDSCs from their normal counterparts (9,43,44). This assay is intended to be employed in the investigation of prognostic correlates and further study of canine MDSC biology.

Table 2.2. Canine leukocyte high-resolution cellular composition.

	All	Healthy	Osteosarcoma	P value*
CD8 T cell	19.9 ± 7.6	18.4 ± 7.2	21 ± 8	0.601
CD8+ Naive	6.1 ± 2.5	5.9 ± 2.8	6.2 ± 2.4	0.536
CD8+ Effector	10.2 ± 5.6	8.7 ± 3.9	11.3 ± 6.5	0.669
CD8+ Memory	3.4 ± 1.3	3.6 ± 1.3	3.2 ± 1.4	0.813
CD8+ gd T cell	0.2 ± 0.2	0.2 ± 0.1	0.3 ± 0.3	0.962
CD4 T cell	36.2 ± 7.6	35 ± 7.5	37 ± 7.9	1
CD4+ Naive	10.3 ± 3.4	10.4 ± 3.1	10.2 ± 3.8	0.962
CD4+ TCM	5.3 ± 1.6	5 ± 1.3	5.4 ± 1.9	1.000
CD4+ TEM	7.4 ± 2.1	7.5 ± 2.2	7.3 ± 2	0.813
CD4+ TEM, Th1-like	3.8 ± 1.8	3.4 ± 1.9	4 ± 1.7	0.475
CD4+ TEM, Th2-like	4 ± 1.6	3.7 ± 1.4	4.3 ± 1.7	0.740
CD4+ TEM, Th17-like	2.1 ± 0.7	1.8 ± 0.4	2.3 ± 0.8	0.161
CD4+ T reg	3 ± 0.7	2.9 ± 0.5	3 ± 0.9	1.000
CD4+, IFN signature	0.3 ± 0.1	0.3 ± 0.1	0.4 ± 0.1	0.813
Monocyte	18.5 ± 8.5	15.1 ± 4.1	20.8 ± 10.1	0.088
Monocyte, CD4-	12.2 ± 5.8	10.2 ± 3.2	13.6 ± 7	0.230
Monocyte, CD4+	4 ± 2.2	3.1 ± 1.1	4.7 ± 2.6	0.230
M-MDSC	1 ± 0.4	0.8 ± 0.2	1.2 ± 0.4	0.019
Monocyte, IFN signature	1.3 ± 0.6	1.1 ± 0.8	1.4 ± 0.5	0.193
Dendritic cell	1.8 ± 1	2 ± 1.3	1.6 ± 0.8	0.813
Pre-DC	0.3 ± 0.2	0.3 ± 0.3	0.2 ± 0.2	0.536
Myeloid cDC1	0.1 ± 0.1	0.1 ± 0.1	0.1 ± 0.1	0.669
Myeloid cDC2	0.9 ± 0.5	0.9 ± 0.5	0.9 ± 0.5	1.000
Plasmacytoid DC	0.3 ± 0.3	0.4 ± 0.4	0.3 ± 0.2	0.601
Unclassified DC	0.1 ± 0.1	0.2 ± 0.2	0.1 ± 0.1	0.962
Granulocyte	11.1 ± 8.1	13.1 ± 12.2	9.8 ± 3.4	1
Neutrophil	6.9 ± 7.6	9.2 ± 11.5	5.4 ± 2.9	0.6691
PMN-MDSC	1 ± 1.2	0.5 ± 0.8	1.4 ± 1.4	0.0431
Eosinophil	2.8 ± 3	3 ± 3.5	2.6 ± 2.7	0.8125
Basophil	0.4 ± 0.3	0.4 ± 0.2	0.4 ± 0.3	0.7396
B cell	9.5 ± 4.5	12.6 ± 4.4	7.4 ± 3.3	0.07
Immature B cell	0.8 ± 0.5	1.1 ± 0.5	0.6 ± 0.4	0.088
Naive B cell	5.4 ± 2.6	7 ± 2.5	4.3 ± 2.1	0.043
Class switched B cell	1.7 ± 0.8	2.4 ± 0.7	1.2 ± 0.5	0.005
Activated B cell	0.7 ± 0.9	1.2 ± 1.2	0.4 ± 0.3	0.014
Plasma cell	0.9 ± 0.6	1 ± 0.7	0.8 ± 0.6	0.887
Miscellaneous	3 ± 1.2	3.7 ± 1.4	2.4 ± 0.7	0.025
DN T cell	1.1 ± 0.5	1.4 ± 0.4	0.8 ± 0.3	0.005
gd T cell	0.7 ± 0.5	0.9 ± 0.7	0.5 ± 0.2	0.740
NK cell	0.5 ± 0.4	0.6 ± 0.6	0.4 ± 0.3	0.364
Cycling T cell	0.3 ± 0.1	0.3 ± 0.2	0.3 ± 0.1	0.887
NK T cell	0.3 ± 0.2	0.3 ± 0.1	0.4 ± 0.2	1.000
CD34+ unclassified	0.1 ± 0.1	0.2 ± 0.2	0.1 ± 0.1	0.261

*P value is the exact two-sided P value obtained by a Wilcoxon Rank Sum test comparing Osteosarcoma column to Healthy.

B cells and miscellaneous cell types

The final immune cell population analyzed was B cells. We were able to identify immature, naïve, activated and class-switched B cells as well as a cluster of plasma cells (Supplemental figure 2.7 a-c). We found minimal differences within B cells when comparing between the cell abundances and transcriptomes of OS versus healthy dogs (Supplemental figure 2.7d). Despite the lack of differences between healthy and diseased states, the data presented here provide transcriptomic signatures for B cell subtypes not previously resolved using traditional characterization methods in dogs.

In addition to the immune cell populations discussed above, we reported the presence of gd T cells, cycling T cells and a CD34+ unclassified population. Initially, we believed the CD34+ unclassified cluster to be common myeloid progenitors. Upon closer examination, we noted the population expressed endothelial cell markers (CD109) (45). Therefore, we state their CD34 positivity, but refrain from assigning a further identity due to the small representation of the cluster and conflicting gene signatures. Ultimately, investigation into each miscellaneous population did not reveal the presence of OS mediated changes, so we solely present their transcriptomic signatures as a resource for future researchers.

Summary of the high-resolution canine leukocyte landscape

Lastly, the cell classifications determined through analysis of individual immune cell populations were compiled to present relative abundances of total circulating leukocytes. We identified 36 unique cell populations believed to each represent a biologically relevant subset of canine immune cells (Table 2.2, Supplemental figure 2.8a/b). Furthermore, we provide the complete transcriptomic signatures for each cell type and present a summary table with short gene lists for each population (Supplemental data 1,2,4, Table 2.3). Comparisons between healthy and OS dogs revealed both PMN-MDSCs and M-MDSCs were expanded in dogs with OS when evaluated as a percentage out of all leukocytes. Additionally, DN T cells, naïve B cells, class-switched B cells, and activated B cells were found to be reduced in dogs with OS, but we were unable to determine if this is a result of age or cancer.

Overall, the breakdown of leukocytes was found to be consistent with previously reported values determined using flow cytometry (46). We used paired flow cytometry data to confirm the relative distribution of major immune cell types. The comparison of cellular percentages between single-cell RNA sequencing and flow cytometry revealed a positive correlation indicating consistent identification of immune cells independent of approach (R^2 = 0.635; Supplemental figure 2.8c). Therefore, the data presented here are largely consistent with traditional classification methods, and act to provide novel insights into the heterogeneity and transcriptomic signatures of canine cell types.

Finally, we integrated our dataset of 7 healthy canine leukocytes with a previously published human reference consisting of 6 healthy adults (18). Cell type homologies between species were then evaluated using hierarchical clustering of SCT normalized data, which evaluated the top 2000 variable features with gene homologies between species. Although the analysis was impacted by differences in the level of cell type annotations between the two studies, the results suggested more similarities than dissimilarities between the species (Supplemental figure 2.9). For instance, human and canine naïve CD4 T cells, neutrophils, and plasmacytoid

DCs paired off 1:1 in terminal clades, suggesting a high degree of similarity. Additionally, human DCs clustered on the same clade as all canine DC subtypes; this was also observed in the case of B cells. Interestingly, we identified subtle differences between species, which included human NK cells clustering with canine CD8 effector T cells and gd T cells from each species clustering into separate clades. Overall, the cross-species analysis emphasizes the similarities in circulating immune cell transcriptional profiles, while also highlighting potential differences between the two species.

Table 2.3. Transcriptional signatures of canine leukocytes

Cell Type	Marker
CD8 T cell	
CD8+ Naive	ITGA1, LEF1, PTGDR, IL2RB, ADGRG1, NBEA
CD8+ Effector	CCL5, TRPM3, IL12RB2, GZMB, KLRB1, GZMA, NCR3, IL2RB, KLRD1, CD96
CD8+ Memory	GZMK, GZMB, PI3, BTBD11, CTSW, CCR5, CCL4, KLRG1, FASLG
CD8+ gd T cell	PTHLH, IGF2BP2, ABTB2, AKAP12, SOX4, CTSW, SLC16A10, PXT1, ZNRF3, SULT2B1
CD4 T cell	
CD4+ Naive	LEF1, CSTA, RGS10, ZNF536, CCR7, COL6A5, LTB, TNFSF8
CD4+ TCM	LEF1, TSHZ2, CD52, CCR7, IL7R, CTPS1, EFHC2, CARMIL1
CD4+ TEM	IL7R, SLC9A9, ICOS, MAF, CD28, SKAP1, CD40LG
CD4+ TEM, Th1-like	IL7R, PTPN13, IL18R1, CD28, RCAN2, CCR9, CCR5, IL12RB2, CD52, PRUNE2
CD4+ TEM, Th2-like	RNF220, ITGA2, GATA3, CCDC3, LGALS3, PTPN13, S100A2, PPEF1, CMA1
CD4+ TEM, Th17-like	NTRK2, PTPN13, ADAM12, NRG2, RGS17, DNAH8, CCR6, NPAS2, RORA, LTBP1
CD4+ T reg	IKZF2, CTLA4, RGS1, ICOS, IL2RA, CD28, ZNF831
CD4+, IFN signature	CXCL10, IFI44, OAS1, ISG15, IFI44L, IFGGB2, CTLA4, STAT1, DDX58, XAF1
Monocyte	
Monocyte, CD4-	LYZ, BPI, LRMDA, MT2A, F13A1, FN1, NRG1, CCDC88A, CD83, RETN
Monocyte, CD4+	IL1B, MAFB, NFKBIA, CXCL8, FN1, BLOC1S6, CD83, S100P, BPI, NRG1
M-MDSC	IL18, IL1B, LTF, MEFV, KCNJ2, CPXM2, S100A12, STEAP4, CSF3R, IL31RA
Monocyte, IFN signature	RSAD2, OAS1, OAS2, DDX58, HERC6, OAS3, RTP4, EIF2AK2, IFIT2

Dendritic cell	
Pre-DC	FGF12, GPHA2, MTUS2, FCER1A, PLCE1, PTPRS, IGF1, NECTIN1, IL3RA, AK8
Myeloid cDC1	ZNF366, SDC2, DISC1, ECRG4, TMEM163, RIMS2, KIT, OTOF, RTKN2, RAB7B
Myeloid cDC2	PKIB, CD300H, SDC2, CD1C, NCAM2, CD86, BATF3, ZNF366, PID1, ECM1
Plasmacytoid DC	COBLL1, RAB3C, IGF1, FCER1A, RYR1, PRKG1, CCND1, STYXL2, ANK1, OCIAD2
Unclassified DC	PLCB4, ZNF366, KCNK13, STRIP2, SDC2, OTOF, HACD1, C5, SLC8A1, CNTLN
Granulocyte	
Neutrophil	S100A12, CD4, SERPINA1, SGK1, S100A8, ALDH1A2, FNDC3B, GGH, SRGN, IL1R2
PMN-MDSC	CAMP, PGLYRP1, CRISP2, MMP9, MMP8, TCN1, CD177, LTF, FADS1, S100A12
Eosinophil	C30H15orf48, TGM2, DACH1, PADI3, SMPD1, CA8, IL5RA
Basophil	DACH1, CA8, IL5RA, DAPK2, TGFA, ANKRD33B, HK2, PRR5L
B cell	
Immature B cell	SYT1, PAX5, VPREB3, ERC2, TMTC2, KLHL14, F8, TEX9, TDRP, ADGRF1
Naive B cell	TNFRSF13C, BANK1, HTR1F, PAX5, EBF1, BTLA, NRIP1, ADAM9
Class switched B cell	TNFRSF13C, GOLM1, BANK1, BTLA, EBF1, DYNC1I1, MTMR2, PAX5
Activated B cell	IGKC, CACNB2, PAX5, TNFRSF13C, IGHM, RASGRF2, AOX2, BCAR3, ADAM32
Plasma cell	JCHAIN, MZB1, TXNDC5, LMAN1, FKBP11, LAP3, DERL3, CCR10, MKI67, TNFRSF13B
Miscellaneous	
DN T cell	KIAA0825, TMEM132D, KANK1, NMB, CTLA4, SYNJ2, BICDL1, SLF1, ID3, KIAA1549
gd T cell	PARD3B, RHEX, IL17RB, CDH4, GATA3, FAT1, TOX2, ADARB1, ZNF683, TGFBR3
NK cell	KLRF1, STMN2, PAX4, NCR3, F2RL3, CD96, IL2RB, IGSF3, FREM1, FASLG
Cycling T cell	TOP2A, MKI67, RRM2, H1-5, DIAPH3, TK1, KIF11, TPX2, ASPM
NK T cell	GPA33, TGFBR3, KLRK1, CD96, SYTL2, MOV10L1, SLA2, DSTN, RARRES1
CD34+ unclassified	TFPI, ZNF521, CD34, NDST3, GUCY1A1, HPGD, CLEC3B, KIT, CD109, DNTT

Discussion

In the present study we generated a reference single-cell RNA sequencing dataset using 7 healthy dogs, then applied the database to investigate how osteosarcoma alters immune cell transcriptomes and abundances. Our analysis revealed the heterogenous canine immune landscape and allowed us to define the transcriptomes of distinct immune cell populations. When comparing healthy dogs and dogs with osteosarcoma (OS), we found that a cancer burden contributes to relative increases in the abundances of polymorphonuclear (PMN-) and monocytic (M-) myeloid-derived suppressor cells (MDSCs). Furthermore, we identified that most transcriptomic changes between healthy and OS dogs resulted from changes to myeloid cell populations. In particular, we noted increases in LTF, CAMP, and S100A12 expression which appear to arise from an expansion of MDSCs. Ultimately, the data presented here sheds light on the diversity of canine immune cells and highlights key changes between healthy dogs and dogs with osteosarcoma.

Overall, there were relatively few aberrations in leukocyte populations isolated from dogs with osteosarcoma. Despite this, the observed differences may represent biologically significant changes in dogs with cancer. For instance, MDSCs have been identified as key immune suppressive populations that dampen antitumor immunity and the relative abundances of these cells have been determined to have prognostic correlates (8). Both monocytic and polymorphonuclear MDSCs have been reported in dogs and were identified using flow-cytometric based approaches (9). Our findings support the classifications schemes used, but further suggest that CD4 expression should be considered when evaluating MDSC burden. We report that M-MDSCs exhibit a phenotype of $CD5^-/CD21^-/CD11b^+/CD14^+/CADO48a^-/MHCII^-/CD4^+$ and that PMN-MDSCs can be defined as $CD5^-/CD21^-/CD11b^+/CD14^{int}/CADO48a^+/MHCII^-/CD4^-$. Because both MDSCs were determined to express LTF, CAMP, and S100A12 it is possible that the use of bulk RNA sequencing or targeted sequencing approaches may be able to provide indirect

assessments of MDSC burdens. Ultimately, further investigation of these cell populations is warranted to determine the clinical relevance of MDSCs in dogs with cancer and other diseases.

Although reductions in B cells and DN T cells were noted, we presented further data to suggest age, in addition to cancer, may be driving this change. Due to this potential confounder and minimal transcriptomic differences between healthy and OS dogs within these cell types, our analysis was limited to characterization of each subtype. Our B cell subtype analysis provides refined B cell classifications in dogs and offers transcriptomic profiles of each population.

Unexpectedly, we did not identify overt OS mediated changes to circulating CD4 T cells. Prior research suggests that regulatory T cells tend to be expanded in tumor bearing humans, mice, and dogs (47,48). The inability to detect an expansion of Tregs is likely a result of our small sample size and lack of power to detect changes in subtly altered cell abundances. As such, our analysis focused on describing the heterogeneity within CD4 T cell subpopulations to provide a comprehensive description of CD4 T cell subtypes. Key contributions include the identification of four distinct CD4 T cell lineages (Th1-like, Th2-like, Th17-like, and Tregs) and the generation of cell type transcriptomic signatures. Our analysis of CD8 T cells and NK cells revealed that, relative to humans, canine CD8 T cells exhibit gene expression patterns that more closely resemble human NK cells than CD8 T cells. This observation was made early on in analysis when using human classifiers to assign cell identities, and again when we completed a direct comparison to a dataset of healthy adult leukocytes. Thus, our findings suggest canine CD8 T cells exhibit a more innate-like transcriptomic signature relative to their human counterparts. A similar conclusion was recently reported in canine single-cell RNA sequencing of αβ T cells (49), providing further evidence of potential differences between T cells in dogs and humans.

Outside of the major immune cell populations, we identified a few distinct cell subpopulations, including gd T cells, cycling T cells, and CD34+ unclassified cells. These cell types were considered to be distinct from the major immune cell populations and were left out of independent reclustering analysis, but their importance should not be discounted given their

documented roles in numerous diseases (50). Thus, their characterization may represent a useful reference for future studies.

The cellular annotations proposed in our analysis represent plausible cellular identities that we determined through use of manual and algorithmic based classification methods. We recognize that unsupervised clustering identifies transcriptionally unique populations, but the boundaries set by clustering do not always correspond with biologically relevant distinctions. As such, in our final annotations presented in Table 2.2, we collapsed some of the clusters into one cell category that we believed to be more biologically relevant. For example, five $CD4^-$ monocyte populations were identified in our myeloid cell analysis, but we could not assign further identities to them, and in turn, annotated them all as one group of $CD4^-$ monocytes. Furthermore, cell identification solely based on canonical cell type markers (i.e., FOXP3 for Tregs and IFNG/TBX21 for Th1-like), proved ineffective in many instances. This is in part due to the low transcript abundance of transcription factors, cytokines, and chemokines in our dataset. Thus, many traditional gene markers for these cell types were not considered upregulated and were largely excluded from gene lists that define each cell population (as defined in Table 2.3 and Supplemental data 1/2). However, when subsetting and using pseudobulk approaches, we were able to better detect defining features, such as FOXP3 in regulatory T cells. This highlights the importance of using independent reclustering on major cell populations and allowed us to provide detailed gene signatures for rare cell populations (available in Supplemental data 3).

Despite our contributions, this study is not without limitations. For instance, granulocytes have been reported to be difficult to study using single-cell transcriptomics due to small transcriptomes, high RNase content, and sensitivity to sample processing (51). As a result, many research groups chose to filter out granulocyte populations. Despite this, we used single-cell RNA sequencing to characterize granulocyte populations and were able to identify distinct cell populations and report differentially expressed genes between healthy and OS dogs. While our data suggests there are biologically relevant differences between granulocyte populations in

healthy and OS dogs, it is important to validate the findings reported here using molecular and functional assays. An additional limitation is the relatively low number of biological replicates and incomplete breed representation used to investigate disease induced changes in cell abundances. Future studies using alternative experimental approaches, such as flow cytometry, should be completed to validate conclusions presented here. Lastly, although we attempted to control for age between the healthy and OS dogs, we observed reductions in B cells and DN T cells which suggests the age difference between groups may have confounded our investigation of cancer associated changes.

We noted that our cell classifications are in discordance with those previously published for canine PBMCs (52). The only consistently annotated cell populations are B cells and platelets, while the remaining classifications differ between studies. For instance, the cluster classified as "monocytes" by Li et al. has high expression of CD3E which suggests a classification of T cells is more appropriate. Furthermore, the "T cell" cluster identified by Li et al. lacks CD3E expression, but has high S100A12/CD4 expression which likely corresponds to a neutrophil cluster (expression levels were determined using the web browser released by Li et al.; http://120.79.46.200:81/Pandora/PBMC.html). The discrepancies between studies likely arose due to classification methods not taking into consideration the well-established phenomenon that, in dogs, neutrophils have high CD4 expression (35). Ultimately, cell classification should be completed using multiple approaches and biological insight when working with non-model species.

In summary, the data presented here serve two purposes. Firstly, the leukocyte subclassifications can act as a valuable resource for the scientific community to use in future research. For example, the proposed classifications provide detailed phenotypes of canine immune cells that can be used to inform the design of flow cytometry assays, bulk RNA sequencing deconvolution algorithms, and more (see Data and software availability section for details on use applications). Secondly, this study identified important differences in leukocyte

abundances and transcriptomes in dogs with OS. Ultimately, our goal is that the database provided here will be used by researchers as a reference dataset as well as shed light on how cancer impacts circulating leukocytes.

References

1. Mak IWY, Evaniew N, Ghert M. Lost in translation: animal models and clinical trials in cancer treatment. Am J Transl Res. 2014;6(2):114.
2. Park JS, Withers SS, Modiano JF, Kent MS, Chen M, Luna JI, et al. Canine cancer immunotherapy studies: linking mouse and human. J Immunother cancer. 2016;4(1):1–11.
3. Schiffman JD, Breen M. Comparative oncology: what dogs and other species can teach us about humans with cancer. Philos Trans R Soc B Biol Sci. 2015;370(1673):20140231.
4. Rzepecka A, Żmigrodzka M, Witkowska-Piłaszewicz O, Cywińska A, Winnicka A. CD4 and MHCII phenotypic variability of peripheral blood monocytes in dogs. PLoS One. 2019;14(7):e0219214.
5. Avery PR, Burton J, Bromberek JL, Seelig DM, Elmslie R, Correa S, et al. Flow Cytometric Characterization and Clinical Outcome of CD 4+ T-Cell Lymphoma in Dogs: 67 Cases. J Vet Intern Med. 2014;28(2):538–46.
6. Horiuchi Y, Nakajima Y, Nariai Y, Asanuma H, Kuwabara M, Yukawa M. Th1/Th2 balance in canine peripheral blood lymphocytes—A flow cytometric study. Vet Immunol Immunopathol. 2007;118(3–4):179–85.
7. Ohki S, Shibata M, Gonda K, Machida T, Shimura T, Nakamura I, et al. Circulating myeloid-derived suppressor cells are increased and correlate to immune suppression, inflammation and hypoproteinemia in patients with cancer. Oncol Rep. 2012;28(2):453–8.
8. Kalathil SG, Thanavala Y. Importance of myeloid derived suppressor cells in cancer from a biomarker perspective. Cell Immunol. 2021;361:104280.
9. Goulart MR, Hlavaty SI, Chang Y-M, Polton G, Stell A, Perry J, et al. Phenotypic and transcriptomic characterization of canine myeloid-derived suppressor cells. Sci Rep. 2019;9(1):3574.
10. Jackson K, Milner RJ, Doty A, Hutchison S, Cortes-Hinojosa G, Riva A, et al. Analysis of canine myeloid-derived suppressor cells (MDSCs) utilizing fluorescence-activated cell sorting, RNA protection mediums to yield quality RNA for single-cell RNA sequencing. Vet Immunol Immunopathol. 2021;231:110144.
11. Hao Y, Hao S, Andersen-Nissen E, Mauck III WM, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. Cell. 2021;184(13):3573–87.
12. McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. Cell Syst. 2019;8(4):329–37.
13. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol. 2019;20(1):1–15.
14. Zappia L, Oshlack A. Clustering trees: a visualization for evaluating clusterings at multiple resolutions. Gigascience. 2018;7(7):giy083.
15. Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. Nat Immunol. 2019;20(2):163–72.

16. Franzén O, Gan L-M, Björkegren JLM. PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. Database. 2019;2019.

17. Bukhari S, Henick BS, Winchester RJ, Lerrer S, Adam K, Gartshteyn Y, et al. Single-cell RNA sequencing reveals distinct T cell populations in immune-related adverse events of checkpoint inhibitors. Cell Reports Med. 2022;100868.

18. Wilk AJ, Rustagi A, Zhao NQ, Roque J, Martínez-Colón GJ, McKechnie JL, et al. A single-cell atlas of the peripheral immune response in patients with severe COVID-19. Nat Med. 2020;26(7):1070–6.

19. Street K, Risso D, Fletcher RB, Das D, Ngai J, Yosef N, et al. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. BMC Genomics. 2018;19:1–16.

20. Squair JW, Gautier M, Kathe C, Anderson MA, James ND, Hutson TH, et al. Confronting false discoveries in single-cell differential expression. Nat Commun. 2021;12(1):1–15.

21. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci. 2005;102(43):15545–50.

22. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innov. 2021;2(3):100141.

23. Cheng S, Li Z, Gao R, Xing B, Gao Y, Yang Y, et al. A pan-cancer single-cell transcriptional atlas of tumor-infiltrating myeloid cells. Cell. 2021;184(3):792–809.

24. Speir ML, Bhaduri A, Markov NS, Moreno P, Nowakowski TJ, Papatheodorou I, et al. UCSC Cell Browser: visualize your single-cell data. Bioinformatics. 2021;37(23):4578–80.

25. Blanter M, Cambier S, De Bondt M, Vanbrabant L, Pörtner N, Salama SA, et al. Method matters: effect of purification technology on neutrophil phenotype and function. Front Immunol. 2022;317.

26. Buurman WA, Vegt PA, Groenewegen G, Linden C, Jeunhomme G. Analysis of buoyant density of canine peripheral blood leukocytes with PVP-Silica (Percoll) density gradients. Vet Immunol Immunopathol. 1982;3(6):547–56.

27. Liang H, Chu X, Zhao J, Xing G, Si Y. Elevated peripheral blood B lymphocytes and CD3+ CD4-CD8-T lymphocytes in patients with non-small cell lung cancer: A preliminary study on peripheral immune profile. Oncol Lett. 2018;15(6):8387–95.

28. Bulati M, Caruso C, Colonna-Romano G. From lymphopoiesis to plasma cells differentiation, the age-related modifications of B cell compartment are influenced by “inflamm-ageing.” Ageing Res Rev. 2017;36:125–36.

29. Holcar M, Goropevšek A, Ihan A, Avčin T. Age-related differences in percentages of regulatory and effector T lymphocytes and their subsets in healthy individuals and characteristic STAT1/STAT5 signalling response in helper T lymphocytes. J Immunol Res. 2015;2015.

30. Zhu J, Yamane H, Paul WE. Differentiation of effector CD4 T cell populations. Annu Rev Immunol. 2009;28:445–89.

31. Schulz EG, Mariani L, Radbruch A, Höfer T. Sequential polarization and imprinting of type 1 T helper lymphocytes by interferon-γ and interleukin-12. Immunity. 2009;30(5):673–83.

32. Wang X, Shen X, Chen S, Liu H, Hong N, Zhong H, et al. Reinvestigation of classic T cell subsets and identification of novel cell subpopulations by single-cell rna sequencing. J Immunol. 2022;208(2):396–406.

33. Cano-Gamez E, Soskic B, Roumeliotis TI, So E, Smyth DJ, Baldrighi M, et al. Single-cell transcriptomics identifies an effectorness gradient shaping the response of CD4+ T cells to cytokines. Nat Commun. 2020;11(1):1801.

34. Collin M, Bigley V. Human dendritic cell subsets: an update. Immunology. 2018;154(1):3–20.

35. Moore PF, Rossitto P V, Danilenko DM, Wielenga JJ, Raff RF, Severns E. Monoclonal antibodies specific for canine CD4 and CD8 define functional T-lymphocyte subsets and high-density expression of CD4 by canine neutrophils. Tissue Antigens. 1992;40(2):75–85.

36. Ong S-M, Teng K, Newell E, Chen H, Chen J, Loy T, et al. A novel, five-marker alternative to CD16–CD14 gating to identify the three human monocyte subsets. Front Immunol. 2019;10:1761.

37. Brandau S, Trellakis S, Bruderek K, Schmaltz D, Steller G, Elian M, et al. Myeloid-derived suppressor cells in the peripheral blood of cancer patients contain a subset of immature neutrophils with impaired migratory properties. J Leukoc Biol. 2011;89(2):311–7.

38. Cassetta L, Bruderek K, Skrzeczynska-Moncznik J, Osiecka O, Hu X, Rundgren IM, et al. Differential expansion of circulating human MDSC subsets in patients with cancer, infection and inflammation. J Immunother Cancer. 2020;8(2).

39. Veglia F, Hashimoto A, Dweep H, Sanseviero E, De Leo A, Tcyganov E, et al. Analysis of classical neutrophils and polymorphonuclear myeloid-derived suppressor cells in cancer patients and tumor-bearing mice. J Exp Med. 2021;218(4).

40. Wu L, Deng Z, Peng Y, Han L, Liu J, Wang L, et al. Ascites-derived IL-6 and IL-10 synergistically expand CD14+ HLA-DR-/low myeloid-derived suppressor cells in ovarian cancer patients. Oncotarget. 2017;8(44):76843.

41. Zhao F, Hoechst B, Duffy A, Gamrekelashvili J, Fioravanti S, Manns MP, et al. S100A9 a new marker for monocytic human myeloid-derived suppressor cells. Immunology. 2012;136(2):176–83.

42. Lim HX, Hong H-J, Cho D, Kim TS. IL-18 enhances immunosuppressive responses by promoting differentiation into monocytic myeloid-derived suppressor cells. J Immunol. 2014;193(11):5453–60.

43. Goulart MR, Pluhar GE, Ohlfest JR. Identification of myeloid derived suppressor cells in dogs with naturally occurring cancer. PLoS One. 2012;7(3):e33274.

44. Sherger M, Kisseberth W, London C, Olivo-Marston S, Papenfuss TL. Identification of myeloid derived suppressor cells in the peripheral blood of tumor bearing dogs. BMC Vet Res. 2012;8(1):1–12.

45. Mancuso P, Calleri A, Gregato G, Labanca V, Quarna J, Antoniotti P, et al. A subpopulation of circulating endothelial cells express CD109 and is enriched in the blood of cancer patients. PLoS One. 2014;9(12):e114713.

46. Platt R, Ng T, Glover S, Roof M, Kimura K, Roth JA. Canine peripheral blood lymphocyte

phenotyping by 7-color multiparameter flow cytometry. Anal Quant Cytopathol Histopathol. 2013;35(4):197–204.

47. Biller BJ, Guth A, Burton JH, Dow SW. Decreased ratio of CD8+ T cells to regulatory T cells associated with decreased survival in dogs with osteosarcoma. J Vet Intern Med. 2010;24(5):1118–23.

48. Strauss L, Bergmann C, Whiteside TL. Functional and phenotypic characteristics of CD4+ CD25highFoxp3+ Treg clones obtained from peripheral blood of patients with cancer. Int J cancer. 2007;121(11):2473–83.

49. Eschke M, Moore PF, Chang H, Alber G, Keller SM. Canine peripheral blood TCRαβ T cell atlas: Identification of diverse subsets including CD8A+ MAIT-like cells by combined single-cell transcriptome and V (D) J repertoire analysis. Front Immunol. 2023;14.

50. Gully BS, Rossjohn J, Davey MS. Our evolving understanding of the role of the γδ T cell receptor in γδ T cell mediated immunity. Biochem Soc Trans. 2021;49(5):1985–95.

51. Wigerblad G, Cao Q, Brooks S, Naz F, Gadkari M, Jiang K, et al. Single-cell analysis reveals the range of transcriptional states of circulating human neutrophils. J Immunol. 2022;209(4):772–82.

52. Li Z, Sun C, Wang F, Wang X, Zhu J, Luo L, et al. Molecular mechanisms governing circulating immune cell heterogeneity across different species revealed by single-cell sequencing. Clin Transl Med. 2022;12(1):e689.

CHAPTER 3: Molecular dissection of primary canine osteosarcoma using single-cell RNA sequencing

Summary

Osteosarcoma (OS) is a heterogeneous, aggressive malignancy of the bone that disproportionally affects children and adolescents. Therapeutic interventions for OS are limited, which is in part due to the complex tumor microenvironment that has proven to be refractory to immunotherapies. Thus, there is a need to study the OS microenvironment to understand tumor biology. To address this need, we used single-cell RNA sequencing to describe the cellular composition of 6 treatment-naïve dogs with spontaneously occurring primary osteosarcoma. We identified 30 distinct immune cell types, 9 unique tumor populations, 1 cluster of fibroblasts, and 1 population of endothelial cells. Deeper analysis revealed the presence of mature regulatory dendritic cells with a predicted role in T cell directed immune suppression. We then completed a direct comparison to human OS, which revealed a high degree of cell type homology between species. Our findings further support the use of spontaneous canine OS tumors to study translational therapeutics as well as provides key insights into canine OS which can be used inform future studies. Generally, the data presented here acts a valuable resource to study canine OS biology which can advance canine translational immuno-oncology research.

Introduction

Osteosarcoma (OS) is an aggressive malignancy of the bone that disproportionally impacts children and teenagers. The propensity of OS to develop in adolescence, indicates that the disease results in a substantial loss of life years in affected individuals. Despite the profound impact, OS therapeutics are lacking and there have been minimal advancements in recent years. Slow advancements in OS therapeutics are, in part, due to the relatively rare occurrence of the

malignancy which hampers accrual into clinical trials. In recent years, there has been an increasing interest in using large animal models to facilitate the evaluation of immunotherapeutics (1). The dog is regarded as an ideal model of OS due to disease prevalence estimated to be up to 75 times that observed in humans, similar disease pathology, and unlike some rodent models, dogs are immune competent and exposed to similar environmental stimuli as humans (2). Although dogs have been identified as a valuable pre-clinical model, reagent limitations have restricted the ability to characterize canine OS and evaluate similarities in the immune landscape between species.

In recent years, single-cell RNA (scRNA) sequencing has emerged as a valuable tool to investigate the transcriptomes of individual cells within heterogenous tissues. The approach overcomes species-specific regent limitations, by relying on a universal transcript capture method that is only limited by the completeness of genome annotations (3). Importantly, the human scRNA landscape of primary, recurrent, and metastatic osteosarcoma has been described and acts as a point of reference for the completion of homology analysis with canine OS (4,5). The aim of this study is to use scRNA sequencing to complete a molecular dissection of canine OS and evaluate transcriptomic homologies between humans and dogs.

Osteosarcoma consists of a complex microenvironment that contains macrophages, osteoclasts, fibroblasts, lymphocytes, malignant osteoblasts as well as other stromal and immune cells. Together, the OS tumor microenvironment (TME) creates a highly immune suppressive milieu that hinders antitumor responses when applying immunotherapy-based interventions (6). Researchers have turned to the TME with the objective of understanding and targeting the cellular constituents that promote immune suppression. Preclinical mouse and canine studies have identified TME modifying therapeutics, such as losartan, to be capable of remodeling tumor stromal cell and reducing rates of metastasis (7–9). Unlike many other cancer types, there have been reports in humans and dogs that suggest increased macrophage abundance in OS reduces metastatic rate and enhances survival (10–12). This unexpected finding and ill-defined

mechanisms of immune suppression in the OS TME highlights the need for a deeper understanding of OS biology.

In the present study we generated a single-cell RNA sequencing atlas of six treatment-naïve dogs with primary osteosarcoma. Our analysis revealed 30 distinct immune cell types, 9 unique tumor populations, 1 cluster of fibroblasts, and 1 population of endothelial cells. Of note, we identified a rare population of mature regulatory DCs, which have been identified in multiple human solid tumors, including OS (13,14). Furthermore, we completed a detailed dissection of tumor-associated macrophage and T cell populations which revealed insights into immune suppressive properties. We completed a human-canine homology analysis and found cell type gene signatures in the TME to be conserved between species. Our findings not only shed light on the heterogeneity in canine OS, but given the similar cellular and transcriptomic compositions, suggest that canine OS may be an ideal model of the human disease. Overall, the data generated here can be used to inform the identification of OS therapeutic targets as well as facilitate the further study of the complex microenvironment.

Methods

Patient selection

Osteosarcoma patients were selected based on the presence of a primary tumor and the absence of previous therapeutic intervention. All dogs presented with radiographic evidence of OS and their first line of intervention was surgical limb amputation. Within 30 minutes of amputation samples were collected for single-cell RNA sequencing processing and for completion of histology to confirm the OS diagnosis. All dogs included in the study were confirmed to have OS by histology. All studies were approved by the Colorado State University (CSU) Institutional Animal Care and Use Committee and the CSU Clinical Review Board. All dog owners provided informed consent prior to sample collection.

Sample preparation

Between 3 and 5 tumor biopsies were collected from the amputated limb and immediately placed in Roswell Park Memorial Institute medium (Thermo Fisher Scientific Inc., Waltham, MA). Each tumor sample was washed with phosphate buffered saline, pH = 7.40 (PBS) then minced using a scalpel and digested with collagenase type II (250 U/mL; Thermo Fisher Scientific Inc.) for 45 minutes at 37˚C with agitation. Samples were passed through a 70-µm cell strainer, washed with PBS, then centrifuged for 5 min at 400 rcf. To enrich for live cells, samples were then pooled into 4-mL Hanks' Balanced Salt Solution, pH 7.4 (HBSS), layered onto 3-mL Ficoll Paque (Cytiva; Marlborough, MA), and centrifuged for 30 min at 400 rcf with acceleration at 9 and brake at 0. The cell interface layer was collected and washed one time with PBS, resuspended in 10-mL of Ammonium-Chloride-Potassium (ACK) lysis buffer for 3-7 minutes at room temperature, then washed an additional time with PBS. To remove small debris and platelets, a final wash at 100 rcf x 15 minutes was completed. Cells were then resuspended in 0.04% molecular grade BSA (Sigma-Aldrich; St. Louis, MO) in PBS and transported to a Chromium iX instrument (10x Genomics; Pleasanton, CA) and captured within 30 minutes of preparation.

Library preparation and sequencing

Single cells were isolated and tagged with unique cell barcodes using a Chromium iX instrument (10x Genomics) with a target of 5,000 cells per sample. Two of the dogs (dogs 1 and 2) had two samples processed each with a 5,000-cell target, for a total target of 10,000 cells from each of those two dogs. Single cells were isolated and processed using a Chromium Next GEM Single Cell 3' Kit v3.1 following manufacture recommended protocols. Individual cell transcriptomes were captured and labeled with molecular barcodes, then a standard Illumina library preparation was completed using a dual index library construction kit (10x Genomics). Samples quality was analyzed using a LabChip (PerkinElmer; Waltham, MA) then submitted for sequencing on an Illumina NovaSeq 6000 sequencer (Novogene Corporation; Sacramento, CA)

with a target of 100,000 150 bp paired-end reads per cell. Raw data was demultiplexed by the sequencing core then transferred for downstream analysis.

Read mapping and quantification

A Cell Ranger analysis pipeline (version 6.1.2, 10x Genomics) was utilized to process raw FASTQ sequencing data, align reads to the canine genome, and generate a count matrix. The default settings were used when running “cellranger count”. The input reference package used for the “cellranger count” command was generated by filtering the CanFam3.1 ensembl annotation (gtf) for protein_coding, lincRNA, antisense, and immunoglobulin gene biotypes. Then the reference package was generated using the “cellranger mkref” command. The output count matrix for each sample was then exported and used for downstream analysis.

Data filtering and integration

For each sample, the count matrix was imported into R using the Read10X() function then converted to a Seurat object using the CreateSeuratObject() function (15). To estimate the number of dead/poor quality cells, the percent mitochondrial reads per cell was calculated using PercentageFeatureSet() to count all reads mapped to features with the prefix "MT-". Each object was then filtered to only retain cells which met the following requirements: 200 < nFeature_RNA < 5500, percent.mt < 12.5, and 100 < nCount_RNA < 75000. Next, DoubletFinder, was used to identify and remove putative cell doublets (16). After completing QC filtering on each sample, all samples were integrated into one object using the SCTransform() and integration workflow (17). During this step, we regressed out the percent mitochondrial reads minimize the impacts on clustering results and used 2000 features as integration anchors. Following data integration, three low quality clusters were identified and removed then the samples were split out to repeat data integration. Ideal clustering parameters (res = 0.8, dims = 45, n.neighbors = 40, min.dist = 0.35) were determined using the R package clustree (18). Dimension reduction and visualization was then completed, and the data were projected using 2-dimensional, non-linear uniform manifold approximation and projection (UMAP) plots.

Cell classification

High level cell type annotations were established using unsupervised clustering results, gene set enrichment analysis, and manual annotation based on the literature (19). Briefly markers included CD3E for T cells, CTSK for osteoclasts, CD68 for macrophages, S100A12 for neutrophils, COL1A1/ALPL/FAP for tumor/fibroblasts, FLT3 for dendritic cells, MS4A1/JCHAIN for B cells, ESAM/PVAP for endothelial cells, TOP2A for cycling cells, and ACE2/MS4A2 for mast cells. Further high-resolution cell identification was completed through independent reclustering on cells within each major (Tumor/Fibroblast, Myeloid, and T cells) cell population.

Feature visualization

Feature expression was visualized using violin plots, feature plots, and dot plots. Selected features were chosen based on prior biological knowledge and features identified using the FindMarkers() function. Y-axis scales for violin plots within a figure are fixed, so they are all on the same scale. Feature plots show normalized expression for each feature and are on variable scales. For all feature plots, grey/light purple color indicates low expression, while positivity is indicated by purple colorization. Dot plots used scaled expression data which depicts deviation from the average value for a gene across the cells being sampled.

Differential gene expression analysis

Differential gene expression (DGE) analysis was completed using pseudobulk conversion followed by a DESeq2 pipeline to evaluate differential gene expression (20). Prior to running DESeq2, low abundance features, defined as features with less than 10 raw counts across all cells sampled, were filtered out. Features that had an adjusted P value of less than 0.01 and a log2fold greater than 0.58 were considered to be statistically significant.

Gene set enrichment/variation analysis

When completing follow-up gene set enrichment analysis (GSEA) on the gene lists generated from DGE analysis, the significantly upregulated and downregulated features were processed separately. The upregulated and downregulated gene lists were used with

clusterProfiler and MSigDB gene sets to infer pathway activity (21,22). Terms which reached a P value of 0.05 or lower were discussed as significantly enriched.

In addition to using GSEA following DGE analysis we also used the R package singleseqgset to complete GSEA on cell type clusters. The tool uses a competitive gene set enrichment test that was based on a Correlation Adjusted MEan RAnk gene set test (23). The log2 fold change and mean expression for every feature within each cell type was calculated and then used to complete GSEA. P values were corrected for multiple comparisons using a false discovery rate (FDR) method and corrected P values were filtered to only retain terms in which at least one cell types had a value less than 0.05. The enrichment values were scaled then the top pathways (weighted by P value) were plotted using a heatmap.

CNV analysis

Copy number variation analysis was completed using CopyKAT on high quality cells that contained more than 2000 unique molecular identifiers (UMIs) (24). Briefly, the approach segments the genome into 220-kb variable genomic bins to establish a genome-wide copy number profile for each single cell at an approximate resolution of 5 Mb. Each sample was run individual because any batch effects in the integrated dataset can interfere with CNV prediction. A known normal cell population consisting of osteoclasts, neutrophils, macrophages, and T cells was used when inferring CNV status. The samples were also run without a known normal and produced similar results, but when run without a known normal population, the aneuploidy/diploidy call had a chance of being reversed, as there is no reference. Individual cell classifications were extracted from each .rds output file, then transferred to the Seurat object containing the integrated data. The CNV status was then visualized on a UMAP. This approach was only used to infer if a cell was aneuploid or diploid and individual chromosome mutations were not evaluated due incompatibilities of software across species.

Regulon activity

Single-cell regulatory network inference and clustering (SCENIC) was used to infer activity of gene regulatory networks by cell types (25). The two regulatory feather files used for analysis were obtained from https://resources.aertslab.org/cistarget/ and were named "hg19-500bp-upstream-7species.mc9nr.feather" and "hg19-tss-centered-10kb-10species.mc9nr.feather". Default settings were then applied to run the analysis pipeline. Regulons specificity scores (rss) were then calculated using AUCell and the rss values were used to infer activity of regulons in cell types analyzed (26).

Human OS homology analysis

Six treatment-naive human OS samples were obtained from GSE162454 (4). The count matrices reported from the study were then loaded in as Seurat object and were filtered using the same parameters as used for the 6 canine OS tumor samples. The human dataset was annotated using the high-level unsupervised clustering while referencing the primary article in attempts to recreate their annotations. Following annotation of each species, the 12 (6 human and 6 canine) OS samples were integrated into one object to using a SCTranscfrom workflow with 3000 variable features used as anchors. Only features with homologues across both species were used for integration. SCT normalized counts were then used to complete hierarchical clustering using the hclust() function with method set to “complete”. Subsequent DGE analysis contrasting cell types within each species was completed each individual species dataset. The adjusted P values obtained from DGE analysis were assigned a sign (+/-) based on the log2 fold change then the signed P values were used to generate a scatter plot.

Data and software availability

Raw sequencing data (FASTQ format) analyzed in this chapter are available upon request and will be released upon publication of the associated manuscript. The analysis code used is available at https://github.com/dyammons/canine_osteosarcoma_atlas.

Results

Establishment of a treatment-naïve canine osteosarcoma reference database

To establish a treatment-naïve canine osteosarcoma reference, we completed single-cell RNA (scRNA) sequencing on 6 dogs and collected data on a total of 35,310 cells. The average number of cells collected per tumor was 5,885 and on average each cell was sequenced to a depth of 72,649 reads per cell. All tumors were confirmed to be osteosarcoma by histology and histological subtyping was completed on each tumor. In total, the curated dataset consisted of 1 fibroblastic, 1 chondroblast, and 4 osteoblastic tumors, with one dog exhibiting radiographical evidence of lung metastasis (Table 3.1).

Table 3.1. Dog demographics.

Dog ID	Sex	Breed	Age (years)	Tumor location	Evidence of metastasis	Histological subtype
Naïve 1	FS	Mixed (Husky)	8	L proximal humerus	No	Osteoblastic
Naïve 2	MC	Catahoula	11.5	R distal femur	Yes	Osteoblastic
Naïve 3	MC	Labrador Retriever	7.8	L distal femur	No	Fibroblastic
Naïve 4	MC	Great Dane	8	R distal radius	No	Osteoblastic
Naïve 5	FS	Mixed	11.3	R distal radius	No	Chondroblastic
Naïve 6	FS	Catahoula	8.4	R distal radius	No	Osteoblastic

Initial low-resolution cell type annotation revealed the presence of 7 major cell types consisting of T cells, B cells, macrophages, dendritic cells (DCs), osteoclasts (OCs), tumor cells, cycling tumor cells, and an additional 3 minor cell populations consisting of neutrophils, mast cells, and endothelial cells (Figure 3.1a). Evaluation of the dataset for evidence of batch effects indicated uniform distribution of cell types between biological replicates. The one exception was that naïve dog 6 had more neutrophils and T cells relative to the other dogs (Figure 3.1b). This skew may be a result of sampling bias in which necrotic tumor, blood, or bone marrow contamination was introduced during sampling. Subsequent analysis of cell type proportions revealed 42.3% of the dataset consisted of tumor or fibroblast, 2.1% was endothelial cells, while the remaining 55.6% were immune cells (Figure 3.1c).

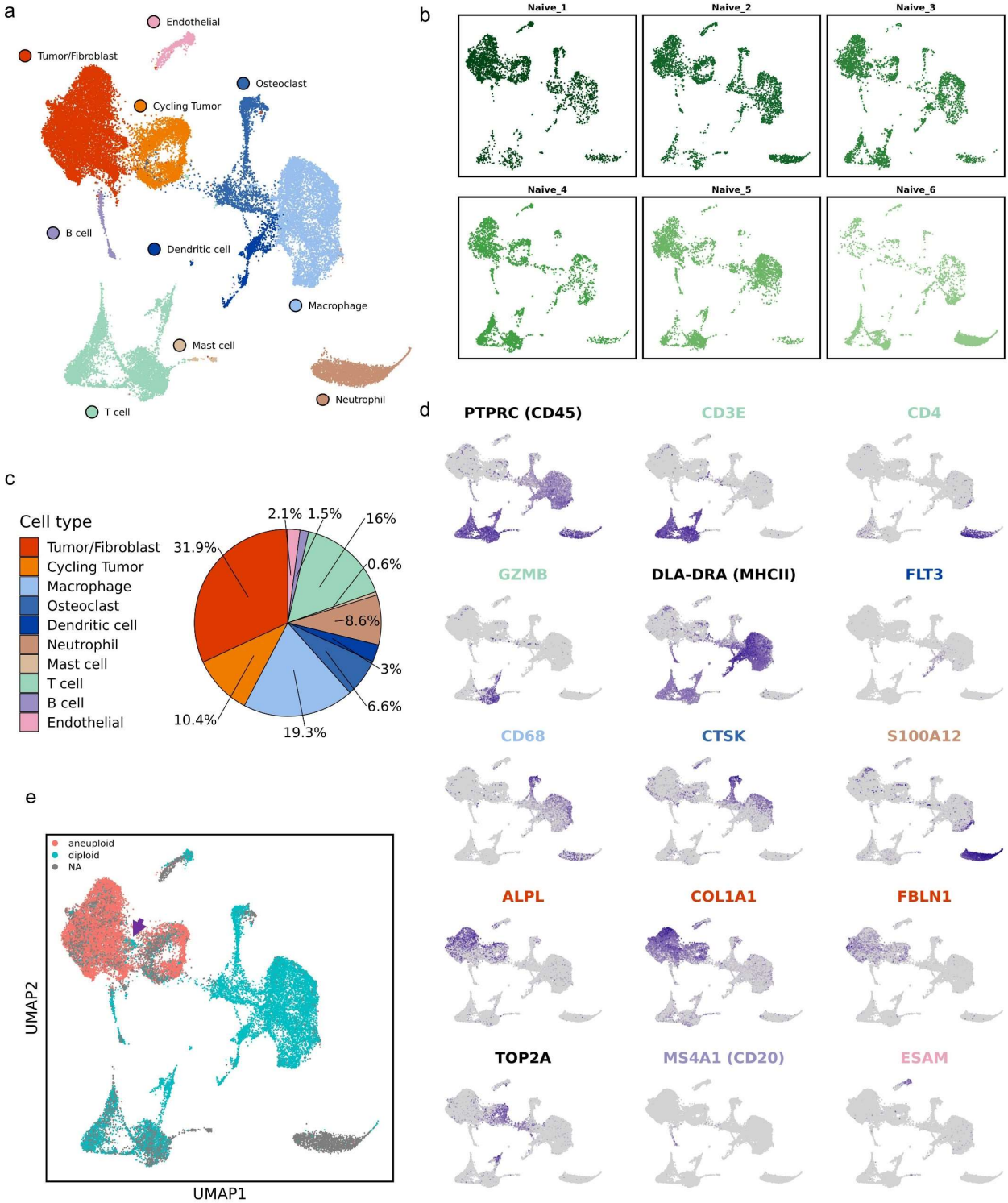

a
Endothelial
Tumor/Fibroblast
Osteoclast
Cycling Tumor
B cell
Dendritic cell
Macrophage
Mast cell
Neutrophil
T cell
b
Naive_1
Naive_2
Naive_3
Naive_4
Naive_5
Naive_6
c
Cell type
Tumor/Fibroblast
Cycling Tumor
Macrophage
Osteoclast
Dendritic cell
Neutrophil
Mast cell
T cell
B cell
Endothelial
31.9%
10.4%
19.3%
6.6%
3%
8.6%
0.6%
16%
2.1%
1.5%
d
PTPRC (CD45)
CD3E
CD4
GZMB
DLA-DRA (MHCII)
FLT3
CD68
CTSK
S100A12
ALPL
COL1A1
FBLN1
TOP2A
MS4A1 (CD20)
ESAM
e
aneuploid
diploid
NA
UMAP2
UMAP1

Figure 3.1: Unsupervised clustering reveals 10 distinct cell types in canine OS. (a) UMAP representation of 40,625 cells obtained from the primary tumors of 6 dogs diagnosed with OS. (b) UMAP faceted by dog and down sampled to depict equal numbers of cell from each sample. (c) Pie chart depicting the cellular composition of the data as a percentage of total cells. (d) Feature plots depicting the log normalized counts of canonical markers used to justify major cell type classifications. (f) UMAP depicting the results of CopyKAT algorithm. Grey (NA) values indicate that the cell did not have a large enough transcriptome to run through the algorithm.

Cell types were annotated based on expression of canonical markers, reference mapping using a human OS dataset, and gene set enrichment analysis (Figure 3.1d). Furthermore, gene lists used by Liu et. al to define cell populations in human OS, were applied using module scoring to provide further support for cell classifications (Supplemental figure 3.1) (4). This approach suggested consistency in the ability of the gene lists to identify T cells, B cells, osteoclasts, and endothelial cells. However, the approach was unable to differentiate fibroblasts from osteoblasts. Ultimately, our high-level unsupervised clustering failed to distinguish between stromal fibroblasts and malignant osteoblasts. This unexpected observation may be in part due to the presence of a fibroblastic osteosarcoma tumor in our dataset and the broad expression of fibroblast markers (FAP, FBLN1) across all tumor cell clusters (Supplemental figure 3.1).

Due to the inability to resolve fibroblasts using feature expression, we used CopyKAT to complete copy number variation analysis to infer which cells exhibited aneuploidy based on their transcriptional properties (Figure 3.1e). The analysis revealed that a majority of cells in the tumor/fibroblast cluster exhibited evidence of CNV aberrations, but there was a small subset of cells that were predicted to be diploid (Figure 3.1e; purple arrow). Thus, the diploid cells are likely a small cluster of fibroblasts, as they over express fibroblast markers relative to the other tumor cells (Supplemental figure 3.1).

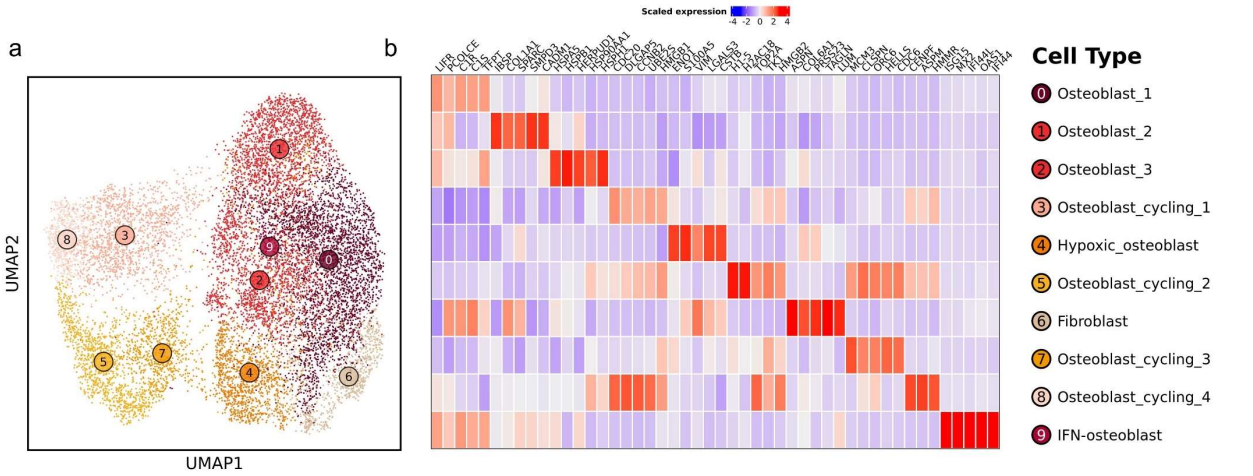
Dissection of the tumor and stromal populations reveals a distinct fibroblast cluster

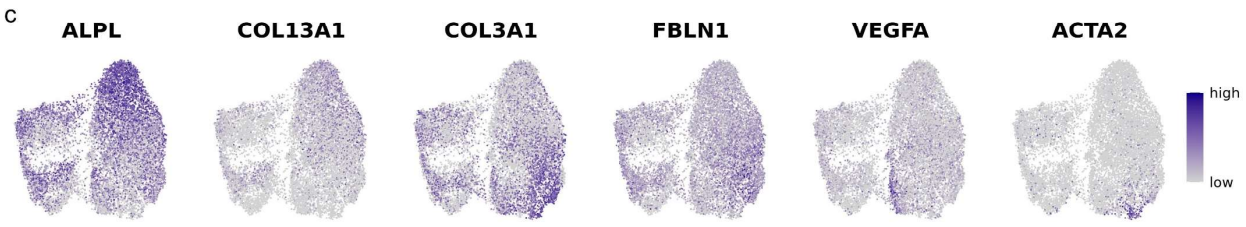
We next used independent reclustering of cycling tumor cells and tumor/fibroblasts which enabled further description of the cellular heterogeneity and the identification of a fibroblast

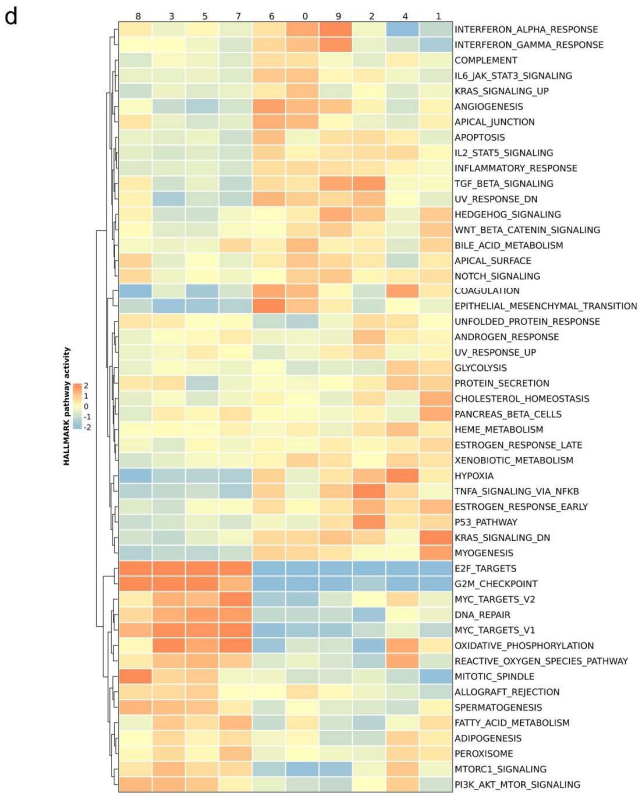
cluster. After filtering out additional low-quality cells, we identified 10 distinct cell clusters which we defined as 4 cycling malignant osteoblasts, 5 malignant osteoblast, and 1 fibroblast cluster (Figure 3.2a). The defining features for each cluster were then identified using a Wilcoxon Rank Sum test and the top 3-5 unique features were visualized (Figure 3.2b/c). Overall, the malignant osteoblast clusters exhibited a unique gene expression profile with collagen-type genes and alkaline phosphatase (ALPL) being major contributors to the gene signatures. We observed a small cluster of tumor cells (c9) that exhibited a gene expression pattern (OAS1, ISG15, OAS2) consistent with an interferon (IFN) gene signature, which was further supported through completion of GSEA using Hallmarks gene set terms (Figure 3.2b/d). Similar IFN signature cells have been reported in immune cells, but the observation of such a cluster in a tumor population has not been reported in human OS (27). Interpretation of GSEA further revealed that fibroblasts (c6) had the most pronounced EMT and angiogenesis signature, which suggests the fibroblasts may be tumor promoting. Additionally, GSEA acted to confirm the annotation of hypoxic osteoblast (c4), as the cluster exhibited the strongest hypoxic transcriptomic signature.

To confirm the identification of fibroblasts, we used module scoring with a human fibroblast gene list and confirmed Cluster 6 to have the strongest signature (Supplemental figure 3.2). After identifying a suspected fibroblast cluster, we completed differential gene expression (DGE) analysis contrasting fibroblasts (c6) and non-hypoxic osteoblasts (c0, c1, and c2) to better define the canine fibroblast gene signature (Figure 3.2e). While key fibroblast markers such as FAP and ACTA2 were identified, the top features consisted of SFRP2 and PRSS23 which have been recently reported to be associated with a fibroblast population involved in wound healing (28). To conclude our analysis of tumor cells, we sought to further investigate the transcriptomic signature of hypoxic osteoblasts (c4) by contrasting with non-hypoxic osteoblasts (c0, c1, and c2). Few differentially expressed gene were identified, suggesting cell types are similar, but subsequent pathway analysis identified enrichment of “response to oxygen levels” to be a top enriched pathway, suggesting the tumor cells are indeed hypoxic (Figure 3.2f). In summary, we were able

to resolve a population of fibroblasts through completion of independent reclustering, as well as define the transcriptional heterogeneity within malignant osteoblasts.







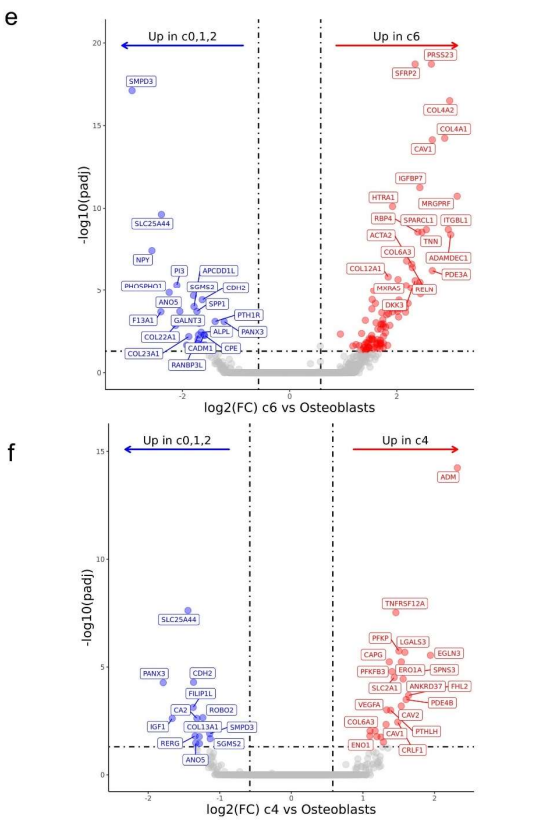


Figure 3.2: Independent reclustering of tumor & stomal cells reveals osteoblast heterogeneity and a distinct fibroblast cluster. (a) UMAP representation of tumor and stomal cells (n = 17,904 cells) depicting the 10 clusters identified though unsupervised clustering. (b) Heatmap depicting expression of the top 3-5 features that define each cluster. (c) Feature plots illustrating the log normalized counts. (d) Heatmap of scaled GSEA pathway activity for HALLMARK terms, with hierarchical clustering of rows and columns. Dendrogram of terms using Euclidean distance show on the left. (e/f) Volcano plots depicting the results of pseudobulk differential gene expression analysis for (e) fibroblasts (c6) versus osteoblasts (c0, c1, c2) and (f) hypoxic osteoblasts (c4) versus non-hypoxic osteoblasts (c0, c1, c2). The top 20 features (weighted by adjusted P value) are labeled.

Independent reclustering reveals a population of CXCL13^{+} exhausted T cells

To ensure we captured all biologically relevant T cell populations, we completed independent reclustering which led to the identification of 10 transcriptomically distinct clusters: 3 CD8 T cell, 4 CD4 T cell, 2 mixed CD4/CD8 T cell, and 1 NK cell clusters (Figure 3.3a/b). Next, we interrogated T cell subtypes using an approach that has been applied in human breast cancer and OS to describe T cell populations (5,29). We modified the gene lists used in previous applications to include signatures for cycling T cells, NK cells, and IFN-signature T cells that we recently established in circulating canine leukocytes (30). Overall, this approach proved to be consistent with annotations assigned using canonical markers (Figure 3.3c). Although the gene signatures were definitive for naïve CD4 T cells and cytotoxic CD8 T cells, other gene signature scores provided weaker support for their corresponding cell type. For instance, regulatory T cells (T_{regs}) and exhausted CD4 T cells ($CD4_{ex}$) both exhibited moderate enrichment for exhausted and costimulatory terms, with minimal distinction between the two T cell types. Additionally, the analysis revealed the presence of a T cell cluster with an interferon (IFN) gene signature.

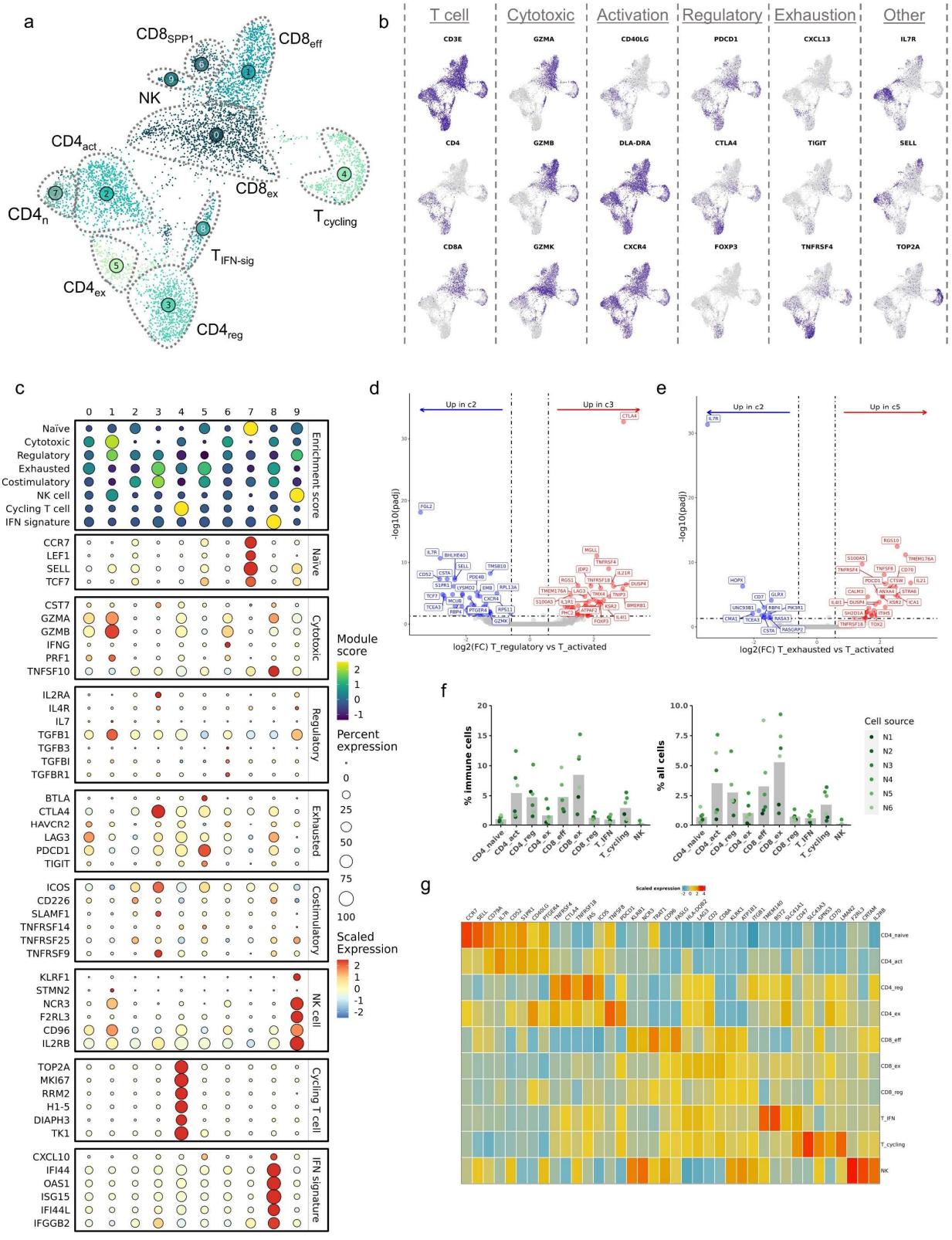

a
CD8SPP1
CD8eff
NK
CD4act
CD8ex
CD4n
Tcycling
TIFN-sig
CD4ex
CD4reg
b
T cell
Cytotoxic
Activation
Regulatory
Exhaustion
Other
CD3E
GZMA
CD40LG
PDCD1
CXCL13
IL7R
CD4
GZMB
DLA-DRA
CTLA4
TIGIT
SELL
CD8A
GZMK
CXCR4
FOXP3
TNFRSF4
TOP2A
c
Naïve
Cytotoxic
Regulatory
Exhausted
Costimulatory
NK cell
Cycling T cell
IFN signature
Enrichment score
CCR7
LEF1
SELL
TCF7
Naïve
CST7
GZMA
GZMB
IFNG
PRF1
TNFSF10
Cytotoxic
IL2RA
IL4R
IL7
TGFB1
TGFB3
TGFBI
TGFBR1
Regulatory
BTLA
CTLA4
HAVCR2
LAG3
PDCD1
TIGIT
Exhausted
ICOS
CD226
SLAMF1
TNFRSF14
TNFRSF25
TNFRSF9
Costimulatory
KLRF1
STMN2
NCR3
F2RL3
CD96
IL2RB
NK cell
TOP2A
MKI67
RRM2
H1-5
DIAPH3
TK1
Cycling T cell
CXCL10
IFI44
OAS1
ISG15
IFI44L
IFGGB2
IFN signature
Module score
Percent expression
Scaled Expression
d
Up in c2
Up in c3
-log10(padj)
log2(FC) T_regulatory vs T_activated
e
Up in c2
Up in c5
-log10(padj)
log2(FC) T_exhausted vs T_activated
f
% immune cells
% all cells
Cell source
N1
N2
N3
N4
N5
N6
CD4_naive
CD4_act
CD4_reg
CD4_ex
CD8_eff
CD8_ex
CD8_reg
T_IFN
T_cycling
NK
g
Scaled expression

Figure 3.3: Analysis of NK and T cells provides gene signatures for intra-tumoral regulatory and exhausted T cells. (a) UMAP representation of NK, CD4, CD8 T cell (n = 5,778 cells) depicting cell subtypes. (b) Feature plots depicting expression of 18 canonical T cell markers. (c) Dot plots depicting the scaled module score ("Enrichment score" facet) and the scaled expression of features associated with each of the 8 gene lists used to calculate the enrichment scores. (d/e) Volcano plots depicting the results of pseudobulk differential gene expression analysis for (e) regulatory T cells (c3) versus activated CD4 T cells (c2) and (f) CD4 exhausted T cells (c7) versus activated CD4 T cells (c2). The top 20 features (weighted by adjusted P value) are labeled for each plot. **(**f) Bar chart of the percent composition of each cluster as a percentage of total immune cells and percentage of total cells. Osteoclasts were included as an immune cell in the calculation. (g) Scaled expression of the top 3-5 features for each cluster. The selected features were chosen based on expression by FindMarkers() and the inclusion of the feature in the surfactome database listed as predicted surface expression.

After identifying each T cell subset, we completed pseudobulk conversion and DGE analysis to further establish the transcriptomic signatures of T_{regs} and $CD4_{ex}$. Comparisons between T_{regs} (c3) and activated CD4 T cells (c0) revealed high expression of IL21R, TNFRSF4, and TNFRSF18, with CTLA4 being the most definitive marker (Figure 3.3d). When repeating this analysis on $CD4_{ex}$ (c6) cells, we identified CXCL13, IL4I1, and TMEM176A to be defining features (Figure 3.3e). The expression of CTLA4 on T_{regs} and CXCL13/IL4I1 on $CD4_{ex}$ is widely reported in human literature, thus our analysis suggests the markers are conversed across species, while also providing complete gene signatures for the canine T cell subtypes (31).

Following comprehensive cell classification, we determined the cellular composition of each cell type as a percentage of immune cells and total cells in each sample (Figure 3.3f). Then we curated a heatmap of defining features predicted to be expressed on the cell surface with the objective of identifying potential cell markers to be used in alternative approaches, such as flow cytometry (Figure 3.4g) (32). With the caveat that transcript presence does not always correlate with protein expression, the analysis suggested that the generation or identification of cross-reactive antibodies specific to TNFRSF4 (OX-40), TNFRSF8 (CD30), and TMEM140 may prove valuable for further investigation of canine T_{regs}, $CD4_{ex}$, and $T_{IFN\text{-}sig}$, respectively. Together, the relative cellular percentages and potential surface markers provide a foundation for further functional study of the cell types identified in our transcriptomic analysis.

Mature regulatory dendritic cells are revealed in canine OS with the potential to modulate T cell mediated immunity

Five dendritic cell (DC) subtypes were identified when completing independent reclustering on $FLT3^+$ cells. The subtypes identified included conventional DC2s (cDC2; c0), cDC1s (c2), mature regulatory DCs (c1; mregDC), precursor DCs (c4; preDC), and plasmacytoid DCs (c3); pDC) (Figure 3.4a). Key features used to assign cell type identities included DNASE1L3 (cDC1), CCR7/IL4I1 (mature regulatory DCs), FCER1A (cDC2), and IGKC (pDC) (Figure 3.4b). The identification of preDCs was convoluted, as the cluster exhibited a gene signature consistent with human pDCs (IL3RA, IGKC). Based on a recent report, the population likely represents the equivalent of human preDC with the tendency to cluster with plasmacytoid DCs (33). Of note, we previously identified cDC2, cDC1, and pDC cell types in canine peripheral blood, however mregDCs (c1) were not observed, suggesting a potential tissue specificity (30). The identification of mregDCs, also reported as migratory (mig) DCs, is of note as this cell type is predicted to play a role in T cell suppression, which indicates the cell type is of therapeutic interest (13). Thus, we provide evidence that a key immune regulatory cell type is present in canine OS.

We next used hierarchical clustering and toll-like receptor expression to investigate differences between preDCs and pDCs. Hierarchical clustering indicated preDCs are more closely related to myeloid cDC2s and cDC1s, which suggests a myeloid lineage, while the pDCs were on their own clade, suggesting a lymphoid origin (Figure 3.4c). In humans, pDCs are reported to exhibit high expression of TLR9 and TLR7, which we identified to be highly expressed on pDCs (Figure 3.4d) (34). To ensure neither of the DCs were of B cell origin, we evaluated MS4A1 (CD20) expression and found it to be minimally expressed (Supplemental figure 3.3). We then used SCENIC to predict active regulons in each DC subtype (Supplemental figure 3.3). This analysis revealed RUNX2, a master regulator of pDC development, to be enriched in both pDCs and

preDCs (35). Overall, the data suggest canine preDCs are closely related to the recently defined plasmacytoid-like human preDCs and may represent useful model to study the cell type.

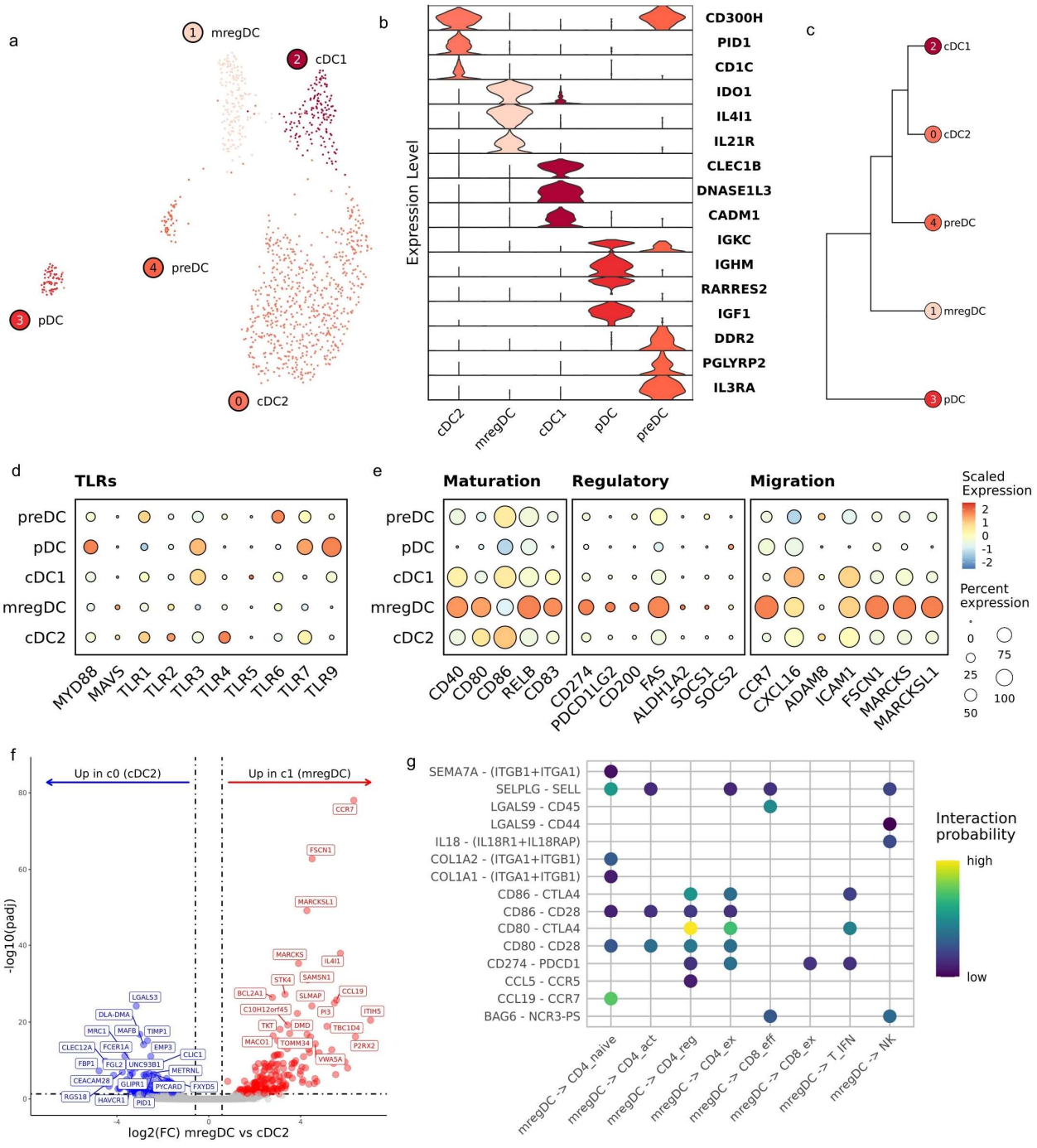


Figure 3.4. Mature regulatory dendritic cells reside in canine OS tumors and are predicted to suppress T cell function. (a) UMAP representation of dendritic cells (DCs) colorized by cell subtype. (b) Violin plots depicting expression of key DC features used for cell identification. (c) Dendrogram depicting results of hierarchical clustering of DC subtypes using log normalized expression. (d/e) Dot plots depicting scaled expression of toll-like receptors (TLRs), DC maturation, regulatory, and migratory features. (f) Volcano plots depicting the results of pseudobulk differential gene expression analysis for mature regulatory (mreg) DCs (c1) versus cDC2s (c0). The top 20 features (weighted by adjusted P value) are labeled for each plot. **(**g) Dot plot depicting the interaction probability (as determined using CellChat) of signaling networks for each mregDC-T cell interaction.

To confirm mregDCs exhibited a mature, immune regulatory transcriptomic signature, we used module scoring with gene lists previously applied to investigate human DC subtypes (13,27). As expected, the analysis revealed mregDCs to have a marked enrichment for migration, regulatory properties, and maturation (Figure 3.4e). Next, we used pseudobulk conversion and DGE analysis to further describe the gene expression profile of canine mregDCs (c1) relative to cDC2s (c0). The analysis revealed a distinct mregDC signature of CCR7, IL4I1, CCL19, and FSCN1 with substantial overlap to the human transcriptional program (Figure 3.4f) (13). With the precedent that mregDCs exhibit immune suppressive properties in human, we wanted to confirm this in dogs using cell-cell interaction inference. We used CellChat to evaluate interactions between mregDCs and T/NK cells (36). As expected, the analysis revealed enriched PD-1/PD-L1 and CTLA4/CD80 interactions between CD4 T regs, exhausted T cells and naïve T cells (Figure 3.4g). In summary, we present the transcriptomic signature of canine mregDCs and provide evidence that the cell type plays a regulatory role in canine OS.

Macrophage transcriptomic states supports a spectrum of cell types

Due to a high degree of transcriptional overlap between macrophages and OCs, we analyzed these two cell types in the same UMAP space. In doing so, our analysis highlighted the relatedness of OCs and macrophages which may have been overlook if analyzed independently. Through independent reclustering we identified 8 transcriptomically distinct

macrophage/monocyte populations which were annotated using modified nomenclature derived from Ma et al. (Figure 3.5a-c) (37). Activated TAMs (c0, TAM_ACT) and intermediate TAMs (c1, TAM_INT) did not fit into any of the macrophage subtypes presented in Ma et al., as such they were annotated based on an activated signature (CD5L, CD40, CD80) and an intermediate polarization signature, respectively. Additionally, unsupervised clustering divided lipid-associated (LA-) TAMs into two subclusters with one defined by C1QC^{hi} (c3) and the other SPP2^{hi} (c2). To better define the distinctions between the two LA-TAM populations we completed pseudobulk-based DGE analysis (Figure 3.5d). The analysis revealed IL2RA, CXCL10, and SERPING1 as key markers of C1QC^{hi} LA-TAM, while ENO1, LGALS3, and RBP4 defined SPP2^{hi} LA-TAMs. Based on the analysis, C1QC^{hi} LA-TAM appear to most closely resemble the definitions of LA-TAMS provided by Ma et al.

In addition to the recently proposed TAM nomenclature, we also used module scoring with pro- and anti-inflammatory gene lists to investigate the macrophage populations in a more traditional dichotomy (38). We identified the C1QC^{hi} LA-TAM (c3) cluster to have the strongest anti-inflammatory transcriptomic signature and M-MDSCs (c11) to exhibit the most prominent pro-inflammatory transcriptomic signature (Figure 3.5e). To further investigate the signatures of Clusters 11 and 3 we completed pseudobulk-based DGE analysis (Figure 3.5f). The genes upregulated in Cluster 11 exhibited overlap with the predefined gene set used to identify the cluster as pro-inflammatory, while also revealing IL-1B, S100A12, LTF, and VCAN as defining features. DGE analysis of the anti-inflammatory cluster exhibited less overlap with the gene list used to identify the cluster as anti-inflammatory (MRC1 was the only overlapping feature), but the analysis revealed APOE, IGF1, and complement receptors C1QA/B/C as key markers. The top features identified when contrasting Clusters 11 and 3 were then used to generate a heatmap to evaluate how the expression of these features varies across the other macrophage clusters (Figure 3.5g). Findings from the analysis suggest that there is a spectrum of macrophage phenotypes, which is consistent with human macrophage literature (39). As such, we next sought

to better define the heterogeneity of the macrophage populations without relying on predefined cell type gene signatures.

Gene set enrichment analysis was used to provide further insight into the inferred functional capacity of each subtype. Clusters 4, 7, and 11 clustered together based on pathway enrichment scores suggesting the three transcriptionally distinct clusters have similar underlying gene signatures. LA-TAMs (c2, c3) and intermediate TAMs (c1) exhibited the strongest scavenger receptor associated activation, suggesting a mature macrophage with immune suppressive properties (40). Several metabolic pathway terms were identified suggesting both $SPP2^{hi}$ LA-TAMS and intermediate TAMs favor oxidative phosphorylation and mitochondrial metabolism. $C1QC^{hi}$ LA-TAMs had a distinct profile suggestive of lipid and polysaccharide metabolism. Lastly, GSEA confirmed c10 to be consistent with IFN-TAMs based on strong enrichment of IFN signaling associated terms. In summary, we provide the transcriptional profiles of macrophages in the OS TME which provides a foundation for further investigation of the functional relevance of each cell type.

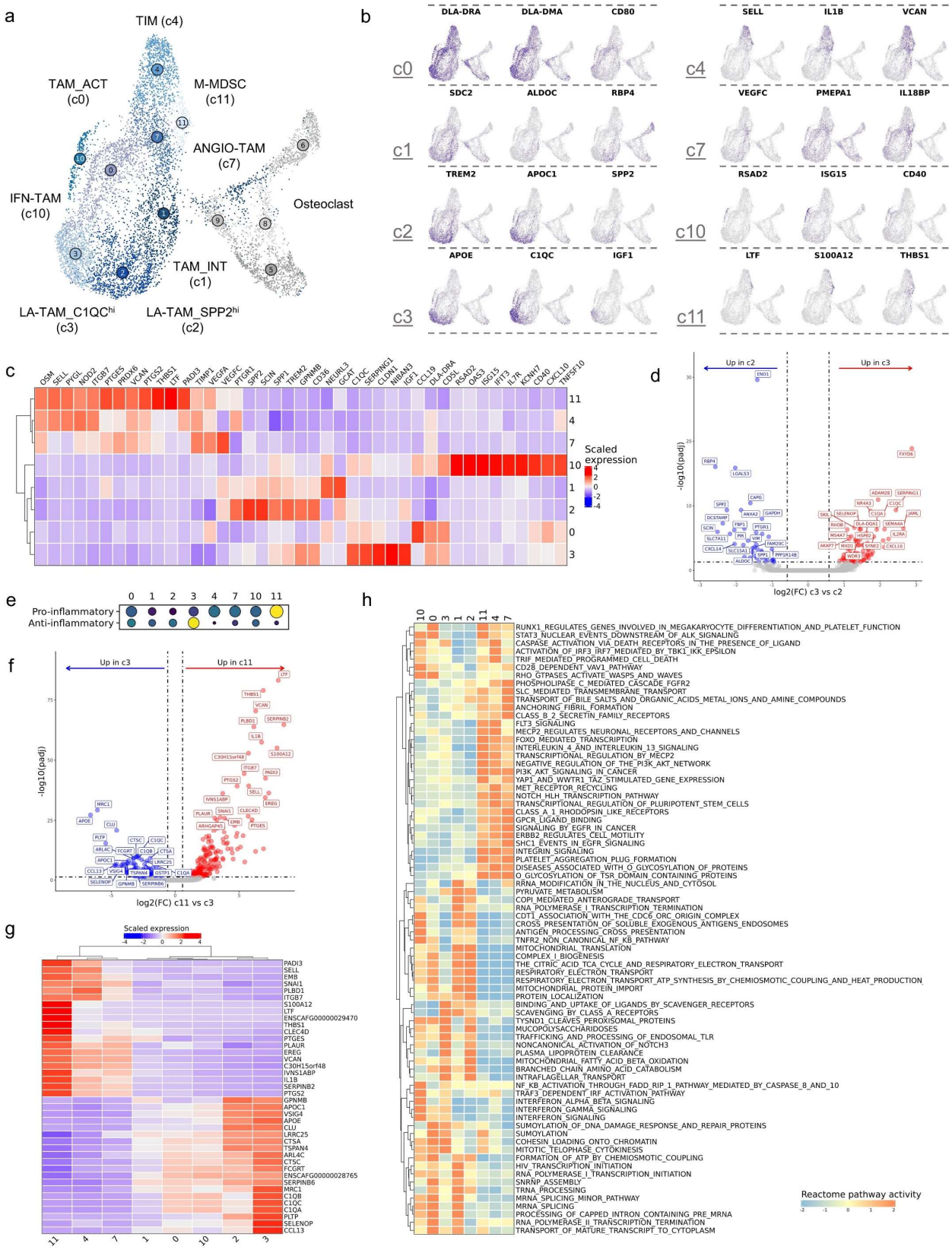

a
TIM (c4)
TAM_ACT (c0)
M-MDSC (c11)
ANGIO-TAM (c7)
IFN-TAM (c10)
Osteoclast
TAM_INT (c1)
LA-TAM_C1QC^hi (c3)
LA-TAM_SPP2^hi (c2)
b
DLA-DRA
DLA-DMA
CD80
SELL
IL1B
VCAN
c0
c4
SDC2
ALDOC
RBP4
VEGFC
PMEPA1
IL18BP
c1
c7
TREM2
APOC1
SPP2
RSAD2
ISG15
CD40
c2
c10
APOE
C1QC
IGF1
LTF
S100A12
THBS1
c3
c11
c
Scaled expression
d
Up in c2
Up in c3
-log10(padj)
log2(FC) c3 vs c2
e
Pro-inflammatory
Anti-inflammatory
f
Up in c3
Up in c11
-log10(padj)
log2(FC) c11 vs c3
g
Scaled expression
h
Reactome pathway activity
RUNX1_REGULATES_GENES_INVOLVED_IN_MEGAKARYOCYTE_DIFFERENTIATION_AND_PLATELET_FUNCTION
STAT3_NUCLEAR_EVENTS_DOWNSTREAM_OF_ALK_SIGNALING
CASPASE_ACTIVATION_VIA_DEATH_RECEPTORS_IN_THE_PRESENCE_OF_LIGAND
ACTIVATION_OF_IRF3_IRF7_MEDIATED_BY_TBK1_IKK_EPSILON
TRIF_MEDIATED_PROGRAMMED_CELL_DEATH
CD28_DEPENDENT_VAV1_PATHWAY
RHO_GTPASES_ACTIVATE_WASPS_AND_WAVES
PHOSPHOLIPASE_C_MEDIATED_CASCADE_FGFR2
SLC_MEDIATED_TRANSMEMBRANE_TRANSPORT
TRANSPORT_OF_BILE_SALTS_AND_ORGANIC_ACIDS_METAL_IONS_AND_AMINE_COMPOUNDS
ANCHORING_FIBRIL_FORMATION
CLASS_B_2_SECRETIN_FAMILY_RECEPTORS
FLT3_SIGNALING
MECP2_REGULATES_NEURONAL_RECEPTORS_AND_CHANNELS
FOXO_MEDIATED_TRANSCRIPTION
INTERLEUKIN_4_AND_INTERLEUKIN_13_SIGNALING
TRANSCRIPTIONAL_REGULATION_BY_MECP2
NEGATIVE_REGULATION_OF_THE_PI3K_AKT_NETWORK
PI3K_AKT_SIGNALING_IN_CANCER
YAP1_AND_WWTR1_TAZ_STIMULATED_GENE_EXPRESSION
MET_RECEPTOR_RECYCLING
NOTCH_HLH_TRANSCRIPTION_PATHWAY
TRANSCRIPTIONAL_REGULATION_OF_PLURIPOTENT_STEM_CELLS
CLASS_A_1_RHODOPSIN_LIKE_RECEPTORS
GPCR_LIGAND_BINDING
SIGNALING_BY_EGFR_IN_CANCER
ERBB2_REGULATES_CELL_MOTILITY
SHC1_EVENTS_IN_EGFR_SIGNALING
INTEGRIN_SIGNALING
PLATELET_AGGREGATION_PLUG_FORMATION
DISEASES_ASSOCIATED_WITH_O_GLYCOSYLATION_OF_PROTEINS
O_GLYCOSYLATION_OF_TSR_DOMAIN_CONTAINING_PROTEINS
RRNA_MODIFICATION_IN_THE_NUCLEUS_AND_CYTOSOL
PYRUVATE_METABOLISM
COPI_MEDIATED_ANTEROGRADE_TRANSPORT
RNA_POLYMERASE_I_TRANSCRIPTION_TERMINATION
CDT1_ASSOCIATION_WITH_THE_CDC6_ORC_ORIGIN_COMPLEX
CROSS_PRESENTATION_OF_SOLUBLE_EXOGENOUS_ANTIGENS_ENDOSOMES
ANTIGEN_PROCESSING_CROSS_PRESENTATION
TNFR2_NON_CANONICAL_NF_KB_PATHWAY
MITOCHONDRIAL_TRANSLATION
COMPLEX_I_BIOGENESIS
THE_CITRIC_ACID_TCA_CYCLE_AND_RESPIRATORY_ELECTRON_TRANSPORT
RESPIRATORY_ELECTRON_TRANSPORT
RESPIRATORY_ELECTRON_TRANSPORT_ATP_SYNTHESIS_BY_CHEMIOSMOTIC_COUPLING_AND_HEAT_PRODUCTION
MITOCHONDRIAL_PROTEIN_IMPORT
PROTEIN_LOCALIZATION
BINDING_AND_UPTAKE_OF_LIGANDS_BY_SCAVENGER_RECEPTORS
SCAVENGING_BY_CLASS_A_RECEPTORS
TYSND1_CLEAVES_PEROXISOMAL_PROTEINS
MUCOPOLYSACCHARIDOSES
TRAFFICKING_AND_PROCESSING_OF_ENDOSOMAL_TLR
NONCANONICAL_ACTIVATION_OF_NOTCH3
PLASMA_LIPOPROTEIN_CLEARANCE
MITOCHONDRIAL_FATTY_ACID_BETA_OXIDATION
BRANCHED_CHAIN_AMINO_ACID_CATABOLISM
INTRAFLAGELLAR_TRANSPORT
NF_KB_ACTIVATION_THROUGH_FADD_RIP_1_PATHWAY_MEDIATED_BY_CASPASE_8_AND_10
TRAF3_DEPENDENT_IRF_ACTIVATION_PATHWAY
INTERFERON_ALPHA_BETA_SIGNALING
INTERFERON_GAMMA_SIGNALING
INTERFERON_SIGNALING
SUMOYLATION_OF_DNA_DAMAGE_RESPONSE_AND_REPAIR_PROTEINS
SUMOYLATION
COHESIN_LOADING_ONTO_CHROMATIN
MITOTIC_TELOPHASE_CYTOKINESIS
FORMATION_OF_ATP_BY_CHEMIOSMOTIC_COUPLING
HIV_TRANSCRIPTION_INITIATION
RNA_POLYMERASE_I_TRANSCRIPTION_INITIATION
SNRNP_ASSEMBLY
TRNA_PROCESSING
MRNA_SPLICING_MINOR_PATHWAY
MRNA_SPLICING
PROCESSING_OF_CAPPED_INTRON_CONTAINING_PRE_MRNA
RNA_POLYMERASE_II_TRANSCRIPTION_TERMINATION
TRANSPORT_OF_MATURE_TRANSCRIPT_TO_CYTOPLASM

Figure 3.5: Tumor-associated macrophages exhibit a spectrum of transcriptional states with C1QChi LA-TAMs exhibiting an immune suppressive transcriptional profile. (a) UMAP representation of macrophage and osteoclast populations (n = 10,512). The four osteoclast (OC) subtypes CD320+ OC (CD320+ OC; c9), cycling OC (c4/c8), and mature OC (c6) are depicted, but greyed out. (b/c) Feature plots and heatmap of canonical features used to define cell types. (d) Volcano plots depicting the results of pseudobulk differential gene expression analysis for C1QChi LA-TAMs (c3) versus SPP2hi LA-TAMs (c2). The top 20 features (weighted by adjusted P value) are labeled for each plot. (e) Dot plot depicting module scoring of "Pro-inflammatory" and "anti-inflammatory" gene signatures. The size of dot indicates percentage of cells enriched for a gene set (larger dot more broadly enriched), and the color indicates strength of enrichment score (brighter color indicates stronger score). (f) Volcano plot depicting results of differential gene expression analysis when contrasting the cluster with the most pro-inflammatory gene signature (c11; M-MDSC) to the cluster with the most anti-inflammatory gene signature (c3; C1QChi LA-TAMs). The top 20 features (weighted by adjusted p-val) are labeled for each plot. (g) Heatmap of the top 20 features identified to be upregulated in in c11 relative to c3 and vice versa. The columns are ordered by hierarchical clustering (Euclidean distance), shown at top. (h) Heatmap of scaled GSEA pathway activity for Reactome terms, with hierarchical clustering of rows and columns. Dendrogram of terms using Euclidean distance show on the left.

Osteoclasts

Using the same UMAP space, we shifted our focus to further characterize osteoclast heterogeneity. Consistent with human and murine reports using single-cell RNA sequencing to characterize OCs, we identified 4 transcriptiomically distinct OC populations (4,5,41). The cycling OCs (c5/c8) in our dataset likely correspond to previously annotated pre/progenitor OCs, while the mature OCs are consistently annotated across previous reports (Figure 3.6a/b). CD320+ OCs (transcobalamin receptor expressing OCs, c9) were not described in the macrophage or osteoclast clusters of any previous reports and may represent a distinct canine specific cell type or more likely a previously unresolved OC subtype (Figure 3.6a/b). Due to the similarity of OCs and macrophages we completed hierarchical clustering of normalized counts to confirm the unsupervised clustering results (Figure 3.6c). The secondary analysis was consistent with unsupervised clustering and further suggests Clusters 5, 6, 8, and 9 represent OCs.

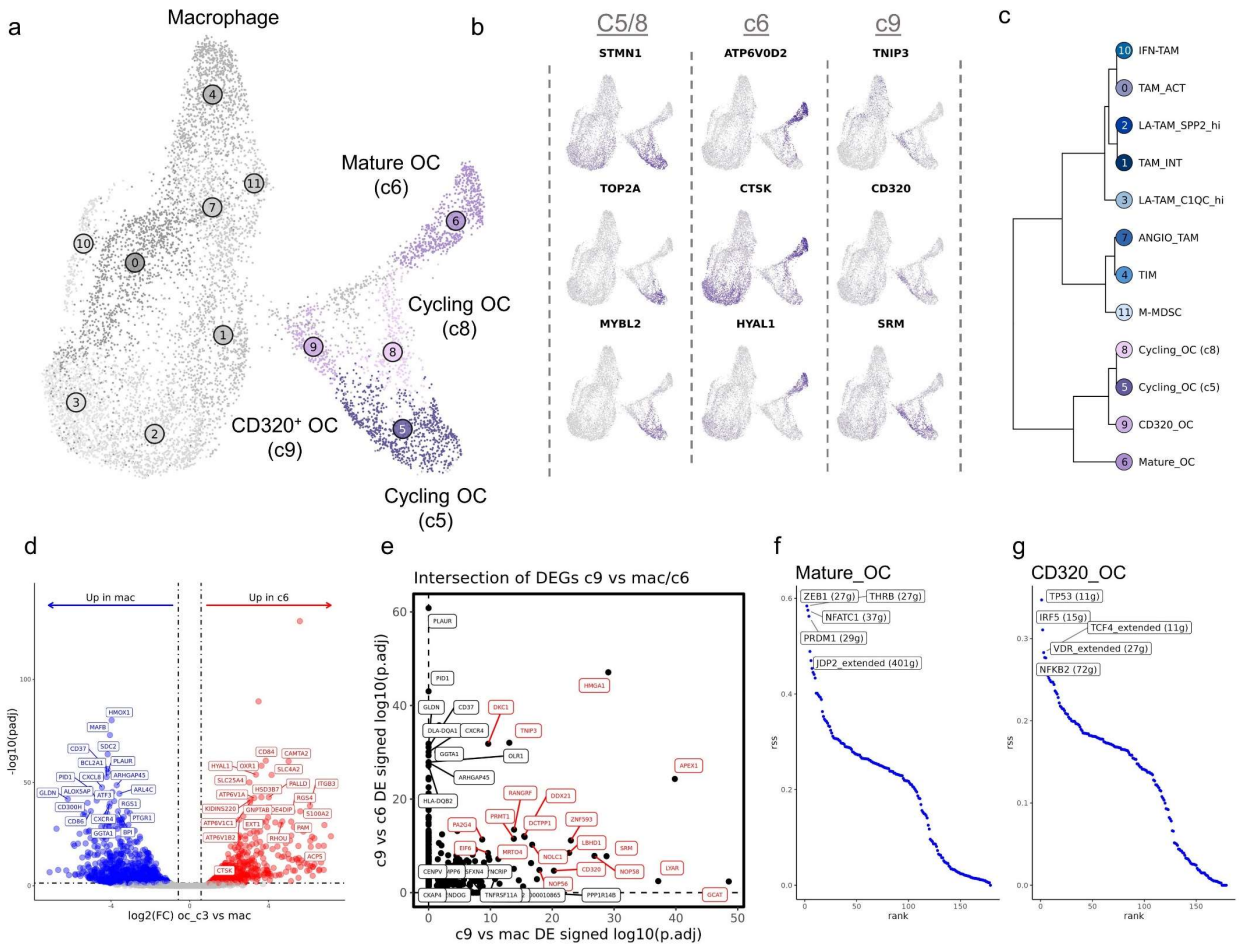


Figure 3.6: Four transcriptional distinct osteoclast subtypes are identified using unsupervised clustering. (a) UMAP representation of macrophage and osteoclast populations (n = 10,512). The 6 macrophage subtypes defined in Figure 5 are greyed out. (b) Feature plots of canonical features used to define cell types. (c) Dendrogram depicting results of hierarchical clustering of OC subtypes using log normalized expression. (d) Volcano plots depicting results of differential gene expression analysis when comparison Mature OC (c6) versus macrophage (c0, c1, c2, and c3). The top 20 features (weighted by adjusted P value) are labeled for each plot, ACP5 and CTSK additionally labeled in (d). (e) Scatter plot depicting the signed log10 adjusted P value when contrasting CD320+ OC versus macrophage (c0, c1, c2, and c3) (x-axis) and CD320+ OCs versus Mature OC (y-axis). The top 20 co-enriched features are labeled in red. (f/g) Scatter plot depicting regulon enrichment score for active transcription factors in mature OC (f) and CD320+ OC clusters, as determined using SCENIC. The parenthetic in the label indicates the number of genes associated with the regulon.

To confirm the mature OC classification and provide a canine specific transcriptomic signature, we completed DGE analysis. When comparing mature OCs (c6) to macrophages (c0,

c1, c2, c3) we identified canine mature OCs to be defined by ATP6V1C1, CD84, HYAL1, and CAMTA2 expression, which subsequent GSEA analysis suggested an association with bone resorption and remodeling (Figure 3.6d). We next completed DGE analysis contrasting CD320^{+} OCs (c9) with macrophages and mature OCs to better define the previously unidentified CD320+ OCs. By evaluating the intersection of the differentially expressed genes upregulated in CD320^{+} OCs we were able to determine CD320+ OCs is defined by HMGA1, TNIP3, and CD320 expression (Figure 3.6e). The analysis provided further evidence that c9 is an OC cluster based on TNFRSF11A (RANK) being identified as upregulated in the macrophage contrast, but not when comparing to mature OCs. Lastly, we use SCENIC's regulon specificity scoring to better understand the master regulators active in mature OCs and CD320^{+} OCs. As an indication of the accuracy of the analysis, ZEB1 and NFATC1, known regulators of OC development, were enriched in mature OCs (Figure 3.6f) (42,43). Therefore, the master regulators identified in CD320^{+} OCs (TCF4, IRF5, and TP53) suggest a unique cell type with a distinct transcriptional profile (Figure 3.6g). IRF5 and TCF4 are both reported to be involved in myeloid cell differentiation, with IRF5 activity associated with a pro-inflammatory macrophage phenotype. IRF5 knockout in macrophages has been reported to reduce the in vitro differentiation of macrophages in osteoclasts (44). Therefore, the CD320^{+} OC cluster may represent a population of macrophage-like OC precursors that have the potential to differentiate into mature OCs.

Transcript abundance of widely used IHC macrophage markers have distinct specificity to myeloid cells

In contrast to other cancer types, there have been multiple reports in humans and dogs suggesting increased TAM infiltrates are correlated with reduced metastasis rates and increased survival (10,11). Contrarily, other groups completing similar analysis have concluded that increased macrophage infiltrates have a negative impact on OS clinical outcomes (45). Given the conflicting nature of previous reports we sought to employ our dataset to investigate which cell

types express the transcript of key markers used in the aforementioned studies. To complete this analysis, we profiled macrophages, dendritic cells, and osteoclasts for the expression of widely used canine (MSR1 aka CD204 and AIF1 aka Iba1) and human (CD163 and CD68) macrophages markers (Figure 3.7a). Although this analysis is limited to transcript abundance and does not evaluate protein expression, we found that CD163 transcript expression is most specific for macrophages. CD68 expression was detected in monocytes, macrophages, and OCs, with a remarkably high expression levels in mature OCs. The expression of CD68 on mature OCs has been reported previously in humans (46). AIF1 (Iba1) was the most non-specific marker with diffuse expression across all cell types, except mature OCs. Lastly, CD204 (MSR1) was largely specific to macrophages, but extended to $CD320^{+}$ OCs and M-MDSCs. Together, the data presented here suggest the specificity of each marker varies widely and is it possible that conflicting reports may be a result of studies evaluating different populations of myeloid cells.

Given the degree of heterogeneity within the myeloid compartment, we used a Wilcoxon Rank Sum test to identify features that define each cell type, then selected for features predicted to be expressed on the cell surface (Figure 3.7b). Overall, the data suggest there is substantial overlap in expression of most features. Despite the overlap, some candidate markers identified include ADAM28 for LA-TAM_C1QC^{hi}, TNFSF13B for IFN-TAMs, and CD84 for mature OCs. Lastly, we calculated the relative percentages of each cell type to further facilitate cell identification (Figure 3.7c). Together the data presented here act as a foundation to further investigate the role of myeloid cells in OS biology.

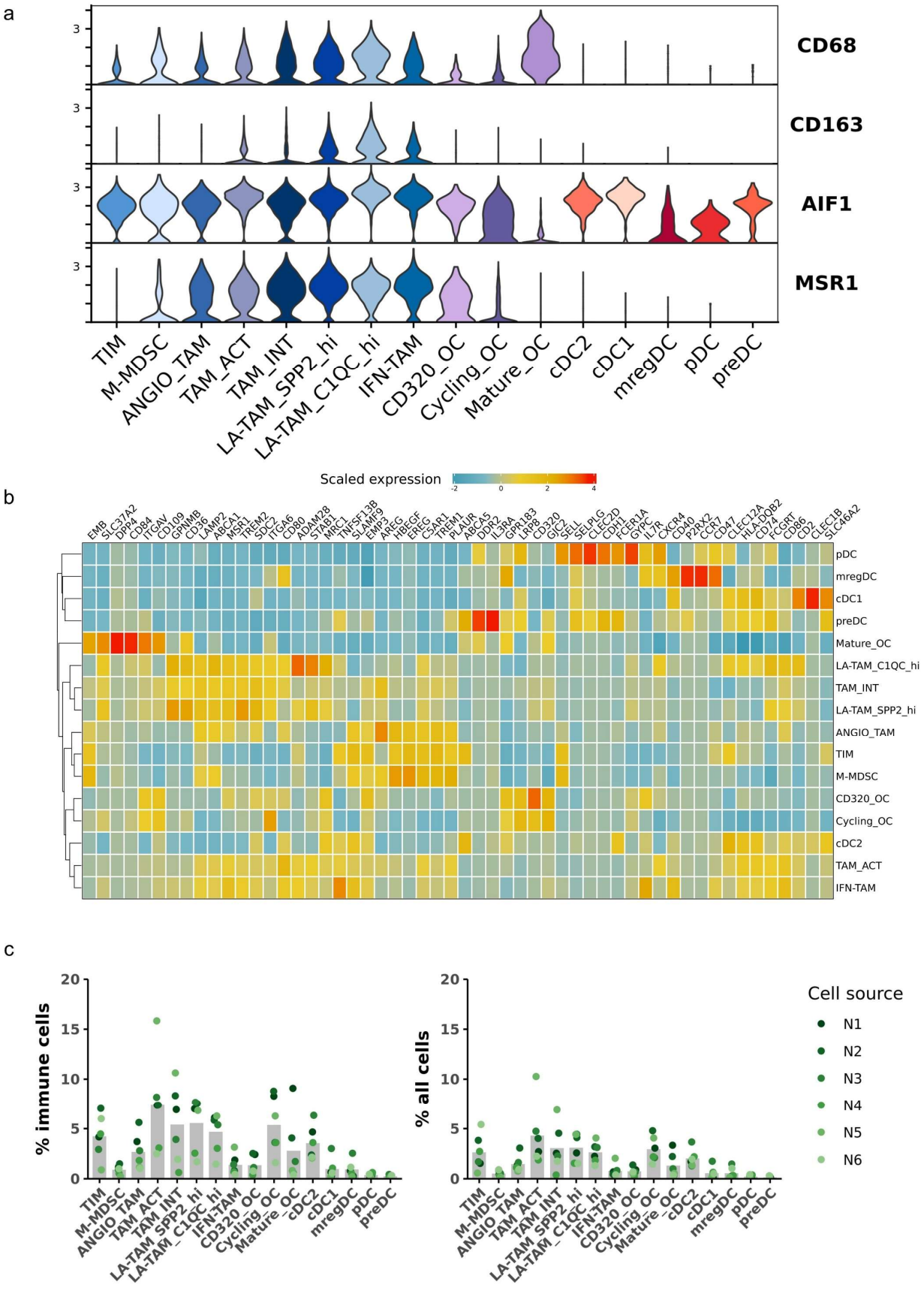

a
CD68
CD163
AIF1
MSR1
TIM
M-MDSC
ANGIO_TAM
TAM_ACT
TAM_INT
LA-TAM_SPP2_hi
LA-TAM_C1QC_hi
IFN-TAM
CD320_OC
Cycling_OC
Mature_OC
cDC2
cDC1
mregDC
pDC
preDC
b
Scaled expression
pDC
mregDC
cDC1
preDC
Mature_OC
LA-TAM_C1QC_hi
TAM_INT
LA-TAM_SPP2_hi
ANGIO_TAM
TIM
M-MDSC
CD320_OC
Cycling_OC
cDC2
TAM_ACT
IFN-TAM
c
% immune cells
% all cells
Cell source
N1
N2
N3
N4
N5
N6

Figure 3.7: Transcript abundance of widely used IHC macrophage markers have distinct specificity to myeloid cells. (a) Violin plots of four commonly used immunohistochemistry macrophage markers widely used to evaluate macrophage infiltrates in canine and human osteosarcoma. (b) Scaled expression of the top 3-5 features for each cluster. The selected features were chosen based on expression by FindMarkers() and the inclusion of the feature in the surfactome database listed as predicted surface expression. (c) Bar chart of the percent composition of each cell type as a percentage of total immune cells and percentage of total cells. Osteoclasts were included as an immune cell in the calculation.

Cell-cell interaction analysis confirms TAMs are key players in immune regulatory pathways

Following cell identification through independent reclustering of major cell types, we evaluated the cell-cell interaction networks using CellChat. Across the 30 cell types included in the analysis, a total of 15,405 inferred interactions across 61 signaling networks were identified. We then used the number of interactions and the predicted interaction strength of incoming (express receptor) versus outgoing (express ligand) signals to infer the activity of cells within the TME (Figure 3.8a). The top three cell types involved in the greatest number of interactions were fibroblasts, mature OCs, and endothelial cells. While these populations were predicted to be highly active in cell-cell communications, we next categorized the significantly enriched networks as “immune specific”, “immune related”, and “non-immune” to investigate if certain cell types were more active in a subset of networks (Figure 3.8b). As expected, the malignant osteoblasts and stromal cells were largely predicted to be involved in “non-immune” interactions, while “immune specific” interactions were largely confined to macrophages and dendritic cells with strong outgoing interactions. This suggests that macrophages and dendritic cells make up key immune influencer populations within the OS TME.

We next subset on immune cells and evaluated the interactions of known immune regulatory pathways. As expected, we identified macrophages and DCs to be the primary cells predicted to be sending signals, while exhausted and regulatory CD4 T cells were predicted to be receiving the most signals (Figure 3.8c). This limited analysis of immune regulatory pathways,

which included PD-1/PD-L1 and CD80-CD86/CTLA4 interactions, revealed TIMs, mregDCs, and IFN-TAMs to have the most interactions, while activated ($CD5L^{+}$) macrophages were predicted to have the strongest outgoing signals. When evaluating the PD-L1 network, mregDCs, TIMs, and IFN-TAMs were determined to express the highest levels of PD-L1 and were predicted to interact with exhausted and regulatory T cells (Figure 3.8d). The CD80/CD86 networks involved a larger portion of myeloid cells, with all CD4 T cells being influenced by the interactions (Figure 3.8e). Overall, TIMs, IFN-TAMs, and mregDCs appear to be key contributors to the suppression of T cell mediated immunity and represent potential immunotherapy targets.

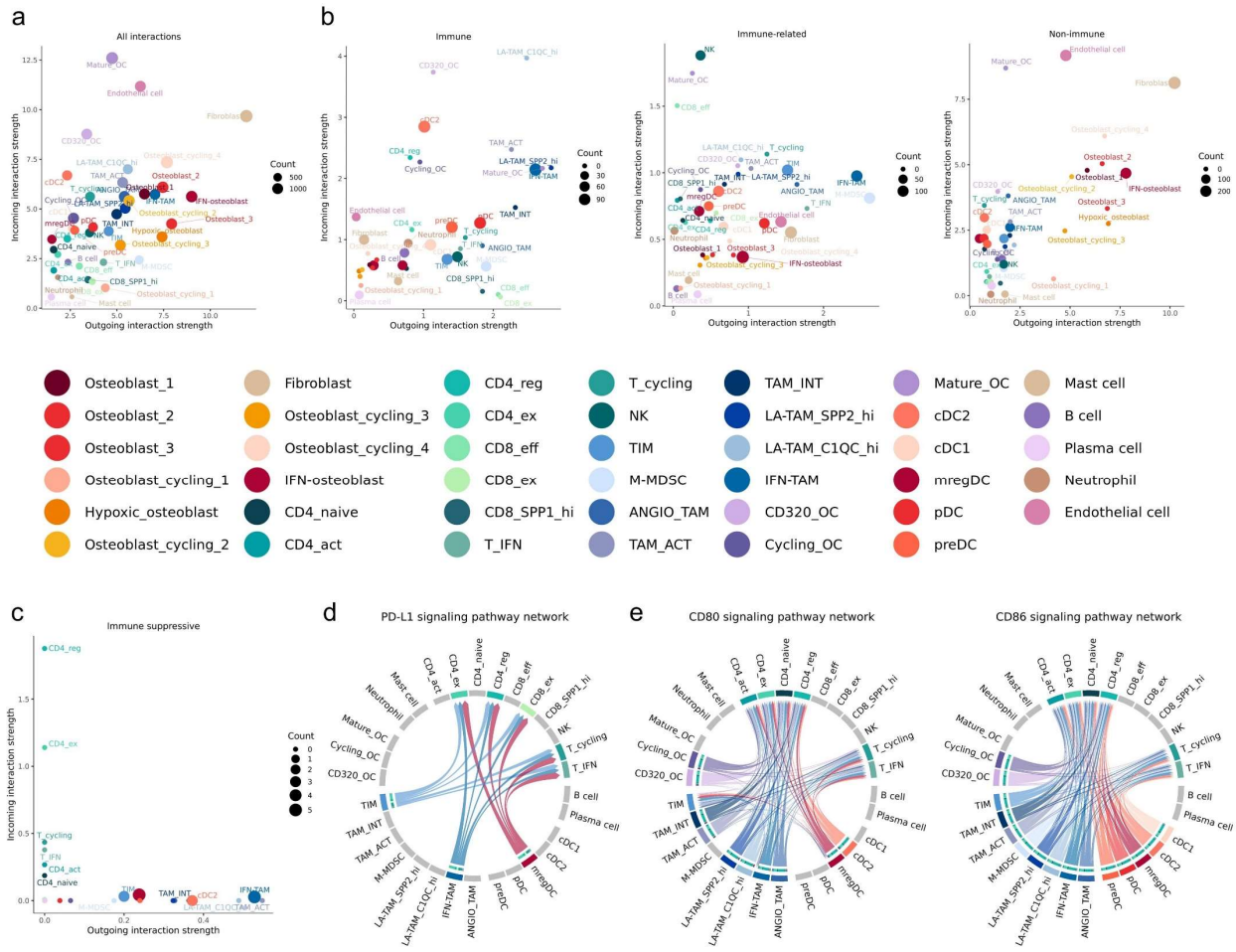


Figure 3.8: Cell-cell interaction analysis reveals fibroblasts as a key communicating cell type and identifies TAMs as immune regulatory. (a) Scatter plot depicting the strength of outgoing (x-axis) and incoming (y-axis) signals for all cell type and calculated using all enriched signaling networks. (b) Scatter plot depicting the strength of outgoing (x-axis) and incoming (y-axis) signals for all cell type and calculated for three subdivision of all enriched networks, “immune”, “immune-related”, and “non-immune”. (c) Scatter plot depicting the strength of outgoing (x-axis) and incoming (y-axis) signals for immune cells and calculated using three immune suppressive signaling networks (PD1/PDL1 & CD80-CD86/CTLA4). (d/e) Circos plots of immune regulatory networks identified using CellChat with PDL1 (d) and CD80/CD86 (e) networks depicted. The arrow origin represents expression of a ligand, while terminal arrow indicates expression of a receptor. Cell types not involved in the network are greyed out.

Comparisons to human OS reveal high degree of similarity in cell type gene signatures between species

Lastly, we obtained 6 publicly available treatment-naïve human OS scRNA datasets to complete a cross-species analysis (GSE162454) (4). The two datasets were integrated using a SCTransform workflow, which is reported to overcome genome annotation differences between species (17). SCT normalized count matrices were then used to complete a direct comparison of canine and human cells. Hierarchical clustering of the SCT normalized data revealed a high degree of similarities between species, with major clades containing similar cell types based on pre-integration annotations (Figure 3.9a). All canine lymphocytes paired 1:1 with their human counterpart, as did endothelial cells and fibroblasts. Intriguing discrepancies between species included the placement of mast cells, which clustered into separate clades. Overall, macrophages clustered in the same clade, but due to differences in annotation levels, many cell types did not pair off into terminal clades.

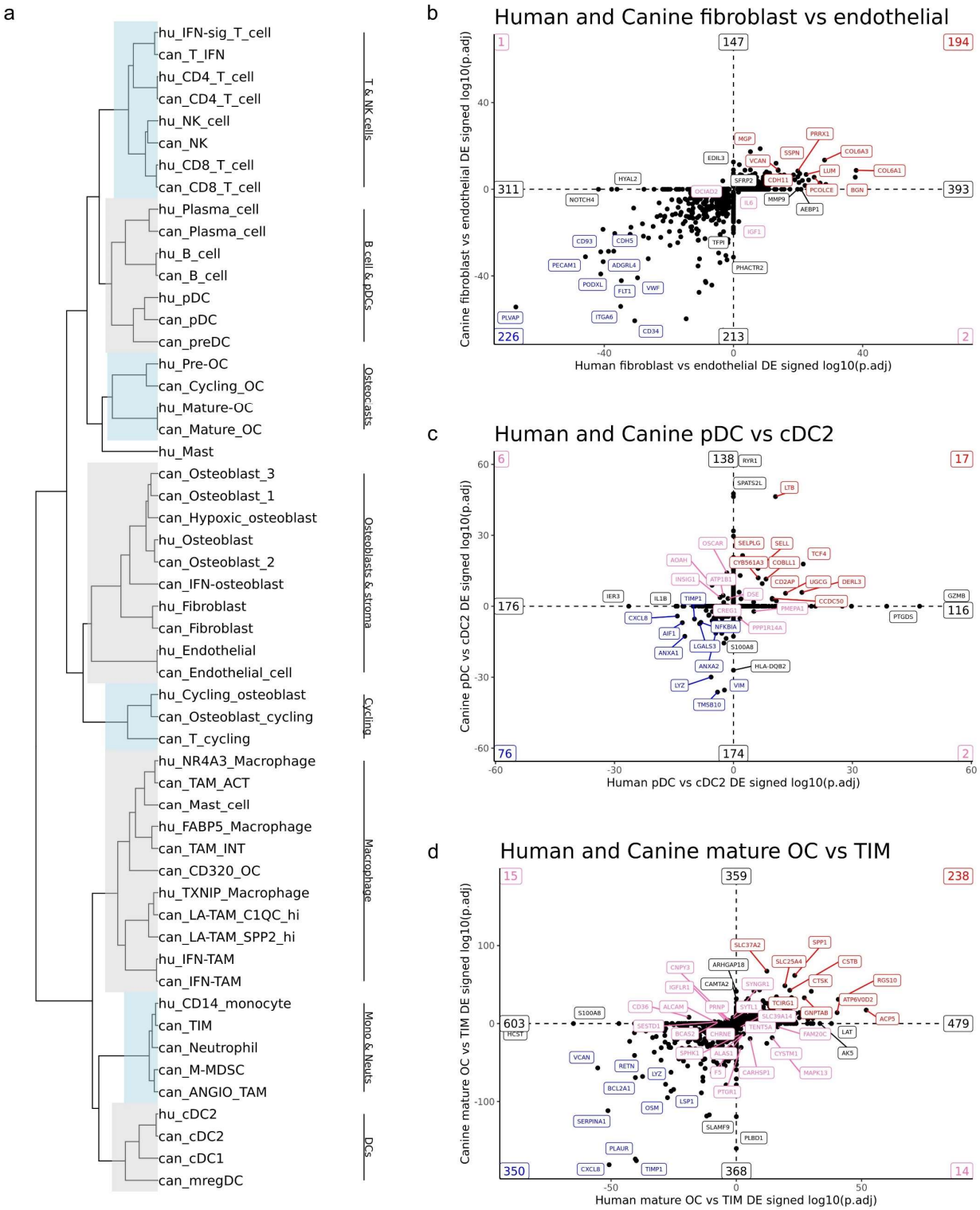

a
hu_IFN-sig_T_cell
can_T_IFN
hu_CD4_T_cell
can_CD4_T_cell
hu_NK_cell
can_NK
hu_CD8_T_cell
can_CD8_T_cell
hu_Plasma_cell
can_Plasma_cell
hu_B_cell
can_B_cell
hu_pDC
can_pDC
can_preDC
hu_Pre-OC
can_Cycling_OC
hu_Mature-OC
can_Mature_OC
hu_Mast
can_Osteoblast_3
can_Osteoblast_1
can_Hypoxic_osteoblast
hu_Osteoblast
can_Osteoblast_2
can_IFN-osteoblast
hu_Fibroblast
can_Fibroblast
hu_Endothelial
can_Endothelial_cell
hu_Cycling_osteoblast
can_Osteoblast_cycling
can_T_cycling
hu_NR4A3_Macrophage
can_TAM_ACT
can_Mast_cell
hu_FABP5_Macrophage
can_TAM_INT
can_CD320_OC
hu_TXNIP_Macrophage
can_LA-TAM_C1QC_hi
can_LA-TAM_SPP2_hi
hu_IFN-TAM
can_IFN-TAM
hu_CD14_monocyte
can_TIM
can_Neutrophil
can_M-MDSC
can_ANGIO_TAM
hu_cDC2
can_cDC2
can_cDC1
can_mregDC
T & NK cells
B cell & pDCs
Osteoclasts
Osteoblasts & stroma
Cycling
Macrophage
Mono & Neuts
DCs
b
Human and Canine fibroblast vs endothelial
Canine fibroblast vs endothelial DE signed log10(p.adj)
Human fibroblast vs endothelial DE signed log10(p.adj)
c
Human and Canine pDC vs cDC2
Canine pDC vs cDC2 DE signed log10(p.adj)
Human pDC vs cDC2 DE signed log10(p.adj)
d
Human and Canine mature OC vs TIM
Canine mature OC vs TIM DE signed log10(p.adj)
Human mature OC vs TIM DE signed log10(p.adj)

Figure 3.9: Integration with human OS scRNA dataset suggests the transcriptomic signatures of key cell types are conserved in dogs. (a) Hierarchical clustering of human ("hu_" prefix) and canine ("can_" prefix) cell types using SCT normalized data of a human-canine integrated dataset. (b-d) Scatter plot comparing the signed log10(adjusted P value) of significantly upregulated or downregulated genes identified when comparing (b) human fibroblasts versus human endothelial cells and canine fibroblasts versus canine endothelial cells, (c) human pDC versus human cDC2s and canine pDC versus canine cDC2s, and (d) human mature OCs versus human CD14 monocytes and canine mature OCs versus canine TIMs. In (b-d) conserved upregulated features are in the top right quadrant (top 10 in red labels) and conserved downregulated are in the bottom left quadrant (top 10 in blue labels). Conflicting features – up in human but down in dog (bottom right quadrant) and down in human but up in dog (top left quadrant) – are labeled in pink (top 10 in each quadrant). Features up or down in one species, but not a DE in the other, fell on the axis and the top 2 on each axis direction are labeled in black. The numbers in the corners and at the ends of axis lines represent how many features fell in that region.

It is important to note that the approach used to complete hierarchical clustering only evaluated the top 3000 variable features with gene homologies between species. As such, the analysis is superficial in nature and likely will not capture subtle differences. To further compare transcriptional programs across species we used an analysis approach adopted from Scheyltjens et al. (47). The approach used pseudobulk conversion and DGE analysis between two cell populations in each species, then signing the adjusted P value to determine if transcriptomic signatures were conserved. When contrasting fibroblasts and endothelial cells in each species, we found an overwhelming amount of overlap in gene expression patterns with key endothelial cell markers (PLVAP, CD34, and PECAM1) enriched in both species (Figure 3.9b). Top features conserved in fibroblasts included VCAN, COL6A1, and LUM, while key features such as FAP and ACTA2 were also conserved. Interesting discrepancies included the expression of HYAL2 and NOTCH being key features in human endothelial cells, but nonsignificant in canine endothelial cells. Completion of the same analysis on plasmacytoid DCs and cDC2s revealed TCF4 to be up in pDCs and BATF expression upregulated in cDC2s, which is consistent with human literature (48). An intriguing distinction between species included the high expression of GZMB and PTGDS (prostaglandin D2 synthase) in human pDCs, but not identified in dogs. Lastly, we applied the approach to compare mature osteoclasts with human CD14 monocytes or canine TIMs. As

expected, mature osteoclasts were defined by CSTK, ACP5, ATP6V0D2, while monocytes in both species were defined by CXCL8, OSM, and LYZ expression. Interestingly, canine monocytes exhibited high expression of SLAMF9 and PLBD1, while human monocytes had high S100A8 and HCST expression. In summary, we present a comprehensive comparison of human and canine OS cell types, which suggests a high degree of consistency across the two species, however we also present evidence of distinct transcriptional programs in pDCs and monocytes.

Discussion

In the present study we complete a comprehensive dissection of canine osteosarcoma (OS) using single-cell RNA sequencing, which revealed the complex network of cells within the tumor microenvironment (TME). Through analysis of 6 treatment-naïve canine OS samples we were able to identify 30 distinct immune cell types, 9 unique tumor populations, 1 cluster of fibroblasts, and 1 population of endothelial cells. We described the transcriptomic heterogeneity within malignant osteoblasts, identified cell types that have not been previously reported in dogs, and applied our data set to investigate CD204 and CD163 expression which has been reported to have prognostic implication in dogs and humans. Importantly, we completed a cross-species analysis with human OS and found a high degree of cell type gene signature homology, which acts to further support the translational relevance of spontaneously occurring canine OS as a model to study translational immunotherapeutics. Ultimately, the data presented here acts as a molecular roadmap of the OS tumor microenvironment, from which further scientific questions can be asked.

The dog has been regarded as an excellent model for human OS, but due to regent limitations and other technical considerations, direct evidence of a conserved OS tumor microenvironment has been limited. By obtaining a publicly available human OS dataset we were able to directly compare the cellular composition and relative relatedness of cell types between

species. We found lymphocytes to exhibit the highest degree of conservation, while mast cells and macrophages exhibited the most pronounce differences. It is possible that the observed differences in macrophages may be a result of discrepancies between cellular annotations, but our data may also suggest there are species-specific distinctions in the myeloid cell compartment of OS. Overall, the data presented here suggest a high degree of conserved cell types within the OS tumor microenvironment and provide further evidence of the validity of canine OS as a model for human disease.

Our analysis revealed the presence of many rare cell populations, including mature regulatory DCs, exhausted CD4/CD8 T cells, and IFN-TAMs which may open avenues for further investigation of these cell populations using the transcriptomic signatures identified in the study. With the caveat that transcript expression does not always correlate with protein expression, we used the surfactome reference database to present possible surface markers for the further study of these rare cell populations. In addition to flow cytometry-based assays, which rely on cross-reactive or canine-specific antibodies, the transcriptomic signatures presented here provide a reference to applied toward the use of deconvolution algorithms (such as CIBERSORTx and TIMER) when evaluating bulk RNA sequencing data obtained from canine OS samples (49,50).

While the single-cell RNA atlas presented here provides key insights into canine OS, the dataset is not without limitations. First, although we sampled male and female dogs across a range of ages, our dataset still only consisted of 6 dogs and may not fully represent all cells found in canine OS. Secondly, the tumor sample obtained from one dog (naïve dog 6) exhibited markedly more neutrophils relative to other samples which may suggest there was blood or bone marrow contamination. Lastly, cellular annotations largely relied on human gene signatures due to the lack of canine specific data available. This may have exaggerated the degree of similarity in OS between species and may have resulted in the forcing of distinct canine-specific cell types into human nomenclature subtypes.

The data presented here represent a valuable resource for comparative oncology research which can be used to shed light on the intricacies of the OS tumor microenvironment. One goal of this project is to make the data accessible to the greater research community, as such, all data is made readily available and multiple avenues are provided for researchers to explore and use the dataset (see data availability statement). Our comparisons between human and canine OS, revealed the conserved nature of tumor cell types and suggest that the dog is a valuable model for the investigation of novel OS therapeutics.

References

1. Tarone L, Barutello G, Iussich S, Giacobino D, Quaglino E, Buracco P, et al. Naturally occurring cancers in pet dogs as pre-clinical models for cancer immunotherapy. Cancer Immunol Immunother. 2019;68:1839–53.

2. Schiffman JD, Breen M. Comparative oncology: what dogs and other species can teach us about humans with cancer. Philos Trans R Soc B Biol Sci. 2015;370(1673):20140231.

3. Zheng GXY, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, et al. Massively parallel digital transcriptional profiling of single cells. Nat Commun. 2017;8(1):14049.

4. Liu Y, Feng W, Dai Y, Bao M, Yuan Z, He M, et al. Single-cell transcriptomics reveals the complexity of the tumor microenvironment of treatment-naive osteosarcoma. Front Oncol. 2021;11:709210.

5. Zhou Y, Yang D, Yang Q, Lv X, Huang W, Zhou Z, et al. Single-cell RNA landscape of intratumoral heterogeneity and immunosuppressive microenvironment in advanced osteosarcoma. Nat Commun. 2020;11(1):6322.

6. Wu C-C, Beird HC, Andrew Livingston J, Advani S, Mitra A, Cao S, et al. Immuno-genomic landscape of osteosarcoma. Nat Commun. 2020;11(1):1008.

7. Li W, Li S, Chen IX, Liu Y, Ramjiawan RR, Leung C-H, et al. Combining losartan with radiotherapy increases tumor control and inhibits lung metastases from a HER2/neu-positive orthotopic breast cancer model. Radiat Oncol. 2021;16(1):1–12.

8. Zhao Y, Cao J, Melamed A, Worley M, Gockley A, Jones D, et al. Losartan treatment enhances chemotherapy efficacy and reduces ascites in ovarian cancer models by normalizing the tumor stroma. Proc Natl Acad Sci. 2019;116(6):2210–9.

9. Regan DP, Chow L, Das S, Haines L, Palmer E, Kurihara JN, et al. Losartan blocks osteosarcoma-elicited monocyte recruitment, and combined with the kinase inhibitor toceranib, exerts significant clinical benefit in canine metastatic osteosarcoma. Clin Cancer Res. 2022;28(4):662–76.

10. Withers SS, Skorupski KA, York D, Choi JW, Woolard KD, Laufer-Amorim R, et al. Association of macrophage and lymphocyte infiltration with outcome in canine osteosarcoma. Vet Comp Oncol. 2019;17(1):49–60.

11. Gomez-Brouchet A, Illac C, Gilhodes J, Bouvier C, Aubert S, Guinebretiere JM, et al. CD163-positive tumor-associated macrophages and CD8-positive cytotoxic lymphocytes are powerful diagnostic markers for the therapeutic stratification of osteosarcoma patients: An immunohistochemical analysis of the biopsies fromthe French OS2006 phase 3 trial. https://doi.org/101080/2162402X20171331193 [Internet]. 2017 Sep 2 [cited 2023 Jun 7];6(9). Available from: https://www.tandfonline.com/doi/abs/10.1080/2162402X.2017.1331193

12. Buddingh EP, Kuijjer ML, Duim RAJ, Bürger H, Agelopoulos K, Myklebost O, et al. Tumor-Infiltrating Macrophages Are Associated with Metastasis Suppression in High-Grade Osteosarcoma: A Rationale for Treatment with Macrophage Activating AgentsImpact of Macrophages on Osteosarcoma Metastases. Clin Cancer Res. 2011;17(8):2110–9.

13. Maier B, Leader AM, Chen ST, Tung N, Chang C, LeBerichel J, et al. A conserved

dendritic-cell regulatory program limits antitumour immunity. Nature. 2020;580(7802):257–62.

14. Liu W, Hu H, Shao Z, Lv X, Zhang Z, Deng X, et al. Characterizing the tumor microenvironment at the single-cell level reveals a novel immune evasion mechanism in osteosarcoma. Bone Res. 2023;11(1):4.

15. Hao Y, Hao S, Andersen-Nissen E, Mauck III WM, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. Cell. 2021;184(13):3573–87.

16. McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. Cell Syst. 2019;8(4):329–37.

17. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol. 2019;20(1):1–15.

18. Zappia L, Oshlack A. Clustering trees: a visualization for evaluating clusterings at multiple resolutions. Gigascience. 2018;7(7):giy083.

19. Thomas DD, Lacinski RA, Lindsey BA. Single-cell RNA-seq reveals intratumoral heterogeneity in osteosarcoma patients: A review. J Bone Oncol. 2023;100475.

20. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):1–21.

21. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innov. 2021;2(3):100141.

22. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci. 2005;102(43):15545–50.

23. Wu D, Smyth GK. Camera: a competitive gene set test accounting for inter-gene correlation. Nucleic Acids Res. 2012;40(17):e133–e133.

24. Gao R, Bai S, Henderson YC, Lin Y, Schalck A, Yan Y, et al. Delineating copy number and clonal substructure in human tumors from single-cell transcriptomes. Nat Biotechnol. 2021;39(5):599–608.

25. Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, et al. SCENIC: single-cell regulatory network inference and clustering. Nat Methods. 2017;14(11):1083–6.

26. Suo S, Zhu Q, Saadatpour A, Fei L, Guo G, Yuan G-C. Revealing the critical regulators of cell identity in the mouse cell atlas. Cell Rep. 2018;25(6):1436–45.

27. Cheng S, Li Z, Gao R, Xing B, Gao Y, Yang Y, et al. A pan-cancer single-cell transcriptional atlas of tumor-infiltrating myeloid cells. Cell. 2021;184(3):792–809.

28. Tabib T, Huang M, Morse N, Papazoglou A, Behera R, Jia M, et al. Myofibroblast transcriptome indicates SFRP 2hi fibroblast progenitors in systemic sclerosis skin. Nat Commun. 2021;12(1):4384.

29. Chung W, Eum HH, Lee H-O, Lee K-M, Lee H-B, Kim K-T, et al. Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. Nat Commun. 2017;8(1):1–12.

30. Ammons DT, Harris RA, Hopkins LS, Kurihara J, Weishaar K, Dow S. A single-cell RNA sequencing atlas of circulating leukocytes from healthy and osteosarcoma affected dogs. Front Immunol. 2023;14.

31. Sadik A, Patterson LFS, Öztürk S, Mohapatra SR, Panitz V, Secker PF, et al. IL4I1 is a metabolic immune checkpoint that activates the AHR and promotes tumor progression. Cell. 2020;182(5):1252–70.

32. Bausch-Fluck D, Goldmann U, Müller S, van Oostrum M, Müller M, Schubert OT, et al. The in silico human surfaceome. Proc Natl Acad Sci. 2018;115(46):E10988–97.

33. See P, Dutertre C-A, Chen J, Günther P, McGovern N, Irac SE, et al. Mapping the human DC lineage through the integration of high-dimensional techniques. Science (80-). 2017;356(6342):eaag3009.

34. Bao M, Liu Y-J. Regulation of TLR7/9 signaling in plasmacytoid dendritic cells. Protein Cell. 2013;4(1):40–52.

35. Sawai CM, Sisirak V, Ghosh HS, Hou EZ, Ceribelli M, Staudt LM, et al. Transcription factor Runx2 controls the development and migration of plasmacytoid dendritic cells. J Exp Med. 2013;210(11):2151–9.

36. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan C-H, et al. Inference and analysis of cell-cell communication using CellChat. Nat Commun. 2021;12(1):1088.

37. Ma R-Y, Black A, Qian B-Z. Macrophage diversity in cancer revisited in the era of single-cell omics. Trends Immunol. 2022;

38. Jablonski KA, Amici SA, Webb LM, Ruiz-Rosado J de D, Popovich PG, Partida-Sanchez S, et al. Novel markers to delineate murine M1 and M2 macrophages. PLoS One. 2015;10(12):e0145342.

39. Mosser DM, Edwards JP. Exploring the full spectrum of macrophage activation. Nat Rev Immunol. 2008;8(12):958–69.

40. Eisinger S, Sarhan D, Boura VF, Ibarlucea-Benitez I, Tyystjärvi S, Oliynyk G, et al. Targeting a scavenger receptor on tumor-associated macrophages activates tumor cell killing by natural killer cells. Proc Natl Acad Sci. 2020;117(50):32005–16.

41. Omata Y, Okada H, Uebe S, Izawa N, Ekici AB, Sarter K, et al. Interspecies single-cell RNA-seq analysis reveals the novel trajectory of osteoclast differentiation and therapeutic targets. JBMR plus. 2022;6(7):e10631.

42. Takayanagi H, Kim S, Koga T, Nishina H, Isshiki M, Yoshida H, et al. Induction and activation of the transcription factor NFATc1 (NFAT2) integrate RANKL signaling in terminal differentiation of osteoclasts. Dev Cell. 2002;3(6):889–901.

43. Zhu L, Tang Y, Li X, Kerk SA, Lyssiotis CA, Feng W, et al. A Zeb1/MtCK1 metabolic axis controls osteoclast activation and skeletal remodeling. EMBO J. 2023;42(7):e111148.

44. Yang J, Park O, Kim J, Kwon Y, Yun C, Han SH. Modulation of macrophage subtypes by IRF5 determines osteoclastogenic potential. J Cell Physiol. 2019;234(12):23033–42.

45. Koirala P, Roth ME, Gill J, Piperdi S, Chinai JM, Geller DS, et al. Immune infiltration and PD-L1 expression in the tumor microenvironment are prognostic in osteosarcoma. Sci Rep. 2016;6(1):30093.

46. Ashley JW, Shi Z, Zhao H, Li X, Kesterson RA, Feng X. Genetic ablation of CD68 results in mice with increased bone and dysfunctional osteoclasts. PLoS One. 2011;6(10):e25838.

47. Pombo Antunes AR, Scheyltjens I, Lodi F, Messiaen J, Antoranz A, Duerinck J, et al. Single-cell profiling of myeloid cells in glioblastoma across species and disease stage reveals macrophage competition and specialization. Nat Neurosci. 2021;24(4):595–610.

48. Collin M, Bigley V. Human dendritic cell subsets: an update. Immunology. 2018;154(1):3–20.

49. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. Nat Biotechnol. 2019;37(7):773–82.

50. Li T, Fu J, Zeng Z, Cohen D, Li J, Chen Q, et al. TIMER2. 0 for analysis of tumor-infiltrating immune cells. Nucleic Acids Res. 2020;48(W1):W509–14.

CHAPTER 4: Molecular characterization of tumor microenvironment induced changes to infiltrating immune cells

Summary

Immune cells are keystones in host defenses against the progression of malignant tumors. The selective pressure that immune cells put on tumors leads to immune escape and modulation of immune cell function to create an environment favorable to tumor growth and progression. The aim of this chapter was to use single-cell RNA sequencing of tumor-infiltrating immune cells and blood immune cells to investigate how the tumor microenvironment modulates the transcriptomic programs of intratumoral immune cells. For this study, blood leukocytes from dogs with primary osteosarcoma tumors were used as a reference to evaluate how the OS TME impacts immune cell transcriptomes. Through completion of differential gene expression analysis and cell type proportion analysis we identified (1) increased relative abundances of exhausted T cells, (2) upregulation of T cell exhaustion markers, (3) increased immune suppressive molecule expression on myeloid cells, and (4) perturb B cell gene expression profiles suggestive or activation or dysfunction. Furthermore, we confirmed that mature regulatory DCs are only found in tumor tissue and not in circulation. Overall, the analysis provides broad insight into transcriptomic changes induced by the tumor microenvironment in each major immune cell type.

Introduction

Systemic and local immune responses have direct impacts on the clinical outcomes of cancer patients. For instance, overall survival and disease-free survival have been demonstrated to be positively corelated with T lymphocyte infiltration across multiple tumor types (1). The infiltration of immune cells into a tumor is dependent on the early recruitment of cells to the tumor and adequate antitumor adaptive immunity has proven to be fundamental for successful

immunotherapy-based treatments (2,3). Although certain tumor microenvironments recruit and maintain enough immune cells to benefit from immunotherapy, other tumors, such as osteosarcoma (OS), are notoriously defined by poor immune cell infiltration (4). The poor immune infiltration has contributed to the stagnation of OS therapies and there is a need to understand mechanisms underlying the inefficient immune infiltration in OS.

The inadequacies of immunotherapy in the treatment of OS have been attributed to an immune suppressive microenvironment that prevents generation of sufficient antitumor immunity. Specifically, OS has been described to have a notable protumoral gene expression pattern consisting of highly expressed programmed cell death protein 1 (PD-L1), indoleamine 2,3-dioxygenase (IDO) and interleukin-10 (IL-10) (5). These programs result in the expansion of immune suppressive myeloid cells as well as promote T cell dysfunction and exhaustion. Despite a basic understanding of underlying immune suppressive mechanisms active in OS, there is a need for a deeper characterization of the OS immune landscape.

The advent of single-cell RNA sequencing has enabled the dissection of complex tissues and the discovery of previously unrecognized mechanisms of immune cell modulation in tumor microenvironments (TME). The robustness of data integration across tissue types has enabled the direct comparison of cell types in different biological tissues (6–8). Using normal tissues, such as circulating leukocytes, transcriptomic changes to immune cells induced by the TME can be investigated in detail.

In the present study we complete a meta-analysis using single-cell RNA datasets of tumor-infiltrating leukocytes and circulating leukocytes in dogs with osteosarcoma. The analysis revealed effector CD8 T cell and CD4 T cells broadly upregulated exhaustion markers (HAVCR1, LAG3, IL4I1) and identified the relative proportion of exhausted and regulatory T cells to be greater in tumors. We demonstrated a recently defined immune suppressive dendritic cell type, mature regulatory DCs, are found in the TME, but not in circulation. Furthermore, we confirmed that, in dogs, neutrophils, monocytes, and dendritic cells upregulate immune modulatory

molecules (increased OSM, PD-L1, and IDO1) relative to their circulating counterparts. Overall, the analysis presented here provides insight into TME induced changes to immune cell gene expression and reveals mechanisms of immune modulation within the canine osteosarcoma tumor microenvironment.

Methods

Read mapping and quantification

Raw FASTQ data were aligned to the canine genome and count matrices generated using a Cell Ranger analysis pipeline (version 6.1.2, 10x Genomics). The reference package used a filtered CanFam3.1 ensembl annotation (gtf) file as described in chapter 2 and 3. The count matrix for each sample was then used for downstream analysis.

Data filtering and integration

For each sample, the count matrix was imported into R using the Read10X() function then converted to a Seurat object using the CreateSeuratObject() function (9). Dead/poor quality cells were filtered out by only retaining cells that met quality control thresholds. For the tumor dataset thresholds used were: 200 < nFeature_RNA < 5500, percent.mt < 12.5, and 100 < nCount_RNA < 15000. For the blood dataset thresholds used were: 200 < nFeature_RNA < 3500, percent.mt < 20, and 500 < nCount_RNA < 20000. After initial filtering, DoubletFinder was used to remove putative cell doublets (10). Samples within each tissue were integrated into separate objects using the SCTransform() and integration workflow while regressing out the percent mitochondrial reads. After each tissue type had all samples integrated into a respective object, previous cell type annotations were transferred and any cells lacking annotation were removed. At this point the tumor dataset was subset to only retain immune cells while also excluding mast cells, osteoclasts, IFN T cells, and macrophage from analysis. The blood dataset was filtered to exclude eosinophils, double negative T cells, gd T cells, IFN signature CD4 T cells, basophils, and CD34+ unclassified

cells. All filtered samples were then integrated into one object using the SCTransform() and integration workflow while regressing out the percent mitochondrial reads. The top 2500 variably expressed features were used as integration anchors then unsupervised clustering was completed using the following parameters: res = 1.2, dims = 40, n.neighbors = 50, min.dist = 0.5). Ideal clustering parameters were identified using the R package clustree. Dimension reduction and visualization was then completed, and the data were projected using 2-dimensional, non-linear uniform manifold approximation and projection (UMAP) plots.

Cell classification

Cell annotations as previously reported were transferred to the integrated dataset using the unique cell barcodes associated with each sample (Chapters 2 & 3). Unsupervised clustering was completed then the composition of cell types (based on the transferred classifications) within each cluster were examined. Clusters in which one cell type made of greater than 75% of the cluster the label was directly transferred. When conflicting cell types fell within a cluster, a new cell identity was assigned to capture the unsupervised clustering results of the current study.

Feature visualization

Feature expression was visualized using feature plots. Selected features were chosen based on the identification of a feature to be differentially expressed when contrasting tumor-infiltrating and blood leukocytes. Feature plots depict normalized expression for each feature and are on variable scales. When visualizing expression between tumor and blood leucocytes on a UMAP, tissues were down sampled to the value equal to the tissue which contributed the fewest number of cells.

Cell abundance analysis

All cell abundance comparisons were made using percentage of parent. To make statistical inferences on changes in cell abundance two-sided Wilcoxon Rank Sum tests were used and adjusted P values were calculated using Holm–Bonferroni correction. The classification of cell abundances as “common”, “skewed”, or “unique” was completed by evaluating the P value

and the abs(log2fc) between the two tissues. If a cell type had an adjusted P value > 0.05 the cell type was considered “common”, if the P value was < 0.05 with an abs(log2fc) less than 3, the cell type was classified as “skewed”, while “unique” cell types were required to have a P value < 0.05 and abs(log2fc) > 3. The classification scheme is designed to identify unique cell types based on the idea that cell types with low to no representation will result in an exaggerated log2-fold change, and in turn pass the high-end cutoff.

Differential gene expression analysis

Differential gene expression analysis (DGE) was completed using pseudobulk conversion followed by a DESeq2 pipeline. Prior to running DESeq2, low abundance features, defined as features with less than 10 raw counts across all cells sampled, were filtered out. Features were determined to be significantly upregulated or downregulated if the P value was less than 0.01 and abs(log2fc) greater than 0.58. Any subsequent pathway analysis was completed using lists of upregulated genes and the clusterProfiler package. Gene ontology and Reactome gene sets were used, and terms were considered enriched if they achieved a P value less than 0.05.

Identification of tissue signatures and removal from analysis

After completing DGE analysis for each cell type, we observed a bias for tumor-infiltrating immune cells to have increased expression of extracellular matrix associated features. Given previous reports documenting the release of mRNA during sample processing and subsequent incorporation with individual cells resulting in confounding background tissue signatures, we devised a strategy to identify and remove features associated with background tissue signatures (6,11). To accomplish this, we completed differential expression analysis within each major immune cell population, then identified found the intersection of across all cell types. This revealed 38 features to be upregulated and three to be downregulated (TXNIP, PPBP, MITD1) across all tumor-infiltrating cell types (Table 4.1). For tissues signature estimation, we considered features with a P value less than 0.05 and abs(log2fc) greater than 0.58 to be significant. We excluded

these 38 features and a list of 108 platelet associated features then repeated DGE analysis (Table 4.2).

Table 4.1. Tumor tissue signature

GADD45B	COQ10B.1	PCOLCE	CRYAB	CREM	GADD45B	ZFAND5
RGS1	PRDM1	SELENOM	CSTB	APOE	RGS1	ATF3
VEGFA	COL1A1	ID1	HMOX1	COL1A2	VEGFA	TUBA1B
NR4A1	CYCS	COL3A1	ENSCAFG00000043646	COL6A1	NR4A1	RAB13
SKIL	IBSP	DNAJB1	ENSCAFG00000012982	LMNA	SKIL	TFPI2
HSP90AA1	SERPINH1	PRDX6	ENSCAFG00000018277	SERPINB6	ZFAND5	SPP1
CTSK	ATF3	DCN				

Table 4.2. Platelet signature

ABRACL	ENSCAFG00000004260	DLA-64	H3-3A	NAA10	SEC11C	TMBIM6
ACOT7	ENSCAFG00000008221	DLA88	HPSE	NAP1L1	SELP	TPM4
ACTB	ENSCAFG00000011134	DYNLRB1	IGF2BP3	NRGN	SEPTIN7	TPT1
ACTG1	ENSCAFG00000015217	EIF1	ITGA2B	PFN1	SERF2	TREM2
ADA	ENSCAFG00000017655	ELOB	ITGA6	PIF1	SH3BGRL2	TUBA1B
AMD1	ENSCAFG00000029155	ENO1	ITGB1	PLEK	SH3GLB1	TUBB1
ARPC3	ENSCAFG00000030286	FERMT3	JMJD6	PPBP	SMDT1	UBB
ATF4	ENSCAFG00000037735	FTH1.1	LGALS1	PRKAR2B	SRP14	VASP
ATG3	ENSCAFG00000042554	FTL	LRRFIP1	RABGAP1L	SSR4	VCL
ATP5PF	ENSCAFG00000042598	FYB1	MARCHF2	RALBP1	SUB1	VDAC3
B2M	ENSCAFG00000043577	GABARAP	METTL7A	RAP1B	SVIP	VPS28
BBLN	ENSCAFG00000046637	GAS2L1	MGST3	RASA3	TBXAS1	YWHAH
BIN2	CD47	GNAS	MSN	RGS10	THBS1	
C9H17orf64	CD63	GNG11	MTDH	RGS18	TIMP1	
CAT	CFL1	GP6	MYH9	RSU1	TKT	
CCL14	GP6	NAP1L1	SH3BGRL2	VASP	TLN1	

Data and software availability

No new data was generated in this study. A portion of data from chapter 2 and all data generated in chapter 3 was used to complete this study. A project specific GitHub page containing all analysis code and software versions used to analyze the data presented in this manuscript is available at https://github.com/dyammons/canine_blood_VS_tils_scRNA.

Results

Overview of study

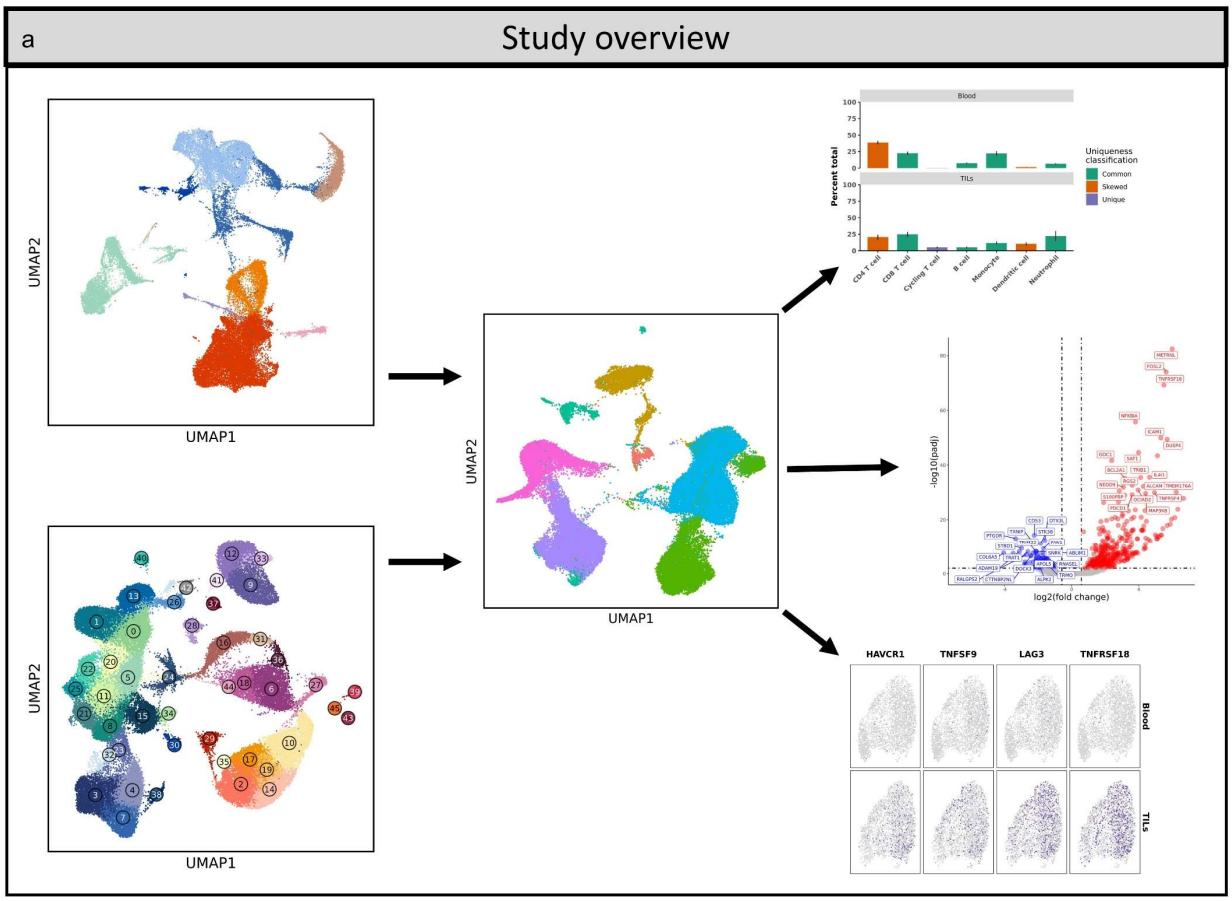
Annotated data obtained from circulating leukocytes of dogs diagnosed with osteosarcoma (OS) and a subset of OS tumor-infiltrating leukocytes (TILs) were integrated into one dataset. Due to the differences in clustering resulting from the inclusion of cells from two different tissue types, we established consensus labels by evaluating the percentage of cell types in each cluster and transferred the labels that best fit clustering of the integrated dataset. In instances where none of the transferred annotations matched the new clustering, we re-annotated cell types with nomenclature based on the original data sources.

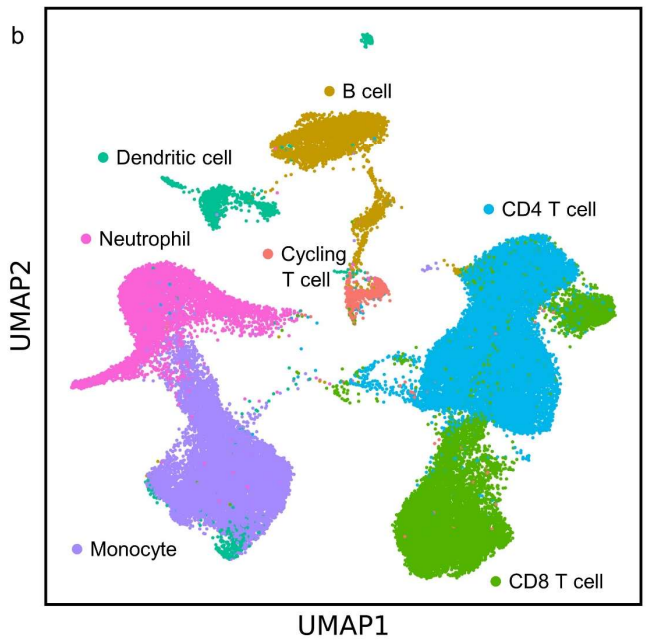
Our analysis approach consisted of subsetting and completing independent reclustering for each major cell population, then evaluating the relative cell type abundances and differentially expressed genes between TILs to blood leukocytes (Figure 4.1a). When we completed differential gene expression (DGE) analysis between TILs and blood leukocytes we observed profound tissue signatures driving differential expression. This tissue signature likely arose from free mRNA present in the solution during cell capture, so we identified tumor and platelet signatures and excluded all associated features from DGE analysis. The overarching goal of this study was to evaluate how the OS tumor microenvironment (TME) modulates immune cell transcriptomes.

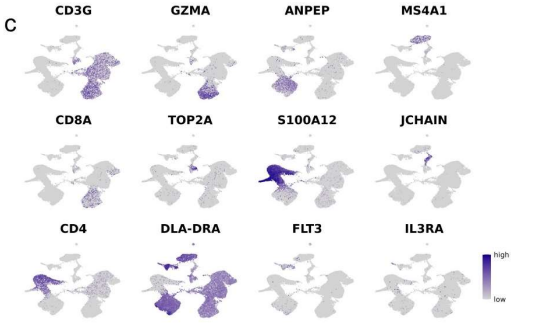
Summary of the integrated dataset

The dataset used for this study consisted of 10 blood leukocyte samples (n = 42,292 cells) and TILs from 6 treatment-naïve OS dogs (n = 11,431 cells). Although macrophages, mast cells, and osteoclasts are immune cells found in the OS TME, they were excluded from this analysis because a normal counterpart is not found in circulation. Therefore, analysis of the OS TME impacts on immune cells was limited to six major cell types: neutrophils, monocytes, CD4 T cells, CD8 T cells, B cells, and dendritic cells (Figure 4.1b/c). Each major cell type was annotated using

transferred labels from the previous annotations (in chapter 2 & 3) and a minor cluster of cycling T cells was also observed.







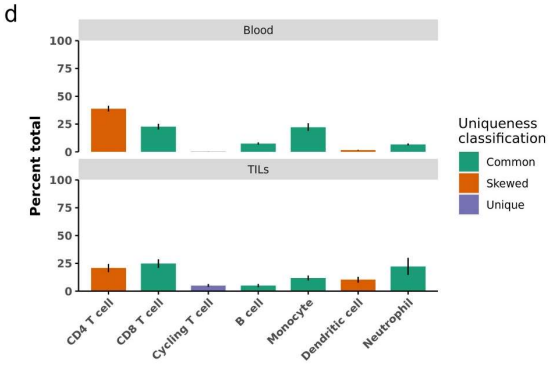


Figure 4.1. Study overview and cell type annotations. (a) Diagram of study design. Tumor infiltrating immune cell and circulating immune cell were integrated into one dataset then transcriptomic and cell type abundances were evaluated. (b) UMAP representation with major cell type annotations of circulating immune cells from 10 dogs with osteosarcoma (n = 42,292 cells) and TILs from 6 treatment-naïve OS dogs (n = 11,431 cells). (c) Feature plots depicting canonical cell type markers. (d) Bar chart depicting average percentage of each cell type with bar color indicating tissue specificity.

Following label transfer of cell types, we evaluated if any population was unique or skewed toward tumor or blood (Figure 4.1d). As expected, all the major cell types were found in both tissue sources, but the relative proportion of CD4 T cells was greater in blood while dendritic cells were disproportionally represented in the tumor. The minor cluster of cycling T cells exhibited a marked increase in the relative proportion of cells in the tumor relative to blood. As such, cycling T cells were classified as unique to the tumor even though they are lowly abundant in the circulation.

Next, we subset on each major cell type and completed a detailed comparison of TILs and blood leukocytes.

Exhausted and regulatory CD4 T cells are preferentially located in the tumor microenvironment

After transferring cell annotations to the CD4 T cell subset, we identified 6 distinct populations. The cell types consisted of naïve, central memory (TCM), effector memory (TEM), Th1-like TEM, Th2-like TEM, and exhausted/regulatory ($T_{ex\text{-}reg}$) cells (Figure 4.2a). The cluster annotated as $T_{ex\text{-}reg}$ did not reach a consensus when transferring cell type labels and was annotated based on the presence of both regulatory and exhausted T cells. Out of the 6 T cell populations, naïve, Th1-like TEM, and $T_{ex\text{-}reg}$ cells were determined to exhibit a skewed distribution (Figure 4.3b). Naïve and Th1-like TEM were more abundant in blood, while $T_{ex\text{-}reg}$ were disproportionally represented in the tumor. Due to the noted heterogeneity in the $T_{ex\text{-}reg}$ cluster, we

evaluated the abundance in CXCL13^{+} cells (normalized count > 0) in the $T_{ex\text{-}reg}$ cluster and found CXCL13^{+} cells to be unique to TILs. Together, this suggests the abundance skew in the $T_{ex\text{-}reg}$ cluster, was in part due to the increased proportion of exhausted T cells in the tumor.

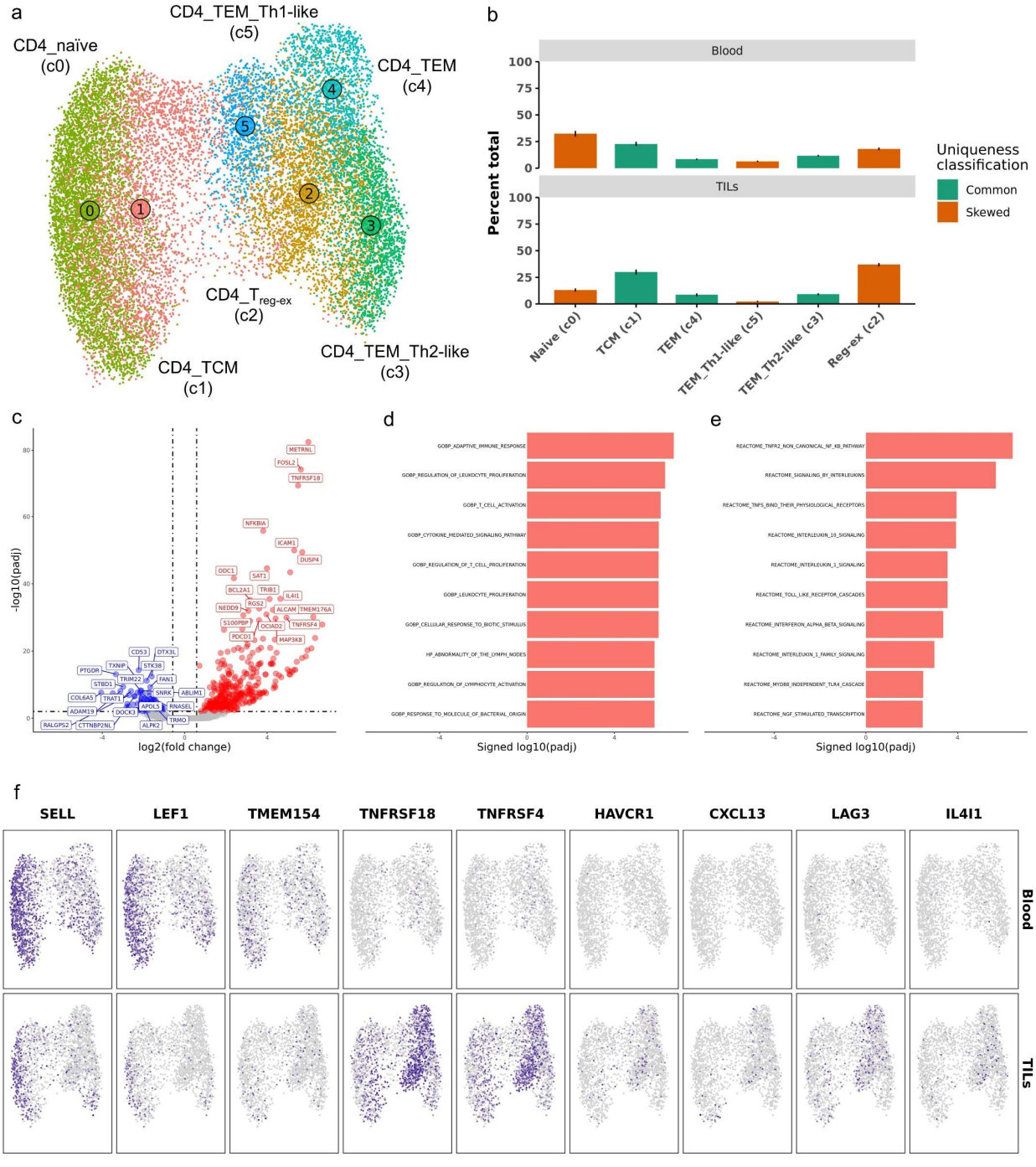


Figure 4.2. Tumor infiltrating cells exhibit higher relative proportions of exhausted CD4 T cell relative to circulating CD4 T cells. (a) UMAP representation of unsupervised clustering of 16,765 CD4 T cells. (b) Bar chart depicting average percentage of each cell type with bar color indicating tissue specificity. (c) Volcano plot depicting results of differential gene expression analysis contrasting all tumor infiltrating CD4 T cells versus all circulating CD4 T cells. Bar charts depicting P value from gene set enrichment analysis using genes upregulated in tumor infiltrating CD4 T cells against gene ontology (d) and Reactome (e) gene sets. (f) Feature plots split by tissue source depicting normalized expression of differentially expressed features.

Next, we completed DGE analysis within CD4 T cells between tissue sources. The analysis revealed 380 genes to be more highly expressed in TILs and 116 genes were over abundant in CD4 blood leukocytes (Figure 4.2c). Subsequent gene set enrichment analysis (GSEA) using gene ontology (GO) terms of the differentially expressed features revealed enrichment for multiple terms associated with leukocyte activation and proliferation (Figure 4.2d). GSEA using Reactome terms suggested CD4 TILs were active in interleukin signaling and interferon responses (Figure 4.3e). Select differentially expressed features were then evaluated to identify which clusters were driving differential expression (Figure 4.3f). The feature plots revealed the overabundance of SELL (CD62L) and LEF1 in blood leukocytes was associated with naïve T cells, while the increases in exhaustion (HAVCR1, CXCL13, LAG3) and activation (TNFRSF4, TNFRSF18) markers was largely confined to TEM and $T_{ex\text{-}reg}$ cell types. Overall, our analysis suggests that in addition to shifts in the abundances of naïve, Th1-like TEM, and $T_{ex\text{-}reg}$ cells, tumor-infiltrating CD4 T cells also exhibit altered transcriptional profiles suggestive of activation and exhaustion relative to their circulating counterparts.

Exhaustion markers are increased in tumor-infiltrating effector CD8 T cells

We next applied the same workflow to investigate tumor associated changes to CD8 T cells and NK cells. Unsupervised clustering of the integrated dataset resulted in an inability to resolve NK cells as annotated in the original data source (Figure 4.3a). In the integrated dataset

NK cells were interspersed in CD8 effector T cell clusters. Abundance analysis revealed naïve CD8 T cells had a skewed distribution, with a preferential abundance in blood (Figure 4.3b).

When comparing gene expression profiles between tumor-infiltrating and blood leukocytes, we identified 67 features more highly expressed in blood leukocytes and 252 more highly expressed in the tumor (Figure 4.3c). Multiple T cell exhaustion markers (LAG3, TNFSF9) and HAVCR1 (TIM-1), a member of the T cell immunoglobulin and mucin domain-containing protein (TIM) family, were identified to be more abundantly expressed in tumor-infiltrating immune cells (12–14). Even with filtering out tumor signature associated features, GSEA analysis still suggested TIL gene expression profiles were associated with extracellular matrix processes (Figure 4.3d). This term may be an artifact driven by tissue signatures or the term may be capturing biologically relevant changes in lymphocytes as they transition from circulating to tissue infiltrating T cell. Further pathway analysis using gene ontology terms revealed an association with T cell activation and recruitment of mononuclear cells. Reactome pathway analysis identified multiple terms associated with NFkB signaling to be enriched in tumor-infiltrating CD8 T cells, which is generally suggestive of increased T cell activation and potential antitumor activity (Figure 4.3e) (15). Visualization of differentially expressed features revealed the increased expression of LEF1 in blood leukocytes was associated with naïve CD8 and double negative T cells, while the increased CX3CR1/PTGDR expression in blood leukocytes was driven by expression on effector CD8 T cells (Figure 4.3f). Tumor-infiltrating effector CD8 T cells were identified as drivers of LAG3, TNFSF9 (4-1BB ligand) and HAVCR1 (TIM-1) overexpression, which suggests effector CD8 T cells are activated and exhausted relative to circulating leukocytes. Consistent with previous findings (7), our overall analysis of tumor-infiltrating CD8 T cells indicates this population is activated and upregulates exhaustion markers when in the OS TME.

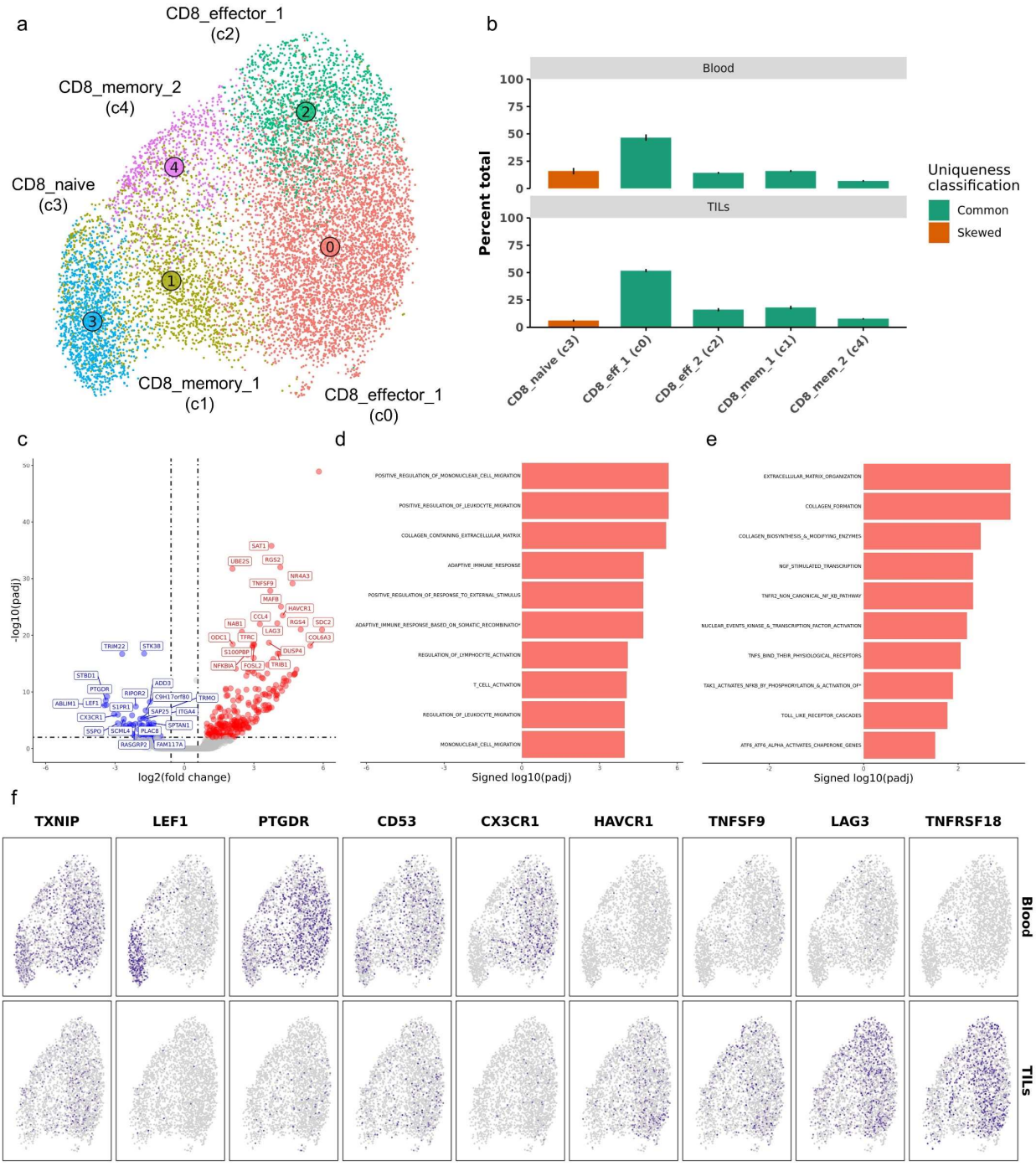


Figure 4.3. Tumor infiltrating CD8 T cells express high levels of exhaustion markers. (a) UMAP representation of unsupervised clustering of 10,962 CD8 T cells. (b) Bar chart depicting average percentage of each cell type with bar color indicating tissue specificity. (c) Volcano plot depicting results of differential gene expression analysis contrasting all tumor infiltrating CD8 T cells versus all circulating CD8 T cells. Bar charts depicting P value from gene set enrichment analysis using genes upregulated in tumor infiltrating CD8 T cells against gene ontology (d) and Reactome (e) gene sets. (f) Feature plots split by tissue source depicting normalized expression of differentially expressed features.

Mature regulatory dendritic cells are present in OS tumors, but not in circulation

All dendritic cell (DC) annotations reached a consensus, which enabled the direct transfer of cell type labels (Figure 4.4a). Of the five DC subtypes, we identified plasmacytoid DCs (pDCs) and precursor (pre) DCs to have a skewed abundance in blood (Figure 4.4b). Mature regulatory DCs (mregDCs) were the only cell types to be classified as unique to TILs, while conventional DC1 and DC2s were determined to be equally abundant in blood and tumor. (Figure 4.4b). The identification of mregDCs as unique to TILs was expected as this cell population has been reported to arise from circulating DC populations that infiltrate the tumor then adopt a mregDC signature (16,17).

To investigate how tumor infiltration impacts DC gene expression, we completed DGE analysis and subsequent GSEA. The analysis revealed 271 features enriched in tumor DCs and 29 in blood DCs (Figure 4.4c). Gene ontology analysis revealed associations with tumor necrosis factor (TNF) responses and vascular development within tumor-infiltrating DCs suggesting DCs may be involved in vascular remodeling of the tumor microenvironment (Figure 4.4d). GSEA using Reactome terms identified increased interleukin activity of tumor-infiltrating DCs with IL4, IL13, and IL10 predicted to elicit the greatest impact on tumor DCs (Figure 4.4e). The observed mregDC abundance skew likely impacted DGE analysis (as evidenced by CCR7, IDO1, IL4I1, and CD274 (PD-L1) upregulation in tumor DCs), so we investigated the localization of differentially expressed genes. We identified IL16 and IRF4 to be broadly down regulated in infiltrating DCs, while CXCR4 was broadly upregulated (Figure 4.4f). Interestingly, VEGFA, a feature determined to be associated with the tumor tissue signature was selectively upregulated in cDC2s, suggesting the differential expression of VEGFA may be a biologically relevant change rather than just an artifact of tissue bias. In summary, we provide evidence that mregDCs are unique to the tumor and that infiltrating DCs are mediators of interleukin signaling in the OS TME.

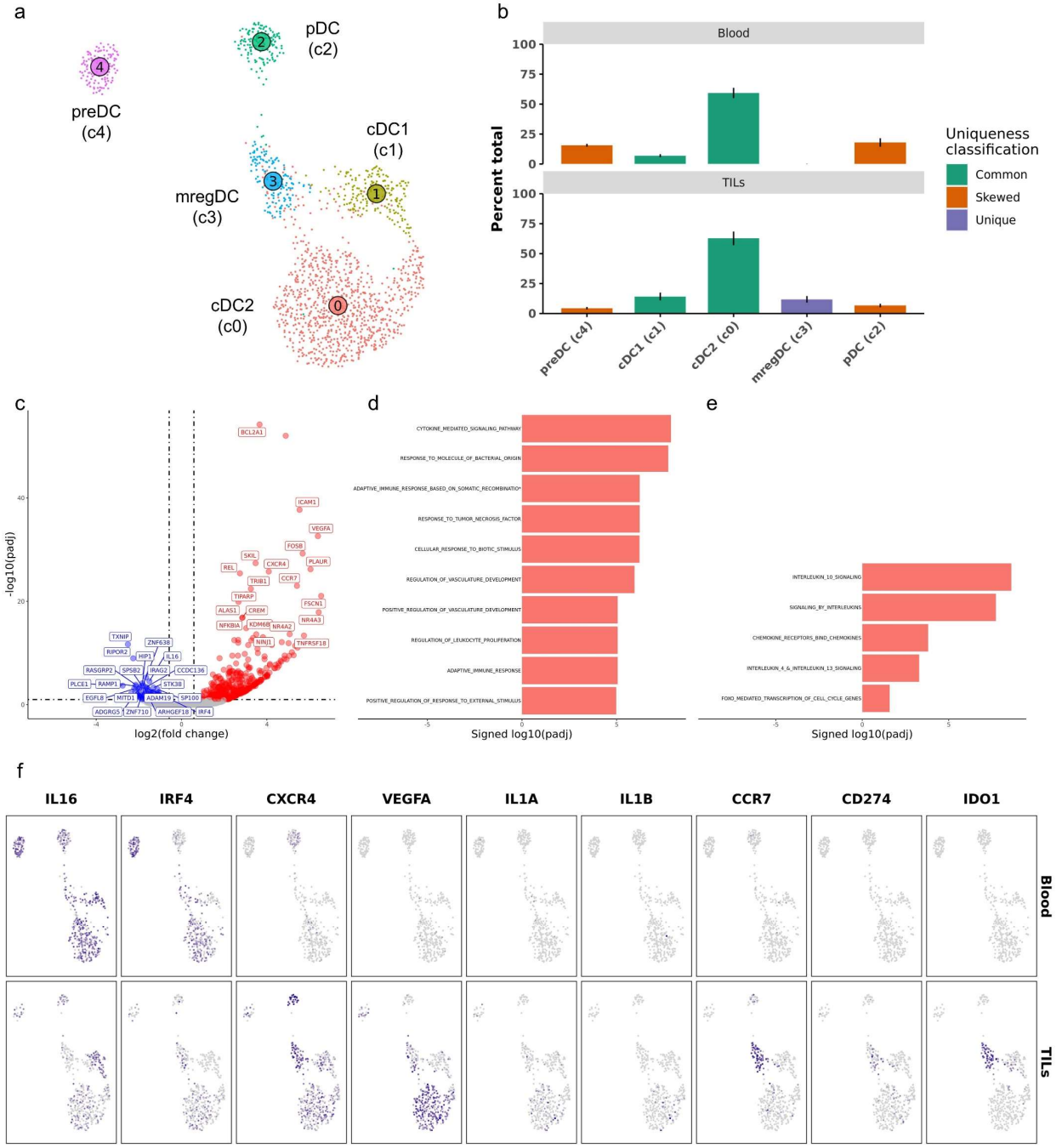


Figure 4.4. Mature regulatory dendritic cells are found in OS tumors, but not in circulation. (a) UMAP representation of unsupervised clustering of 1,556 dendritic cells. (b) Bar chart depicting average percentage of each cell type with bar color indicating tissue specificity. (c) Volcano plot depicting results of differential gene expression analysis contrasting all tumor infiltrating dendritic cells versus all circulating dendritic cells. Bar charts depicting P value from gene set enrichment analysis using genes upregulated in tumor infiltrating dendritic cells against gene ontology (d) and Reactome (e) gene sets. (f) Feature plots split by tissue source depicting normalized expression of differentially expressed features.

Tumor-infiltrating monocytes exhibit a macrophage-like gene signature and upregulate chemokine expression

Tumor-infiltrating and circulating monocytes integrated uniformly and no tissue preferences were observed for any of the four monocytes subtypes (Figure 4.5a/b). Despite uniform integration, DGE analysis revealed marked transcriptomic changes induced by the tumor with 356 features enriched in TILs and 69 enriched in blood monocytes (Figure 4.5c). Several macrophage-associated features (MSR1, CTSK, APOC1) were enriched in TIMs which suggests that although monocyte characteristics are apparent relative to TAMs (as presented in chapter 3), the cells have a macrophage signature relative to circulating monocytes. Therefore, it is likely that the TIMs are in an intermediate state differentiating toward a macrophage phenotype.

Subsequent pathway analysis revealed similar terms observed in other cell types which implicates TIMs in general immune activation, adhesion, and interleukin signaling (Figure 4.5d/e). Key features identified to define circulating monocytes included IL16 and FGL2 (Fibrinogen-like protein 2), which were broadly expressed, whereas LTF (lactotransferin) expression was specific to M-MDSCs (Figure 4.5f). TIMs exhibited increased expression of multiple chemokines (CXCL10, CXCL16, CCL5, CCL8, CCL19, CCL7) suggesting that infiltration into the tumor rapidly drives monocytes to recruit additional monocytes. Furthermore, circulating monocytes exhibited higher levels of CCR2 expression which may interact with the CCL7 and CCL8 secreted by TIMs to promote further infiltration (18). Immune modulatory molecules, IL1A and OSM, were determined to be broadly upregulated in TIMs, while PTGES was specifically upregulated in M-MDSCs. Lastly, C1QC expression was evaluated and was found to be localized to IFN associated cells and may represent unintentional inclusion of TAMs in the analysis.

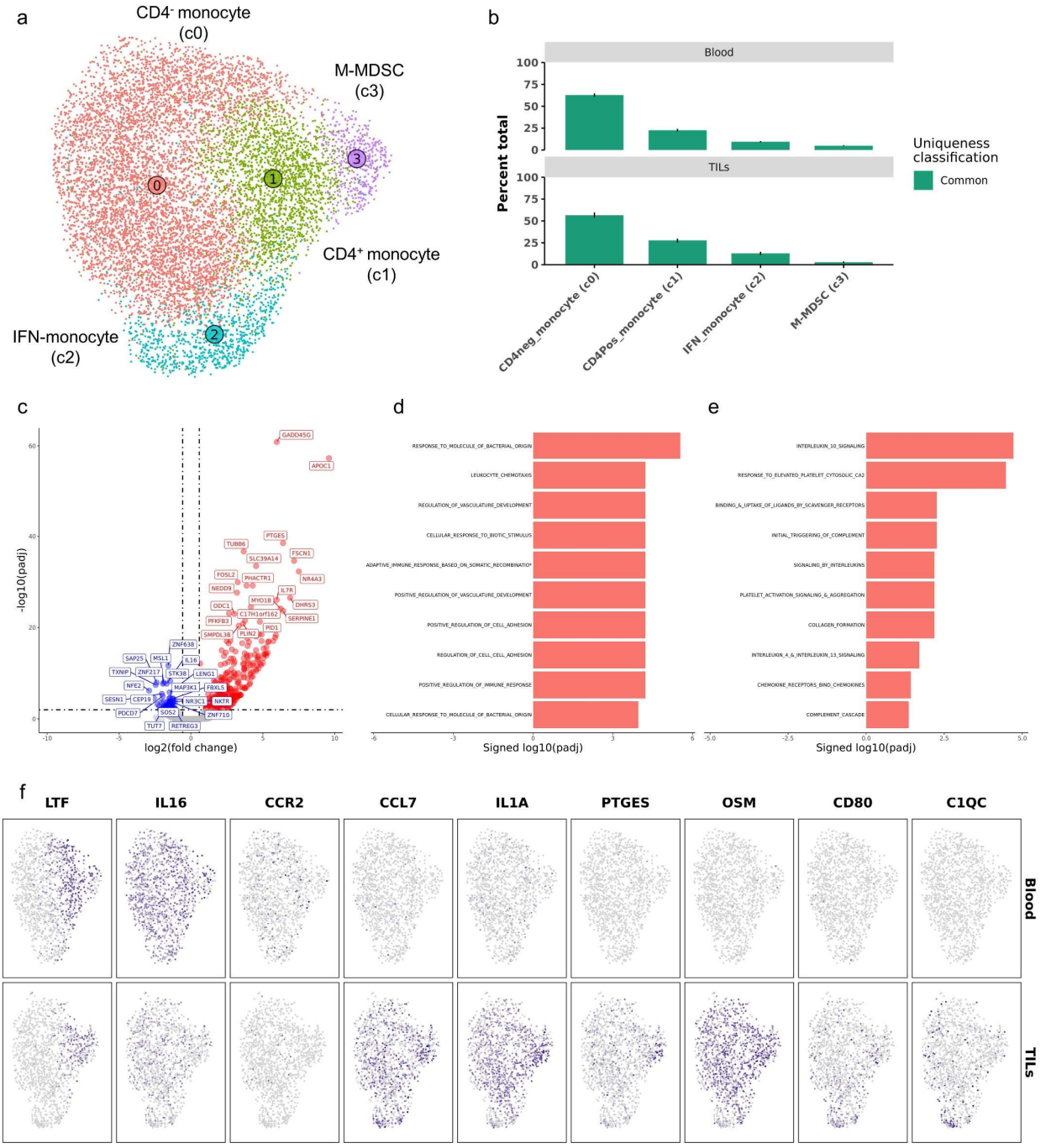


Figure 4.5. Tumor infiltrating monocytes upregulate inflammatory and immune modulatory molecules relative to circulating monocytes. (a) UMAP representation of unsupervised clustering of 10,004 monocytes. (b) Bar chart depicting average percentage of each cell type with bar color indicating tissue specificity. (c) Volcano plot depicting results of differential gene expression analysis contrasting all tumor infiltrating monocytes versus all circulating monocytes. Bar charts depicting P value from gene set enrichment analysis using genes upregulated in tumor infiltrating monocytes against gene ontology (d) and Reactome (e) gene sets. (f) Feature plots split by tissue source depicting normalized expression of differentially expressed features.

Tumor-infiltrating B cells upregulate FOS and have gene expression patterns suggestive of increased endoplasmic reticulum activity

Integration of circulating and infiltrating B cells revealed a preferential distribution of immature B cells in blood (Figure 4.6a/b). Due to distinct transcriptional profiles of plasma cells, we excluded plasma cells then completed DGE analysis between tissues in non-plasma cell clusters (c0, c1, and c3) (Figure 4.6c). The analysis revealed the upregulation of FOS and FOSB in tumor-infiltrating B cells, which suggests infiltrating B cells are terminally differentiated and have interacted with a cognate antigen (19). The upregulation of FOS expression may also indicate infiltrating B cells are undergoing apoptosis (20,21). Apoptosis is further supported by GSEA revealing ERAD (Endoplasmic reticulum (ER)-associated degradation) pathway is enriched in infiltrating B cells (Figure 4.6d) (22).

GSEA further associated infiltrating B cell gene signatures with protein modification, endoplasmic reticulum activity, and antigen presentation (Figure 4.6d/e). These terms suggest the infiltrating B cells may be processing tumor antigens and presenting antigen to infiltrating T cells or that the B cells may be transitioning to plasma cells to secrete antibodies. Broad reduction in IGHM on infiltrating B cells, relative to circulating B cells, further suggests the cells are undergoing class switching and potentially differentiating into short-lived plasma cells (Figure 4.6f). Lastly, GSEA revealed endoplasmic reticulum stress is enriched in infiltrating B cells which could be suggestive of dysfunction. Overall, B cell transcriptional profiles appear to be remarkably impacted by the tumor microenvironment and further investigation is needed to determine if the B cells are activated or dysfunctional.

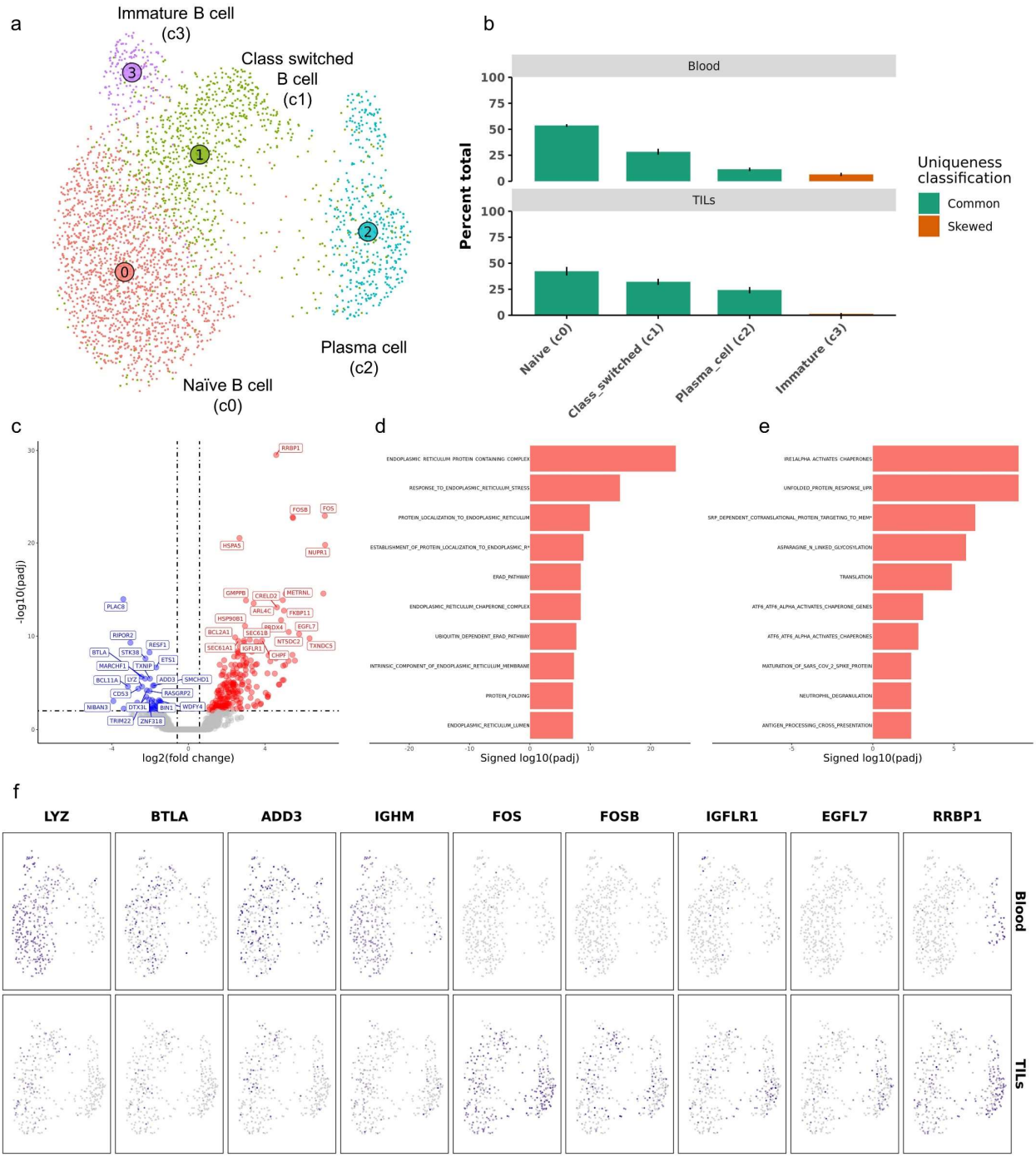


Figure 4.6. Tumor infiltrating B cells exhibit gene expression patterns suggestive of antigen presentation and maturation. (a) UMAP representation of unsupervised clustering of 3,108 B cells. (b) Bar chart depicting average percentage of each cell type with bar color indicating tissue specificity. (c) Volcano plot depicting results of differential gene expression analysis contrasting all tumor infiltrating B cells versus all circulating B cells. Bar charts depicting P value from gene set enrichment analysis using genes upregulated in tumor infiltrating B cells against gene ontology (d) and Reactome (e) gene sets. (f) Feature plots split by tissue source depicting normalized expression of differentially expressed features.

Tumor-associated neutrophils increase Oncostatin M and chemokine expression relative to circulating neutrophils

Evaluation of tumor-associated neutrophils (TANs) and blood neutrophils revealed no skews in cell abundances between the two tissues (Figure 4.7a/b). Differential gene expression analysis identified up regulation of 139 features in TANs, with only 4 features preferentially expressed in the blood (Figure 4.7c). Subsequent pathway analysis suggested TAN transcriptomic signatures are associated general neutrophil activation, responses to interleukins, and response to endoplasmic reticulum stress (Figure 4.d/e). Together the analysis indicated TANs exhibit a shift toward an activation state with increased involvement in the regulation of cell migration upon tumor infiltration.

Oncostatin M (OSM), a member of the IL-6 family with reported immune suppressive properties, was found to be broadly upregulated in TANs (Figure 4.7f) (23). TANs overexpressed multiple chemokines CCL5, CCL7, and CXCL8 relative to circulating neutrophils, which indicates upon infiltration, neutrophils increase signals to further recruit myeloid cell infiltration. Although not significantly differentially expressed, cathelicidin antimicrobial peptide (CAMP) was observed to be slightly reduced with a specificity for PMN-MDSCs. IL1R2, a decoy receptor for IL1A/B, was broadly upregulated in TANs which suggests they may function to dampen inflammatory responses (24). Lastly, CD274 (PD-L1) was enriched in TANs which implicates TANs as suppressors of adaptive T cell responses.

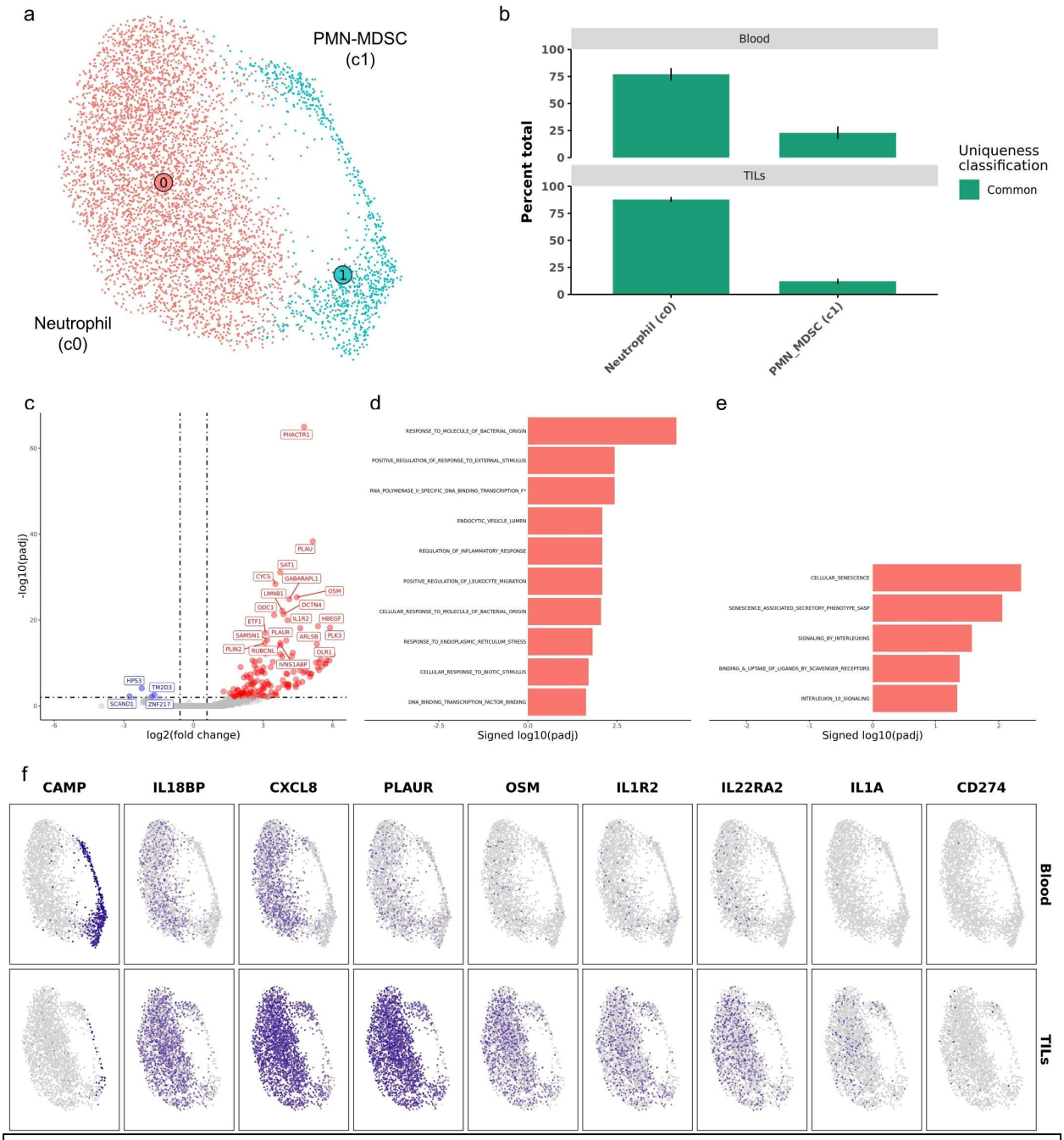


Figure 4.7. Tumor infiltrating neutrophils increase expression of activation and inflammatory markers relative to circulating neutrophils. (a) UMAP representation of unsupervised clustering of 5,470 neutrophils. (b) Bar chart depicting average percentage of each cell type with bar color indicating tissue specificity. (c) Volcano plot depicting results of differential gene expression analysis contrasting all tumor infiltrating neutrophils versus all circulating neutrophils. Bar charts depicting P value from gene set enrichment analysis using genes upregulated in tumor infiltrating neutrophils against gene ontology (d) and Reactome (e) gene sets. (f) Feature plots split by tissue source depicting normalized expression of differentially expressed features.

Discussion

In the present study we investigated how the osteosarcoma tumor microenvironment modulates the transcriptomic signatures of infiltrating immune cells. Using circulating immune cells as a point of reference, we identified tumor-infiltrating T cells exhibited increased expression of exhaustion markers and that mature regulatory cells were unique to the tumor microenvironment. As expected, our analysis of monocytes indicated tumor-infiltrating monocytes have a stronger macrophage signature than blood monocytes and upregulate immune regulatory features including PD-L1 (CD274), CD80, and CD86. We identified Oncostatin M (OSM) to be more highly expressed on tumor-infiltrating monocytes/neutrophils relative to their circulating counterparts, implicating myeloid cells as the primary producer of this potentially protumor cytokine. The data presented here suggests many features of tumor-induced immune suppression and exhaustion are conserved between humans and dogs, while also providing data to identify potential species-specific differences.

Immune suppression and evasion are hallmarks of cancer and understanding the mechanisms in which immune cells are impacted is the foundation for the development of effective therapeutics. In the context of osteosarcoma, a tumor type notoriously unresponsive immunotherapy, it can be valuable to identify which cell types drive immune suppression, to then target them therapeutically (25). Although we did not directly investigate functional changes in this study, we completed gene set enrichment analysis to make inferences about potential functional states of cell. Our analysis revealed tumor-infiltrating B cells and neutrophils to have increased endoplasmic reticulum and increased oxidative stress which may be suggestive of dysfunctional properties. Further investigation into the functional relevance of the observed transcriptomic changes is warranted.

Although our findings provide insight into how TILs differ from blood leukocytes, many of the differences in gene expression may be due to the cells being in circulation rather than tissue

resident normal cells. Because bone is largely acellular with minimal to no leukocytes present in normal bone, the use of circulating leukocytes represents one of the best avenues to explore how infiltrating immune cells are modulated by the OS TME. In attempts to minimize the impact of tissue signatures that arose from mRNA released during sample processing we identified a tumor tissue signature and excluded those features from differential gene expression analysis. In doing so we attempted to filter out background signal, but the features may also represent tumor-induced gene expression changes observed across all cell types. The 38 excluded features were associated with extracellular matrix GSEA terms suggesting the genes are implicated in a tissue signature, not immune-specific changes. Future work incorporating lymph nodes or other non-circulating immune cells may more effectively distinguishing between tissue and tumor associated changes.

The data presented here act to confirm canine tumor-infiltrating immune cells exhibit transcriptomic changes similar to those reported in humans (6–8). We observed abundance skews in which naive T cells and plasmacytoid DCs are preferentially found in circulation. Furthermore, we identified mature regulatory dendritic cells to be unique to the tumor which provides indirect evidence that this cell type likely travels via lymphatics to bring tumor antigens to draining lymph nodes. Ultimately, the analysis completed here sheds light on immune suppression and dysfunction present in tumor-infiltrating immune cells.

References

1. Jochems C, Schlom J. Tumor-infiltrating immune cells and prognosis: the potential link between conventional cancer therapy and immunity. Exp Biol Med. 2011;236(5):567–79.
2. Liu R, Yang F, Yin J-Y, Liu Y-Z, Zhang W, Zhou H-H. Influence of tumor immune infiltration on immune checkpoint inhibitor therapeutic efficacy: a computational retrospective study. Front Immunol. 2021;12:685370.
3. Yin Y, Sakakibara R, Honda T, Kirimura S, Daroonpan P, Kobayashi M, et al. High density and proximity of CD8+ T cells to tumor cells are correlated with better response to nivolumab treatment in metastatic pleural mesothelioma. Thorac Cancer. 2023;
4. Fritzsching B, Fellenberg J, Moskovszky L, Sápi Z, Krenacs T, Machado I, et al. CD8+/FOXP3+-ratio in osteosarcoma microenvironment separates survivors from non-survivors: a multicenter validated retrospective study. Oncoimmunology. 2015;4(3):e990800.
5. Tian H, Cao J, Li B, Nice EC, Mao H, Zhang Y, et al. Managing the immune microenvironment of osteosarcoma: the outlook for osteosarcoma treatment. Bone Res. 2023;11(1):11.
6. Szabo PA, Levitin HM, Miron M, Snyder ME, Senda T, Yuan J, et al. Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. Nat Commun. 2019;10(1):4706.
7. Zheng L, Qin S, Si W, Wang A, Xing B, Gao R, et al. Pan-cancer single-cell landscape of tumor-infiltrating T cells. Science (80-). 2021;374(6574):abe6474.
8. Mulder K, Patel AA, Kong WT, Piot C, Halitzki E, Dunsmore G, et al. Cross-tissue single-cell landscape of human monocytes and macrophages in health and disease. Immunity. 2021;54(8):1883–900.
9. Hao Y, Hao S, Andersen-Nissen E, Mauck III WM, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. Cell. 2021;184(13):3573–87.
10. McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. Cell Syst. 2019;8(4):329–37.
11. Young MD, Behjati S. SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. Gigascience. 2020;9(12):giaa151.
12. Wolf Y, Anderson AC, Kuchroo VK. TIM3 comes of age as an inhibitory receptor. Nat Rev Immunol. 2020;20(3):173–85.
13. Yang Z-Z, Kim HJ, Villasboas JC, Chen Y-P, Price-Troska T, Jalali S, et al. Expression of LAG-3 defines exhaustion of intratumoral PD-1+ T cells and correlates with poor outcome in follicular lymphoma. Oncotarget. 2017;8(37):61425.
14. Li Y, Wang Z, Jiang W, Zeng H, Liu Z, Lin Z, et al. Tumor-infiltrating TNFRSF9+ CD8+ T cells define different subsets of clear cell renal cell carcinoma with prognosis and immunotherapeutic response. Oncoimmunology. 2020;9(1):1838141.
15. Barnes SE, Wang Y, Chen L, Molinero LL, Gajewski TF, Evaristo C, et al. T cell-NF-κB activation is required for tumor control in vivo. J Immunother cancer. 2015;3(1):1–11.

16. Gordon JR, Ma Y, Churchman L, Gordon SA, Dawicki W. Regulatory dendritic cells for immunotherapy in immunologic diseases. Front Immunol. 2014;5:7.

17. Liu W, Hu H, Shao Z, Lv X, Zhang Z, Deng X, et al. Characterizing the tumor microenvironment at the single-cell level reveals a novel immune evasion mechanism in osteosarcoma. Bone Res. 2023;11(1):4.

18. Hao Q, Vadgama J V, Wang P. CCL2/CCR2 signaling in cancer pathogenesis. Cell Commun Signal. 2020;18:1–13.

19. Ohkubo Y, Arima M, Arguni E, Okada S, Yamashita K, Asari S, et al. A role for c-fos/activator protein 1 in B lymphocyte terminal differentiation. J Immunol. 2005;174(12):7703–10.

20. Inada K, Okada S, Phuchareon J, Hatano M, Sugimoto T, Moriya H, et al. c-Fos induces apoptosis in germinal center B cells. J Immunol. 1998;161(8):3853–61.

21. Liebermann DA, Gregory B, Hoffman B. AP-1 (Fos/Jun) transcription factors in hematopoietic differentiation and apoptosis. Int J Oncol. 1998;12(3):685–1385.

22. Kong S, Yang Y, Xu Y, Wang Y, Zhang Y, Melo-Cardenas J, et al. Endoplasmic reticulum-resident E3 ubiquitin ligase Hrd1 controls B-cell immunity through degradation of the death receptor CD95/Fas. Proc Natl Acad Sci. 2016;113(37):10394–9.

23. Wallace PM, MacMaster JF, Rouleau KA, Brown TJ, Loy JK, Donaldson KL, et al. Regulation of inflammatory responses by oncostatin M. J Immunol. 1999;162(9):5547–55.

24. Dinarello CA. Overview of the IL-1 family in innate inflammation and acquired immunity. Immunol Rev. 2018;281(1):8–27.

25. Wu C-C, Beird HC, Andrew Livingston J, Advani S, Mitra A, Cao S, et al. Immuno-genomic landscape of osteosarcoma. Nat Commun. 2020;11(1):1008.

CHAPTER 5: Single-cell RNA sequencing reveals the impacts of losartan, ladarxin, and toceranib combination therapy on the canine osteosarcoma microenvironment

Summary

Therapeutic options for patients diagnosed with osteosarcoma (OS) are limited and prognosis is dismal with a 5-year survival rate of 20%. As such, this chapter aims to evaluate the impacts of an OS combination therapeutic (TDI; losartan, toceranib, ladarixin) designed to deplete tumor-infiltrating myeloid cells. Specifically, we evaluated the impacts of TDI on the TME and circulating immune cells. Through completion of single-cell RNA (scRNA) sequencing of 6 dogs receiving treatment for 14 days and 6 treatment-naïve dogs we were able to describe how intervention impacted the OS tumor microenvironment (TME). We coupled the tumor data with scRNA sequencing of circulating leukocyte of 5 dogs pre- and 14 days post-treatment. In total, we completed scRNA sequencing on just under 100,000 cells which provided key insights into the local and systemic impacts of treatment. Notable findings include treatment associated increases in circulating M-MDSCs, broad depletion of infiltrating immune cells, and profound alteration to the transcriptional profiles of malignant osteoblast and stromal fibroblasts. Overall, this study provides unique insights into the heterogeneity within canine osteosarcoma tumors and highlights the impacts of myeloid targeted adjuvant immunotherapy.

Introduction

Osteosarcoma (OS) is an aggressive malignancy of the bone with a complex tumor microenvironment that has proven to be impervious to widely used immunotherapeutics. Immune suppressive cell populations within the heterogenous tumor microenvironment (TME) are thought to contribute to the inefficacy of immunotherapy (1). Recent advances in single-cell RNA (scRNA) sequencing have enabled the molecular dissection of the OS TME in humans and dogs

(2,3,chapter 3). While these studies have revealed key cell types and interactions in treatment-naive tumors, there is a need to understand how therapeutic intervention modulates the cellular constituents in the TME.

Dogs spontaneously develop OS and exhibit similar disease pathology to humans. OS in dogs occurs at a rate up to 75 times that observed in humans and when coupled with the shorter life span of dogs, makes the model appealing for the rapid evaluation of clinical trial outcomes (4). To this end, we have recently reported that losartan is able to block monocyte migration to primary canine OS tumors through a CCR2 dependent mechanism (5). A multi-center study of oral toceranib phosphate (receptor tyrosine kinase inhibitor) led to the approval of the drug for treatment of mast cell tumors in dogs and is thought to act through depletion of regulatory T cells and myeloid-derived suppressor cells (6). While we recently reported that the combination of losartan and toceranib was able to induce a clinical benefit rate of 50% (7), the addition of ladarixin, an inhibitor of the interleukin-8 signaling, may work synergistically in a triple-drug immunotherapeutic (TDI) protocol (8).

Single-cell RNA sequencing is one approach to study the biology of complex tissues. Application of this technology has revealed how therapeutic intervention influences systemic immune responses as well as intratumoral changes following therapeutic intervention (9–11). Due to limited availability of OS tumor samples, the impacts of therapeutic intervention in the setting of OS remains largely unknown.

In the current study, we used single-cell RNA sequencing to investigate how treatment with triple-drug immunotherapy (losartan, toceranib, ladarixin) impacts the TME and circulating immune cells. Our analysis revealed reductions in tumor-infiltrating myeloid cells, with an unexpected reduction in the numbers of tumor-infiltrating T cells. Analysis of circulating leukocytes identified TDI treatment to drive increases in monocytic myeloid-derived suppressor cells. Furthermore, we noted relative increases in fibroblasts and modifications to osteoblast

transcriptional programs suggesting TDI had a broad impact on the OS TME. Ultimately, our findings provide insight into how immunotherapy shapes the TME in the context of osteosarcoma.

Methods

Patient selection

Osteosarcoma patients were selected based on the presence of an appendicular primary tumor and the absence of previous therapeutic intervention. All dogs presented with radiographic evidence of OS. Within 30 minutes of amputation samples were collected for single-cell RNA sequencing processing and for completion of histology to confirm the OS diagnosis. All dogs included in the study were confirmed to have OS by histology. All studies were approved by the Colorado State University (CSU) Institutional Animal Care and Use Committee and the CSU Clinical Review Board. All dog owners provided informed consent prior to sample collection.

Therapeutic protocol

The therapeutic arm of the study consisted of 6 dogs that has suspected osteosarcoma and were brought to CSU's Flint Animal Cancer Center for treatment (Figure 5.1a). After diagnostic measures were taken patients were enrolled into the clinical trial in which they were to be treated daily with oral losartan (10 mg/kg, BID), toceranib (2.75 mg/kg every other day), and ladarixin (15 mg/kg, BID) then undergo amputation approximately 14 days following treatment initiation. At the time of amputation tumor samples were collected for single cell RNA sequencing. Treatment-naïve tumor samples were obtained from 6 dogs that has suspected osteosarcoma and were brought to CSU's Flint Animal Cancer Center for treatment.

In addition to tumor biopsies, paired blood samples were collected from 5 dogs pre- and 14 days post-treatment (Figure 5.1a). Whole blood was collected prior to treatment initiation and after 14 days on treatment protocol just prior to limb amputation. A summary of patient demographics is provided in Table 5.1.

Sample preparation

Tumor samples were process as previously described in chapter 3, while whole blood samples were process as previously described in chapter 2. Briefly, for each tumor sample 3 to 5 biopsies were collected, digested with collagenase type II (250 U/mL; Thermo Fisher Scientific Inc., Waltham, MA) then enriched for live cells by completing density centrifugation. Red blood cells were depleted using Ammonium-Chloride-Potassium (ACK) lysis buffer, washed with PBS, then prepared for single cell isolation using a Chromium iX instrument (10x Genomics; Pleasanton, CA).

Table 5.1. Dog demographics.

Breed	Sex	Age (years)	Tumor location	Histological subtype	Tumor	Blood
Mixed (Staffordshire Terrier)	FS	6.8	R proximal humerus	osteoblastic	Yes	Yes
Goldendoodle	FS	10.4	L proximal humerus	osteoblastic	Yes	Yes
Plott Hound	MC	7.8	L radius	osteoblastic	Yes	Yes
Australian Cattle Dog/German Shorthaired Pointer	MC	7.7	R distal radius	osteoblastic	Yes	No
Chesapeake Bay Retriever + Aussie	MC	7.0	R distal radius	osteoblastic	Yes	No
American Bulldog	MC	6.9	L ulna	osteoblastic	Yes	No
Mixed (Husky)	FS	8.0	L proximal humerus	osteoblastic	Yes	No
Catahoula	MC	11.5	R distal femur	osteoblastic	Yes	No
Labrador Retriever	MC	7.8	L distal femur	fibroblastic	Yes	No
Great Dane	MC	8.0	R distal radius	osteoblastic	Yes	No
Mixed	FS	11.3	R distal radius	chondroblastic	Yes	No
Catahoula	FS	8.4	R distal radius	osteoblastic	Yes	No
Bernese Mountain Dog/Great Pyrenees	MI	7.3	R distal radius	unknown	No	Yes
Saint Bernard	MI	5.7	L distal tibia	unknown	No	Yes

Library preparation and sequencing

Single cells were isolated and tagged with unique cell barcodes using a Chromium iX instrument (10x Genomics) with a target of 5,000 cells per sample. Three of the dogs (naïve dog 1, naïve dog 2, and TDI treated dog 1) had two samples processed each with a 5,000 cell targets, for a total target of 10,000 cells from each of those dogs. Single cells were isolated and processed

using a Chromium Next GEM Single Cell 3' Kit v3.1 following manufacture recommended protocols. Individual cell transcriptomes were captured and labeled with molecular barcodes, then a standard Illumina library preparation was completed using a dual index library construction kit (10x Genomics). Samples quality was analyzed using a LabChip (PerkinElmer; Waltham, MA) then submitted for sequencing on an Illumina NovaSeq 6000 sequencer (Novogene Corporation; Sacramento, CA) with a target of 100,000 150 bp paired-end reads per cell. Raw data was demultiplexed by the sequencing core then transferred for downstream analysis.

Read mapping and quantification

A Cell Ranger analysis pipeline (version 6.1.2, 10x Genomics) was utilized to process raw FASTQ sequencing data, align reads to the canine genome, and generate a count matrix as described in chapter 2. The output count matrix for each sample was then exported and used for downstream analysis.

Data filtering and integration

For each sample, the count matrix was imported into R using the Read10X() function then converted to a Seurat object using the CreateSeuratObject() function (12). Dead/poor quality cells were removed by evaluating the percent mitochondrial reads per cell was calculated using PercentageFeatureSet() to count all reads mapped to features with the prefix "MT-". For the tumor dataset thresholds used were: 200 < nFeature_RNA < 5500, percent.mt < 12.5, and 100 < nCount_RNA < 15000. For the blood dataset thresholds used were: 200 < nFeature_RNA < 3500, percent.mt < 12.5, and 500 < nCount_RNA < 20000. Next, DoubletFinder, was used to identify and remove putative cell doublets (13). After completing QC filtering on each sample, all samples from tumors were integrated into one object, while the pre- and post- treatment blood samples were integrated into a separate object. Both datasets were integrated using a SCTransform() and integration workflow. During integration, we regressed out the percent mitochondrial reads minimize the impacts on clustering results and used the top 2500 variable features as integration anchors. Following data integration, ideal clustering parameters (tumor: res = 0.8, dims = 45,

min.dist = 0.3, n. neighbors = 30; blood: res = 1.2, dims = 45, min.dist = 0.3, n. neighbors = 30) were determined using the R package clustree. Dimension reduction and visualization was then completed, and the data were projected using 2-dimensional, non-linear uniform manifold approximation and projection (UMAP) plots.

Cell classification

For the tumor and blood datasets, annotations were generated using the results of unsupervised clustering of the integrated datasets while cross referencing the cell type annotations presented in chapter 3 and chapter 2, respectively. Cross referencing was completed by using module scoring with short gene lists, reference mapping to the annotated naïve atlas presented in chapter 3 (tumor) or the combined healthy and OS blood atlas in chapter 2 (blood), and by using canonical markers.

Feature visualization

Feature expression was visualized using feature plots and heatmaps. Selected features were chosen based on the identification of a feature to be differentially expressed when contrasting TDI treated and naïve cells. Feature plots depict normalized expression for each feature and are on variable scales. Datasets were down sampled to the value equal to the condition (treated or naïve) which contributed the fewest number of cells. Heatmaps represent the scaled expression for each feature within the dataset. Scaled expression values represent the deviation in expression values relative to the average across the cells being studied. As such values greater than 0 indicated higher than average expression and values less than 0 indicate lower than average expression. To evaluate statistically significant changes in expression between treated and naïve cells within each cell type, we used two-sided Wilcoxon Rank Sum tests and significance was noted if the adjusted P value was less than 0.01. Significant changes within a cell type are indicated by an asterisk with the higher expression group marked by an as asterisk.

Cell abundance analysis

All cell abundance comparisons were made using percentage of parent. To make statistical inferences on changes in cell abundance two-sided Wilcoxon Rank Sum tests were used and P values less than 0.1 were discussed at significantly different.

Differential gene expression analysis

Differential gene expression analysis was completed using the FindMarkers() function (two-sided Wilcoxon Rank Sum test) and the results were plotted on a scatter plot depicting average normalized count of each feature grouped by classification (Y versus X). Values that fell below a line of y-intercept = 0; slope = 1 and were determined to be higher in the X category using the FindMarkers() function (log2(fold change) < -0.58 and adjusted P value < 0.01), were discussed as increased in X or decreased in Y. Alternatively, values that fell above a line of y-intercept = 0; slope = 1 and were determined to be higher in the Y category using the FindMarkers() function (log2(fold change) > 0.58 and adjusted P value < 0.01), were discussed as increased in Y or decreased in X.

Gene set enrichment analysis

The R package singleseqgset was used to complete GSEA on cell type clusters. The tool uses a competitive gene set enrichment test that was based on a Correlation Adjusted MEan RAnk gene set test (14). The log2fold change and mean expression for every feature within each cell type was calculated and then used to complete GSEA. P values were corrected for multiple comparisons using a false discovery rate (FDR) method and corrected P values were filtered to only retain terms in which at least one cell types had a value less than 0.05. The enrichment values using the HALLMARKS gene set terms were scaled then plotted using a heatmap (15).

Data and software availability

Raw sequencing data (FASTQ format) generated in this Chapter (6 TDI treated tumors and 5 TDI treated blood samples) are available upon request and will be released upon publication of the associated manuscript. The 5 pretreatment blood samples made up a portion

of the OS blood atlas in the chapter 2 dataset, while the 6 naïve OS tumor samples were generated for the analysis completed in chapter 3. The analysis code and software versions are available at https://github.com/dyammons/canine_OS_tx_scRNA.

Results

Losartan, ladarixin, and toceranib combination immunotherapy broadly reduces tumor immune infiltrates

To evaluate the effects of losartan, ladarixin, toceranib combination immunotherapy (TDI), we integrated 6 tumor samples obtained from treated dogs with 6 treatment-naïve dogs. Cell type annotations of the integrated dataset were established using our recently reported canine OS atlas as the reference to identify anchors and transfer annotations (chapter 3). In total, the integrated tumor dataset consisted of 70,179 cells with 46% tumor, 8% stromal cells, and 46% immune cells (Figure 5.1b). We next evaluated how cell type percentages were modified with treatment. The analysis revealed reductions in tumor-infiltrating monocytes (TIMs; P value = 0.065) and tumor-associated macrophages (TAMs; P value = 0.093) as well as unexpected treatment induced reductions in T cells (P value = 0.026) and dendritic cells (P value = 0.041) (Figure 5.1c). Together this suggests the TDI regiment may have broader impacts on the tumor immune microenvironment than expected. Because treatment reduced the number of infiltrating immune cells, we also observed a relative increased in the proportion of fibroblasts and malignant osteoblasts. Overall, the modulation of the tumor immune compartment suggests that the treatment was able to have an impact on the TME within the 14-day treatment window.

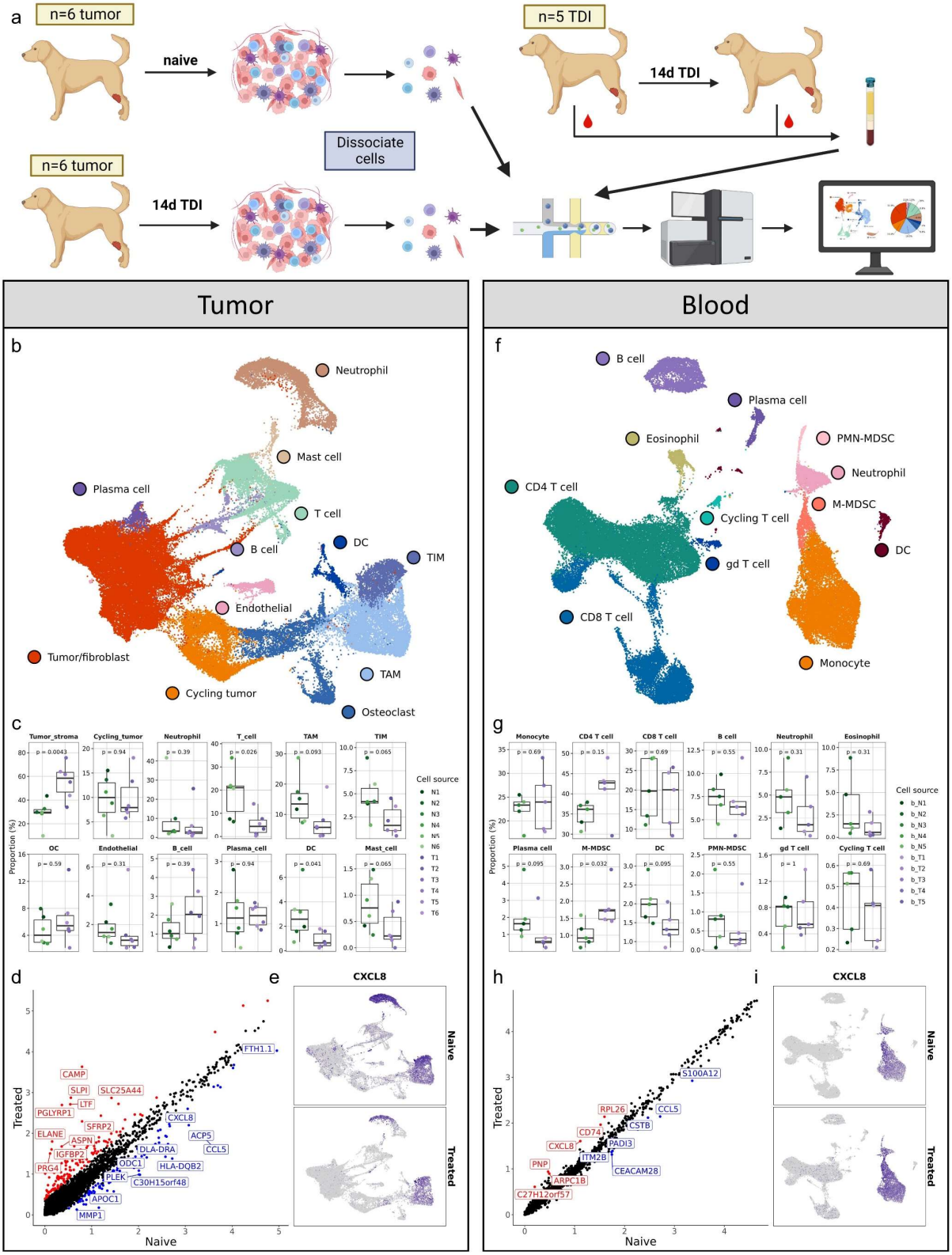

a
n=6 tumor
naive
n=5 TDI
14d TDI
Dissociate cells
n=6 tumor
14d TDI
Tumor
Blood
b
Neutrophil
Mast cell
Plasma cell
T cell
B cell
DC
TIM
Endothelial
Tumor/fibroblast
TAM
Cycling tumor
Osteoclast
c
Tumor_stroma
Cycling_tumor
Neutrophil
T_cell
TAM
TIM
OC
Endothelial
B_cell
Plasma_cell
DC
Mast_cell
Proportion (%)
Cell source
d
e
CXCL8
Naive
Treated
f
B cell
Plasma cell
Eosinophil
PMN-MDSC
Neutrophil
CD4 T cell
M-MDSC
Cycling T cell
DC
gd T cell
CD8 T cell
Monocyte
g
Monocyte
CD4 T cell
CD8 T cell
B cell
Neutrophil
Eosinophil
Plasma cell
M-MDSC
DC
PMN-MDSC
gd T cell
Cycling T cell
Proportion (%)
Cell source
h
i
CXCL8
Naive
Treated

Figure 5.1: Triple drug immunotherapy depletes tumor immune cells and modulates chemokine expression. (a) Diagram of study design. (b) UMAP representation of 70,179 cells obtained from the primary tumors of 6 treatment-naive dogs and 6 dogs that received triple drug immunotherapy (TDI). (c) Box plots of cell percentages in TDI (Treated) and naive dogs. (d) Scatter plot depicting average normalized expression of each feature in naive (x-axis) and TDI (y-axis) dogs. Differentially expressed features (as determined by Wilcoxon Rank Sum test) are red (up in treated) or blue (down in treated). (e) Feature plots split by treatment status depicting normalized expression of differentially expressed features. Grey/light purple indicates low expression, while positivity is indicated by dark purple colorization. Scales are variable between plots. (f) UMAP representation of 29,320 cells obtained from blood of 5 dogs pre- and 14 days post-TDI. (g) Box plots comparing cell percentages of circulating leukocytes between TDI (b_T) and pre-TDI (b_N) dogs. (h) Scatter plot depicting average normalized expression of each feature in pre- (Naïve) and post-TDI (Treated) dogs. (i) Feature plots split by treatment status depicting normalized expression of differentially expressed features.

Following cell abundance analysis, we next evaluated how treatment impacted the transcriptomes of cells in the TME. Using a Wilcoxon Rank Sum test, we identified 181 features to be upregulated by treatment and 134 downregulated (Figure 5.1d). Of the differentially expressed genes, CXCL8 was determined to be downregulated following treatment. This was unexpected as we anticipated increases in CXCL8 following treatment, as ladarixin blocks the CXCL8-CXCR1/2 signaling axis. To investigate further, we visualized CXCL8 expression in naïve and TDI treated samples which confirmed expression to be largely specific to TIMs/TAMs and neutrophils (Figure 5.1e). In addition to CXCL8 modulation, DGE analysis revealed CCL4, CCL5, CCL7, CCL14, and CXCL14 were differentially expressed. Treatment induced reductions in CCL4 and CCL5 were largely mediated by reductions in expression levels within TIMs/TAMs and T cells. The remaining chemokines exhibited increased expression in treated tumors with CCL7/CXCL14 largely driven by tumor/fibroblast and CCL14 increases mediated by TIMs/TAMs. Together this analysis revealed chemokine modulation was broad and had impacts on non-immune cells within the tumor.

Triple drug immunotherapy increases proportion of circulating monocytic myeloid-derived suppressor cells

Due to potential of TDI to exhibit systemic effects we also completed single-cell RNA sequencing of 5 dogs pre- and 14 days post-treatment. Integration of the 10 blood samples (29,320 cells) revealed 12 major cell types, with the only detected shift in cell abundances being associated with increases in the proportions of monocytic myeloid-derived suppressor cells (M-MDSCs) in treated samples (Figure 5.1e/f). Although a high burden of circulating M-MDSCs is typically associated with poorer clinical outcomes (16), it is possible that the increase represents the inability of M-MDSCs to migrate to the tumor due to blockade of the CXCR1/2-CXCL8 or the CCR2/CCL2 axis. It is also possible that the increase in M-MDSCs is due to disease continuing to progress over the 14-day treatment window. Differential gene expression analysis of all circulating cells revealed minimal changes in gene expression (6 upregulated; 6 downregulated in treated). Despite few differentially expressed features, CXCL8 expression was identified to be upregulated in circulating immune cells (Figure 5.1g). This observation was expected as signaling blockade by ladarixin should induce a compensatory increase in CXCL8 production due to competition for CXCR1/2 binding. The distribution of CXCL8 was largely specific to monocytes and neutrophils, but treatment also induced lymphocytes to contribute to increased CXCL8 transcript abundance (Figure 5.1h). Overall, analysis of circulating leukocytes indicates that the impacts of TDI extends outside of the tumor, as systemic effects to immune cells were also observed.

TDI treatment drives broad chemokine axis modulation

With the evidence that TDI impacted the CXCR1/2-CXCL8 and other chemokine signaling axes, we next wanted to complete a broader screen to assess if treatment drove cells to use alternative recruitment mechanisms. First, we evaluated expression of all CC- and CXC-chemokine receptors/ligands annotated in the canine genome within immune cells of the tumor

samples. The analysis revealed 7 chemokine ligands and 3 chemokine receptors to be differentially expressed (Figure 5.2). The previously noted reduction in CXCL8 in the tumor was further identified to be statistically driven by reduced transcript levels in neutrophils, T cells, and B cells. Three chemokines (CCL14, CCL7, and CCL13) were upregulated in tumor-associated macrophages in TDI treated cells, which may represent an alternative recruitment pathway used to compensate for the inability to efficiently recruit with CXCL8. When completing the same analysis in circulating leukocytes, fewer chemokine associated features were modulated, but the increase in CXCL8 observed earlier was determined to be driven by increased expression in monocytes (Figure 5.3). Together, these findings suggest TDI was able to modulate chemokine signaling, with neutrophils and macrophages being particularly impacted by the intervention.

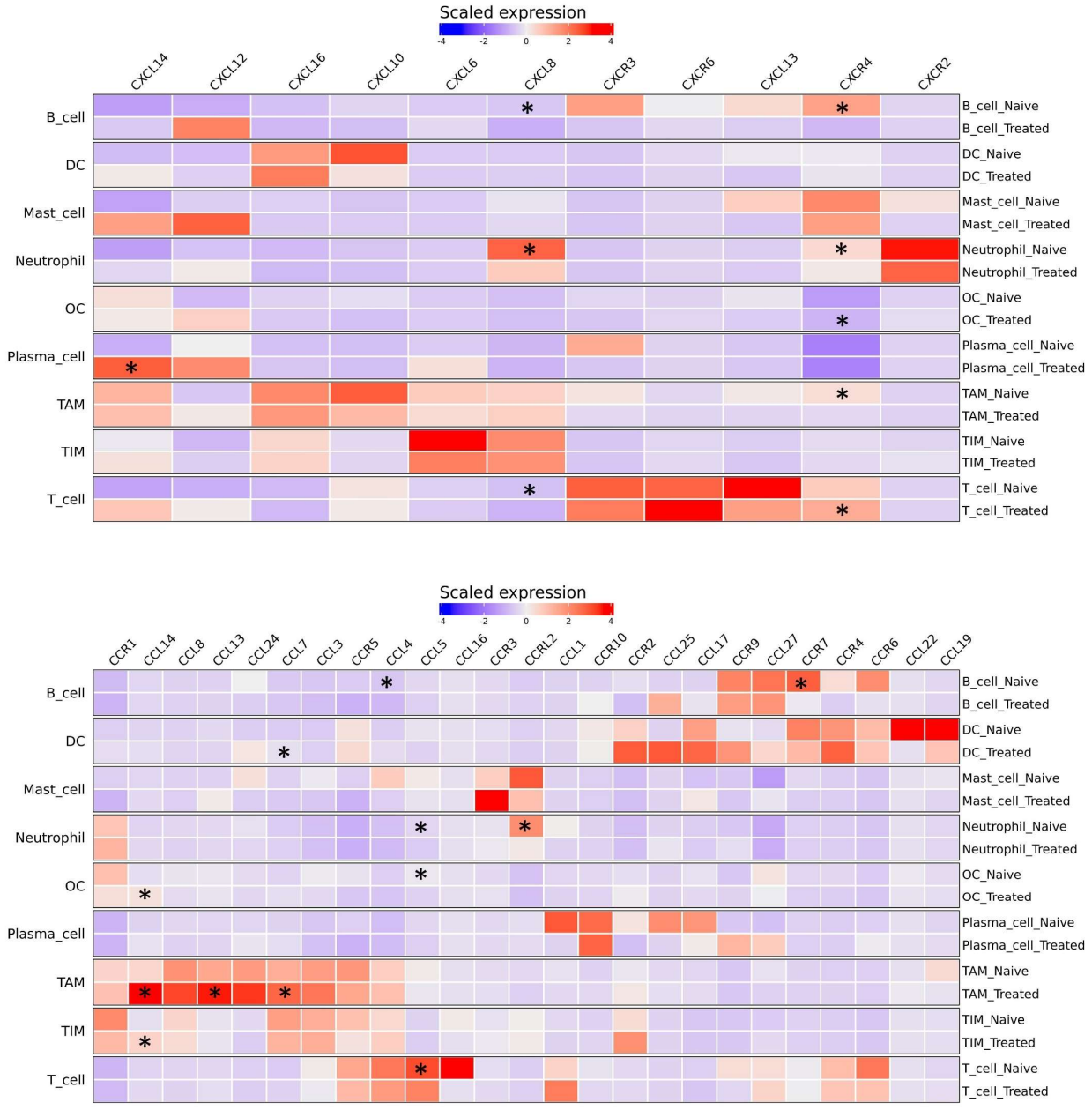


Figure 5.2: Impact of TDI on chemokine expression in tumor immune cells. Heatmap depicting the scaled expression of CXC- (top) and CC- (bottom) chemokine receptors and ligands. Expression levels were contrasted between Treated and Naïve cells within each cell type using two-sided Wilcoxon Rank Sum tests. Significant changes (P value < 0.01) are indicated by an asterisk with the higher expression group marked by an asterisk.

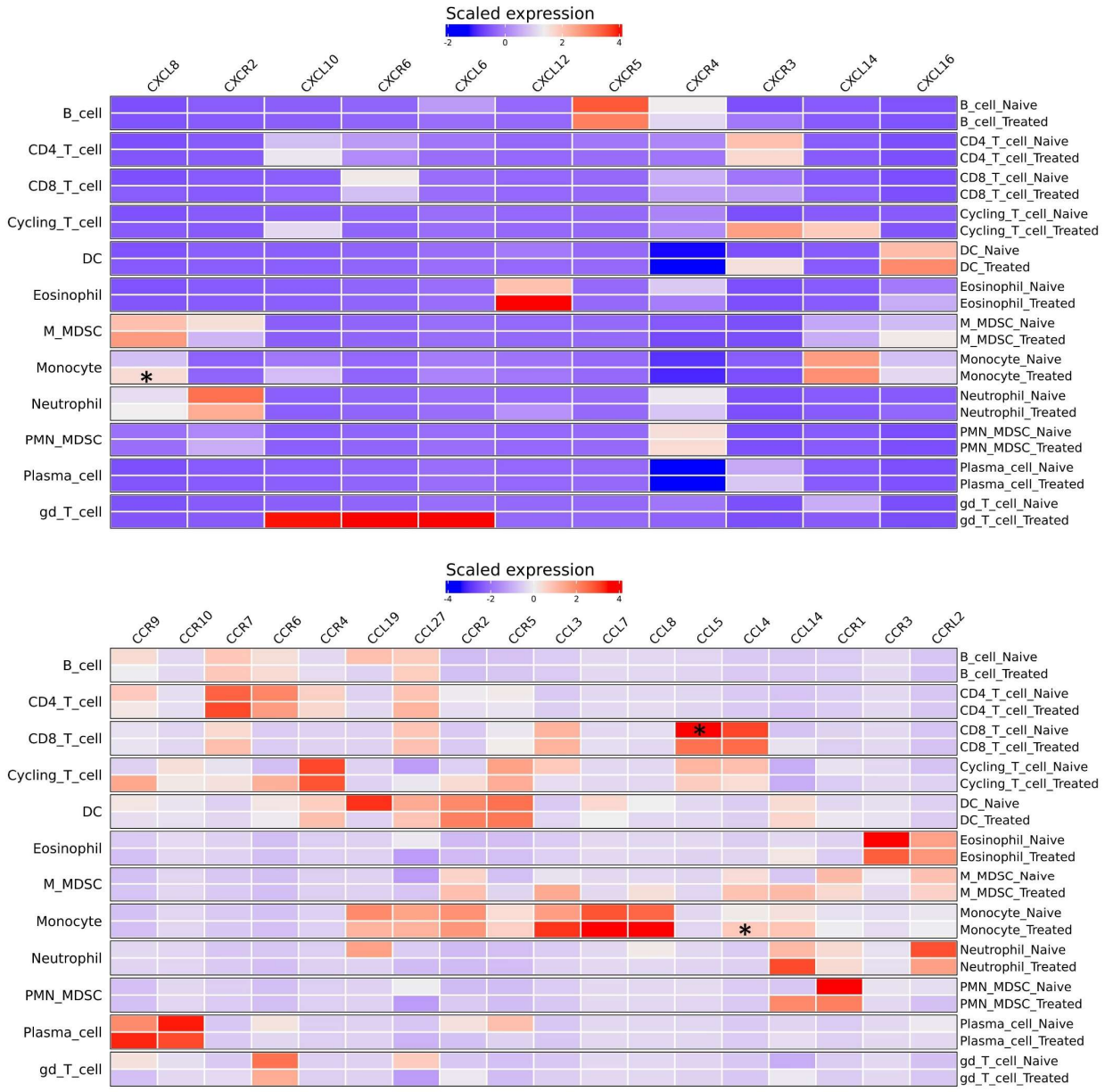


Figure 5.3: Impact of TDI on chemokine expression in circulating immune cells. Heatmap depicting the scaled expression of CXC- (top) and CC- (bottom) chemokine receptors and ligands. Expression levels were contrasted between Treated and Naïve cells within each cell type using two-sided Wilcoxon Rank Sum tests. Significant changes (P value < 0.01) are indicated by an asterisk with the higher expression group marked by an asterisk.

Triple drug immunotherapy drives relative fibroblast expansion and promotes hypoxic gene signatures

Aside from broad immune cell depletion and modulation to chemokine expression, there were minimal changes to the immune compartment. As such, we next focused on TDI induced changes observed in the tumor and fibroblast compartments. Through completion of independent reclustering we identified two malignant osteoblast clusters (c1, c6), one malignant osteoblast cluster with a hypoxic gene expression pattern (c5), two cycling malignant osteoblast clusters (c3, c4), and three transcriptionally distinct fibroblast clusters (c0, c2, c7) (Figure 5.4a). Due to previously reports indicating a low number of fibroblasts in treatment-naïve OS tumors and difficulty distinguishing malignant osteoblasts from fibroblasts (as described in chapter 3), we used a fibroblast gene signature to determine which clusters most closely resembled fibroblasts. Using this approach, we identified the largest cluster (c0), c2, and c7 to exhibit the most pronounced fibroblast signature, which acted to confirm the cell type annotation (Figure 5.4b). Next, we used a two-sided Wilcoxon Rank Sum test to evaluate the cellular composition of TDI treated and naïve tumors. The analysis revealed that fibroblasts are overrepresented in treated tumors, while cycling malignant osteoblasts are underrepresented (Figure 5.4c). This suggests that TDI either drove an expansion of fibroblasts within the tumor or that the cycling compartment of tumor cells was depleted with TDI intervention.

To further investigate the impacts of TDI we completed DGE analysis using a two-sided Wilcoxon Rank Sum test which revealed 166 features to be overexpressed and 90 to be under expressed in TDI treated tumors relative to naïve. Of note the expression of insulin growth factor (IGF) associated features was increased in TDI treated tumors, with a specificity for fibroblasts (Figure 5.4d). IGF activation in tumor microenvironment has been reported to promote tumor growth by activating STAT3 and enhancing the invasiveness of tumor cells (17). We then used GSEA and hierarchical clustering of the 8 clusters split out by treatment status (TDI treated or naïve). We unexpectedly found cell type clusters did not pair off, which suggests treatment status

drove hierarchical clustering more than cell type signatures (Figure 5.4e). With the exception of cycling cells, TDI treated clusters exhibited strong enrichment for hypoxia, angiogenesis, epithelial to mesenchymal transition (EMT), and apoptosis. Evaluation of cell metabolism associated terms indicated naïve tumors favored oxidative phosphorylation and fatty acid metabolism while TDI treated cells were moderately enriched for glycolytic pathways. Together, the analysis suggests transcriptomic signatures of TDI treated tumors are more hypoxic and glycolytic with the potential for increased invasiveness.

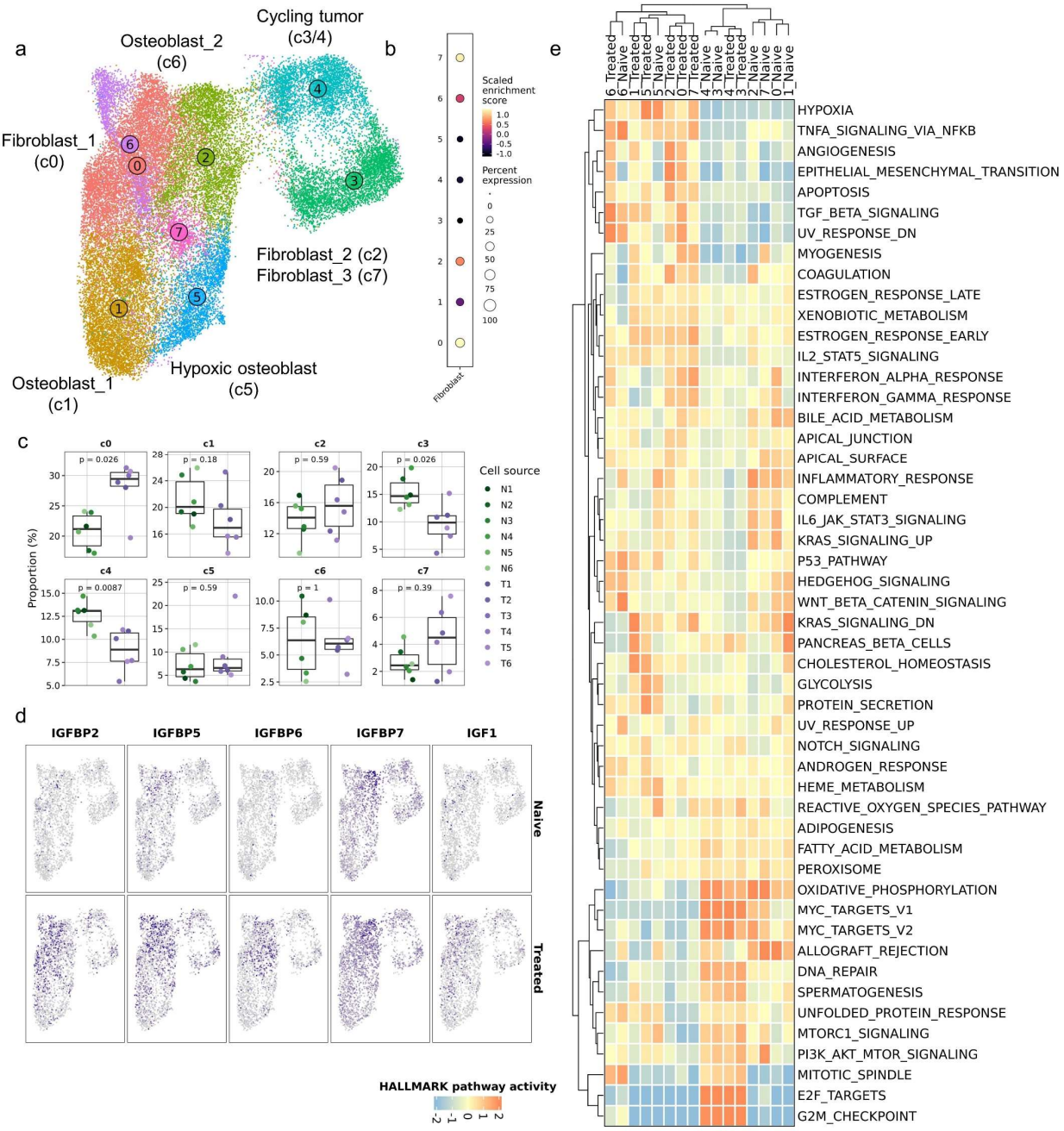


Figure 5.4: TDI induces relative increase in fibroblasts and skews tumor/fibroblasts toward hypoxic gene signatures. (a) UMAP representation of 34,027 tumor and fibroblasts. (b) Dot plot depicting scaled enrichment score of a fibroblast gene signature. (c) Box plots of cell percentages in TDI (T) and naïve (N) dogs. (d) Feature plots split by treatment status depicting normalized expression of differentially expressed features. Grey/light purple indicates low expression, while positivity is indicated by dark purple colorization. Scales are variable between plots. (e) Heatmap of scaled GSEA pathway activity for HALLMARK terms, with hierarchical clustering of cell types (columns) and terms (rows) using Euclidean distance.

Discussion

Using single-cell RNA sequencing we were able to interrogate how a myeloid targeted oral immunotherapy consisting of losartan, ladarixin, and toceranib (TDI) altered the OS tumor microenvironment. Our analysis indicated that TDI was able to deplete tumor-infiltrating myeloid cells, but also resulted in a reduction of T cells and dendritic cells. We observed modulation to key chemokine pathways targeted by the combination intervention strategy and found CCL14, CCL7, and CCL13 as potential adaptive chemokine pathway to circumvent the blockade of CXCR1/2 and CCR2 by ladarixin and losartan, respectively. Intervention may have disrupted M-MDSC migration to the tumor, as the cell abundances of circulating M-MDSC increased after 14 days of treatment. Furthermore, we noted a marked relative increase in fibroblast abundances and modifications to osteoblast transcriptional programs suggesting TDI impact osteoblasts and fibroblasts in addition to the targeted immune cells. Overall, the findings presented here provide novel insights of how myeloid targeted therapeutics impact the OS TME which may reveal mechanisms of treatment resistance in OS.

Through use of scRNA sequencing we were able to identify unexpected treatment induced changes in the TME and circulating immune cells. In circulation, we observed increases in the relative abundance of monocytic myeloid-derived suppressor cells. It is possible that blockade of chemokine signaling resulted in the retention of M-MDSCs in circulation due to inefficient recruitment to the tumor, which could suggest a positive change. However, there is substantial evidence that increased pre-treatment M-MDSC burden is correlated with poor clinical outcomes (16,18), so the significance in the context of post-TDI treatment requires further investigation.

In the tumor we observed broad depletion of immune cell cells which included reductions in the relative proportions of T cells and dendritic cells. Under naïve conditions, OS is typically described a “cold” tumor with few numbers of immune infiltrates, so the reduction may act to further reduce immune reactivity to the tumor. This discovery suggests further combination with

immune stimulatory therapies, such as radiation therapy, may be valuable to enhance immune cell infiltration. Outside of immune cells we identified TDI to have profound impacts on malignant osteoblasts and stromal fibroblasts. The changes observed included increased enrichment for gene signatures consistent with epithelial to mesenchymal transition, hypoxia, and angiogenesis, as well as marked increases in insulin-like growth factor-1 and associated binding proteins. Together these changes are suggestive of malignant changes in the tumor (19,20), so additional clinical data regarding patient outcomes is necessary to further interpret these findings.

Although this study provides key insight into how myeloid targeted interventions modulated the OS TME, the small sample size and incompletely described clinical outcomes of treated dogs limit the conclusions that can be made. The intratumoral treatment effects were impacted by the inability to obtain pre- and post-treatment samples which increased variability. However, our systemic treatment effects were able to use paired samples which reduced biological confounding, but also introduced the possibility of disease progression confounding the post-treatment sample.

This study largely focused on treatment induced changes within the primary tumor, but the clinical impact from intervention is likely to come from modulation to pre-metastatic and micro-metastatic disease in distant sites. Thus, the changes reported in the primary tumor may be consistent with changes observed in metastatic tissues, but further work is required to describe the effects of treatment on metastatic tissue. Despite these limitations we were able to detect treatment induced changes to all cell types in the tumor and we provided the first characterization of treatment-induced changes in OS using scRNA sequencing.

In summary, we applied single-cell RNA sequencing of losartan, ladarixin, and toceranib treated primary canine osteosarcoma to provide first description of how the tumor microenvironment is modulated following treatment with a myeloid targeted immunotherapy. Findings from a concurrent clinical trial investigating the impacts of TDI on survival and disease

progression will reveal additional insights regarding the biological relevance of the molecular changes reported here.

References

1. Wu C-C, Beird HC, Andrew Livingston J, Advani S, Mitra A, Cao S, et al. Immuno-genomic landscape of osteosarcoma. Nat Commun. 2020;11(1):1008.
2. Liu Y, Feng W, Dai Y, Bao M, Yuan Z, He M, et al. Single-cell transcriptomics reveals the complexity of the tumor microenvironment of treatment-naive osteosarcoma. Front Oncol. 2021;11:709210.
3. Zhou Y, Yang D, Yang Q, Lv X, Huang W, Zhou Z, et al. Single-cell RNA landscape of intratumoral heterogeneity and immunosuppressive microenvironment in advanced osteosarcoma. Nat Commun. 2020;11(1):6322.
4. Schiffman JD, Breen M. Comparative oncology: what dogs and other species can teach us about humans with cancer. Philos Trans R Soc B Biol Sci. 2015;370(1673):20140231.
5. Regan DP, Coy JW, Chahal KK, Chow L, Kurihara JN, Guth AM, et al. The angiotensin receptor blocker losartan suppresses growth of pulmonary metastases via AT1R-independent inhibition of CCR2 signaling and monocyte recruitment. J Immunol. 2019;202(10):3087–102.
6. London CA, Malpas PB, Wood-Follis SL, Boucher JF, Rusk AW, Rosenberg MP, et al. Multi-center, placebo-controlled, double-blind, randomized study of oral toceranib phosphate (SU11654), a receptor tyrosine kinase inhibitor, for the treatment of dogs with recurrent (either local or distant) mast cell tumor following surgical excision. Clin Cancer Res. 2009;15(11):3856–65.
7. Regan DP, Chow L, Das S, Haines L, Palmer E, Kurihara JN, et al. Losartan blocks osteosarcoma-elicited monocyte recruitment, and combined with the kinase inhibitor toceranib, exerts significant clinical benefit in canine metastatic osteosarcoma. Clin Cancer Res. 2022;28(4):662–76.
8. Piro G, Carbone C, Agostini A, Esposito A, De Pizzol M, Novelli R, et al. CXCR1/2 dual-inhibitor ladarixin reduces tumour burden and promotes immunotherapy response in pancreatic cancer. Br J Cancer. 2023;128(2):331–41.
9. Zhang L, Li Z, Skrzypczynska KM, Fang Q, Zhang W, O'Brien SA, et al. Single-cell analyses inform mechanisms of myeloid-targeted therapies in colon cancer. Cell. 2020;181(2):442–59.
10. Wang L, Sfakianos JP, Beaumont KG, Akturk G, Horowitz A, Sebra RP, et al. Myeloid Cell–associated Resistance to PD-1/PD-L1 Blockade in Urothelial Cancer Revealed Through Bulk and Single-cell RNA SequencingMyeloid Cell-Associated Resistance to PD-1/PD-L1 Blockade. Clin Cancer Res. 2021;27(15):4287–300.
11. Bukhari S, Henick BS, Winchester RJ, Lerrer S, Adam K, Gartshteyn Y, et al. Single-cell RNA sequencing reveals distinct T cell populations in immune-related adverse events of checkpoint inhibitors. Cell Reports Med. 2023;4(1).
12. Hao Y, Hao S, Andersen-Nissen E, Mauck III WM, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. Cell. 2021;184(13):3573–87.
13. McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. Cell Syst. 2019;8(4):329–37.

14. Wu D, Smyth GK. Camera: a competitive gene set test accounting for inter-gene correlation. Nucleic Acids Res. 2012;40(17):e133–e133.

15. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database hallmark gene set collection. Cell Syst. 2015;1(6):417–25.

16. Wang P-F, Song S-Y, Wang T-J, Ji W-J, Li S-W, Liu N, et al. Prognostic role of pretreatment circulating MDSCs in patients with solid malignancies: A meta-analysis of 40 studies. Oncoimmunology. 2018;7(10):e1494113.

17. Jentzsch T, Robl B, Husmann M, Bode-Lesniewska B, Fuchs B. Worse prognosis of osteosarcoma patients expressing IGF-1 on a tissue microarray. Anticancer Res. 2014;34(8):3881–9.

18. Draghiciu O, Lubbers J, Nijman HW, Daemen T. Myeloid derived suppressor cells—an overview of combat strategies to increase immunotherapy efficacy. Oncoimmunology. 2015;4(1):e954829.

19. MacEwen EG, Pastor J, Kutzke J, Tsan R, Kurzman ID, Thamm DH, et al. IGF-1 receptor contributes to the malignant phenotype in human and canine osteosarcoma. J Cell Biochem. 2004;92(1):77–91.

20. Tam SY, Wu VWC, Law HKW. Hypoxia-induced epithelial-mesenchymal transition in cancers: HIF-1α and beyond. Front Oncol. 2020;10:486.

CHAPTER 6: Reprogramming the canine glioma microenvironment with tumor vaccination plus oral losartan and propranolol induces objective responses[3]

Summary

Malignant gliomas have a highly immune-suppressive tumor microenvironment (TME) which renders them largely unresponsive to conventional therapeutics. Therefore, the current study evaluated a therapeutic protocol designed to overcome the immune barrier by combining myeloid cell–targeted immunotherapy with tumor vaccination. We utilized a spontaneously occurring canine glioma model to investigate an oral TME modifying immunotherapy in conjunction with cancer stem cell (CSC) vaccination. Dogs were treated daily with losartan (monocyte migration inhibitor) and propranolol (myeloid-derived suppressor cell depleting agent) plus anti-CSC vaccination on a biweekly then monthly schedule. Tumor volume was monitored by MRI and correlated with patient immune responses. We enrolled 10 dogs with histologically confirmed gliomas into a prospective, open-label clinical trial to evaluate the immunotherapy protocol. Partial tumor regression was observed in 2 dogs, while 6 dogs experienced stable disease, for an overall clinical benefit rate of 80%. Overall survival times (median = 351 days) and progression-free intervals (median = 163 days) were comparable with prior studies evaluating surgical debulking followed by immunotherapy. Dogs with detectable anti-CSC antibody responses had an increased overall survival time relative to dogs that did not generate antibody responses (vaccine responder MST = 500 days; vaccine nonresponder MST = 218 days; P = 0.02). These findings suggest that combining myeloid cell–targeted oral immunotherapy with tumor vaccination can generate objective tumor responses, even in the absence of conventional therapy. Overall, this approach has promise as a readily implemented therapeutic strategy for use in patients with brain cancer.

[3]Ammons, Dylan T., et al. "Reprogramming the canine glioma microenvironment with tumor vaccination plus oral losartan and propranolol induces objective responses." Cancer Research Communications 2.12 (2022): 1657-1667. https://doi.org/10.1158/2767-9764.CRC-22-0388

Introduction

Malignant gliomas are an aggressive tumor of the brain that carry a dismal prognosis for long-term survival. For instance, humans diagnosed with a grade III glioma exhibit a median progression-free interval of 11 weeks and a median overall survival time of 39 weeks (1). Of note, there are remarkable similarities in survival time, clinical symptoms, and histological properties between human and canine malignant gliomas (2,3). The similarities in disease progression and the accessibility of canine patients makes the dog spontaneous glioma model valuable for investigation of novel therapeutics, with the potential for accelerated translation to human medicine (4,5).

In recent years immunotherapy has come to the forefront of anti-cancer therapeutics, with the blockade of immune suppressive checkpoint pathways being adopted as a leading treatment for a variety of cancers. Such checkpoint blocking immunotherapies are highly effective in a subset of cancers but have proven to be largely ineffective for treatment of adult and pediatric gliomas (6). The lackluster responses in brain tumors are partly due to the immunologically cold nature of the malignancy, which is a result of a highly immune suppressive tumor microenvironment (TME) that inhibits antitumor responses (7). Therefore, recent efforts for brain cancer treatment have focused on strategies to modify the TME, with the goal of enhancing the efficacy of existing immunotherapies.

In the present study, we targeted myeloid cells for depletion from the TME using two repurposed drugs: losartan and propranolol. The choice of these two drugs was based on our previous demonstration that losartan can block inflammatory monocyte migration and the findings of other groups which indicate that propranolol can inhibit myeloid-derived suppressor cell (MDSC) mobilization and immune suppression (8–10). When losartan is administered in a continuous manner, its off-target activity as a C-C chemokine receptor type 2 (CCR2) antagonist blocks inflammatory monocyte migration and can eventually lead to depletion of tumor-associated

macrophages (8). The continuous administration of propranolol has been demonstrated to interact with the β-2 adrenergic receptor on myeloid cells to inhibit STAT3 signaling and suppress the mobilization and activation of MDSCs (9,10). Importantly, it has been reported that both drugs reach effective concentrations in cerebrospinal fluid and brain tissues (11,12). Therefore, to disrupt the immune suppressive TME, we used daily losartan and propranolol to deplete tumor resident myeloid cell populations in dogs with glioma.

In addition to TME modification, we used an allogeneic vaccination strategy to generate tumor specific immune responses. The allogeneic vaccine consisted of pooled lysates derived from three non-glioma canine cancer cell lines. The cell lines were screened for cancer stem cell (CSC) antigen expression and selected based on their enriched CSC antigen levels. This vaccine strategy was selected because gliomas are reported to have a high abundance of CSCs and to be dependent on this population for growth and spread (13,14). To upregulate CSC antigen expression, the cell lines used in this study were cultured in nonadherent conditions with cytokine supplementation and serum free growth medium (15). Finally, to maximize vaccine immunogenicity, we delivered the lysate using cationic liposomes complexed with TLR-3/9 agonists (polyinosinic-polycytidylic acid and CpG oligonucleotides), which we have previously shown effectively cross-primes CD8 T cell responses to protein antigens (16,17).

In the present study, 10 dogs with histologically confirmed gliomas were enrolled into a trial evaluating the impact of combined TME disruption and therapeutic cancer vaccination. All animals in the study were continuously treated with losartan and propranolol, while also receiving bi-weekly and then monthly immunizations with a canine CSC vaccine. Throughout the study, serial magnetic resonance imaging (MRI) and clinical assessments were completed to monitor the impact of combined immunotherapy on tumor growth and animal survival. Key study findings were that the immunotherapy protocol led to an 80% clinical benefit rate, with 2 animals achieving partial tumor regression and 6 dogs experiencing stable disease. Importantly, overall survival times of study dogs were comparable to those reported in canine glioma patients that underwent

surgical tumor debulking prior to vaccination (18). Thus, our findings provide guidance for the use of repurposed TME modifying drugs combined with other immunotherapies for the treatment of gliomas.

Methods

Study design and enrollment criteria

This study was designed as a prospective, open-label clinical trial to enroll dogs with newly diagnosed gliomas. All studies were approved by the Colorado State University (CSU) Institutional Animal Care and Use Committee and the CSU Clinical Review Board. All dog owners provided informed consent prior to enrollment into the trial.

Inclusion criteria included a diagnostic MRI indicating the presence of a supra-tentorial lesion consistent with a glioma. For inclusion in this report, histologic confirmation of glioma was required with 4 study dogs being histologically diagnosed post-mortem (trial 1A, dogs 1-4), and 6 study dogs being histologically diagnosed before treatment using biopsy tissues (trial 1B, dogs 5-10) acquired with a neuronavigation system (Brainsight; Rogue Research; Montreal, Canada; NICO Myriad, NICO Corporation, Indianapolis, IN). Study dogs were screened and confirmed to be free of concurrent medical conditions, including hypertension and renal dysfunction.

Exclusion criteria included the presence of other cancers or prior treatment of the glioma. To avoid immunosuppressive effects, any dog that was actively receiving corticosteroids was not enrolled until their corticosteroid dosage was reduced to less than 0.5 mg/kg/day for at least 72 hours prior to vaccination.

Losartan and propranolol dosing schedule

All study dogs were treated with orally administered high-dose losartan (10 mg/kg) and propranolol (1 mg/kg); both administered every 12 hours (Figure 6.1). Blood pressure was monitored throughout the study to ensure that hypotension did not develop. Complete blood count

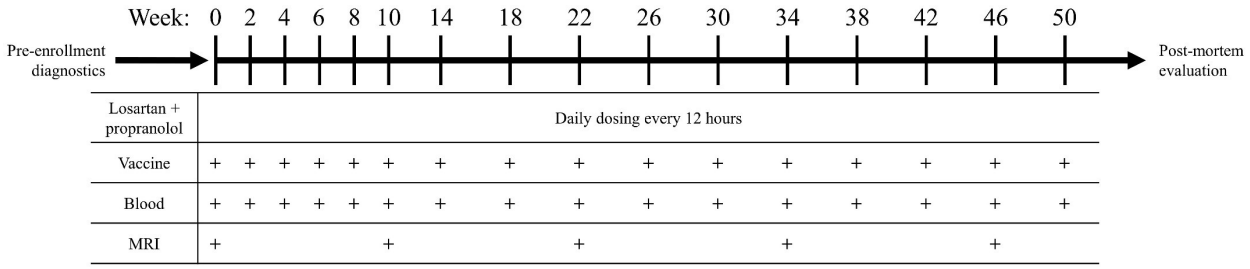


Figure 6.1. Study design. Dogs diagnosed with a histologically confirmed glioma were selected for inclusion into the trial. A total of 10 dogs were enrolled and received daily losartan plus daily propranolol with biweekly vaccination for the first 10 weeks then monthly through the duration of survival. After the first 10 weeks, blood collection and MRI were completed every three months throughout the duration of the study. Following death, a necropsy was completed.

and serum biochemical panels were performed at study initiation, 1 month, and 3 months to assess renal and hepatic function.

Tumor cell lines and screening for high CSC marker expression

A series of 13 canine tumor cell lines were screened for upregulation of CSC markers following culture in tumor spheroid medium. The cell lines were obtained from the Flint Animal Cancer Center tumor cell line bank or were established in our laboratory (15). All tumor cell lines were confirmed to be of canine origin and mycoplasma free. Tumor cell lines were initially maintained in Gibco Minimum Essential Medium (MEM) containing 10% fetal bovine serum (FBS), 1% L-glutamine, and 1% penicillin/streptomycin (Thermo Fisher, Waltham, MA; Peak Serum Inc., Wellington, CO). After 7 days of expansion, cells were transferred to ultra-low attachment plates (Corning, Kennebunk, ME) and cultured in spheroid medium. Spheroid generating medium consisted of Dulbecco's Modified Eagle Medium/Ham's F-12 Nutrient Mixture (Thermo Fisher), B-27™ Supplement (Thermo Fisher), 20 ng/ml basic fibroblast growth factor (Peprotech, Rocky Hill, NJ), 20 ng/ml epidermal growth factor (Peprotech), 2 µg/ml heparin (Sigma-Aldrich, St. Louis, MO) and penicillin/streptomycin (Thermo Fisher). Medium changes were completed every other day for a total of 7 days in spheroid medium.

To quantify CSC antigen expression, each cell line was screened by flow cytometry for expression of 8 CSC markers (CD90, CD44, CD133, CD34, CD24, Sca1, CD117, and Oct3/4),

as described previously (19). From this screen, 3 cell lines (Jenny {melanoma}, Gracie {osteosarcoma}, and Bliley {transitional cell carcinoma}) with highest CSC marker expression were selected for use in the allogeneic CSC vaccine (Supplemental figure 6.1).

Preparation of CSC cell lysate vaccines

Following a 7-day period of growth in spheroid medium, the spheroids were collected by gravity settling, washed with PBS, then lysed using distilled water. Tumor cell lysates were generated through 5 freeze/thaw cycles in liquid nitrogen, and then 3 sonication cycles on ice. Lysates were passed through 0.2 µm sterile filters (Pall, Port Washington, NY) and resuspended in distilled water. The protein concentration was quantified using a bicinchoninic acid kit (Thermo Fisher) and the vaccine was prepared by pooling equivalent amounts of spheroid lysate (166.6 µg) generated from each of the three cell lines. The pooled tumor lysate was combined with a cationic liposomal-TLR3/9 vaccine adjuvant, as described previously, to generate the vaccine (17,20). Briefly, each vaccine consisted of 500 µg pooled tumor lysate complexed to 500 µl liposome-TLR3/9 adjuvant, in a total volume of 2 ml diluent. Vaccines were then lyophilized and stored at -80°C until administered.

Immunization schedule

Vaccines were reconstituted with 2 ml of sterile water then delivered via two 1 ml sub-cutaneous injections split between each flank. After the initial immunization, vaccines were administered on weeks 2, 4, 6, 8, and 10 then once monthly until death or withdrawal from study (Figure 6.1). Vaccine reactions were monitored by evaluation of the injection site and an owner questionnaire.

Tumor monitoring by MRI and clinical evaluation

Serial MRI evaluations were conducted prior to treatment (trial 1A) or after tumor biopsy (trial 1B) to obtain baseline tumor volume measurements and subsequently at 3-month intervals throughout the duration of the study period. MRI data were acquired on a 1.5 T GE Scanner (Signa HDxt) running 24.x software, using a standard human knee coil. Acquisition protocols

followed the consensus recommendations on standardized magnetic resonance imaging protocols for multicenter canine brain tumor clinical trials (4). This included transverse, sagittal, and dorsal plane T1-weighted pre- and post-contrast sequences; transverse and sagittal T2-weighted sequences; transverse GRE, FLAIR, and DTI sequences; and 3D isotropic T1-weighted images. All T1-weighted sequences were evaluated pre- and post-contrast administration. Three-dimensional T1 anatomical images were performed using the FSPGR BRAVO sequence (FOV = 25.6 cm, matrix = 256x256, flip angle = 12°, TR = 8.19 ms, TE = 3.39 ms, TI = 450 ms, slice thickness = 1.0 mm, NEX = 1).

Image measurements were made by veterinary radiologists, using either Philips IntelliSpace (v.4.4.551, Philips Medical, Cleveland, OH) or Osirix MD (v.10.0.4, Pixmeo SARL, Bernex, Switzerland). The evaluating radiologists were blinded to study date and dog identity when analyzing the imaging data and the software platform used for analysis was consistent over time for an individual dog. Linear measurements were made in all three dimensions on the post contrast T1-weighted images with width measured on the transverse plane, height measured on the sagittal plane, and length measured on the dorsal plane including both contrast enhancing and non-contrast enhancing lesion components. The largest diameter and its orthogonal dimension were recorded. Volumetric measurements were made in transverse plane post-contrast T1-weighted images. When applicable, the volume calculation was corrected for inter-slice gap. Progression of disease was determined based on tumor volume calculations according to the RANO/iRANO criteria, using a 40% volume increase as disease progression, a 65% reduction of tumor volume as a partial response, and all responses in between were classified as stable disease (21,22). Progression-free interval was calculated based on the date in which the MRI measured tumor volume was greater than 40% larger than baseline tumor volume. Dogs that achieved stable disease for 3 months on protocol were counted toward the clinical benefit rate. For cases in which the dog did not survive to the first follow-up MRI (n = 1) the progression-free interval was recorded as the date of death.

Measurement of vaccine-induced antibody responses to canine glioma cells

A canine glioma cell line (J3T, a generous gift of Dr. Peter Dickinson, University of California Davis College of Veterinary Medicine) was propagated under standard plastic adherence conditions (parental) using complete MEM (10% FBS, 1% L-glutamine, and 1% penicillin/streptomycin) or under spheroid forming conditions, as described above. Parental cells were detached by trypsinization, while spheroids were disaggregated using Accumax (Innovative Cell Technologies, Inc., San Diego, CA) to generate single cell suspensions.

For quantification of anti-glioma antibody titers, serum from study dogs obtained prior to vaccination and 4-weeks after the first vaccination was evaluated using a flow cytometric assay. Serial dilutions were made at 4-fold steps then serum solutions were incubated for 60 minutes at room temperature in FACS buffer (5% FBS plus 0.1% sodium azide in PBS). A rabbit anti-dog IgG H&L-FITC antibody (Jackson Laboratories, Bar Harbor, ME, Cat# 304-095-003, RRID:AB_2339350) was added at a dilution of 1:200 and incubated for 30 minutes at room temperature. Finally, 5 µl of 7-aad viability dye (Thermo Fisher) was added, then cells were run on a Beckman Coulter Gallios Flow Cytometer (Indianapolis, Indiana). All serum samples were analyzed by two independent endpoint titers. Endpoint cut off was determined to be a value less than the average plus 3.5 times the standard deviation of at least 6 blanks.

Western blotting to assess antigen recognition by serum from vaccinated dogs

Protein lysate used in the CSC vaccine (15 µg per lane) was electrophoresed under reducing conditions into a 4–15% Mini-PROTEAN® TGX™ gel (BioRad, Hercules, California) then transferred to a polyvinylidene difluoride (PVDF) membrane and blocked with 5% bovine serum albumin (BSA) in TBST (BioRad). Serum obtained prior to and 1-month after the first immunization was used at a dilution of 1:100,000 in 5% BSA to probe the membrane. An HRP conjugated rabbit anti-dog IgG H&L (Jackson Laboratories, Cat# 304-035-003, RRID: AB_2339344) antibody at 1:100,000 was used to detect bound canine antibodies. Finally, the

membrane was imaged using Clarity™ ECL substrate on a ChemiDoc gel imaging system (BioRad). All paired samples were analyzed under the same exposure conditions.

Assessment of antibody recognition of canine glioma cells by confocal microscopy

J3T spheroid cells were fixed for 30 minutes with 4% paraformaldehyde (Thermo Fisher) then embedded in OCT embedding medium as previously described (23). Sectioned spheroids were blocked with 5% normal rabbit serum (Jackson Laboratories) in immunofluorescence buffer (0.2% Triton X-100, 0.1% BSA, 0.05% Tween 20 in PBS). Trial dog serum obtained 4 weeks after enrollment was then diluted 1:500 in immunofluorescence buffer and incubated overnight at 4°C. The following day, a rabbit anti-dog IgG H&L-FITC antibody (Jackson Laboratories) was diluted 1:200 and added for 1 hour at room temperature. Background fluorescence was quenched with a 30-minute incubation in 10 mM copper sulfate + 50 mM ammonium acetate solution then counter stained with DAPI (1 µg/ml) (Sigma-Aldrich). After immunolabeling, slides were visualized using an Olympus IX3 confocal microscope and processed using Olympus cellSens software (Waltham, Massachusetts).

Evaluation of immune infiltrates in tumor biopsies from study dogs and untreated dogs with glioma

Formalin-fixed, paraffin embedded (FFPE) tumor tissues were obtained from 4 study dogs (all study dogs that had tissue samples available) and from a reference group of 14 dogs with untreated gliomas. Immunohistochemistry was performed on FFPE tissues using a Leica BOND-MAX Fully Automated IHC Staining System, with canine cross-reactive antibodies: mouse monoclonal anti-human CD3 (Leica Biosystems, Lincolnshire, IL, Cat# PA0554, RRID:AB_10554454, clone LN10) and mouse monoclonal anti-human Myeloid/Histiocyte antigen (Bio-Rad Cat# MCA874GA, RRID:AB_324314, clone MAC387). Antigen retrieval was performed using Leica Epitope Retrieval Solution 2 (Tris-EDTA buffer, pH 9) for 20 minutes at 95°C. Detection was performed with PowerVision IHC Detection Systems (Leica Biosystems, Inc.),

using a polymeric horseradish peroxidase anti-mouse IgG and DAB chromogen (CD3) or a polymeric alkaline phosphatase anti-mouse IgG and Fast Red chromogen (MAC387).

Whole slide brightfield images of IHC stained slides were digitally captured using an Olympus VS120 slide scanner at 20x magnification and fixed exposure times for all samples. Quantitative image analysis was performed using Visiopharm software (Hesholm, Denmark) and visually confirmed by a veterinary pathologist (DPR). The density of tumor-infiltrating immune cells was calculated as number of immune cells per mm^2 of viable tumor tissue.

Gene expression analysis by NanoString

Tumor biopsies were evaluated histologically to confirm that at least 50% of the sample contained tumor cells. Four 10-µm sections from each sample were used for RNA extraction with an RNeasy FFPE Mini kit (QIAGEN, Hilden, Germany) then evaluated with RNA High Sensitivity assays on the Qubit 2.0 Fluorometer (Invitrogen/LifeTechnologies, Carlsbad, CA, USA) and 5200 Fragment Analyzer Automated CE System (Agilent, Santa Clara, CA, USA). NanoString (NanoString Technologies, Seattle, WA, USA) gene expression analysis was performed using the NanoString Canine IO panel (24). NanoString data was collected using the nCounter Analysis System at the University of Arizona Genetics Core, software 4.0.1.8. Gene expression count data were pre-processed using nSolver software (NanoString) then normalized to housekeeping genes and batch effect regressed using panel standards. Differential gene expression analysis was completed using DESeq2 with corrections for multiple comparisons (25). For corrected p-values, a threshold of 0.1 was used to determine statistical significance.

Statistical analysis

Comparisons of overall survival times and progression-free intervals between dog signalment and tumor parameters were completed using multiple univariate tests, with statistical significance determined using a likelihood-ratio test. For comparisons between two groups, Wilcoxon two sample t-tests were utilized. When making comparisons between pre- and post-vaccine clinical samples we used paired analysis, while all other comparisons used non-paired

tests. Unless specified otherwise, statistical testing that resulted in a p-value of 0.05 or lower was considered statistically significant. All analysis was completed using R software and data were visualized using the R packages ggplot2 and survminer (26–28).

Data availability

The data presented in this study are available from the corresponding author, S.D., upon reasonable request. The R code used to analyze data in the current study is available on GitHub https://github.com/dyammons/CanineGliomaManuscript (DOI: 10.5281/zenodo.7309262).

Results

Study dog demographics

Ten dogs, consisting of 8 oligodendrogliomas and 2 astrocytomas, were enrolled into the clinical trial. Tumors were present in various brain regions with both high (n = 6) and low (n = 4) grade gliomas represented (Table 6.1, Supplemental table 6.1). The average age at enrollment was 7.5 years old, and there were 6 male and 4 female dogs. The breed distribution of the study dogs was skewed towards brachycephalic breeds, which is consistent with the literature, as this population is known to have a predilection to develop gliomas (29,30).

Two study dogs were withdrawn due to progressive clinical signs but were followed to determine date of death (dog 7 withdrawn on day 272; dog 8 withdrawn on day 264). Both of the dogs that were withdrawn early, transitioned to a therapeutic protocol which included radiation therapy with continued daily losartan and propranolol plus monthly vaccination. Survival data from these two dogs was censored at the date of first radiation therapy. On follow-up, dog 8 survived for 13 months and dog 7 survived for 15 months. All other dogs remained on protocol until succumbing to disease.

Table 6.1: Study demographics and objective response to therapeutic intervention

Dog ID	Age (years)	Sex	Breed	Tumor Location	Glioma subclassification	Grade	Trial[b]	Time to maximal response (months)	Maximal response (%)	Objective response
1	7	FS	Boxer	Parietal	Oligodendroglioma	High	1A	3	-58	Stable
2	8	MI	Boxer	Temporal	Astrocytoma	Low	1A	3	-18	Stable
3	8	FS	Boxer	Temporal	Oligodendroglioma	Low	1A	3	-44	Stable
4[a]	8	MC	English Bulldog	Parietal	Oligodendroglioma	High	1A	0	Progression	Progression
5	5	MC	French bulldog	Frontal, Parietal	Oligodendroglioma	High	1B	3	-28	Stable
6	6	FS	Mixed	Parietal, Occipital	Oligodendroglioma	High	1B	3	+43	Progression
7	9	MC	Boston Terrier	Temporal	Oligodendroglioma	High	1B	6	-92	Partial
8	8	MI	Swiss Mountain	Parietal	Astrocytoma	Low	1B	3	-60	Stable
9	8	FS	Boxer	Right Pyriform	Oligodendroglioma	High	1B	6	-88	Partial
10	8	MC	Boxer	Right Pyriform	Oligodendroglioma	Low	1B	3	-9	Stable

[a]Denotes the dog succumbed to disease before follow-up MRI.
[b]Trial 1A dogs had a postmortem histological diagnosis, while 1B received a biopsy pre-enrollment to confirm the diagnosis.

Clinical responses to treatment

Using the MRI based RANO/iRANO classification scheme of objective tumor responses, we determined that 2 dogs exhibited rapid disease progression, 6 had stable disease, and 2 demonstrated a partial tumor response (Table 6.1). All dogs reached their maximal response within 6 months of starting the therapeutic intervention. Overall survival time was calculated based on the time that the dogs survived following administration of the first vaccine. The median survival time for study dogs was 351 days, with a 1-year survival rate of 42% (Figure 6.2A). In addition to overall survival, the progression-free interval was determined to be 163 days from time of initial vaccination to MRI tumor volume progression or death (Figure 6.2B). Tumor growth curves revealed that four dogs (dogs 3, 7, 8, and 9) exhibited a tumor volume reduction (44%, 92%, 60%, and 88%; respectively) following the first immunization and maintained their positive response. Although dogs 7 and 8 had a sustained tumor volume reduction, they showed signs of neurological progression and were withdrawn to receive supplemental radiation therapy. Three other dogs (dogs 1, 2, and 10) exhibited a rapid relapse after an initial tumor volume reduction (Figure 6.2C/D). Additionally, one dog (dog 4) rapidly succumbed to progressive disease. Lastly, we observed that dogs with an initial tumor volume of less than 5 cm^3 had a tendency to exhibit sustained tumor regression following immunotherapy. This observation suggests that dogs with smaller tumor burdens may be better candidates for the immunotherapy only protocol. Overall, these data indicate that combination immunotherapy led to stabilization of tumor growth in a majority of dogs, while a smaller subset achieved partial tumor regression.

A secondary objective of this study was to evaluate the safety of the immunotherapy protocol. Notably, there were no detectable adverse events associated with daily oral administration of losartan or propranolol throughout the duration of the study, and we did not observe hypotension in any of the study dogs. We observed grade 1 adverse events in 5 dogs following at least one immunization which were limited to minor skin irritation at the injection site

that resolved without treatment. All other dogs exhibited no detectable adverse reactions to vaccination.

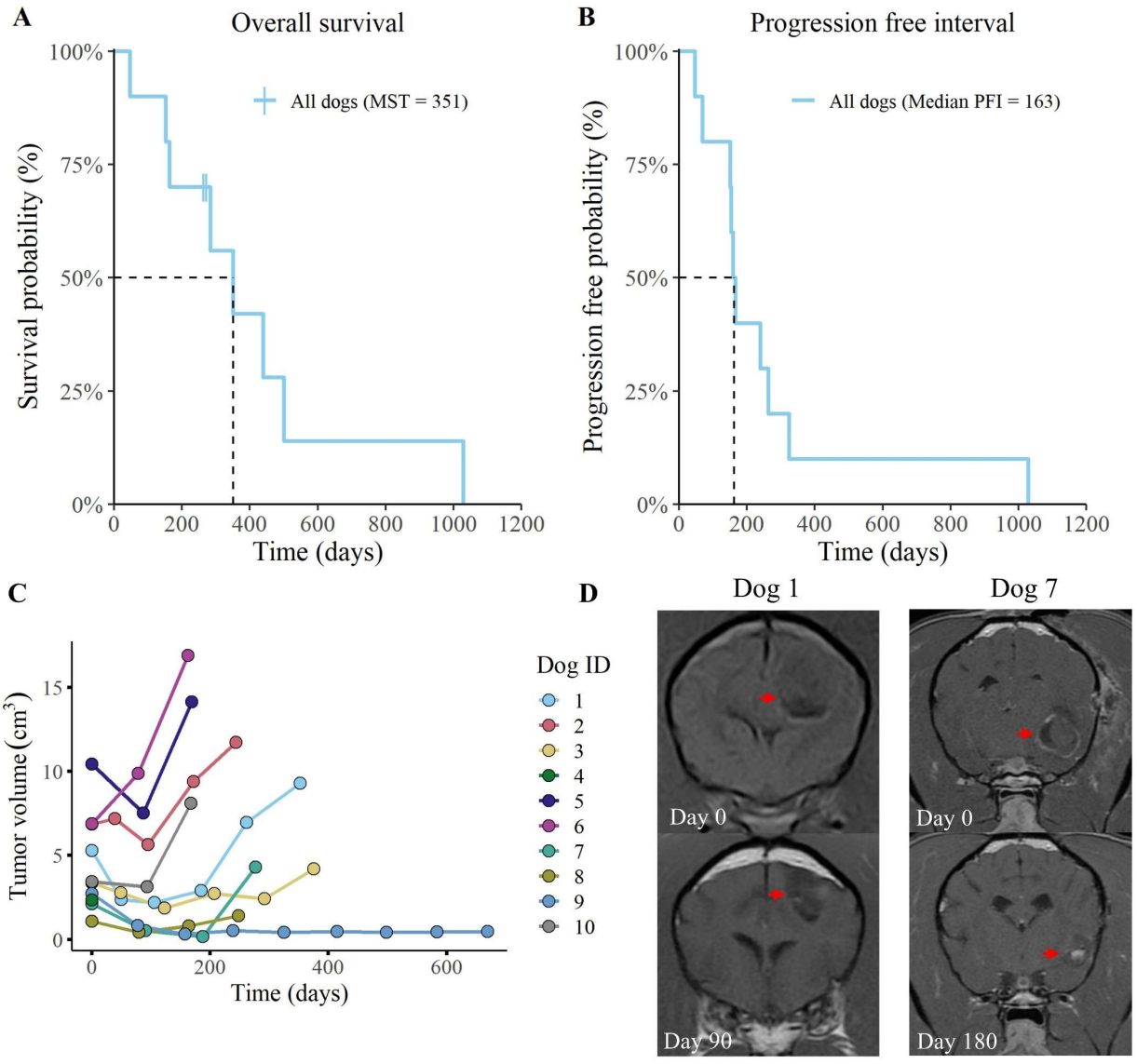


Figure 6.2. Overall survival and PFI curves for all enrolled dogs. Overall survival (A) and PFI (B) curves for all 10 dogs are depicted. Two dogs were censored from the overall survival data following deviation from the treatment protocol (time of censorship is noted with a vertical hash mark). Plot (C) represents tumor volume measurements, as determined by MRI. Tumor volume was evaluated every 3 months and volume was determined using transverse plane post-contrast T1-weighted images. Measurements were recorded until death or withdrawal from study. Images (D) are representative of changes in tumor volume for one dog with stable disease (dog 1) and one dog with partial tumor regression (dog 7).

Evaluation of CSC tumor vaccine immunogenicity

We next evaluated antibody titers following CSC vaccination, as a means of assessing overall immunotherapy responses. The primary evaluation of vaccine response was completed using flow cytometry-based quantitative endpoint titers to detect the binding of circulating immunoglobulins to J3T (a canine glioma cell line) spheroids (31). Evaluation of J3T spheroid endpoint titers revealed that 4 dogs were non-responders (undetectable anti-J3T antibodies) and that 6 dogs exhibited detectable increases in anti-J3T antibodies post-vaccination (Figure 6.3A). We also determined that J3T cells cultured under spheroid conditions were more immunogenic than the same cells cultured as plastic adherent cells, which indicates specificity to antigens only present on spheroid cells (Supplemental figure 6.2).

Pre- and post-immunization serum samples were also screened by western blotting to evaluate reactivity to proteins in the vaccine. The banding patterns from 5 humoral responders

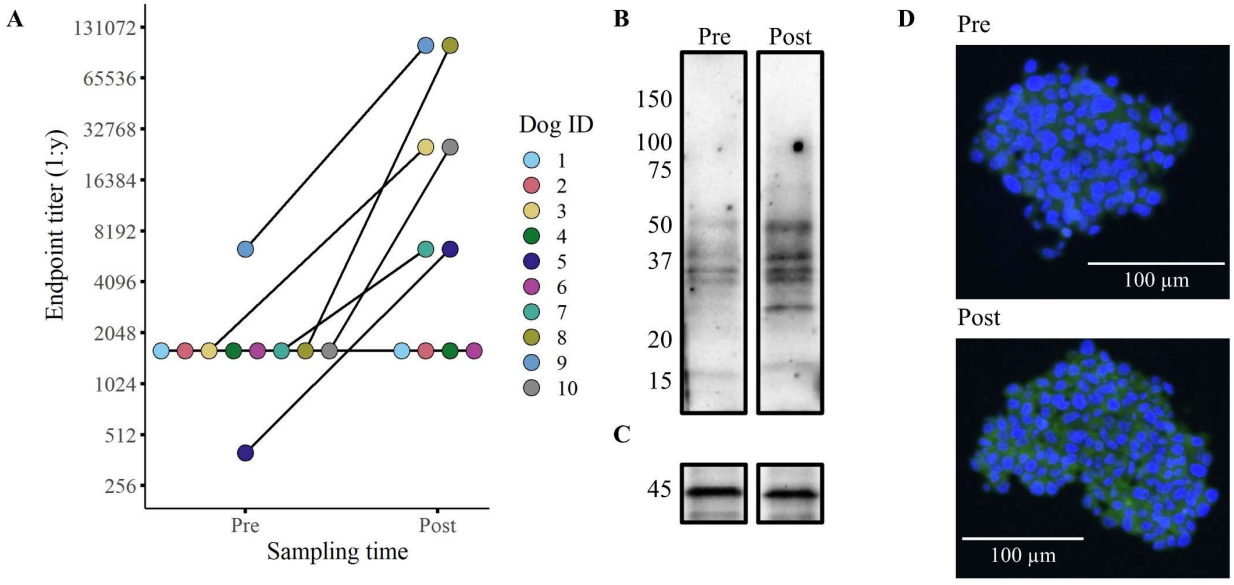


Figure 6.3. Six of ten dogs exhibited detectable anti-glioma humoral responses. Humoral responses to vaccination were quantified using flow cytometry-based endpoint titers with disaggregated J3T spheroids (A). Representative western blots (B/C) and confocal micrographs (D) depicting the serum antibody binding in a dog before (pre) immunization and 1-month post-vaccination (post). The western blots show binding to proteins found in the vaccine lysate (serum from dog 7) with the lower blot (C) depicting β-actin loading control. The confocal images tested binding of antibodies to J3T spheroids (D; serum from dog 3). All serum samples used in the assays were obtained after two vaccinations (week 4).

revealed consistent protein recognition patterns between dogs (Figure 6.3B/C, Supplemental figure 6.3). The most highly conserved protein bands appeared at molecular weights of 50 and 37 kDa. To further investigate antibody responses, we complete immunocytochemistry of J3T spheroids which revealed recognition of both intracellular and surface proteins (Figure 6.3D, Supplemental figure 6.4). These findings indicate that the tumor lysate vaccine induced broad recognition of multiple glioma antigens.

Finally, the duration of antibody responses was evaluated in a subset of 5 study dogs using serum samples obtained throughout the study. Notably, the antibody titers remained consistent with minimal fluctuation in antibody binding over the months evaluated (Supplemental figure 6.5). Together these data indicate that vaccination produced durable anti-glioma humoral immune responses, and that evaluation of antibody titers could be used to predict clinical responses to the combined immunotherapy protocol.

Correlation of immune response to immunotherapy with clinical outcomes

To determine if antibody titers correlated with tumor responses, we stratified overall survival and progression-free interval data by vaccine antibody status (i.e., vaccine responder vs non-responder) then completed univariate analysis. The analysis revealed that dogs with a detectable anti-glioma antibody response had enhanced overall survival times relative to non-responders ($p = 0.02$; responder MST = 500 days; non-responder MST = 218 days; Figure 6.4A). When progression-free interval was subjected to the same analysis, we found that the correlation with vaccine status did not reach a level of statistical significance, though numerically the median progression-free interval (mPFI) was longer in the responder cohort ($p = 0.15$; responder mPFI = 212 days; non-responder mPFI = 118 days; Figure 6.4B).

To further evaluate clinical correlates with humoral responses, we compared the maximal tumor response (% change in tumor volume) between dogs classified as vaccine responders and non-responders (Supplemental figure 6.6). This analysis failed to reach statistical significance (p

= 0.11) but provided further data to correlate clinical outcomes with humoral status. Lastly, we evaluated other factors that could have influenced study dog outcomes. This analysis was completed using multiple univariate analysis on key strata (breed, sex, tumor grade, and trial) and none of the tests revealed evidence of statistically significant correlations with clinical outcome measures (Supplemental table 6.2). Therefore, vaccine antibody response was the only factor identified in this study that correlated significantly with clinical outcomes. Overall, these findings suggest that vaccine response was the key variable associated with increased survival in dogs treated with this novel immunotherapy combination.

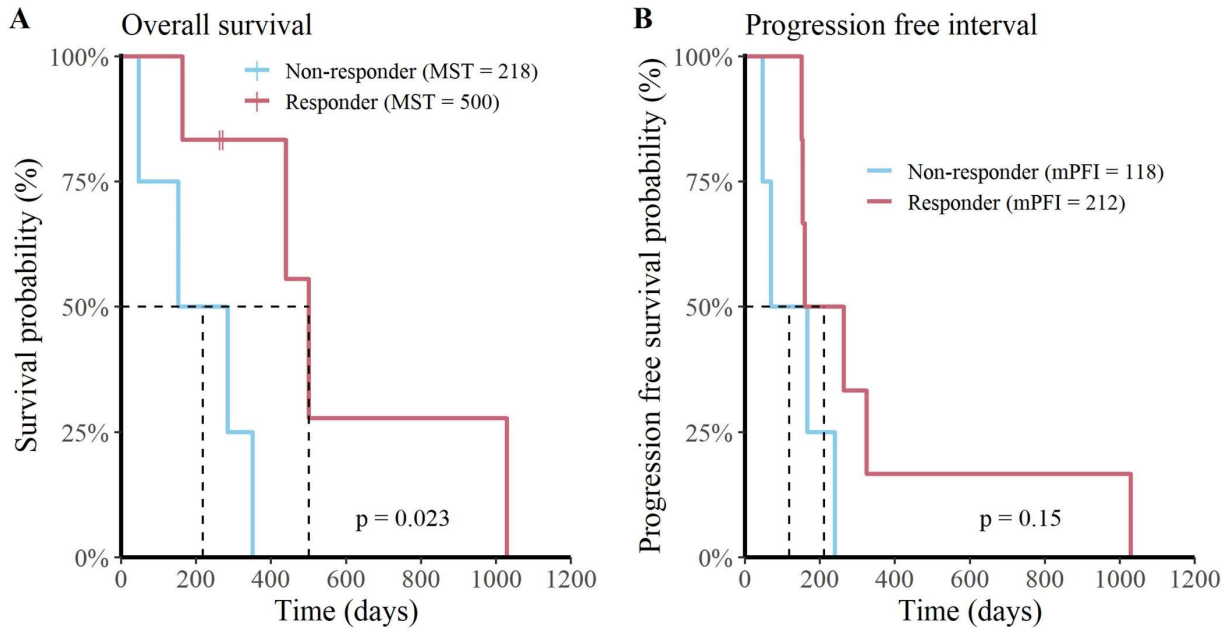


Figure 6.4. Dogs with detectable humoral responses have an increased overall survival rate. Depiction of (A) overall survival and (B) progression-free interval curves stratified for vaccine responders (n = 6; red) and non-responders (n = 4; blue) for all 10 dogs.

Assessment of tumor immune infiltrates in study dogs compared to untreated canine gliomas

Finally, we used post-mortem tumor biopsy tissues to compare the densities of immune infiltrates and the transcriptomes of trial dogs relative to an untreated cohort of dogs with gliomas. In total, we evaluated 4 dogs from the trial (dogs 1, 2, 3, and 6) and 14 dogs from an untreated

reference group. The reference group consisted of 10 oligodendrogliomas, 2 astrocytomas, and 2 undefined glioma tumor types. Seven of the tumors were classified as high-grade and 7 were classified as low-grade. The 4 study dogs used for analysis consisted of one dog who exhibited a vaccine response and had partial tumor regression while the other three dogs did not appear to respond to intervention. Therefore, the tissues collected from the treated dogs are likely not fully representative of the immunological changes observed in dogs who responded to intervention.

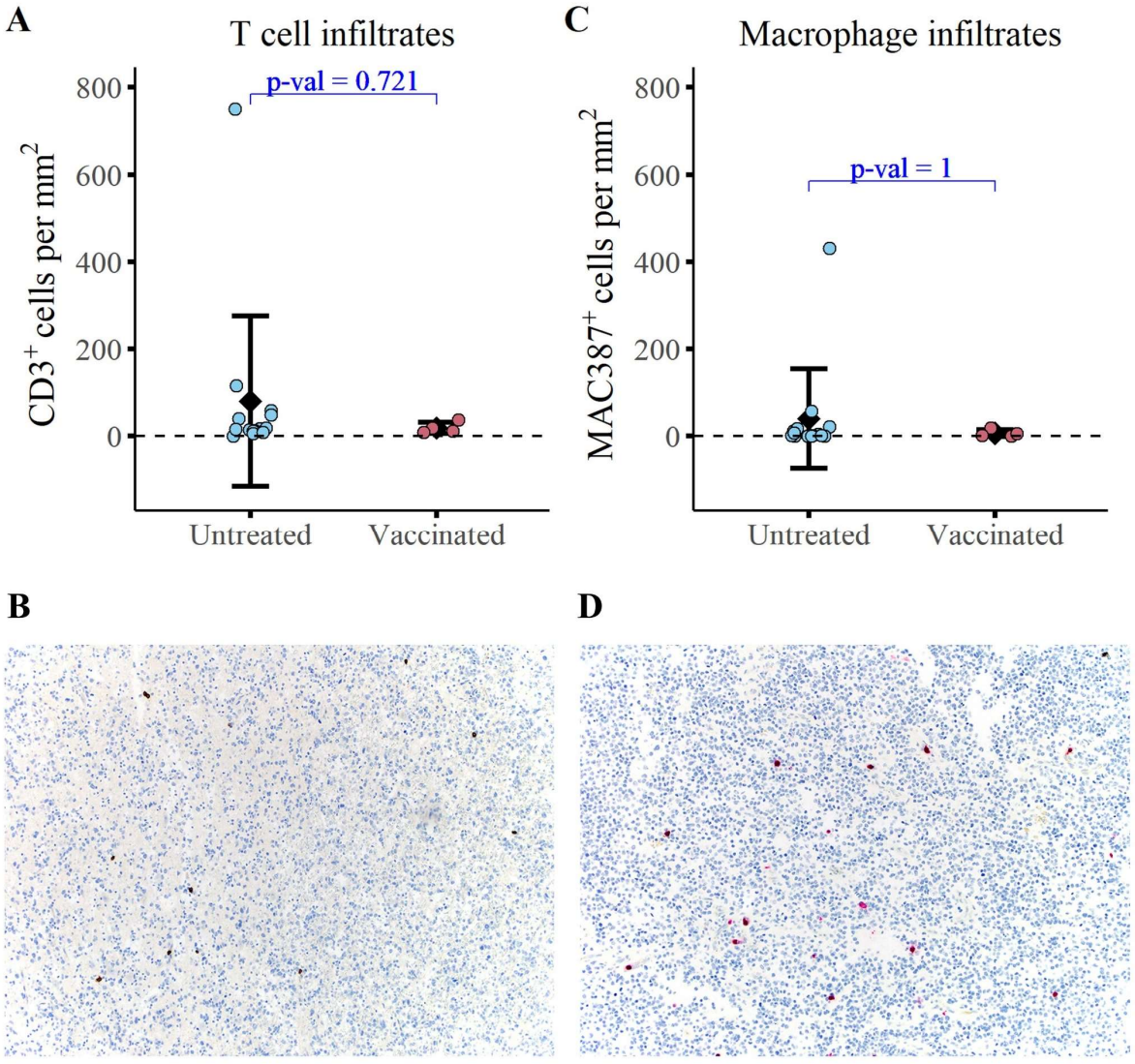


Figure 6.5. Immunized dogs exhibit no detectable change in tumor infiltrating immune cells relative to untreated controls at the time of necropsy. Quantification of (A) T cell infiltrates ($CD3^+$) and (C) macrophage infiltrates ($MAC387^+$) at the time of necropsy in 4 study dogs and 14 untreated dogs with gliomas. Representative images of CD3 immunolabeling (B) and MAC387 immunolabeling (D) are depicted. A Wilcoxon two sample t-test was conducted to evaluate statistical significance.

Immunohistochemical analysis revealed that the mean density of tumor-infiltrating $CD3^+$ T cells was 19.0±12.6 cells/mm^2 in study dogs compared to 79.9±195 cells/mm^2 in reference dogs (Figure 6.5A/B; ±SD). The average density of macrophage tumor infiltrates ($MAC387^+$) was 6.4±8.2 cells/mm^2 in study dogs compared to 39.9±114 cells/mm^2 in the reference dogs (Figure 6.5C/D; ±SD). Statistically, the mean tumor-infiltrating T cell and macrophage densities were not significantly different between study dogs and control dogs. Notably, the sparse densities of immune infiltrates are consistent with the densities typical of human glioma (32,33). Overall, these findings support the classification of canine gliomas as immunologically “cold” tumors.

We next completed immune transcriptome analysis to investigate whether there were changes in immune gene expression profiles between treated and untreated gliomas. These studies employed the canine immune-oncology (IO) NanoString panel and used DESeq2 to complete differential gene expression analysis (24). Analysis indicated there were no statistically significant differences in the transcriptome between treated and untreated tumors. To further analyze the dataset, we completed co-variate analysis to investigate how sex, tumor type, and tumor grade impacted gene expression. This analysis revealed statistically significant changes in gene expression between dogs with high-grade and low-grade tumors (28 genes upregulated, 3 genes downregulated; Supplemental figure 6.7). Of note, gene set enrichment analysis revealed that genes associated with PD-1 and CD28 signaling were increased in high-grade tumors relative to low-grade gliomas. The enrichment of these pathways provides evidence that there may be more exhausted T cells present in high-grade gliomas compared to low-grade tumors. To investigate further, we compared the abundances of T cell and macrophage tumor infiltrates and found no detectable differences between high- and low-grade gliomas (Supplemental figure 6.8). These findings indicate that the transcriptomic changes are likely due to functional differences in T cells rather than changes in T cell abundance. Overall, these data provide a better understanding of the immune landscape of canine gliomas and suggest that canine gliomas recapitulate human gliomas in terms of their overall immune responses.

Discussion

In the present study, we assessed a novel combination immunotherapy for generation of antitumor activity in dogs with spontaneously arising gliomas. Importantly, the canine glioma model is considered to be an excellent large animal model for human glioma and is valuable in the investigation of translational therapeutics (34,35). Key study findings include generation of partial tumor regression in 20% of dogs, along with durable tumor responses (median progression-free interval 163 days) in most animals. The median overall survival time (351 days) reported in this study is comparable to survival times previously reported in dogs treated with debulking surgery plus vaccination (MST = 212 days), or with debulking surgery plus chemotherapy (MST = 240 days) (18,36). In the present study, tumor responses correlated with vaccine antibody responses, as dogs that mounted detectable vaccine responses experienced significantly longer overall survival times than dogs with undetectable antibody responses (responder MST = 500 days; non-responder MST = 218 days). Taken together, these results indicate that the combination of TME disruption using losartan and propranolol with an allogeneic tumor vaccination, can elicit substantial antitumor activity in dogs with gliomas.

The prognosis for patients diagnosed with glioma is dismal and there is a need for effective therapeutic interventions in both human and canine patients. In dogs, definitive treatment for glioma consists of surgical debulking, radiation therapy, and/or chemotherapy (37,38). Dogs that receive some form of definitive treatment have been reported to have a median survival time of 84 days, while dogs that only receive palliative treatment have a 26-day MST (39). Recently, novel immunotherapy interventions have been explored in the canine glioma model. One such immunotherapy for canine glioma used a CD200 agonist, autologous vaccination, and conjunctive surgical debulking to achieve a median survival time of 12.7 months (approximately 387 days) (40). The CD200 trial only evaluated high-grade glioma cases, which limits direct comparison with the present trial results. Nonetheless, the median survival time observed in the present study (351

days) obtained with immunotherapy alone compares favorably with the results of combination immunotherapy plus surgery.

Limitations of the current study include a small sample size, a mix of low- and high-grade tumors, and a lack of a control arm consisting of dogs with untreated gliomas. However, prior literature indicates that the survival times for untreated dogs with glioma are short, thus all animals were treated in the present study (40,41). Furthermore, due to the limited availability of frozen blood leukocyte samples, our study did not evaluate T cell responses to vaccination. Nonetheless, prior studies have found that antitumor antibody responses following tumor vaccination typically correlate with T cell responses, and therefore antibody titers can be used as a proxy for overall T cell immune responses (42,43). Lastly, the immune density and transcriptome studies only included tumor samples from 4 study dogs, and only one was obtained from a vaccine responder, thereby skewing the analysis toward non-responders (n = 3) versus untreated dogs (n = 14).

Analysis of tumor-infiltrating immune cell densities and transcriptomes failed to detect any statistically significant differences between glioma samples from treated (n = 4) and naïve (n = 14) dogs. As noted above, the lack of vaccine responding animals is a major limitation to this analysis. However, the failure to observe significant T cell infiltrates in tumor samples from study dogs could also reflect the possibility that humoral immune responses may have been more important than cellular responses in this study. For example, in the 6 vaccine responding dogs tumor control may have been primarily mediated through antibody-dependent cellular cytotoxicity, which might not be reflected in T cell and macrophage infiltrates (44). Alternatively, it is possible that the 6 dogs who generated antibody responses were more immune competent than the non-responders and may have exhibited enhanced clinical outcomes regardless of intervention. Nonetheless, further investigation of the mechanism(s) of action associated with combined TME modification plus vaccination is warranted.

Further analysis of immune transcriptome responses in tumor tissues revealed important grade-related changes in gene expression. We identified upregulation of key immune regulatory

pathways in high-grade tumors compared to low-grade tumors, including lymphoid cell interactions, integrin interactions, and cell surface interactions at endothelial surface. Furthermore, we also found that the densities of tumor immune cell infiltrates were not significantly different between high- and low-grade tumors. Together, these findings suggest that tumor-infiltrating immune cells in high-grade tumors are likely more activated and potentially exhausted relative to low-grade tumors.

We have previously demonstrated the antitumor activity of losartan as a TME modifying drug in dogs, and other groups have noted the efficacy of propranolol as an MDSC modifying agent in rodent studies and as an anti-angiogenic agent in dogs (9,45,46). In addition, prior studies have found that losartan and propranolol can enhance antitumor immunity and contribute to increased survival times (47,48). Therefore, the current pilot study was not designed to investigate the antitumor activity of each component used in the combined protocol. Lastly, a valuable aspect of using repurposed drugs (i.e., losartan and propranolol) as cancer immunotherapy agents is that the drugs are FDA approved, have strong safety records, and can be readily obtained for use in clinical trials (49).

Our study employed a whole cell lysate vaccine derived from non-glioma canine cancer cell lines enriched for CSC antigens. We have previously reported the efficacy of similar allogeneic cell line lysate cancer vaccines in dogs with hemangiosarcoma (50). At the time of study initiation, we did not have access to canine glioma cell lines for inclusion in the vaccine. Since then, several canine glioma cell lines, including the J3T cell line used for immune assays in the current study, have been made available (31,51). Thus, the use of canine glioma cell lines in the preparation of future allogeneic vaccines would be predicted to generate greater antitumor activity.

In summary, a novel immunotherapy protocol consisting of daily treatment with losartan and propranolol combined with tumor vaccination was well-tolerated, induced partial tumor regression in 2 of 10 dogs, led to stable disease in 6 of 10 dogs, and resulted in a median overall

survival time of 351 days in dogs with glioma. The overall positive results of this pilot study, obtained in a spontaneous canine glioma model, indicate the potential for such an approach to be incorporated into the design of future glioma immunotherapy trials in human patients.

References

1. Lamborn, K. R., Yung, W. A., Chang, S. M., Wen, P. Y., Cloughesy, T. F., DeAngelis, L. M. *et al.* Progression-free survival: an important end point in evaluating therapy for recurrent high-grade gliomas. *Neuro. Oncol.* **10**, 162–170 (2008).

2. Herranz, C., Fernández, F., Martín-Ibáñez, R., Blasco, E., Crespo, E., De la Fuente, C. *et al.* Spontaneously arising canine glioma as a potential model for human glioma. *J. Comp. Pathol.* **154**, 169–179 (2016).

3. Hubbard, M. E., Arnold, S., Bin Zahid, A., McPheeters, M., Gerard O'Sullivan, M., Tabaran, A. F. *et al.* Naturally occurring canine glioma as a model for novel therapeutics. *Cancer Invest.* **36**, 415–423 (2018).

4. Packer, R. A., Rossmeisl, J. H., Kent, M. S., Griffin IV, J. F., Mazcko, C., & LeBlanc, A. K. Consensus recommendations on standardized magnetic resonance imaging protocols for multicenter canine brain tumor clinical trials. *Vet. Radiol. Ultrasound* **59**, 261–271 (2018).

5. Koehler, J. W., Miller, A. D., Miller, C. R., Porter, B., Aldape, K., Beck, J. *et al.* A revised diagnostic classification of canine glioma: towards validation of the canine glioma patient as a naturally occurring preclinical model for human glioma. *J. Neuropathol. Exp. Neurol.* **77**, 1039–1054 (2018).

6. Brahm, C. G., van Linde, M. E., Enting, R. H., Schuur, M., Otten, R. H., Heymans, M. W. *et al.* The current status of immune checkpoint inhibitors in neuro-oncology: a systematic review. *Cancers (Basel).* **12**, 586 (2020).

7. Tong, N., He, Z., Ma, Y., Wang, Z., Huang, Z., Cao, H. *et al.* Tumor Associated Macrophages, as the Dominant Immune Cells, Are an Indispensable Target for Immunologically Cold Tumor—Glioma Therapy? *Front. Cell Dev. Biol.* **9**, (2021).

8. Regan, D. P., Coy, J. W., Chahal, K. K., Chow, L., Kurihara, J. N., Guth, A. M. *et al.* The angiotensin receptor blocker losartan suppresses growth of pulmonary metastases via AT1R-independent inhibition of CCR2 signaling and monocyte recruitment. *J. Immunol.* **202**, 3087–3102 (2019).

9. Mohammadpour, H., MacDonald, C. R., Qiao, G., Chen, M., Dong, B., Hylander, B. L. *et al.* β2 adrenergic receptor–mediated signaling regulates the immunosuppressive potential of myeloid-derived suppressor cells. *J. Clin. Invest.* **129**, 5537–5552 (2019).

10. MacDonald, C., Ministero, S., Pandey, M., Robinson, D., Hong, E. F., Hylander, B. *et al.* Comparing thermal stress reduction strategies that influence MDSC accumulation in tumor bearing mice. *Cell. Immunol.* **361**, 104285 (2021).

11. Walle, U. K., Thibodeaux, H., Privitera, P. J. & Walle, T. Stereochemistry of tissue distribution of racemic propranolol in the dog. *Chirality* **1**, 192–196 (1989).

12. Li, Z., Bains, J. S. & Ferguson, A. V. Functional evidence that the angiotensin antagonist losartan crosses the blood-brain barrier in the rat. *Brain Res. Bull.* **30**, 33–39 (1993).

13. Zhai, Y., Li, G., Li, R., Chang, Y., Feng, Y., Wang, D. *et al.* Single-cell RNA-sequencing shift in the interaction pattern between glioma stem cells and immune cells during tumorigenesis. *Front. Immunol.* **11**, 2475 (2020).

14. Piper, K., DePledge, L., Karsy, M. & Cobbs, C. Glioma stem cells as immunotherapeutic targets: advancements and challenges. *Front. Oncol.* **11**, 92 (2021).

15. Ishiguro, T., Ohata, H., Sato, A., Yamawaki, K., Enomoto, T., & Okamoto, K. Tumor-derived spheroids: relevance to cancer stem cells and clinical applications. *Cancer Sci.* **108**, 283–289 (2017).

16. U'ren, L., Kedl, R. & Dow, S. Vaccination with liposome–DNA complexes elicits enhanced antitumor immunity. *Cancer Gene Ther.* **13**, 1033–1044 (2006).

17. Zaks, K., Jordan, M., Guth, A., Sellins, K., Kedl, R., Izzo, A. *et al.* Efficient immunization and cross-priming by vaccine adjuvants containing TLR3 or TLR9 agonists complexed to cationic liposomes. *J. Immunol.* **176**, 7335–7345 (2006).

18. Merickel, J. L., Pluhar, G. E., Rendahl, A. & O'Sullivan, M. G. Prognostic histopathologic features of canine glial tumors. Vet. Pathol. **58**, 945–951 (2021).

19. Guth, A. M., Deogracias, M. & Dow, S. W. Comparison of cancer stem cell antigen expression by tumor cell lines and by tumor biopsies from dogs with melanoma and osteosarcoma. *Vet. Immunol. Immunopathol.* **161**, 132–140 (2014).

20. Boss, M. K., Watts, R., Harrison, L. G., Hopkins, S., Chow, L., Trageser, E. *et al.* Immunologic Effects of Stereotactic Body Radiotherapy in Dogs with Spontaneous Tumors and the Impact of Intratumoral OX40/TLR Agonist Immunotherapy. *Int. J. Mol. Sci.* **23**, 826 (2022).

21. Okada, H., Weller, M., Huang, R., Finocchiaro, G., Gilbert, M. R., Wick, W. *et al.* Immunotherapy response assessment in neuro-oncology: a report of the RANO working group. *Lancet Oncol.* **16**, e534–e542 (2015).

22. Ellingson, B. M., Wen, P. Y. & Cloughesy, T. F. Modified criteria for radiographic response assessment in glioblastoma clinical trials. *Neurotherapeutics* **14**, 307–320 (2017).

23. Gomes, I. C., Acquarone, M., de Moraes Maciel, R., Erlich, R. B. & Rehen, S. K. Analysis of pluripotent stem cells by using cryosections of embryoid bodies. *J. Vis. Exp. JoVE* (2010).

24. Lenz, J. A., Assenmacher, C. A., Costa, V., Louka, K., Rau, S., Keuler, N. S. *et al.* Increased tumor-infiltrating lymphocyte density is associated with favorable outcomes in a comparative study of canine histiocytic sarcoma. *Cancer Immunol. Immunother.* **71**, 807–818 (2022).

25. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 1–21 (2014).

26. R Core Team. R: A Language and Environment for Statistical Computing. (2020).

27. Wickham, H. Elegant graphics for data analysis. *Media* **35**, 10–1007 (2009).

28. Kassambara, A., Kosinski, M., Biecek, P. & Fabian, S. Package 'survminer'. *Draw. Surviv. Curves using 'ggplot2'(R Packag. version 0.4.9)* (2017).

29. Miller, A. D., Miller, C. R. & Rossmeisl, J. H. Canine primary intracranial cancer: a clinicopathologic and comparative review of glioma, meningioma, and choroid plexus tumors. *Front. Oncol.* **9**, 1151 (2019).

30. Dobson, J. M. Breed-predispositions to cancer in pedigree dogs. *Int. Sch. Res. Not.* **2013**, (2013).

31. Rainov, N. G., Koch, S., Sena-Esteves, M. & Berens, M. E. Characterization of a canine glioma cell line as related to established experimental brain tumor models. *J. Neuropathol. Exp. Neurol.* **59**, 607–613 (2000).

32. Krane, G. A., O'Dea, C. A., Malarkey, D. E., Miller, A. D., Miller, C. R., Tokarz, D. A. *et al.* Immunohistochemical evaluation of immune cell infiltration in canine gliomas. *Vet. Pathol.* **58**, 952–963 (2021).

33. Robinson, M. H., Vasquez, J., Kaushal, A., MacDonald, T. J., Vega, J. E. V., Schniederjan, M., & Dhodapkar, K. Subtype and grade-dependent spatial heterogeneity of T-cell infiltration in pediatric glioma. *J. Immunother. cancer* **8**, (2020).

34. Dickinson, P. J., LeCouteur, R. A., Higgins, R. J., Bringas, J. R., Larson, R. F., Yamashita, Y. *et al.* Canine spontaneous glioma: a translational model system for convection-enhanced delivery. *Neuro. Oncol.* **12**, 928–940 (2010).

35. Hicks, J., Platt, S., Kent, M. & Haley, A. Canine brain tumours: a model for the human disease? *Vet. Comp. Oncol.* **15**, 252–272 (2017).

36. Hidalgo Crespo, E., Farré Mariné, A., Pumarola i Battle, M., Borrego Massó, J. F. & Luján Feliu-Pascual, A. Survival Time after Surgical Debulking and Temozolomide Adjuvant Chemotherapy in Canine Intracranial Gliomas. Vet. Sci. **9**, 427 (2022).

37. Dickinson, P. J. Advances in diagnostic and treatment modalities for intracranial tumors. *J. Vet. Intern. Med.* **28**, 1165–1185 (2014).

38. Debreuque, M., De Fornel, P., David, I., Delisle, F., Ducerveau, M. N., Devauchelle, P. & Thibaud, J. L. Definitive-intent uniform megavoltage fractioned radiotherapy protocol for presumed canine intracranial gliomas: retrospective analysis of survival and prognostic factors in 38 cases (2013–2019). *BMC Vet. Res.* **16**, 1–14 (2020).

39. José-López, R., Gutierrez-Quintana, R., de la Fuente, C., Manzanilla, E. G., Suñol, A., Pi Castro, D. *et al.* Clinical features, diagnosis, and survival analysis of dogs with glioma. *J. Vet. Intern. Med.* (2021).

40. Olin, M. R., Ampudia-Mesias, E., Pennell, C. A., Sarver, A., Chen, C. C., Moertel, C. L. et al. Treatment combining CD200 immune checkpoint inhibitor and tumor-lysate vaccination after surgery for pet dogs with high-grade glioma. Cancers (Basel). 11, 137 (2019).

41. Rossmeisl, J. H., Jones, J. C., Zimmerman, K. L. & Robertson, J. L. Survival time following hospital discharge in dogs with palliatively treated primary brain tumors. *J. Am. Vet. Med. Assoc.* **242**, 193–198 (2013).

42. Pluhar, G. E., Grogan, P. T., Seiler, C., Goulart, M., SantaCruz, K. S., Carlson, C. *et al.* Antitumor immune response correlates with neurological symptoms in a dog with spontaneous astrocytoma treated by gene and vaccine therapy. *Vaccine* **28**, 3371–3378 (2010).

43. Jäger, E., Nagata, Y., Gnjatic, S., Wada, H., Stockert, E., Karbach, J. *et al.* Monitoring CD8 T cell responses to NY-ESO-1: correlation of humoral and cellular immune responses. *Proc. Natl. Acad. Sci.* **97**, 4760–4765 (2000).

44. Andersen, B. M., Pluhar, G. E., Seiler, C. E., Goulart, M. R., SantaCruz, K. S., Schutten, M. M. *et al.* Vaccination for invasive canine meningioma induces in situ production of antibodies capable of antibody-dependent cell-mediated cytotoxicity. *Cancer Res.* **73**, 2987–2997 (2013).

45. Regan, D. P., Chow, L., Das, S., Haines, L., Palmer, E., Kurihara, J. N. *et al.* Losartan blocks osteosarcoma-elicited monocyte recruitment, and combined with the kinase inhibitor toceranib, exerts significant clinical benefit in canine metastatic osteosarcoma. *Clin. Cancer Res.* (2021).

46. Amaya, C. N., Perkins, M., Belmont, A., Herrera, C., Nasrazadani, A., Vargas, A. *et al.* Non-selective beta blockers inhibit angiosarcoma cell viability and increase progression free-and overall-survival in patients diagnosed with metastatic angiosarcoma. *Oncoscience* **5**, 109 (2018).

47. Wrobel, L. J., Bod, L., Lengagne, R., Kato, M., Prévost-Blondel, A., & Le Gal, F. A. Propranolol induces a favourable shift of antitumor immunity in a murine spontaneous model of melanoma. *Oncotarget* **7**, 77825 (2016).

48. O'Rawe, M., Wickremesekera, A. C., Pandey, R., Young, D., Sim, D., FitzJohn, T. et al. "Treatment of glioblastoma with re-purposed renin-angiotensin system modulators: Results of a phase I clinical trial." Journal of Clinical Neuroscience. **95**, 48-54 (2022).

49. Murphy, J. E., Wo, J. Y., Ryan, D. P., Clark, J. W., Jiang, W., Yeap, B. Y. *et al.* Total neoadjuvant therapy with FOLFIRINOX in combination with losartan followed by chemoradiotherapy for locally advanced pancreatic cancer: a phase 2 clinical trial. *JAMA Oncol.* **5**, 1020–1027 (2019).

50. U'Ren, L. W., Biller, B. J., Elmslie, R. E., Thamm, D. H. & Dow, S. W. Evaluation of a novel tumor vaccine in dogs with hemangiosarcoma. *J. Vet. Intern. Med.* **21**, 113–120 (2007).

51. Empl, M. T., Macke, S., Winterhalter, P., Puff, C., Lapp, S., Stoica, G. *et al.* The growth of the canine glioblastoma cell line D-GBM and the canine histiocytic sarcoma cell line DH82 is inhibited by the resveratrol oligomers hopeaphenol and r2-viniferin. *Vet. Comp. Oncol.* **12**, 149–159 (2014).

CONCLUSIONS

The dog represents a valuable translational animal model with the potential to inform therapeutic strategies in human medicine. Despite the translational value of the dog, reagent limitations have hampered the use of canine disease models. Therefore, one of the overarching aims of this dissertation was to establish reference databases for canine cell types. We used single-cell RNA (scRNA) sequencing to characterize circulating immune cells in healthy and osteosarcoma tumor burdened dogs as well as describe the heterogeneity of OS tumors. From there we used the references to investigate how the tumor microenvironment (TME) modulates immune cell transcriptomes. The second aim of this body of work was to use canine cancer models to investigate the utility of myeloid targeted immunotherapeutics. As such we used scRNA sequencing to investigate the systemic and intratumoral impacts of losartan, toceranib, and ladarixin combination therapy in the setting of canine OS. Furthermore, we applied a combination vaccination, losartan, and propranolol therapeutic in canine glioma to investigate the clinical impact and immunological correlates with intervention. Overall, the work presented here contributes much needed insight into the intricacies of canine cell types and demonstrates the utility of employing canine cancer patients to investigate translational therapeutics.

The contributions in chapter 2 represents a substantial advancement in canine immunology, as it is the first comprehensive characterization of all circulating leukocytes. In publication of the work, we took extra measures to ensure the utility of the dataset extends beyond the write-up. In addition to the bare minimum of depositing raw data into the NCBI GEO database, we also uploaded annotated processed data files, provided detailed documentation of potential uses/applications, deposited the processed data in an online cell browser for easy data exploration (https://canine-leukocyte-atlas.cells.ucsc.edu), and provided all code used to complete the analysis. As the single cell ‘omics field continues to develop it is essential that

veterinary medical researchers capitalize on the data available to them. To ensure we utilize the data to its fullest potential it is imperative that we practice transparent reporting, use consistent cell type annotations, and work toward establishing consensus cell type gene signatures. This will be particularly relevant for the collection of data on canine leukocytes, as blood is one of the most accessible clinical samples that can be collected to understand a patient's immunological healthy. Although scRNA sequencing will likely never have utility as a diagnostic, it does have the potential to inform the development of clinically accessible assays. To this end, we proposed a modified myeloid-derived suppressor cell flow cytometry-based assay that built off previous literature and was refined using information obtained from the scRNA dataset.

Therapeutic strategies for osteosarcoma have been stagnant for decades and there is a need to use novel approaches to investigate the disease. The rarity of the human OS is part of the reason for slow clinical advancement, so the increased incidence of OS in dogs makes canine OS a useful model for understanding OS biology and identifying potential therapeutic targets. In chapter 3 we sought to address this need by using scRNA sequencing to characterize the heterogeneity in treatment-naïve tumors. Through detailed analysis of the dataset, we were able to dissect the heterogenous tumor and immune cell populations and draw comparisons to human OS. Through direct comparisons between human and canine OS scRNA dataset we observed substantial cell type homologies which provides further evidence that canine OS has the potential to faithfully recapitulate the human disease.

The utility of chapter 4 was to demonstrate one way to maximize data extraction from scRNA datasets. As such, we completed a meta-analysis of the references generated in chapters 2 and 3 to investigate mechanisms of immune modulation mediated by the OS tumor microenvironment. Analysis of the integrated dataset confirmed changes reported in human tumors are also present in canine tumors. Key tumor induced changes included T cell activation, overexpression of immune suppressive markers in myeloid cells, and an increased relative proportion of exhausted T cells. Overall, the analysis presented in chapter 4 highlights the

mechanisms in which tumors modulate immune cell function and opens up the possibility to directly compare with human datasets, something we wish to do in the future.

In chapter 5 we harnessed the power of a canine clinical trial and scRNA sequencing to investigate the mechanisms underlying a novel myeloid target immunotherapy. We evaluated systemic effects, using circulating leukocytes, and local tumor effects, using cells isolated from OS tumor biopsies, to describe the effect of treatment with losartan, toceranib, and ladarixin. We found treatment to deplete myeloid cells, but also unexpected deplete dendritic cells and T cells. The broad depletion of immune cells resulted in the relative increase in fibroblasts and tumors, which upon closer investigation displayed evidence of marked transcriptional changes. We observed malignant osteoblasts and fibroblasts to have increased enrichment of hypoxic and epithelial to mesenchymal transition gene signatures. Currently, the chapter acts as a description of what is observed from a molecular standpoint, but an ongoing canine clinical trial will shed light on the clinical impact of treatment and enable deeper analysis of the dataset.

Finally, in chapter 6 we evaluated the immunological correlates of canine glioma patients treated with vaccination and a myeloid targeted drug regimen. Although we focused on a different tumor type than in the previous chapters, this clinical trial provides evidence of the clinical impact of myeloid targeted therapeutics, with losartan being a common agent in chapters 5 and 6. We provided evidence of clinical benefit following intervention and were able to identify a subset of immunological responders that exhibited an enhanced overall survival rate. Overall, myeloid targeted therapeutic approaches warrant further investigation as treatment strategy for glioma and osteosarcoma.

In summary, the body of work presented in this dissertation represents a considerable advancement in veterinary medical research with the potential to enhance the utility of the dogs as translation immune-oncology model. The references generated can be used to inform development of targeted approaches and used as tool when employing the dog in translational research, while the clinical trials provide insight into the utility of myeloid targeted therapeutics.