Impact of metastatic prostate cancer on human bone marrow

Ninib Baryawno [1,2,3,4,*,§], Youmna Kfoury [1,2,3,*], Nicolas Severe [1,2,3,*], Shenglin Mei [5,*], Taghreed Hirz [1,2,3,*], Karin Gustafsson [1,2,3,*], Thomas Brouse [1], Elizabeth Scadden [1], Bryan Choi [6], Nikolas Barkas [5], John H. Shin [6,†], Philip J. Saylor [7,†], David T. Scadden [1,2,3,†], David B. Sykes [1,2,3,†], Peter V. Kharchenko [2,5,†, §], as part of the Boston Bone Metastasis Consortium

[1] Center for Regenerative Medicine, Massachusetts General Hospital, Boston, MA, USA
[2] Harvard Stem Cell Institute, Cambridge, MA, USA
[3] Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA, USA
[4] Childhood Cancer Research unit, Dep. Of Children's and Women's Health, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden
[5] Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA
[6] Department of Neurosurgery, Harvard Medical School, Boston, MA, USA
[7] Massachusetts General Hospital Cancer Center, Harvard Medical School, Boston, USA

[*,†] Contributed equally

[§] Correspondence should be addressed to Ninib Baryawno (n.baryawno@ki.se) and Peter V. Kharchenko (peter_kharchenko@hms.harvard.edu)

Running title

Impact of prostate metastasis on human bone marrow

Abstract

Cancer-related mortality due to solid tumor malignancies is overwhelmingly due to the development and progression of metastases. In advanced prostate cancer, metastases most often involve the bone[1] and generally represent incurable disease. It remains unclear what aspects of the bone marrow microenvironment make it hospitable to metastatic dissemination. Similarly, the impact of the metastatic tumor on hematopoiesis and the marrow immune response is poorly understood. Here we take advantage of rare spinal cord decompression surgeries to profile marrow and metastatic tumors from men with advanced prostate cancer at single-cell resolution. We contrast the cellular composition and transcriptional states in matched samples of tumor and liquid bone marrow collected at adjacent vertebral body levels, as well as bone marrow of orthopedic patients without malignancy. Metastatic prostate cancer was associated with hematopoietic suppression and multifaceted immune distortion. There was exhaustion of specific T-cell subsets, appearance of inflammatory lymphocytes and macrophages, and alteration of cytokine profiles. The chemokine CCL20 was notably overexpressed by myeloid cells, as was its cognate CCR6 receptor on T-cells. This dual overexpression was associated with repressed immune responses. We used a syngeneic mouse model of bone-metastatic prostate cancer to explore this observation, and demonstrated that disruption of the CCL20-CCR6 axis resulted in significant prolongation of survival. Overall, comparative high-resolution analysis of bone marrow reveals distinct alterations associated with prostate cancer bone metastases that may be amenable to therapeutic targeting with the goal of altering cancer progression.

Manuscript

Prostate cancer frequently metastasizes to bone where, in the castration-resistant setting, 70 to 90% of patients have radiographically-detectable skeletal involvement[1-5]. To date, bone metastases represent a generally incurable form of prostate cancer and contribute significantly to disease-specific morbidity[6]. In particular, metastatic involvement of the spinal column can result in pathologic fracture and spinal cord compression leading to significant neurologic and functional disability. Surgical decompression/stabilization and targeted radiation are used selectively. Approved bone-targeted systemic therapies offer some benefit. Zoledronic acid and denosumab delay skeletal morbidity but do not improve overall survival[7]. The radiopharmaceutical radium-223 offers a modest improvement in overall survival[8]. The immune checkpoint therapies targeting CTLA4 and PD-1/PD-L1 have been unsuccessful in clinical trials in unselected prostate cancer populations[9] despite the high abundance of T lymphocytes in the tissue. In particular, emerging evidence indicates that patients with bone metastases are at a disadvantage when it comes to response to immune checkpoint therapies[10]. Altogether, this indicates that there is a clear need for new and more effective approaches for our patients with bone metastatic prostate cancer based on a better understanding of the microenvironment that is permissive for its growth and is a critical player in response to therapy.

Availability of fresh clinical prostate cancer bone metastatic samples is essential to the study of relationships between cancer cells and the marrow microenvironment. Bone sampling by needle biopsy can be technically challenging and provides scant quantities of tissue. We therefore developed an inter-disciplinary workflow to obtain tissue from emergent clinically-indicated surgeries for those rare cases of spinal cord compression due to epidural extension of tumor or pathologic fracture of a vertebral body. The collection of tissue samples was integrated into the standard workflow of surgery.

We collected matched sets of tissue fractions from each patient: solid metastatic tissue (*Tumor* fraction), liquid bone marrow at the vertebral level of spinal cord compression (*Involved*), as well as liquid bone marrow from a different vertebral body distant from the tumor site but within the surgical field (*Distal*) (Fig. 1a, Supplementary Table 1). This allowed for a comparison within the same individual, controlling for inter-individual variation. Correlative clinical, pathologic and radiographic information was prospectively recorded. Bone marrow samples from patients undergoing hip replacement surgery (*Benign*) served as a non-malignant comparator group. The transcriptional composition of all samples was assessed using single-cell RNA-seq (scRNA-seq), yielding on average of ~2,000 cells per sample (Supplementary Table 2).

We observed notable inter-individual variation of expression patterns (Extended Data Fig. 1a) and used a recently-developed approach[11] to integrate the full collection of samples (Fig. 1b,c). Joint analysis revealed a rich repertoire of immune cells and ongoing hematopoiesis with HSC/progenitor populations giving rise to B cell, monocyte, and erythroid lineages (Extended Data Fig. 1b,c, Supplementary Table 3). The T cell population was disconnected, which is expected as T cell maturation occurs outside of the bone marrow with the thymus. Tumor cells were detected in some of the *Distal* fractions, reflecting diffuse marrow involvement in certain patients (Extended Data Fig.

1d,e). The non-immune subpopulations, mostly captured within the *Tumor* fraction, included endothelial cells, pericytes, osteoblasts, osteoclasts as well as the tumor cells themselves (Extended Data Fig. 1f). Granulocytes, whose abundance was quantified using flow cytometric analysis (Extended Data Fig. 1g), were absent in the single-cell data, reflecting a known cell lineage limitation of the single-cell protocol[12].

The presence of metastases significantly altered the immune cell composition of the bone marrow (Extended Data Fig. 1f). Relative to *Benign* controls, the three cancer fractions showed a striking depletion of B cells and B cell progenitors (Fig. 1d,f, Extended Data Fig. 1h). The *Tumor* fraction had an increase in the proportion of macrophages compared to the *Involved/Distal* fractions (Fig. 1e,f). As expected, the solid *Tumor* fraction also showed more abundant endothelial, pericyte, osteoblast, and tumor cell populations (Extended Data Fig. 1f).

Complementary to the shifts in the frequency of cell types, scRNA-seq data also captured the transcriptional state permitting an examination of expression differences (see Methods). With the exception of the *Benign* controls, the similarity of cell states reflected patient-specific signatures (Extended Data Fig. 1j). The extent of inter-patient differences was smallest for the *Benign* controls, and incrementally increased towards the site of metastasis (*Distal* to *Involved* to *Tumor* fraction) (Fig. 1g). This significant increase in the inter-patient variability in cancer, and the *Tumor* fraction in particular, demonstrates the divergent impact of metastasis on the bone marrow across individuals and makes it more challenging to identify recurrent trends. In this regard, comparisons of different bone marrow fractions from the same individual increase statistical power. For example, to determine which cell types were most affected in terms of their transcriptional state by the presence of metastasis, we compared *Tumor* fraction with other fractions from the same individual (Extended Data Fig. 1j). While macrophage and HSC/progenitor populations showed the greatest median difference in expression states, a broad range of other cell types also showed changes of comparable magnitude, including T cells and myeloid cells illustrating the broad impact of metastasis across different cellular compartments.

The myeloid cells have been implicated in supporting tumor progression in certain cancers[13,14]. Focused examination of the myeloid compartment revealed substantial population shifts between *Benign* and malignant fractions (Fig. 2a,b, Extended Data Fig. 2a,b). While *Benign* controls primarily contained resting monocytes, the monocytes of both *Involved* and *Distal* fractions expressed genes indicative of an activated and proliferative state in cancer patients (Mono-2, myeloid cell activation, GO:0002275 $P<10^{-16}$). Striking changes were observed in the *Tumor* compartment, driven by the appearance of prominent populations of Tumor Inflammatory Monocytes (TIMs), and Tumor Associated Macrophages (TAMs) (Fig. 2b,c, Extended Data Fig. 2b,c, Supplementary Table 4), a finding that was validated by flow cytometric analysis of independent samples (Fig. 2d, Extended Data Fig. 2d and extended Fig. 1g). The TAM population, had an expression pattern characteristic of M2 macrophages[15] (Fig. 2e,f) which have been shown to suppress anti-tumor immune responses

across a broad range of tumors[16]. Accordingly, TAM cells express anti-inflammatory cytokines (TGFB1, IL-10, Interleukin 1 receptor antagonist IL1rn)[17] as well as factors implicated in cancer growth and invasiveness (CCL2, CCL3, CCL4, MMPs)[18,19]. Interestingly, TGFB1 has been recently shown to be a molecular mediator for T helper cell polarization toward Th17 in response to immune checkpoint therapy in the bone metastatic microenvironment and blocking it in combination with the immune checkpoint inhibitors improved survival in a mouse model of prostate cancer [20]. Further, a recent study by an independent group suggest that exposure to the anti-androgen drug enzaluatamide is associated with an increased tumor-intrinsic activation of TGFβ signaling and EMT signature in metastatic castration resistant prostate cancer[21], suggesting that resistance to anti-androgen treatment may be induced by TGFB1. TIM cells had a pro-inflammatory monocyte signature with high expression of pro-inflammatory cytokines such as IL1-B and TNF and scoring high for the M2 pro-tumoral signature (Fig. 2e,f) - a pattern that has been previously observed in other cancers[15,22,23].

The TIM and TAM populations that we observed in primary samples closely paralleled *ex vivo* effects on peripheral blood monocytes of tumor-secreted factors[24]. The cells expressed epidermal growth factor-like ligands (epiregulin in TIMs, HB-EGF in TAMs) and a common repertoire of interleukin-6-like JAK/STAT3 pathway activators (IL-6, Oncostatin-M). Both signaling modalities support tumor growth directly as well as indirectly by suppressing the activation of CD8+ T cells, expanding suppressive Treg cells, promoting M2 macrophage polarization and differentiating myeloid cells into myeloid derived suppressor cells (MDSCs)[25]. TIMs and TAMs also expressed inhibitors of the pro-inflammatory nuclear factor-*k*B pathway (NFKBIA, NFKBIZ, TNFAIP3)[26] (Fig. 2f). The presence of these TIM and TAM expression signatures are associated with significantly worse prognosis for patients across a broad range of cancer types (Extended Data Fig. 2e).

Interestingly, the distribution of TIMs and TAMs among our patients mirrored two distinct clinical entities. Patients who developed cord compression in the setting of cancer progression and a rising PSA despite treatment had tumors that were enriched with TAMs. In comparison, patients with pathologic fracture of the vertebral body and cord compression despite stable disease with low PSA and no tumor cells detected at the time of surgery and had samples enriched with TIMs (Supplementary Table 1, Extended Data Fig. 2f).

Detailed analysis of the lymphocyte compartment revealed the expected subsets, $CD4^+$ T helper and T regulatory (Treg) cells, $CD8^+$ cytotoxic T lymphocytes (CTLs), as well as NK and NKT cells (Fig. 3a, Extended Data Fig. 3a, Supplementary Table 3, 4). The smallest population of $CD8^+$ T cells expressed high levels of *CCR7*, *LEF1* and *SELL*, indicating that these are antigen-inexperienced, naïve cells[27-29] (Extended Data Fig. 3a). The CTLs could be further subdivided into two subtypes: CTL-1 expressing effector T cell genes such as *KLRG1*, *GZMK* and other cytotoxicity mediators,

whereas CTL-2 are characterized by an effector/memory-like transcriptional profile including *IL7R* and *KLRB1* (Fig. 3b)[30,31]. We also identified CCR7 expressing naïve CD4+ cells whereas the mature T helper cell compartment appeared to be a mix of T_H1 and T_H17 cells expressing *CXCR3* and *CCR6* respectively[31].

The relative abundance of the CTL, Treg and T_H populations followed a common pattern, gradually increasing from the *Distal* fraction towards the site of metastasis (*Tumor* fraction) (Fig. 3c,d, Extended Data Fig. 3b). The abundance of these cells was also high in the *Benign* controls, likely reflecting the known inflammation in the osteoarthritis patients from whom the samples were collected. An additional set of bone marrow samples from *healthy* individuals[32] revealed significantly lower abundance of these populations, offset by higher proportion of $CD4^+$ naïve cells (Fig. 3d). The functional state of these T cell populations was also impacted, in particular CTL-2 in the *Tumor* fraction exhibited reduced cytotoxicity expression signature in combination with a pronounced T cell exhaustion signature, characteristic of dysfunctional CTL commonly observed in tumors[33,34] (Fig. 3e, Extended Data Fig. 3c-e). The NK cell population showed a complementary pattern, with high abundance in the *Involved* and *Distal* samples, but not in *Benign* and *Tumor* fractions (Fig. 3c, d) - an observation that was further validated by flow cytometric analysis (Fig. 3f). This suggests that although NK cells can be recruited to the general area of the bone metastasis, they may fail to infiltrate the tumor.

The observed deficiencies of cytotoxicity in the proximity of the tumor may arise through repressive actions by other immune populations[35]. For instance, we find that the Treg cells, which typically act to suppress immune response[36], showed increased activity signatures at the site of the metastasis (Extended Fig. 3f-h). The complex patterns of immune signaling are also likely to involve the myeloid compartment. Indeed, considering variation of TAM and TIM abundance among patients (Extended Data Fig. 2f) we find that increased proportion of TAMs at the site of metastasis is correlated with CTL-2 exhaustion (Fig. 4a). Similarly, increased abundance of TIMs coincides with a reduction in the CTL-2 cytotoxicity signature (Fig. 4a). Analysis of expressed receptor and ligand combinations suggests an extensive number of potential channels through which TIM and TAM populations could potentially affect T lymphocytes (Extended Data Fig. 4a).

Of the cytokines that were differentially expressed within the TIM and TAM cells of the *Tumor* fraction, CCL20 was one of the most highly expressed (Fig. 2f, 4b,c, Extended Data Fig. 4b), and its cognate receptor CCR6 was predominantly expressed on Treg and $T_H1/17$ cells (Fig. 4c, Extended Data Fig. 4b,c). Systematic expression analysis of ligand-receptor pairs confirmed the potential interaction of CCL20-CCR6 axis in TIM/TAM and Treg, $T_H1/17$ (Extended Data Fig. 4a, Supplementary Table 5). Studies of other cancers have demonstrated that CCL20 signaling can promote tumor growth, invasiveness and chemoresistance[37-39] by recruitment of Tregs and/or T_H17 cells[39,40]. To confirm CCL20 protein expression, ELISA was performed on bone marrow plasma from the *Involved* and *Distal* fractions, as well as *Benign* samples. The plasma concentration of CCL20

was detected mostly in the samples of *Involved* fraction (Supplementary Table 6), whereas the protein levels in the other fractions were below the detection level of the ELISA. CCL20 and CCR6 have been showed to be expressed on primary prostate cancer tissue and in stromal fibroblasts within the tumor microenvironment[39,41]. However, the direct role of the CCL20-CCR6 axis in prostate bone metastases is unclear.

To investigate the impact of CCL20-CCR6 axis in prostate bone metastasis, we developed a syngeneic model of prostate bone metastasis, based on the RM1-BM cell line derived from C57BL/6 RM1 prostate cancer cells[42,43]. These cells were sequentially injected into C57BL/6J wild-type mice to generate a highly penetrant bone tropic prostate cancer cell line, RM1-BoM3, that induced osteolytic and osteoblastic lesions as shown by micro-computed tomography (microCT) and radiography (Extended Data Fig. 4d-g).

Using this model, we found that the absence of the CCR6 receptor improved survival of bone metastases bearing mice when compared to controls (Fig. 4d,e). Single-cell analysis on myeloid and T cells from the bone metastasis isolated from wild-type and CCR6KO mice showed average reduction in Active Tregs and Exhausted CTLs with a concomitant increase in Naive CTLs. These results suggest that blockade of the CCR6-CCL20 axis counteracts immunosuppression in the tumor microenvironment, though variability between animals was noticeable (Fig. 4f, Extended Data Fig. 4h-l). Together, these data reveal a mechanism by which myeloid cells in the bone metastasis microenvironment can shape the activity of T cell populations through CCR6-coupled signaling.

In addition to suggesting potential therapeutic avenues, our results demonstrate a recurrent pattern of widespread disruption of the human bone marrow by prostate cancer metastasis. These data provide a foundational resource for exploring other potential signaling axes and cellular relationships that inhibit immune response and remodel the bone marrow resulting in a tumor-permissive environment.

Acknowledgements

We are particularly indebted to our patients and their clinical care teams. We acknowledge funding from NIH CA193481 and DK103074 (to DTS and PVK), National Cancer Institute CA 163191 (To DTS), HL131768 (to PVK), Dana-Farber / Harvard Cancer Center Nodal Award (CCSG grant P30CA006516). We gratefully acknowledge support from Bill & Cheryl Swanson and from Gunther & Maggie Buerman. N.B. was funded by the Swedish Research Council. N.S was a recipient of the AACR-Millennium Fellowship in Prostate Cancer Research. Y.K was supported by a grant from the STARR cancer consortium. The Benign bone marrow material was provided by Mark A Randolph from the plastic surgery research lab at Massachusetts General Hospital. Olga Kharchenko designed the medical illustration in Figure 1. Patient samples were sorted at the HSCI/CRM flow cytometry core facility at MGH. We thank Forsyth Multiplex Core for performing the multiplex assay. We thank the Center for Skeletal Research at Massachusetts General Hospital, supported by a grant from NIAMS (P30 AR075042) at NIH.

Author contributions

N.B., Y.K., N.S., D.B.S. and P.J.S. conceived the study. P.J.S. coordinated the multi-disciplinary teams and the IRB-approved protocol. N.B., N.S., Y.K., P.S., J.H.S., D.B.S. D.T.S. and P.V.K. directed the study. N.B., Y.K., N.S. designed experiments. Sample collection methodology and surgeries were performed by J.H.S. Human malignant samples were collected, isolated and libraries prepared by N.B. Y.K. collected human benign samples. FACS was performed by Y.K. Flow cytometry was performed by Y.K., T.H. and K.G. T.H. and K.G. collected human samples for flow cytometry analysis. ELISA measurements were performed by T.H. N.S. established the bone metastatic mouse model. N.S. and T.H. performed in vivo experiments. T.H. and K.G. performed FACS and library preparations on samples form mouse experiments. S.M. and P.V.K. performed the computational analysis. N.B., Y.K., N.S., S.M., T.H., K.G., J.H.S., P.J.S., D.T.S., D.B.S. and P.V.K. interpreted the data and wrote the manuscript. All authors read, edited and approved the manuscript.

Competing interests

P.V.K. serves on the Scientific Advisory Board to Celsius Therapeutics Inc. D.T.S is a director and shareholder for Agios Therapeutics and Editas Medicines; a founder, director, shareholder and scientific advisory board member for Magenta Therapeutics and LifeVault Bio, a shareholder and founder of Fate Therapeutics, and a director, founder, shareholder for Clear Creek Bio and Red Oak Medicines. D.B.S. is a founder and shareholder for Clear Creek Bio.

References

1 Halabi, S. *et al.* Meta-Analysis Evaluating the Impact of Site of Metastasis on Overall Survival in Men With Castration-Resistant Prostate Cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **34**, 1652-1659, doi:10.1200/JCO.2015.65.7270 (2016).

2 de Bono, J. S. *et al.* Abiraterone and increased survival in metastatic prostate cancer. *The New England journal of medicine* **364**, 1995-2005, doi:10.1056/NEJMoa1014618 (2011).

3 de Bono, J. S. *et al.* Prednisone plus cabazitaxel or mitoxantrone for metastatic castration-resistant prostate cancer progressing after docetaxel treatment: a randomised open-label trial. *Lancet* **376**, 1147-1154, doi:10.1016/S0140-6736(10)61389-X (2010).

4 Petrylak, D. P. *et al.* Docetaxel and estramustine compared with mitoxantrone and prednisone for advanced refractory prostate cancer. *The New England journal of medicine* **351**, 1513-1520, doi:10.1056/NEJMoa041318 (2004).

5 Tannock, I. F. *et al.* Docetaxel plus prednisone or mitoxantrone plus prednisone for advanced prostate cancer. *The New England journal of medicine* **351**, 1502-1512, doi:10.1056/NEJMoa040720 (2004).

6 Xiang, L. & Gilkes, D. M. The Contribution of the Immune System in Bone Metastasis Pathogenesis. *International journal of molecular sciences* **20**, doi:10.3390/ijms20040999 (2019).

7 Saad, F. *et al.* Long-term efficacy of zoledronic acid for the prevention of skeletal complications in patients with metastatic hormone-refractory prostate cancer. *Journal of the National Cancer Institute* **96**, 879-882, doi:10.1093/jnci/djh141 (2004).

8 Parker, C. *et al.* Alpha emitter radium-223 and survival in metastatic prostate cancer. *The New England journal of medicine* **369**, 213-223, doi:10.1056/NEJMoa1213755 (2013).

9 Laccetti, A. L. & Subudhi, S. K. Immunotherapy for metastatic prostate cancer: immuno-cold or the tip of the iceberg? *Current opinion in urology* **27**, 566-571, doi:10.1097/MOU.0000000000000433 (2017).

10 Beer, T. M. *et al.* Randomized, Double-Blind, Phase III Trial of Ipilimumab Versus Placebo in Asymptomatic or Minimally Symptomatic Patients With Metastatic Chemotherapy-Naive Castration-Resistant Prostate Cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **35**, 40-47, doi:10.1200/JCO.2016.69.1584 (2017).

11 Barkas, N. *et al.* Joint analysis of heterogeneous single-cell RNA-seq dataset collections. *Nature methods* **16**, 695-698, doi:10.1038/s41592-019-0466-z (2019).

12 Chen, J. *et al.* PBMC fixation and processing for Chromium single-cell RNA sequencing. *Journal of translational medicine* **16**, 198, doi:10.1186/s12967-018-1578-4 (2018).

13 Coffelt, S. B., Wellenstein, M. D. & de Visser, K. E. Neutrophils in cancer: neutral no more. *Nature reviews. Cancer* **16**, 431-446, doi:10.1038/nrc.2016.52 (2016).

14 Lewis, C. E., Harney, A. S. & Pollard, J. W. The Multifaceted Role of Perivascular Macrophages in Tumors. *Cancer cell* **30**, 365, doi:10.1016/j.ccell.2016.07.009 (2016).

15 Azizi, E. *et al.* Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. *Cell* **174**, 1293-1308 e1236, doi:10.1016/j.cell.2018.05.060 (2018).

16 DeNardo, D. G. & Ruffell, B. Macrophages as regulators of tumour immunity and immunotherapy. *Nature reviews. Immunology* **19**, 369-382, doi:10.1038/s41577-019-0127-6 (2019).

17 Sanjabi, S., Zenewicz, L. A., Kamanaka, M. & Flavell, R. A. Anti-inflammatory and pro-inflammatory roles of TGF-beta, IL-10, and IL-22 in immunity and autoimmunity. *Current opinion in pharmacology* **9**, 447-453, doi:10.1016/j.coph.2009.04.008 (2009).

18 Merchant, N. *et al.* Matrix metalloproteinases: their functional role in lung cancer. *Carcinogenesis* **38**, 766-780, doi:10.1093/carcin/bgx063 (2017).

19 Pellikainen, J. M. *et al.* Expression of matrix metalloproteinase (MMP)-2 and MMP-9 in breast cancer with a special reference to activator protein-2, HER2, and prognosis. *Clinical cancer research : an official journal of the American Association for Cancer Research* **10**, 7621-7628, doi:10.1158/1078-0432.CCR-04-1061 (2004).

20 Jiao, S. *et al.* Differences in Tumor Microenvironment Dictate T Helper Lineage Polarization and Response to Immune Checkpoint Therapy. *Cell* **179**, 1177-1190 e1113, doi:10.1016/j.cell.2019.10.029 (2019).

21 He, M. X., Cuoco M. S., Crowdis, J., Bosma-Moody, A., Zhang, Z., Bi, K., Kanodia, A., Su, M. J., Rodman, C., DelloStritto, L., Shah, P., Burke, K. P., Izar, B., Bakouny, Z., Tewari, A. K., Liu, D., Camp, S. Y., Vokes, N. I., Park, J., Vigneau, S., Fong, L., Rozenblatt-Rosen, O., Regev, A., Rotem, A., Taplin, M. E., Van Allen, E. M. Transcriptional mediators of treatment resistance in lethal prostate cancer (manuscript submitted).
22 Chevrier, S. *et al.* An Immune Atlas of Clear Cell Renal Cell Carcinoma. *Cell* **169**, 736-749 e718, doi:10.1016/j.cell.2017.04.016 (2017).
23 Muller, S. *et al.* Single-cell profiling of human gliomas reveals macrophage ontogeny as a basis for regional differences in macrophage activation in the tumor microenvironment. *Genome biology* **18**, 234, doi:10.1186/s13059-017-1362-4 (2017).
24 Vlaicu, P. *et al.* Monocytes/macrophages support mammary tumor invasivity by co-secreting lineage-specific EGFR ligands and a STAT3 activator. *BMC cancer* **13**, 197, doi:10.1186/1471-2407-13-197 (2013).
25 Huynh, J., Etemadi, N., Hollande, F., Ernst, M. & Buchert, M. The JAK/STAT3 axis: A comprehensive drug target for solid malignancies. *Seminars in cancer biology* **45**, 13-22, doi:10.1016/j.semcancer.2017.06.001 (2017).
26 Lawrence, T. The nuclear factor NF-kappaB pathway in inflammation. *Cold Spring Harbor perspectives in biology* **1**, a001651, doi:10.1101/cshperspect.a001651 (2009).
27 Campbell, J. J. *et al.* 6-C-kine (SLC), a lymphocyte adhesion-triggering chemokine expressed by high endothelium, is an agonist for the MIP-3beta receptor CCR7. *The Journal of cell biology* **141**, 1053-1059, doi:10.1083/jcb.141.4.1053 (1998).
28 Picker, L. J. *et al.* Control of lymphocyte recirculation in man. I. Differential regulation of the peripheral lymph node homing receptor L-selectin on T cells during the virgin to memory cell transition. *Journal of immunology* **150**, 1105-1121 (1993).
29 Willinger, T. *et al.* Human naive CD8 T cells down-regulate expression of the WNT pathway transcription factors lymphoid enhancer binding factor 1 and transcription factor 7 (T cell factor-1) following antigen encounter in vitro and in vivo. *Journal of immunology* **176**, 1439-1446, doi:10.4049/jimmunol.176.3.1439 (2006).
30 Fergusson, J. R. *et al.* CD161(int)CD8+ T cells: a novel population of highly functional, memory CD8+ T cells enriched within the gut. *Mucosal immunology* **9**, 401-413, doi:10.1038/mi.2015.69 (2016).
31 Zhang, L. *et al.* Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. *Nature* **564**, 268-272, doi:10.1038/s41586-018-0694-x (2018).
32 Oetjen, K. A. *et al.* Human bone marrow assessment by single-cell RNA sequencing, mass cytometry, and flow cytometry. *JCI insight* **3**, doi:10.1172/jci.insight.124928 (2018).
33 Lee, P. P. *et al.* Characterization of circulating T cells specific for tumor-associated antigens in melanoma patients. *Nature medicine* **5**, 677-685, doi:10.1038/9525 (1999).
34 Li, H. *et al.* Dysfunctional CD8 T Cells Form a Proliferative, Dynamically Regulated Compartment within Human Melanoma. *Cell* **176**, 775-789 e718, doi:10.1016/j.cell.2018.11.043 (2019).
35 Munn, D. H. & Bronte, V. Immune suppressive mechanisms in the tumor microenvironment. *Current opinion in immunology* **39**, 1-6, doi:10.1016/j.coi.2015.10.009 (2016).
36 Woo, E. Y. *et al.* Cutting edge: Regulatory T cells from lung cancer patients directly inhibit autologous T cell proliferation. *Journal of immunology* **168**, 4272-4276, doi:10.4049/jimmunol.168.9.4272 (2002).
37 Lee, S. K. *et al.* Human antigen R-regulated CCL20 contributes to osteolytic breast cancer bone metastasis. *Scientific reports* **7**, 9610, doi:10.1038/s41598-017-09040-4 (2017).
38 Lu, E., Su, J., Zhou, Y., Zhang, C. & Wang, Y. CCL20/CCR6 promotes cell proliferation and metastasis in laryngeal cancer by activating p38 pathway. *Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie* **85**, 486-492, doi:10.1016/j.biopha.2016.11.055 (2017).
39 Walch-Ruckheim, B. *et al.* Stromal Fibroblasts Induce CCL20 through IL6/C/EBPbeta to Support the Recruitment of Th17 Cells during Cervical Cancer Progression. *Cancer research* **75**, 5248-5259, doi:10.1158/0008-5472.CAN-15-0732 (2015).
40 Wang, D. *et al.* Colorectal cancer cell-derived CCL20 recruits regulatory T cells to promote chemoresistance via FOXO1/CEBPB/NF-kappaB signaling. *Journal for immunotherapy of cancer* **7**, 215, doi:10.1186/s40425-019-0701-2 (2019).

41 Beider, K. *et al.* Interaction between CXCR4 and CCL20 pathways regulates tumor growth. *PloS one* **4**, e5125, doi:10.1371/journal.pone.0005125 (2009).
42 Power, C. A. *et al.* A novel model of bone-metastatic prostate cancer in immunocompetent mice. *The Prostate* **69**, 1613-1623, doi:10.1002/pros.21010 (2009).
43 Thompson, T. C., Southgate, J., Kitchener, G. & Land, H. Multistage carcinogenesis induced by ras and myc oncogenes in a reconstituted organ. *Cell* **56**, 917-930, doi:10.1016/0092-8674(89)90625-9 (1989).
44 Nunamaker, E. A., Artwohl, J. E., Anderson, R. J. & Fortman, J. D. Endpoint refinement for total body irradiation of C57BL/6 mice. *Comparative medicine* **63**, 22-28 (2013).
45 Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G. & Hacohen, N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* **160**, 48-61, doi:10.1016/j.cell.2014.12.033 (2015).
46 Liu, H., Shi, B., Huang, C. C., Eksarko, P. & Pope, R. M. Transcriptional diversity during monocyte to macrophage differentiation. *Immunology letters* **117**, 70-80, doi:10.1016/j.imlet.2007.12.012 (2008).
47 Martinez, F. O., Gordon, S., Locati, M. & Mantovani, A. Transcriptional profiling of the human monocyte-to-macrophage differentiation and polarization: new molecules and patterns of gene expression. *Journal of immunology* **177**, 7303-7311, doi:10.4049/jimmunol.177.10.7303 (2006).
48 Vento-Tormo, R. *et al.* Single-cell reconstruction of the early maternal-fetal interface in humans. *Nature* **563**, 347-353, doi:10.1038/s41586-018-0698-6 (2018).
49 Cancer Genome Atlas Research, N. *et al.* The Cancer Genome Atlas Pan-Cancer analysis project. *Nature genetics* **45**, 1113-1120, doi:10.1038/ng.2764 (2013).

Figure Legends

Figure 1. Transcriptional impact of prostate cancer metastasis on the bone marrow

a. Schematic illustration of study design, including bone marrow sampling strategy (left), scRNA-seq (middle), and mouse model validation (right).

b-c. The analysis integrates scRNA-seq from three fractions (*Tumor, Involved, Distal*) of 9 metastatic prostate cancer patients (rows), as well as 7 *Benign* bone marrow controls from hip replacement surgery patients. The datasets were integrated using *Conos* [11] to establish joint embedding and annotation (c). Projections of individual samples are shown in b.

d. The difference in frequency of major subpopulations between *Benign* controls and cancer patients (all three fractions) is visualized on the joint embedding. The *Benign* controls show notably higher proportion of B cell and non-cytotoxic T cell populations.

e. The compositional differences between *Tumor* fraction and *Involved/Distal* fractions is visualized on the joint embedding. The *Tumor* fraction shows higher frequency of macrophage and specific T cell subpopulations.

f. Compositional shifts for the most affected cell types are shown as boxplots comparing proportion of major cell populations between *Benign* controls and different cancer fractions (see Extended Data Fig. 1f for all populations).

g. Average cell-type-specific expression distances between different patients is shown for *Benign* controls and the three cancer fractions. The analysis shows that controlling for the compositional differences, the expression state of different cell types in *Benign* controls is significantly more similar than in the cancer fractions. The expression differences between individuals increase incrementally towards the site of metastasis, with the highest expression variation observed in the *Tumor* fraction.

Figure 2. Myeloid populations shift towards inflammatory monocytes and repressive macrophages in the Tumor fraction

a. Detailed annotation of the myeloid subpopulations is shown on a myeloid-specific joint embedding (left), together with select markers (right).

b. Changes in the composition of the myeloid compartment between sample fractions is visualized as cell density on the joint embedding. The composition of the *Distal* and *Involved* fractions is dominated by an activated Mono-2 population, while *Tumor* fraction shows appearance of Tumor Inflammatory Monocyte (TIM) and Tumor Associated Macrophages (TAM) populations.

c. Average expression of macrophage and monocyte signature genes (see Methods, GSE5099) in different myeloid subpopulations.

d. Boxplot showing the percent of monocytes/macrophages within CD11b+ cells (Supplementary Table 7) by flow cytometry in *Benign*, *Involved* and *Tumor* fractions from three patients. Statistical analysis was performed using Student's t-test (*$p<0.05$, ****$p<0.0001$).

e. Average expression of the M2 macrophage signature gene panel (left), and monocyte inflammatory gene panel (right) across different monocyte populations.

f. Heatmap shows average expression of select genes from different categories (rows) across different cell populations (top color bar, colors matching panel a) for each patient (columns).

Figure 3. Tumor fraction shows increased abundance of exhausted cytotoxic populations and helper T cells

a. Detailed annotation of the T cell subpopulations across *Benign* controls and cancer fractions, shown on a T-cell specific joint embedding.

b. Expression of key genes from different categories (rows) in CTL-1 and CTL-2 populations is shown across different sample types.

c. Shifts in the T cell populations across Benign and cancer fractions is visualized as cell density on the joint embedding.

d. Box plots of cell frequency changes of different subpopulations between sample fractions (including additional *Healthy* controls, grey) show increased frequency of CTL and Treg populations in the *Tumor* faction, decreased $CD4^+$ naive. NK and NK-T peaks in the *Involved* and *Distal* fractions. Unlike *Healthy* samples, the *Benign* controls, collected from hip replacement surgeries, also show high fraction of CTL cell populations, significantly higher than bone marrow of healthy individuals.

e. Cytotoxicity and exhaustion scores for CTL-1 in *Benign* and *Tumor* fraction. The scores are defined as normalized average expression of the signature genes (Supplementary Table 4). The size of the points are indicative of the proportion of the CTL-2 cells relative to all T cells in the sample.

f. Flow cytometry validation of the NK cell abundance (Supplementary Table 7) in *Benign*, *Involved*, and *Tumor* fractions from three patients. Statistical analysis was performed using a Student's t-test (*$p<0.05$).

Figure 4. Crosstalk between myeloid and T cell compartments benefits tumor growth

a. The abundance of TIM and TAM populations (x axis) is associated with the increased exhaustion and decreased cytotoxicity of the CTL-2 populations (y axis).

b. Expression of CCL20 across different monocyte subpopulations.

c. CCR6 and CCL20 expression levels in T cell and myeloid cells subpopulations, respectively.

d. Representative images of bone metastases localization as shown by bioluminescence optical imaging in the mandible or in long bones of syngeneic wild-type and CCR6KO mice one week post injection of RM1-BoM3 cells.

e. CCR6 knockout results in a statistically significant survival benefit in syngeneic mouse model of bone marrow metastasis (n=9 wild-type and n=13 CCR6KO). Two independent experiments.

f. CD3+ T cells were isolated from wild-type and CCR6KO mice and analyzed by single-cell RNA-seq (n=2 mice per group). Changes in the composition of the T cell compartment between wild-type and CCR6-KO is visualized as cell density on the joint embedding.

Extended data Figure legends

Extended Data Figure 1

a. Inter-individual variation between samples is illustrated on a tSNE embedding (left) of all *Involved* fraction samples, processed without any additional corrections. The color shows sample identity of each cell. The cells separate by sample. *Conos* graphs (right) allow to align corresponding cell populations for downstream comparisons.

b. Expression of genes specific to major cell populations are shown as a heatmap.

c. Expression of marker genes for the different cell types visualized on joint graph embedding.

d. Heterogeneity of tumor cells from different samples (color) visualized on a tSNE embedding. While some samples show notable intra-tumoral heterogeneity, no common axes of intra-tumoral variation could be established between individuals

e. Number of tumor cells detected in *Distal*, *Involved* and *Tumor* fractions is shown for different patients.

f. Boxplots comparing relative abundance of different cell populations between *Benign* controls and the three cancer fractions.

g. Box plots showing flow cytometry analysis of granulocytes (Supplementary Table 7) proportion in *Benign*, *Involved* and *Tumor* fractions (left) from three patients. Statistical analysis was performed using a Student's t-test (*$p<0.05$; **** $p < 0.0001$). Gating strategy for flow validation of myeloid populations. Labels above the flow plots refer to the parent population and the percentages are of parent gate (right).

h. Box plots showing flow cytometry analysis of B cells (Supplementary Table 7) proportion in *Benign*, *Involved* and *Tumor* fractions from three patients. Statistical analysis was performed using a Student's t-test (*$p<0.05$). Gating strategy for flow validation of lymphoid cell types. Labels above the flow plots refer to the parent population in the percentages are of the parent gate (right).

i. A t-SNE embedding of different samples, based on their overall expression distance. The similarity measure measures the magnitude of expression change for each subpopulation, using size-weighted average to combine them into an overall expression distance that controls for compositional differences. Colors demarcate patients (all Benign controls are shown in one color), symbols indicate different sample fractions.

j. Boxplots showing average expression distance for each cell subpopulation between *Tumor* and *Involved*/*Distal* fractions of the same patient.

Extended Data Figure 2

a. Heatmap of the highly expressed genes in different myeloid populations.

b. Comparison of relative abundance of different myeloid cell clusters in *Benign*, *Distal*, *Involved* and *Tumor* fractions.

c. Average expression of monocytes and macrophage signature genes selected based on in vitro differentiation of monocytes to macrophages (see Methods).

d. Box plot showing flow cytometry analysis of Myeloid dendritic cells (Supplementary Table 7) in *Benign*, *Involved* and *Tumor* fractions from three patients. Statistical analysis was performed using a Student's t-test (*$p<0.05$).

e. Kaplan–Meier survival curves for TCGA cancer types: Esophageal Carcinoma (ESCA), cervical squamous carcinoma (CESC), kidney renal clear cell carcinoma (KIRC), lung squamous carcinoma (LUSC), liver hepatocellular carcinoma (LIHC). Patients were stratified into two groups based on the average expression (binary: top 25% versus bottom 25%) of TIM and TAM expression signatures as annotated by key marker genes in Supplementary Table 4.

f. Relative abundance of TIM and TAM cells in different tumor samples, measured as a fraction of myeloid population in each sample.

Extended Data Figure 3

a. Expression of marker genes for different T cell subpopulations, shown on a joint T cell embedding.

b. Comparison of proportions of CD8+ naïve, NKT and Treg in in *Benign*, *Distal*, *Involved* and *Tumor* fractions.

c. Similar to Figure 3e of the main manuscript, the relationship between exhaustion and cytotoxicity is shown for the CTL-1 population.

d. Cytotoxicity of CD8+ T cells. Boxplots illustrate significant decrease in the cytotoxicity of the CTL-2 subpopulation in the *Tumor* fraction.

e. Exhaustion of CD8+ cytotoxic T cells. Boxplots show composite exhaustion scores of CTLs in different fractions, highlighting significant increase of exhaustion signature in the *Tumor* fraction. See Supplementary Table 4 for the genes defining exhaustion signature.

f. Boxplots illustrate significant increase of Treg activity of Treg subpopulation in the *Tumor* fraction. Treg activity was measured as an average expression of genes shown in Supplementary Table 4.

g. Focused re-analysis of Treg compartment reveals activated and resting Treg subpopulations. Expression of key marker genes a visualized on the same tSNE embedding.

h. Volcano plot showing differentially expressed genes between active Treg cells in *Benign* and *Tumor* fractions. The plot shows negative $\log_{10}$ transformed Q values (y axis), as a function of the average $\log_2$ fold changes in expression (x axis). The genes exhibiting significantly higher expression in *Tumor* faction are shown on the top (red), while those significantly higher in *Benign* fraction are shown on the bottom (blue).

Extended Data Figure 4

a. Ligand-receptor interactions between myeloid and T cell subpopulations. Each cell indicates potential channels of communication from ligand (row) to receptor (column). The right panel lists candidate ligand-receptor pairs between TIM/TAM and T cell subpopulations. (see Methods).

b. Expression of CCL20 and CCR6 are illustrated on the joint embedding (left), with grey color indicating zero expression, and red color indicating high expression.

c. CCR6 expression levels in Treg and T_H cells.

d. Illustration of a third-generation syngeneic bone tropic prostate bone metastasis cell line expressing luciferase and Tdtomato allowing to monitor experimental bone metastases in live syngeneic mice by bioluminescence imaging.

e. Micro-computed tomography of long bones from wild-type mice without bone metastases and wild-type mice injected with RM1-BoM3 cells 3 weeks post intracardiac injection (3 representative mice per group).

f. Quantification of the osteolytic lesions induced by the RM1-BoM3 cells in femur of mice (n=4 to 5 mice per group). Statistical analysis was performed using an unpaired t-test (*$p<0.05$).

g. Radiography imaging of femurs from wild-type mice without bone metastases and wild-type mice with RM1-BoM3 cells 3 weeks post intracardiac injection representing osteolytic lesions (solid arrows) and osteoblastic lesions (dotted arrow).

h. Selected marker gene expression in T cells of bone metastatic model.

i. Myeloid cells were isolated from wild-type and CCR6KO mice and analyzed by single-cell RNA-seq (n=2 mice per group). Changes in the composition of the myeloid cell compartment between wild-type and CCR6-KO is visualized as cell density on the joint embedding.

j. Selected marker gene expression in myeloid cells of bone metastatic model.

k. Bar plot showing fraction of all myeloid cells. The error bar gives the 95% confidence interval.

l. Bar plot showing fraction of all T cells. The error bar gives the 95% confidence interval of the proportion estimate for each animal.

Methods

Patient material

All human-subjects tissue collection was carried out with institutional review board (IRB) approval (Dana Farber/Harvard Cancer Center protocol 13-416 and Partners protocol 2017P000635/PHS).

Surgical approach and collection of tumor and bone marrow specimens

Tumor specimens

In each case, the patient was brought to the operating room for clinically indicated decompression and stabilization in the setting of spinal cord compression related to metastatic prostate cancer. Each patient consented to use of their tissue for research purposes. Bone marrow (*e.g. Involved*) and tumor samples were taken with the patient in the prone position under general anesthesia as the spine was approached posteriorly. After the levels of the spine were exposed and identified, a Jamshidi needle was malleted directly into the desired vertebral body and a syringe connected to the needle was used to aspirate bone marrow immediately upon cannulation. Cannulation of the vertebral body is standard prior to placing stabilizing instrumentation into the bone. By placing the Jamshidi needle directly into the bone, this ensures that the aspirate from deep within the vertebral body is not diluted by surrounding blood in the surgical field or surgical irrigation. The marrow aspirate taken from the vertebral body is then directly stored into collection tubes and transferred to the laboratory for further sample preparation. Similar technique was utilized for the distant vertebral body level samples (*e.g. Distal*). During each surgery, several vertebral body levels above and below the primary site of spinal cord compression are instrumented for stabilization. As such, there is access to numerous vertebral body levels through a single surgical approach. After the vertebral body marrow had been collected and the spine instrumented, the spinal cord is decompressed through a laminectomy. Epidural tumor that is circumferentially surrounding and compressing the spinal cord is taken directly from the field and transferred to the laboratory for further processing. Specimens were also submitted to pathology for standard confirmation of diagnosis of metastatic prostate cancer.

Animals

CCR6 KO mouse models were developed by Deltagen, Inc and ordered from The Jackson Laboratory and (B6.129P2-*Ccr6*^{tm1Dgen}/J, #005793) and compared to age and gender match control mice C57BL6/J mice (#000664). All mice were maintained in pathogen-free conditions and all procedures were approved by the institutional Animal Care and Use Committee of Massachusetts General Hospital.

Mice with signs of bone metastasis, as shown by bioluminescence imaging, were included in the survival study during which the health condition of the mice was scored. We adapted this scoring method from Nunamaker et al., to refine the endpoint for mice bearing bone metastasis to minimize

the pain and distress associated with bone metastasis progression[44]. For each category, mice are scored from 0 to 3: Body posture, eye appearance and activity level. Mice with a score at 3 in one of the categories or a cumulative score >6 are euthanized.

Development of syngeneic prostate bone metastases cell line:

The RM1-BoM1 cell line was obtained from the Power laboratory[42] and was established from the parental prostate cancer mouse model RM1 (RAS and c-Myc oncogenes were overexpressed in normal epithelial prostate cell and injected in syngeneic mice to form tumors[42]. We enriched RM1-BoM1 cells negative for GFP and transduced with Tdtomato (LeGO-T2, Addgene #27342) and luciferase (pENTR-LUC, Addgene, #17473). We injected 2.10^5 RM1-BoM1 cells into the left ventricle of C57Bl/6 mice and monitored bone metastases development by sequential bioluminescence imaging, bone metastatic cells were harvested and selected for the Luciferase and Tdtomato expression. We used the mouse as a bioreactor and repeated a second round of injection and isolation of bone metastases cells in order to establish the RM1-BoM3 bone metastatic cells. The RM1-BoM3 sells were maintained in DMEM (Corning, 15-013-CV) complemented with 10% FBS (GIBCO by Life Technologies, A31605-01) and 1% Penicillin-Streptomycin (GIBCO by Life Technologies, 15140-122).

In-vivo Bioluminescence imaging

Mice were anesthetized by isoflurane inhalation with 2% O_2 and 4.5mg/mouse of D-luciferin K salt (RR labs Inc., San Diego) was administered by intraperitoneal injection. After 5 minutes, mice are imaged using a SPECTRAL Ami X allowing to detect the localization of the cancer cells in the mice and to measure the luciferase activity.

Sample preparation

Dissociation of tissues into single cells

All samples were collected in Media 199 supplemented with 2% (v/v) FBS. Single cell suspensions of the tumors were obtained by cutting the tumor in to small pieces ($1mm^3$) followed by enzymatic dissociation for 45 minutes at 37°C with shaking at 120 rpm using Collagenase I, Collagenase II, Collagenase III, Collagenase IV (all at a concentration of 1 mg/ml) and Dispase (2mg/ml) in the presence of RNase inhibitors (RNasin (Promega) and RNase OUT (Invitrogen). Erythrocytes were subsequently removed by ACK Lysing buffer (Quality Biological) and cells resuspended in Media 199 supplemented with 2% (v/v) FBS for further analysis.

Bone marrow processing

Bone marrow samples were filtered using 70 micron filter then centrifuged at 600 g for 7 minutes at 4°C. Plasma were collected followed by erythrocytes removal using ACK Lysing buffer (Quality Biological). Cells were resuspended in Media 199 supplemented with 2% (v/v) FBS for further analysis.

FACS sorting of samples for single-cell RNA-sequencing

Single cells from tumor and bone marrow samples subjected to RBC lysis were surface stained with anti-CD235-PE (Biolegend) for 30 min at 4°C. Cells were washed twice with 2% FBS-PBS (v/v) followed by DAPI staining (1 ug/ml). Flow sorting for live and non-erythroid cells (DAPI-neg/CD235-neg) was performed on a BD FACS Aria III equipped with a 100um nozzle (BD Biosciences, San Jose, CA) instrument. All flow cytometry data were analyzed using FlowJo software (Treestar, San Carlos, CA).

FACS analysis

Cells from bone marrow and tumor samples were surface stained with lymphoid, myeloid and hematopoietic stem and progenitor antibody panels after blocking with anti human fc block (BD Pharmingen 564219) for 10 minutes at room temperature (Supplementary Table 7) for 30 min at 4°C. Cells were washed twice with 2% FBS-PBS (v/v). Samples that were to eventually be fixed and permeabilized were stained with LIVE/DEAD fixable viability dye (Thermo Fisher Scientific, Waltham, MA). For the intracellular staining, to determine Treg infiltration, cells stained with lymphoid surface markers were fixed and permeabilized with Cytofix/Cytoperm (BD Biosciences, San Jose, CA) for 20 min at 4°C. Cells were then washed twice with 1× Perm/Wash buffer (BD Biosciences, San Jose, CA) and incubated overnight at 4°C with anti-FoxP3-PE. On the following day, cells were washed twice in Perm/Wash buffer and resuspended in 2% FBS-PBS (v/v) for analysis.

Flow cytometry was performed on a BD FACS Aria III (BD Biosciences, San Jose, CA) instrument. All flow cytometry data were analyzed using FlowJo software (Treestar, San Carlos, CA) and Prism software. Statistical analyses were performed as indicated and *P* values of ≤0.05 considered significant.

Analysis

Quality control and preprocessing of single-cell RNA-seq data.

FASTQ files were processed with the CellRanger software (10x Genomics, Inc., version 2.0). Human genome hg19 was used as the reference genome (10x Genomics, Inc.) to generate the matrix files containing cell barcodes and transcript counts. Only cells with total UMI exceeding 600 were included in the downstream analysis. Statistics on the sequencing results are available in Supplementary

Table 2. Samples were analyzed using Pagoda2 (https://github.com/hms-dbmi/pagoda2) with multiple level clustering and visualized by t-SNE.

Joint Clustering Algorithm for analysis

We used Conos (https://github.com/hms-dbmi/conos) to integrate multiple scRNA-seq datasets together and to align our data with other public scRNA-seq data, closely following the Conos tutorial. Sample pre-processing was performed using default pagoda2 settings. The integration was performed using CPCA space and the mutual nearest neighbors (mNN) mapping. Louvain clustering was used to build to determine joint cell clusters across the entire dataset collection. A 2D embedding was generated using largeVis. Annotation of the cluster communities was done using marker gene expression and co-alignment with published datasets. To compare with published scRNA-seq data, the raw count matrix was fetched from healthy bone marrow[32], and integrated into the joint graph using the same Conos settings. Conos label transfer was used to examine cell annotation mapping.

Cell subpopulation identification, and magnitude of expression differences

To identify myeloid and T cells subpopulation, we extracted all myeloid and all T cell populations (CD8+ T cells, CD4+ T cells, NK and NKT cells), and realigned separately using Conos. Leiden community detection method (as implemented in Conos) was used to determine refined joint clusters, providing higher resolution than the initial analysis. New embedding was calculated for the T cell subpopulations (Fig. 2), while myeloid subpopulations were visualized on a part of the original 2D embedding.

Expression differences between matching subpopulations (Fig. 1f, Extended Data Fig. 1f) were determined by first estimating “mini-bulk” (or meta-cell) RNA-seq measurements for each subpopulation in each sample. Briefly, in each dataset, the molecules from all cells belonging to a given subpopulation were summed for each gene (i.e. forgetting cellular barcodes). The distance between the resulting high-coverage RNA-seq vectors was calculated using Jensen-Shannon divergence (JS). Expression distances between samples (Extended Data Fig. 1i, Fig. 1g) were determined as a normalized weighted sum of JS divergence distances across all cell subpopulations contained in both samples, with the weight equal to the subpopulation proportion (measured as a minimal proportion that the given cell subpopulation represents among the two samples being compared).

Score signature gene modules

To account for the complex gene expression patterns in M2 tumor associated macrophage, monocytes differentiation to macrophage, Treg activity, CD8+ T cell dysfunction and cytotoxicity, The signature scores were estimated as average expression values of the genes in a given set.

Specifically, we first calculated signature score for each cell as an average normalized (for cell size) gene expression magnitudes, then the signature score for each sample was computed as the mean across all cells. All signature gene modules are listed in the Supplementary Table 4. The M2 signature genes were from Aziz et al[15]. CD8+ T cell cytotoxicity were measured by expression of CD8A, CD8B, GZMA, GZMM, GZMB, GZMK, GZMB, PRF1[45]. The CD8+ T cell dysfunction genes were taken from Li et al[34]. To define gene signature of monocytes to macrophage differentiation, we took top 100 differentially expressed genes based on published *in vitro* studies of monocytes to macrophage differentiation[46,47]. For this, the microarray data was analyzed by using affy[45] R package, and the differentially expressed genes were identified using limma[46] R package. Wilcoxon rank-sum test was used to identify statistically significant changes in the resulting gene signature scores between cell subsets or different sample fractions.

Differentially expressed genes

Differential expression (DE) and marker gene detection was performed using Wilcoxon rank sum test, implemented by the getDifferentialGenes() function from Conos R package. The genes were considered differentially expressed if the p-value determined Z score was greater than 3. Since the genes were used primarily for pathway enrichment analysis, the Z score was not corrected for multiple hypothesis testing. getDifferentialGenes function was used to identify difference between similar subpopulations (CTL-1 vs. CTL-2; Mono1 vs. Mono2 vs. Mono3). For comparisons between sample fractions (for example *Tumor* Treg vs. *Benign* Treg), getPerCellTypeDE() function in Conos was utilized. As described previously, it first forms “mini-bulk” (or meta-cell) RNA-seq measurements by combining all molecules measured for each gene in each subpopulation in each sample. This results in a collection of bulk-like RNA-seq samples, and the differential expression problem is then reformulated as a standard bulk RNA-seq differential expression problem DESeq2 to compare these “bulk-like” meta-cell samples, using appropriate design models (i.e. *Tumor* vs. *Benign* factor on samples).

Ligand and Receptor analysis

To screen for potential channels of communications between different cell types (Extended Data Fig. 4d), we looked at expression of previously annotated reciprocal ligand-receptor pairs. The annotated list containing 1307 pairs of well-annotated receptors and ligands was downloaded from CellPhoneDB[48]. Both ligands and corresponding receptors were required to be detected in more than 10% of the cells of a given type. Further filtering for ligand-receptor channels potentially connecting TIM/TAM with T cell populations was performed by requiring ligands to be upregulated in TIM or TAM cells (compared to other myeloid cells), and requiring the corresponding receptor to be upregulated (log fold change >0) in T cells in *Tumor* compared to *Benign* (Supplementary Table 5). Differentially

expressed genes were identified using Wilcoxon rank sum test, implemented by the getDifferentialGenes function in Conos.

Survival analysis on bulk data

To test if a given gene signature is associated with differential survival of cancer patients [49] (Extended Data Fig. 2e), we first calculated average expression of the signature in each cancer type based on the bulk RNA-seq data. The bulk patient samples were then stratified into two groups, based on average signature expression, separating patients with top 25% of scores and bottom 25%. A standard Kaplan-Meier survival analysis was then used to determine the association of these groups with survival rate. Kaplan-Meier survival analysis in Fig. 4e and Extended Data Fig. 2e were performed using survival R package.

Gene ontology and gene set enrichment analysis.

We used DAVID to perform Gene Ontology analysis of top differential expressed genes. Pre-ranked analysis module in GSEA was used for Gene Set Enrichment Analysis. The gene sets were downloaded from MSigDB.

Massively parallel single cell RNA-sequencing

Single cells were encapsulated into emulsion droplets using Chromium Controller (10x Genomics). scRNA-seq libraries were constructed using Chromium Single Cell 3' v2 Reagent Kit according to the manufacturer's protocol. Briefly, post sorting sample volume was decreased and cells were examined under a microscope and counted with a hemocytometer. Cells were then loaded in each channel with a target output of ~4,000 cells. Reverse transcription and library preparation was done on C1000 Touch Thermal cycler with 96-Deep Well Reaction Module (Bio-Rad). Amplified cDNA and final libraries were evaluated on an Agilent BioAnalyzer using a High Sensitivity DNA Kit (Agilent Technologies). Individual libraries were diluted to 4nM and pooled for sequencing. Pools were sequenced with 75 cycle run kits (26bp Read1, 8bp Index1 and 55bp Read2) on the NextSeq 500 Sequencing System (Illumina) to ~70-80% saturation level.

Data availability

The expression datasets generated in this study are available at dbGaP, GEO and SRA through their respective accession numbers GSEXXX and SRA: PRJNAXXX.

Figure 1

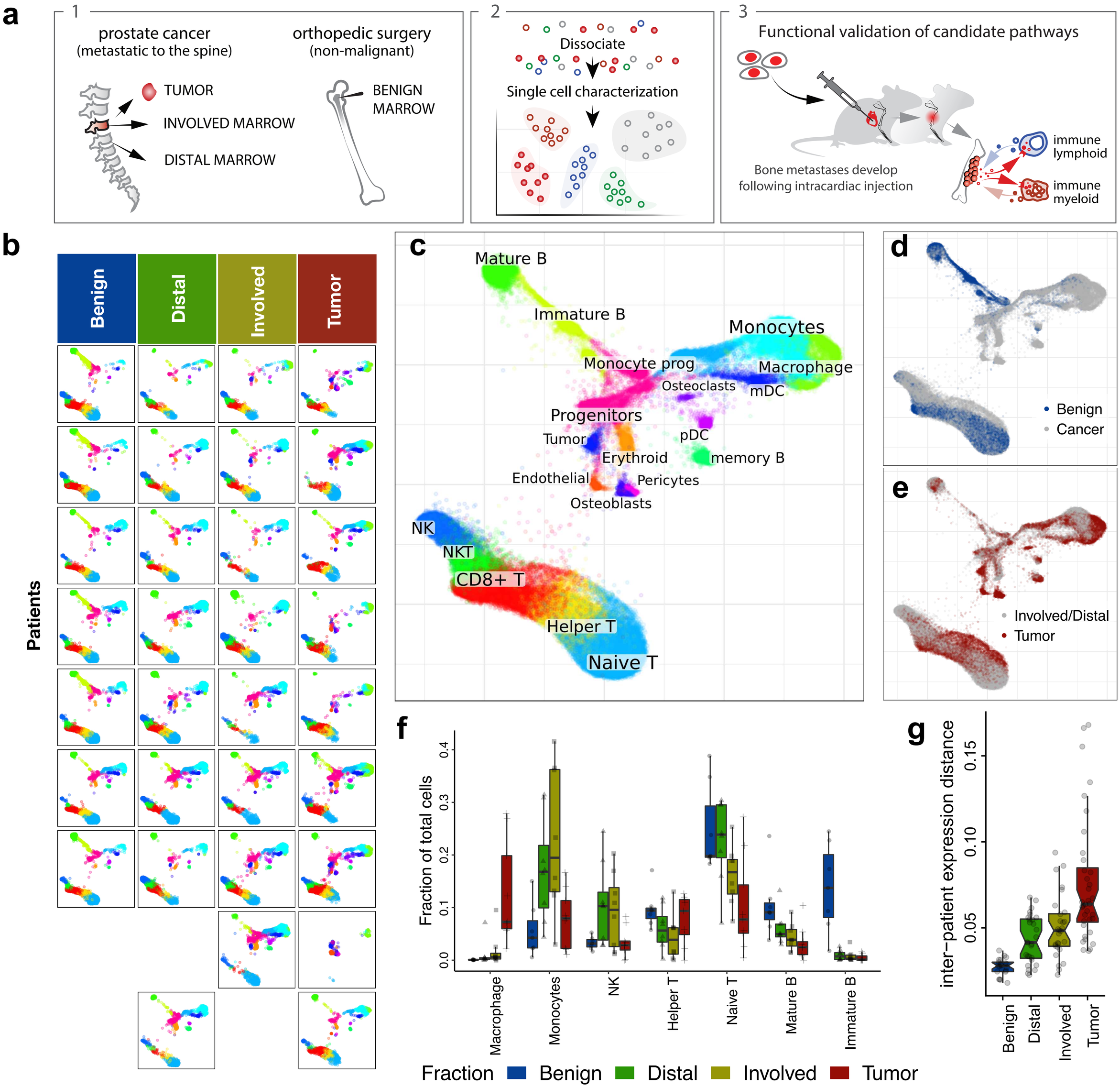

a
1
prostate cancer
(metastatic to the spine)
orthopedic surgery
(non-malignant)
TUMOR
INVOLVED MARROW
DISTAL MARROW
BENIGN MARROW
2
Dissociate
Single cell characterization
3
Functional validation of candidate pathways
Bone metastases develop following intracardiac injection
immune lymphoid
immune myeloid
b
Benign
Distal
Involved
Tumor
Patients
c
Mature B
Immature B
Monocytes
Monocyte prog
Macrophage
Osteoclasts
mDC
Progenitors
Tumor
pDC
Erythroid
memory B
Endothelial
Pericytes
Osteoblasts
NK
NKT
CD8+ T
Helper T
Naive T
d
Benign
Cancer
e
Involved/Distal
Tumor
f
Fraction of total cells
0.0
0.1
0.2
0.3
0.4
Macrophage
Monocytes
NK
Helper T
Naive T
Mature B
Immature B
Fraction
Benign
Distal
Involved
Tumor
g
inter−patient expression distance
0.05
0.10
0.15
Benign
Distal
Involved
Tumor

Figure 2

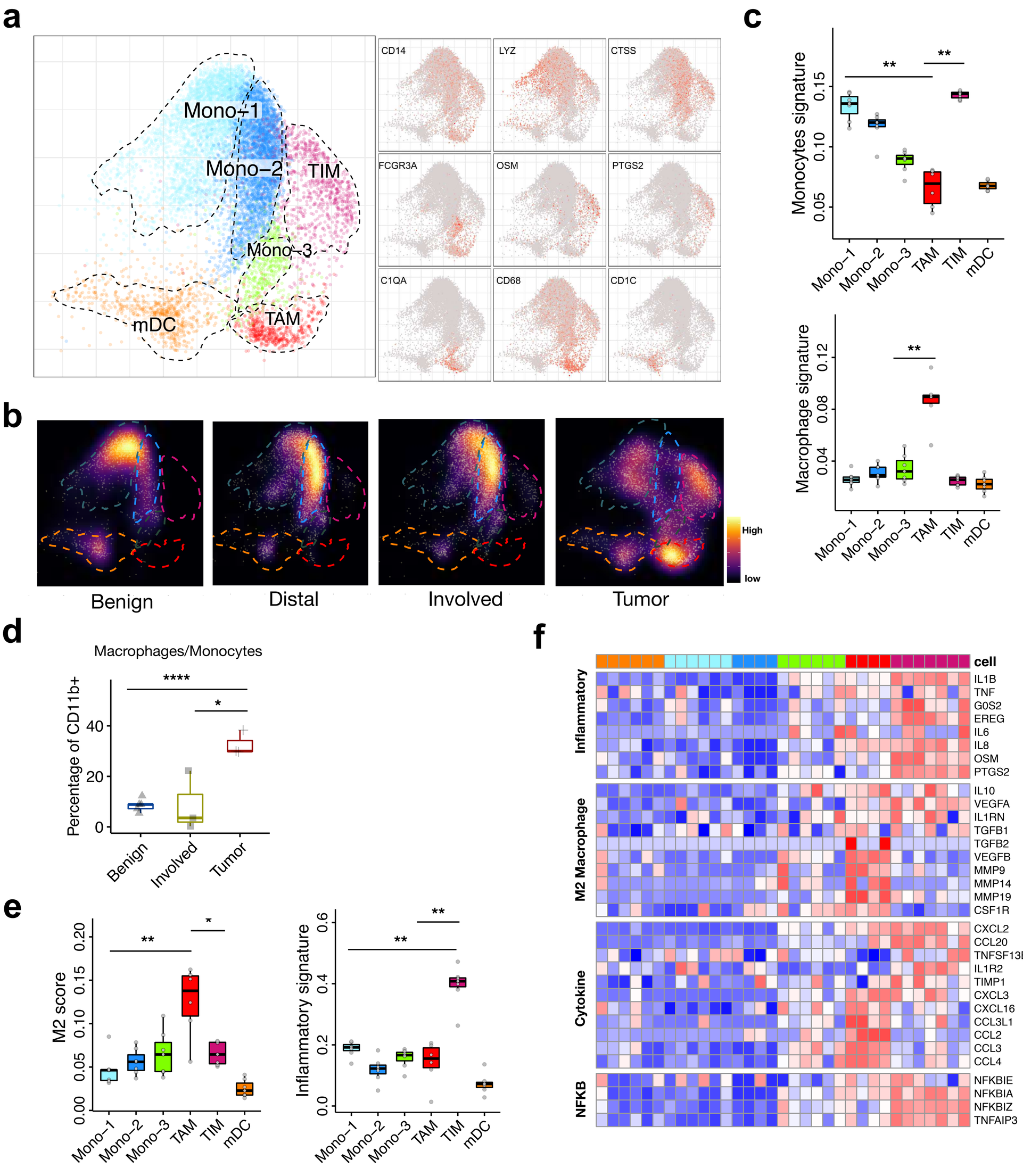

a
Mono−1
Mono−2
TIM
Mono−3
mDC
TAM
CD14
LYZ
CTSS
FCGR3A
OSM
PTGS2
C1QA
CD68
CD1C
b
High
low
Benign
Distal
Involved
Tumor
c
Monocytes signature
Mono−1
Mono−2
Mono−3
TAM
TIM
mDC
Macrophage signature
d
Macrophages/Monocytes
Percentage of CD11b+
Benign
Involved
Tumor
e
M2 score
Inflammatory signature
f
cell
Inflammatory
IL1B
TNF
G0S2
EREG
IL6
IL8
OSM
PTGS2
M2 Macrophage
IL10
VEGFA
IL1RN
TGFB1
TGFB2
VEGFB
MMP9
MMP14
MMP19
CSF1R
Cytokine
CXCL2
CCL20
TNFSF13B
IL1R2
TIMP1
CXCL3
CXCL16
CCL3L1
CCL2
CCL3
CCL4
NFKB
NFKBIE
NFKBIA
NFKBIZ
TNFAIP3

Figure 3

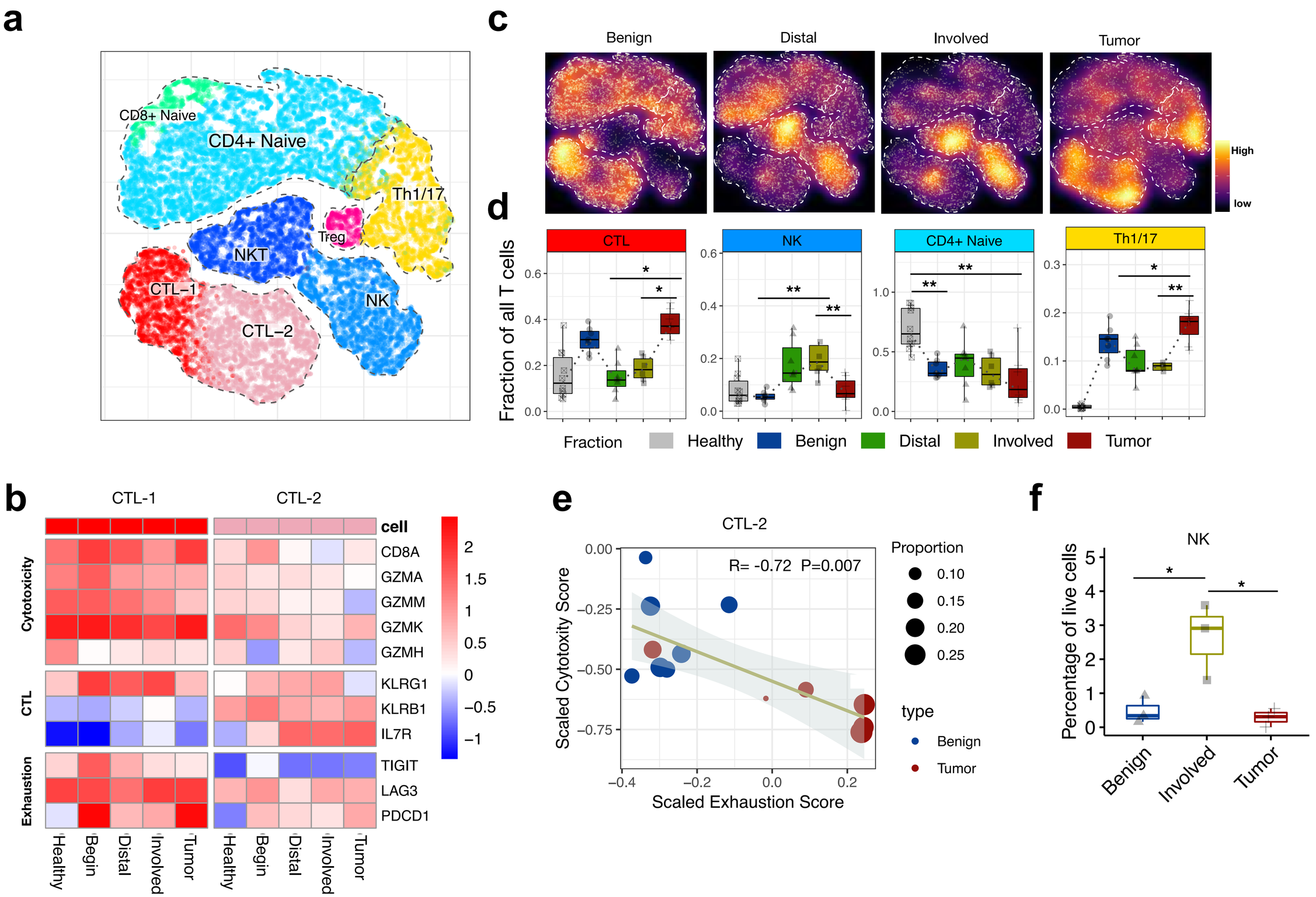
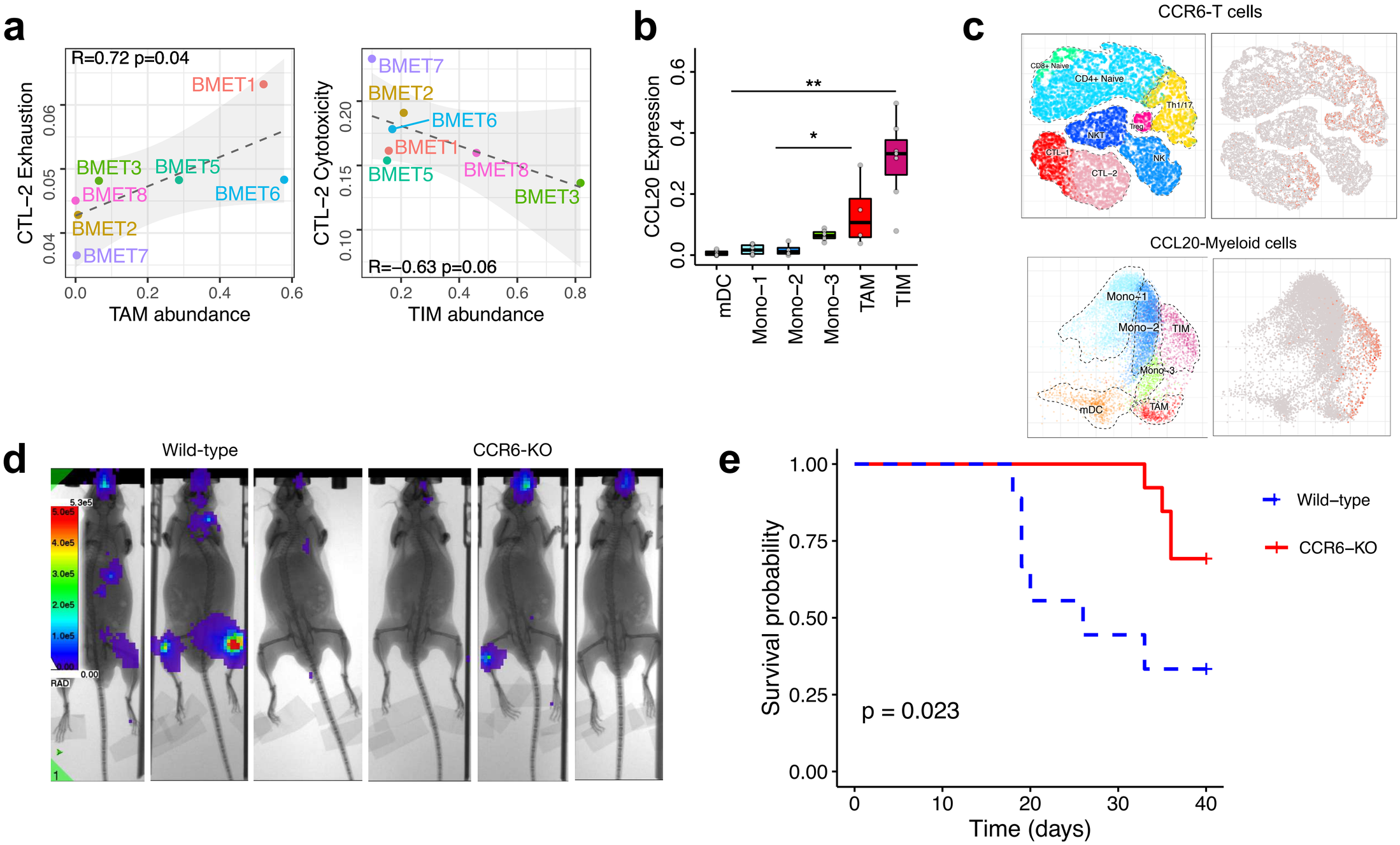
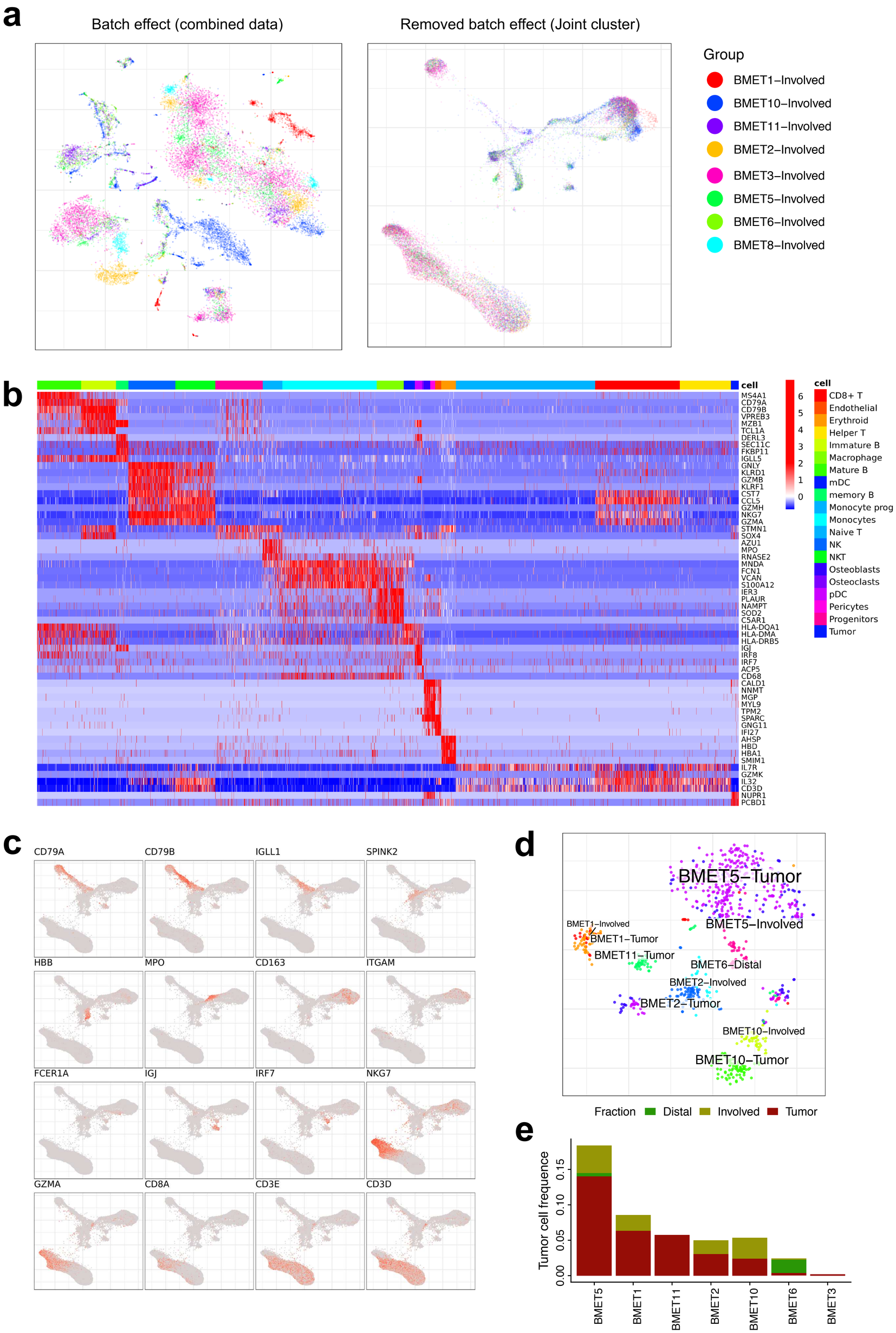
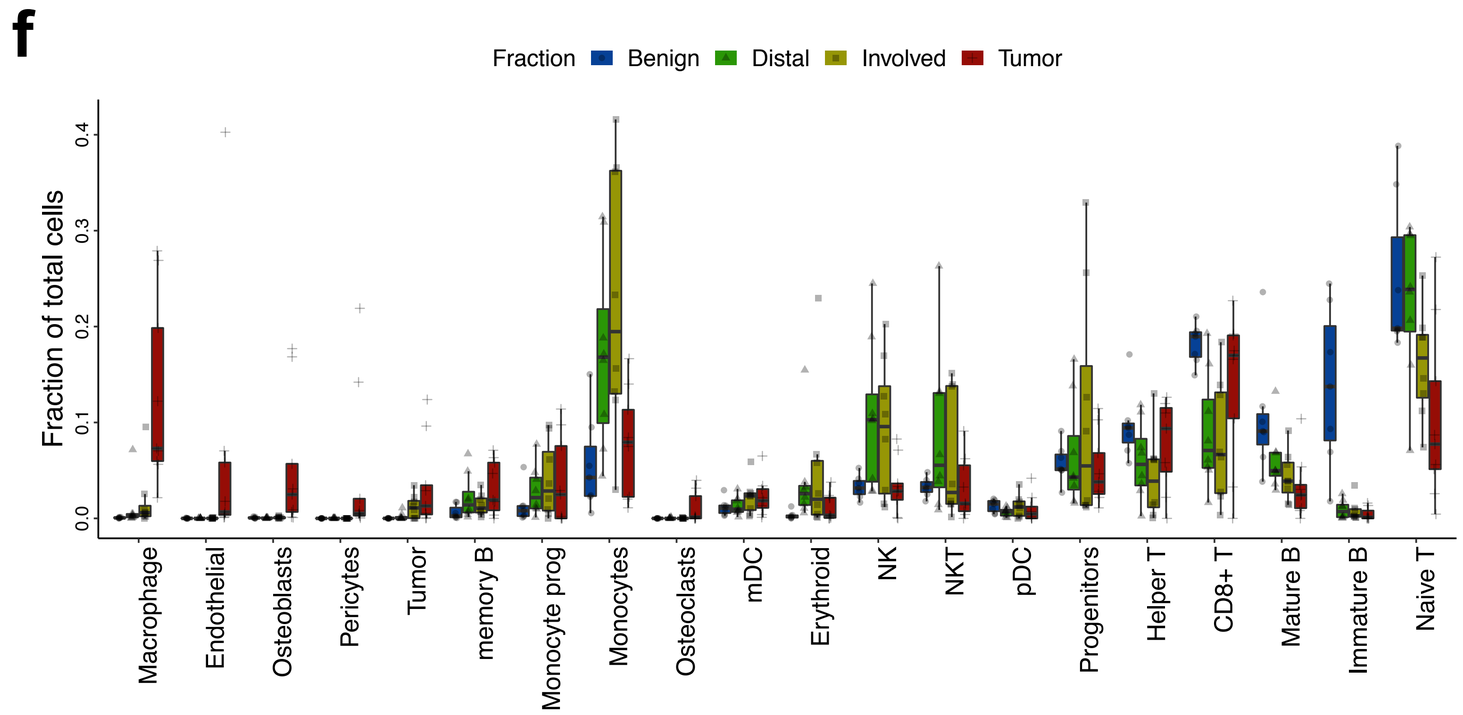

a
CD8+ Naive
CD4+ Naive
Th1/17
Treg
NKT
CTL−1
NK
CTL−2
c
Benign
Distal
Involved
Tumor
High
low
d
Fraction of all T cells
CTL
NK
CD4+ Naive
Th1/17
Fraction
Healthy
Benign
Distal
Involved
Tumor
b
CTL-1
CTL-2
cell
Cytotoxicity
CTL
Exhaustion
CD8A
GZMA
GZMM
GZMK
GZMH
KLRG1
KLRB1
IL7R
TIGIT
LAG3
PDCD1
Healthy
Begin
Distal
Involved
Tumor
e
CTL-2
R= -0.72 P=0.007
Scaled Cytotoxity Score
Scaled Exhaustion Score
Proportion
type
Benign
Tumor
f
NK
Percentage of live cells
Benign
Involved
Tumor

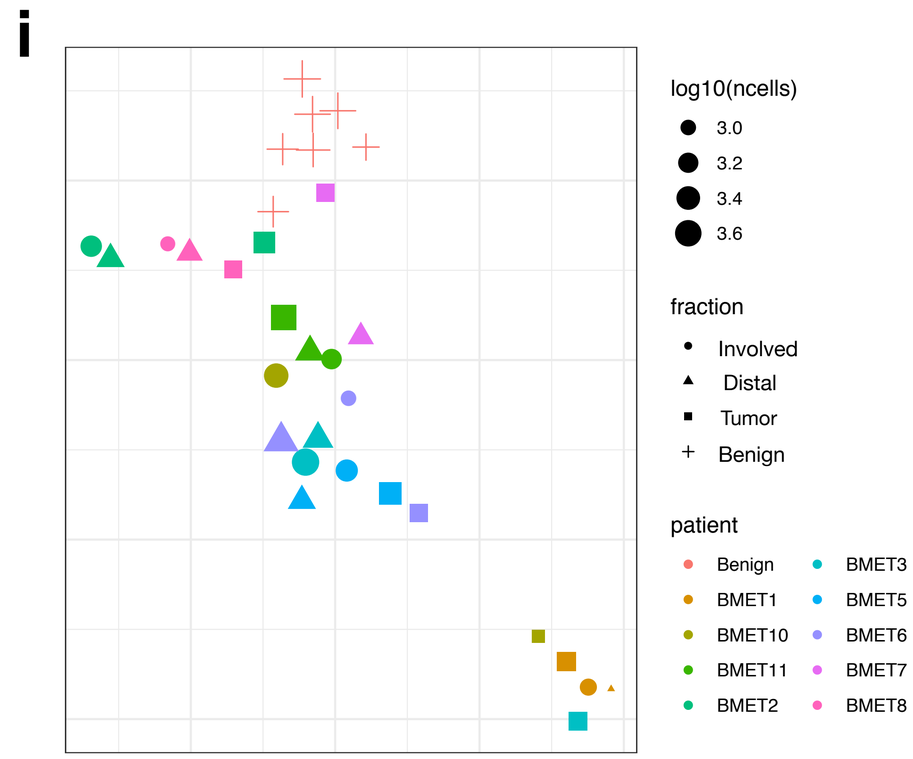
Figure 4

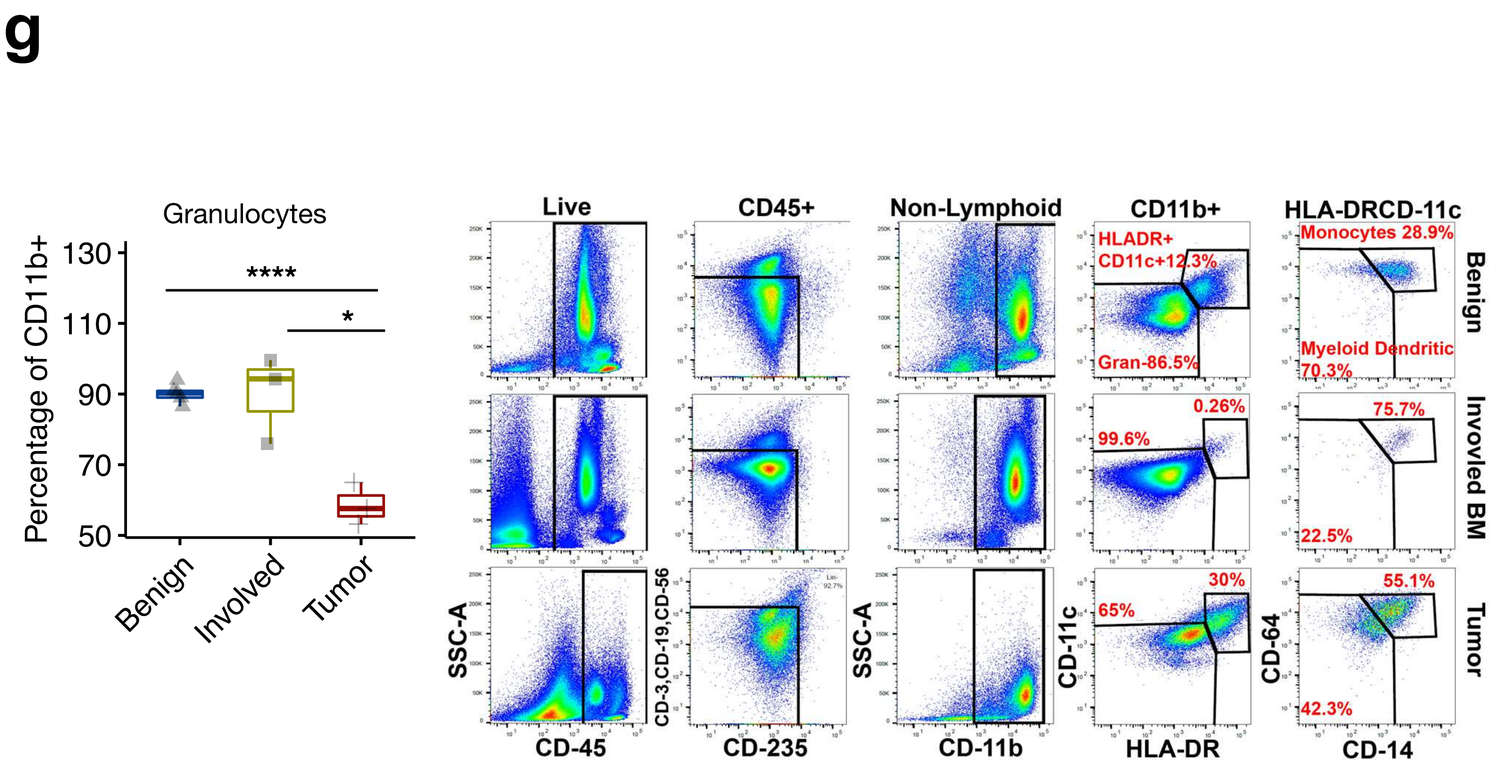

a
R=0.72 p=0.04
CTL−2 Exhaustion
TAM abundance
BMET1
BMET3
BMET5
BMET8
BMET6
BMET2
BMET7
R=−0.63 p=0.06
CTL−2 Cytotoxicity
TIM abundance
b
CCL20 Expression
mDC
Mono−1
Mono−2
Mono−3
TAM
TIM
c
CCR6-T cells
CD8+ Naive
CD4+ Naive
Th1/17
Treg
NKT
CTL−1
NK
CTL−2
CCL20-Myeloid cells
Mono−1
Mono−2
TIM
Mono−3
mDC
TAM
d
Wild-type
CCR6-KO
e
Survival probability
p = 0.023
Time (days)
Wild−type
CCR6−KO
f
Resting Treg
Active Treg
Less exhausted CTL
T helper
Exhausted CTL
transition exhausted
Naive CTL
NK
Wild-type
CCR6-KO

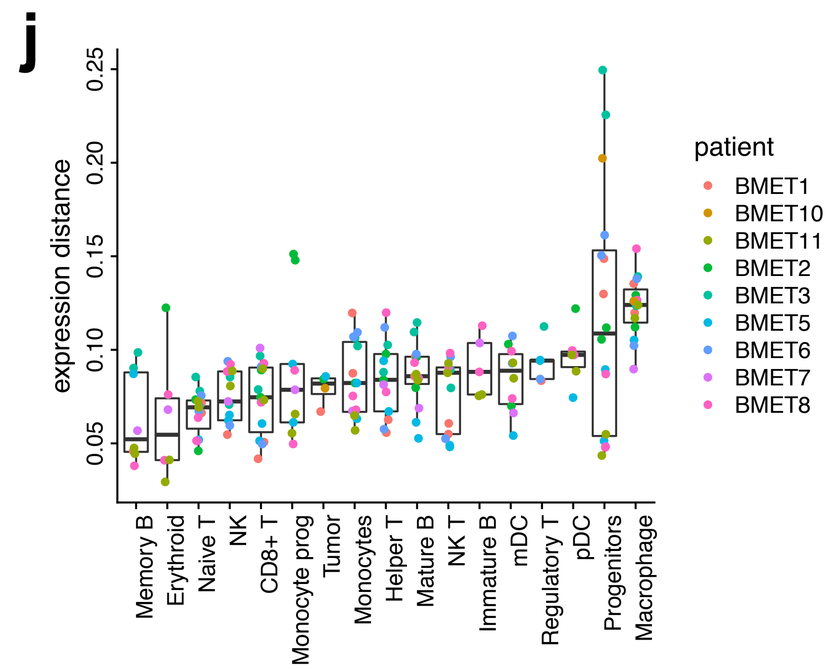
Extended Data Fig. 1

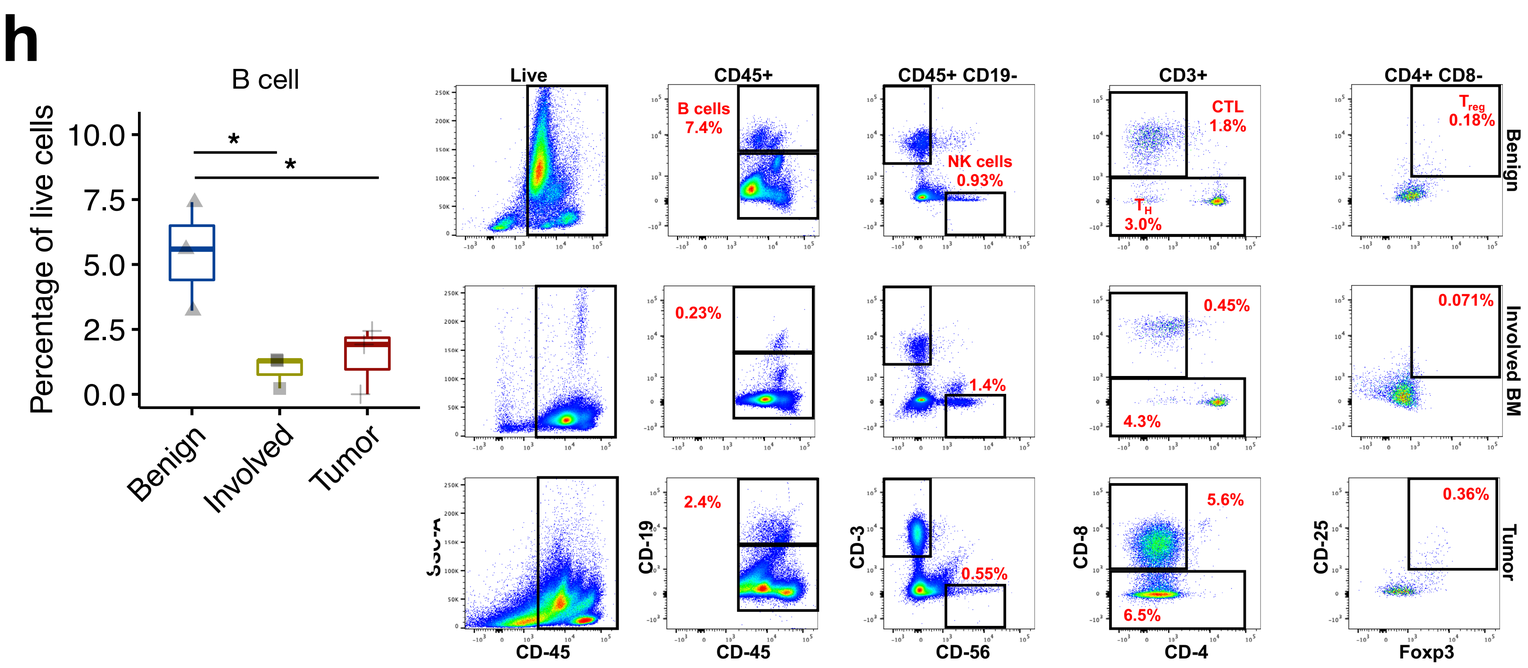



f
Fraction Benign Distal Involved Tumor
Fraction of total cells
0.0
0.1
0.2
0.3
0.4
Macrophage
Endothelial
Osteoblasts
Pericytes
Tumor
memory B
Monocyte prog
Monocytes
Osteoclasts
mDC
Erythroid
NK
NKT
pDC
Progenitors
Helper T
CD8+ T
Mature B
Immature B
Naive T

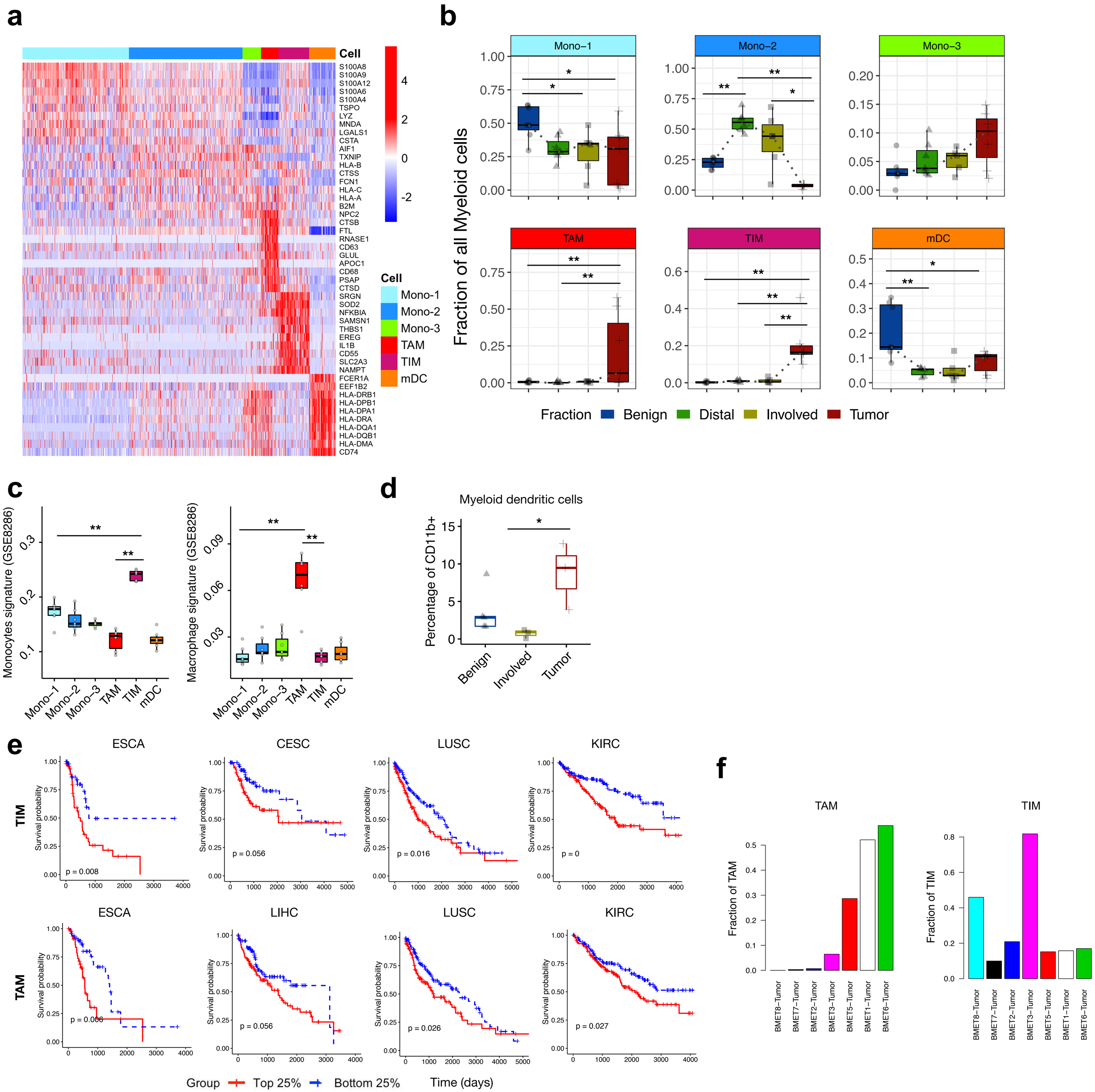

i
log10(ncells)
3.0
3.2
3.4
3.6
fraction
Involved
Distal
Tumor
Benign
patient
Benign
BMET1
BMET10
BMET11
BMET2
BMET3
BMET5
BMET6
BMET7
BMET8


g
Granulocytes
Percentage of CD11b+
130
110
90
70
50
****
*
Benign
Involved
Tumor
Live
CD45+
Non-Lymphoid
CD11b+
HLA-DRCD-11c
HLADR+
CD11c+12.3%
Gran-86.5%
Monocytes 28.9%
Myeloid Dendritic
70.3%
Benign
99.6%
0.26%
75.7%
22.5%
Invovled BM
65%
30%
55.1%
42.3%
Tumor
SSC-A
CD-45
CD-3,CD-19,CD-56
CD-235
SSC-A
CD-11b
CD-11c
HLA-DR
CD-64
CD-14

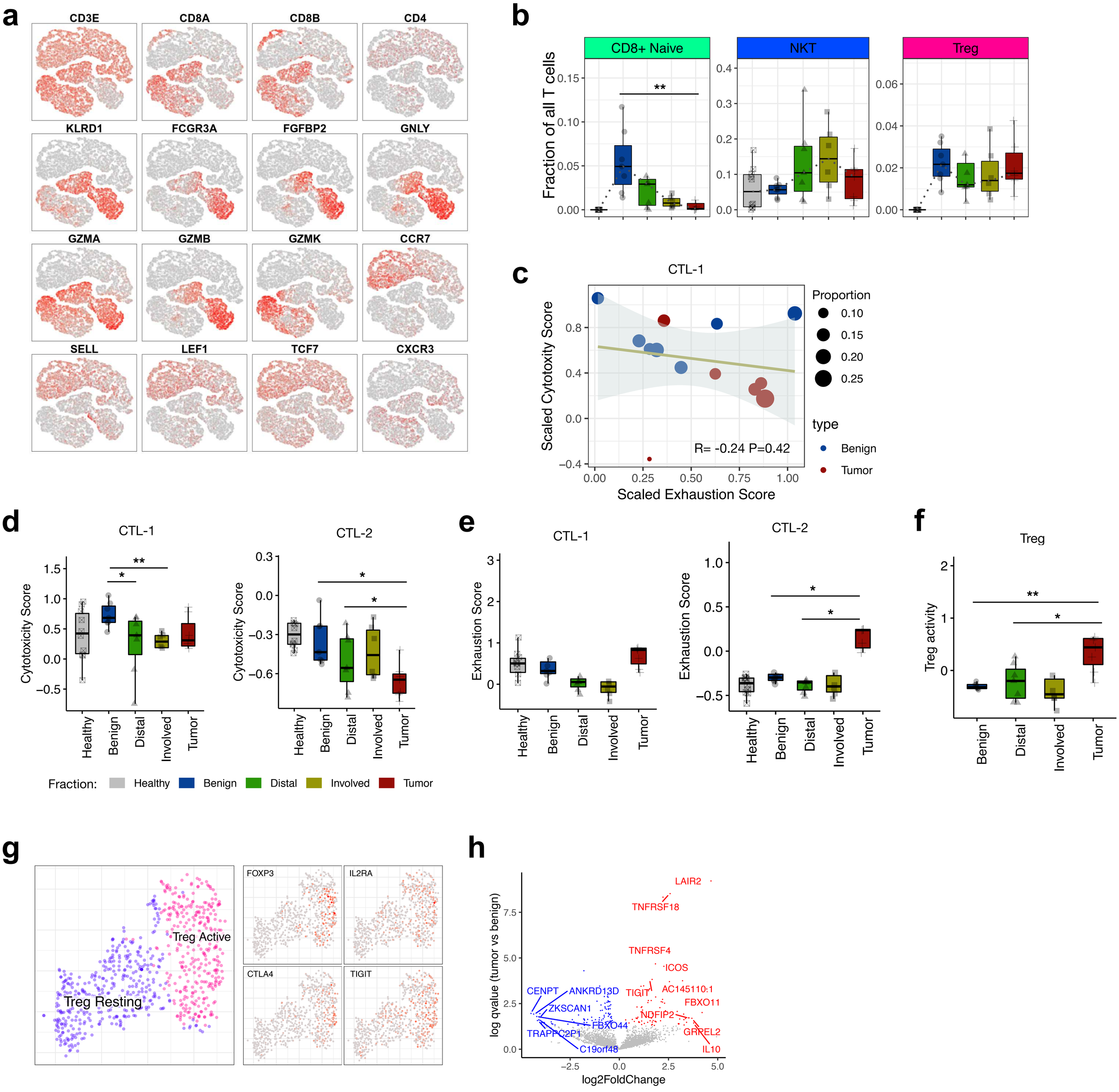
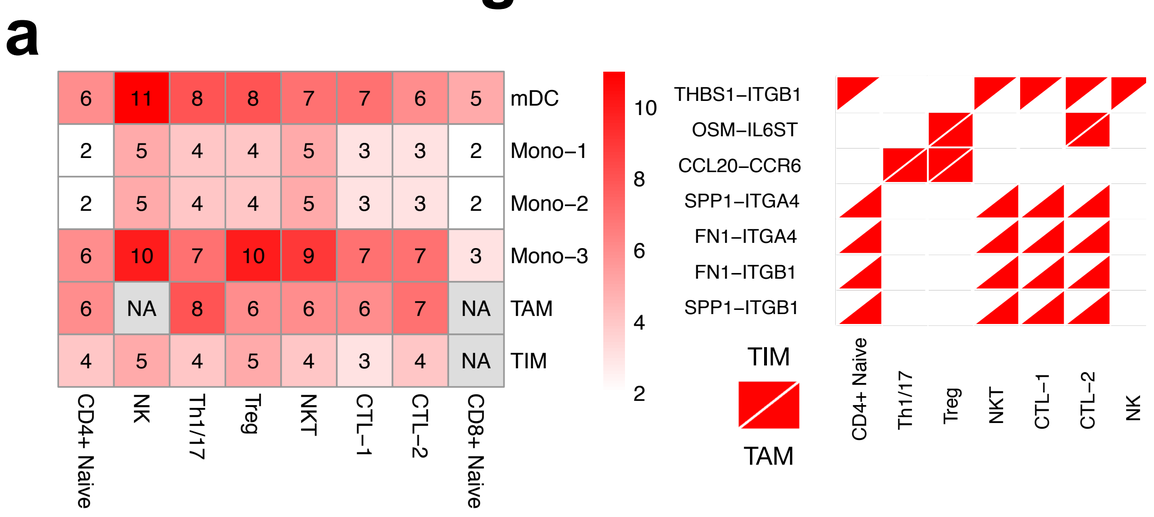
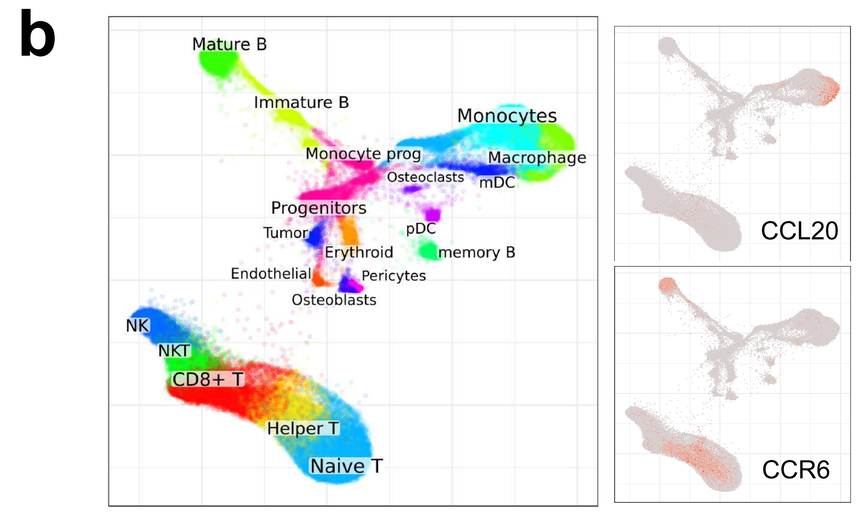
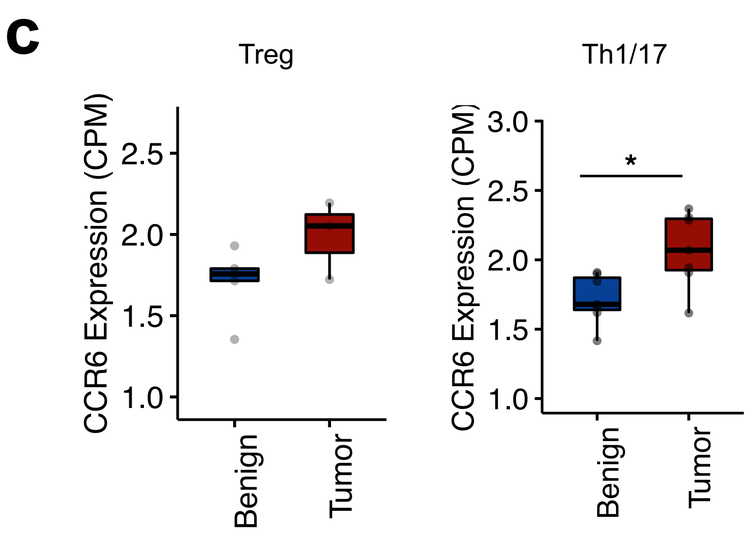
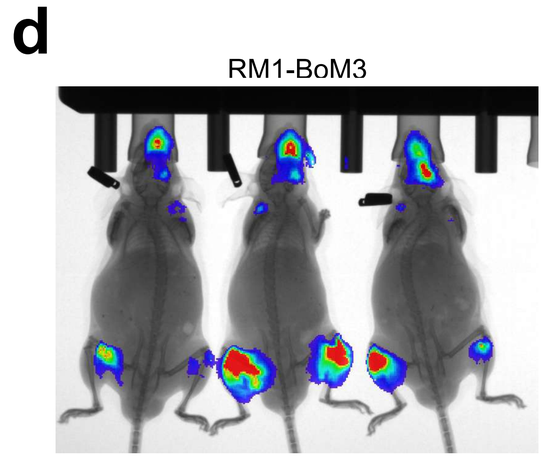
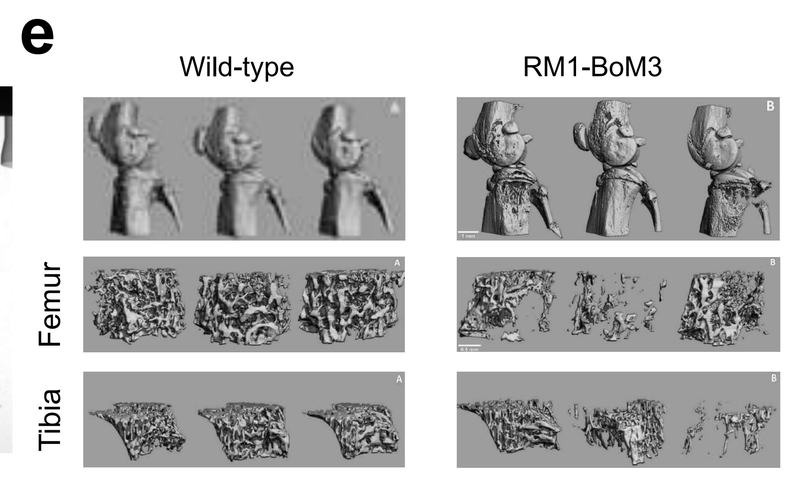
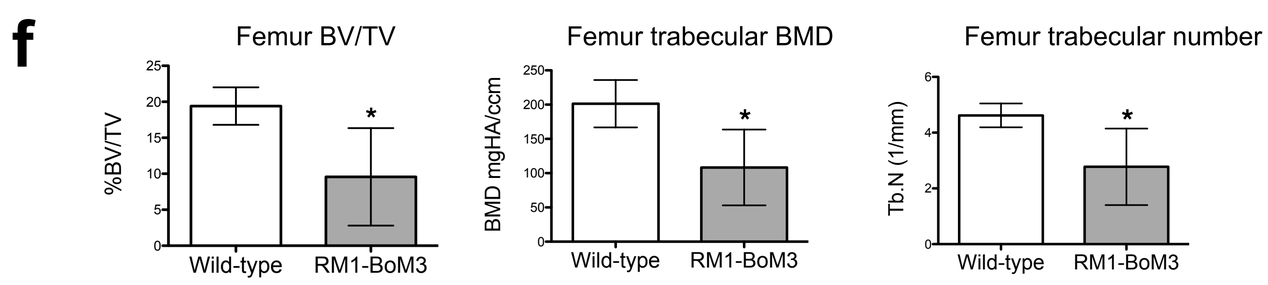
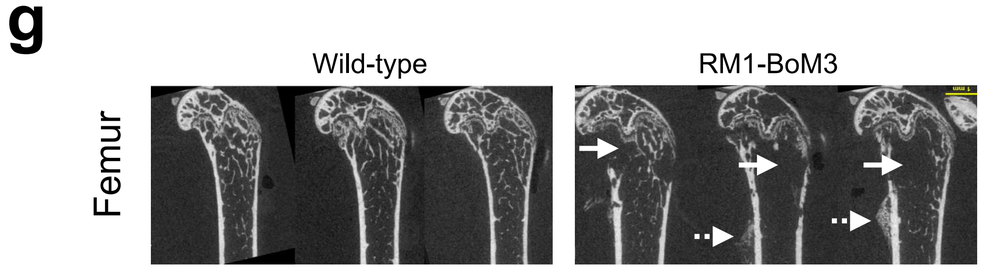
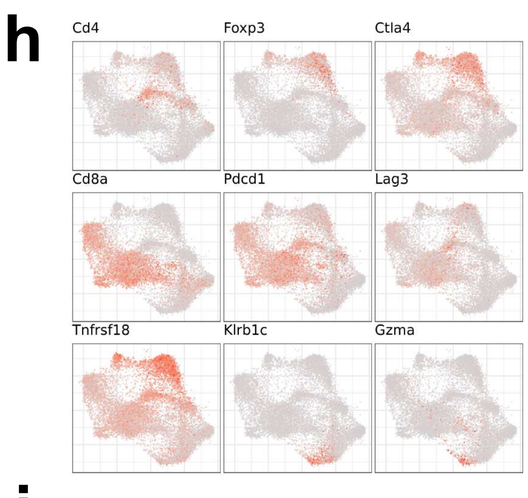
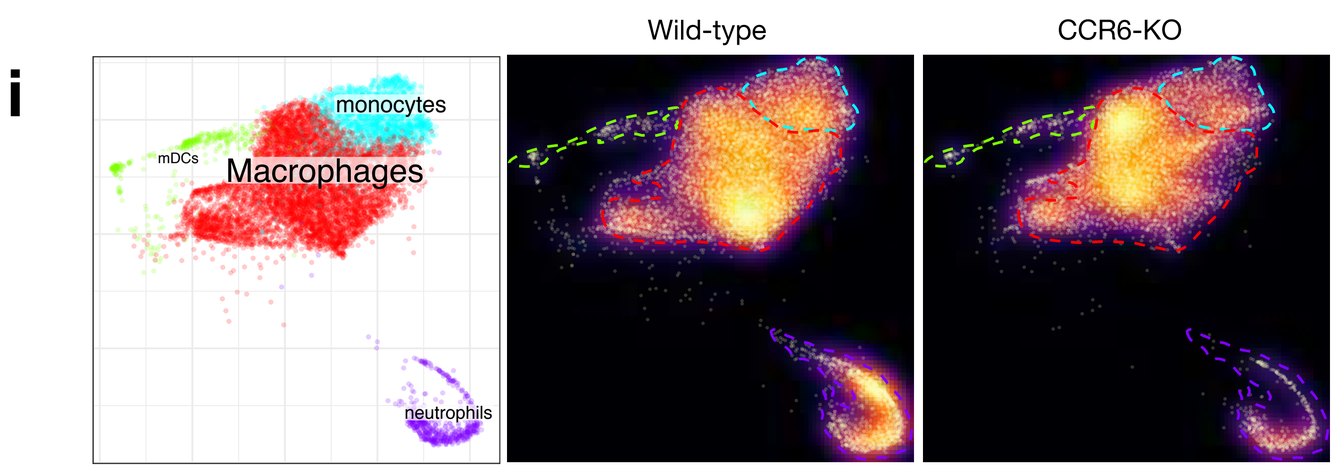
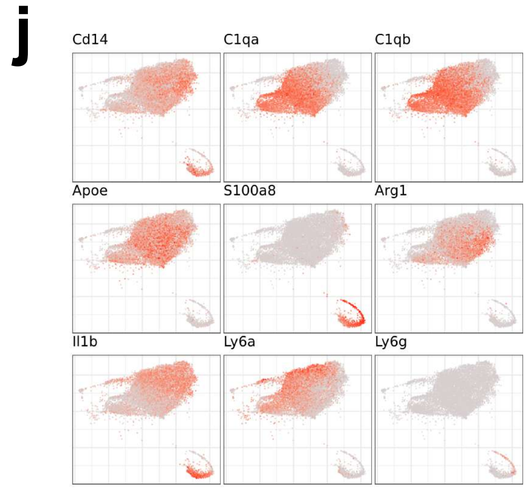
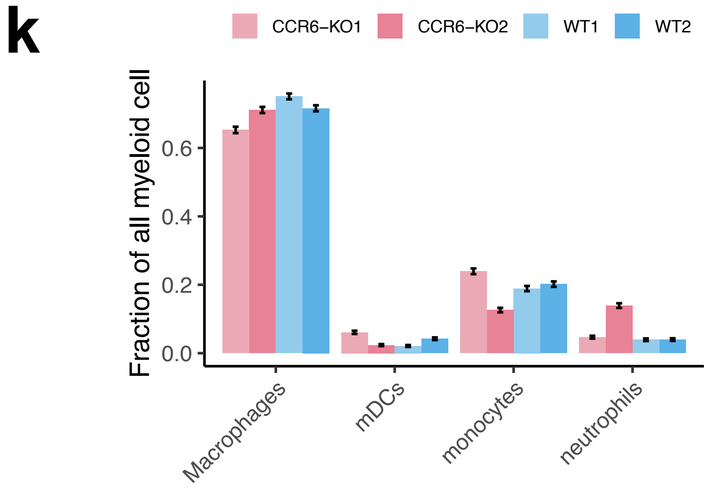
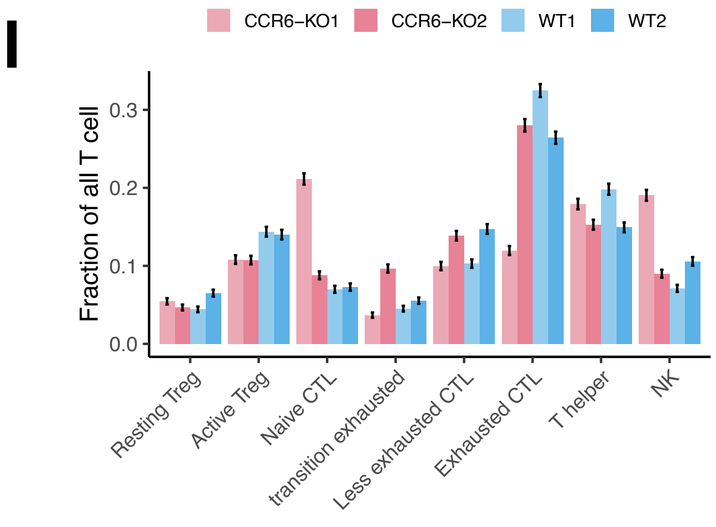

j
expression distance
0.05
0.10
0.15
0.20
0.25
patient
BMET1
BMET10
BMET11
BMET2
BMET3
BMET5
BMET6
BMET7
BMET8
Memory B
Erythroid
Naive T
NK
CD8+ T
Monocyte prog
Tumor
Monocytes
Helper T
Mature B
NK T
Immature B
mDC
Regulatory T
pDC
Progenitors
Macrophage


h
B cell
Percentage of live cells
10.0
7.5
5.0
2.5
0.0
*
*
Benign
Involved
Tumor
Live
CD45+
CD45+ CD19-
CD3+
CD4+ CD8-
B cells
7.4%
NK cells
0.93%
CTL
1.8%
TH
3.0%
Treg
0.18%
Benign
0.23%
1.4%
0.45%
4.3%
0.071%
Involved BM
2.4%
0.55%
5.6%
6.5%
0.36%
Tumor
CD-45
CD-19
CD-45
CD-3
CD-56
CD-8
CD-4
CD-25
Foxp3

Extended Data Fig. 2



Extended Data Fig. 3



Extended Data Fig. 4