

# UCLA

## UCLA Previously Published Works

### Title

Dissecting the treatment-naive ecosystem of human melanoma brain metastasis

### Permalink

<https://escholarship.org/uc/item/3087t3dx>

### Journal

Cell, 185(14)

### ISSN

0092-8674

### Authors

Biermann, Jana  
Melms, Johannes C  
Amin, Amit Dipak  
et al.

### Publication Date

2022-07-01

### DOI

10.1016/j.cell.2022.06.007

Peer reviewed



Published in final edited form as:

Cell. 2022 July 07; 185(14): 2591–2608.e30. doi:10.1016/j.cell.2022.06.007.

## Dissecting the treatment-naïve ecosystem of human melanoma brain metastasis

*A full list of authors and affiliations appears at the end of the article.*

### SUMMARY

Melanoma brain metastasis (MBM) frequently occurs in patients with advanced melanoma, yet our understanding of the underlying salient biology is rudimentary. Here, we performed single-cell/nucleus RNA-seq in 22 treatment-naïve MBM and 10 extracranial melanoma metastases (ECM), and matched spatial single-cell transcriptomics and T cell receptor (TCR)-seq. Cancer cells from MBM were more chromosomally unstable, adopted a neuronal-like cell state, and enriched for spatially variably expressed metabolic pathways. Key observations were validated in independent patient cohorts, patient-derived MBM/ECM xenograft models, RNA/ATAC-seq, proteomics, and multiplexed imaging. Integrated spatial analyses revealed distinct geography of putative cancer immune evasion, and evidence for more abundant intra-tumoral B to plasma cell differentiation in lymphoid aggregates in MBM. MBM harbored larger fractions of monocyte-derived macrophages and dysfunctional TOX<sup>+</sup>CD8<sup>+</sup> T cells with distinct expression of immune checkpoints. This work provides comprehensive insights into MBM biology and serves as a foundational resource for further discovery and therapeutic exploration.

### Graphical Abstract

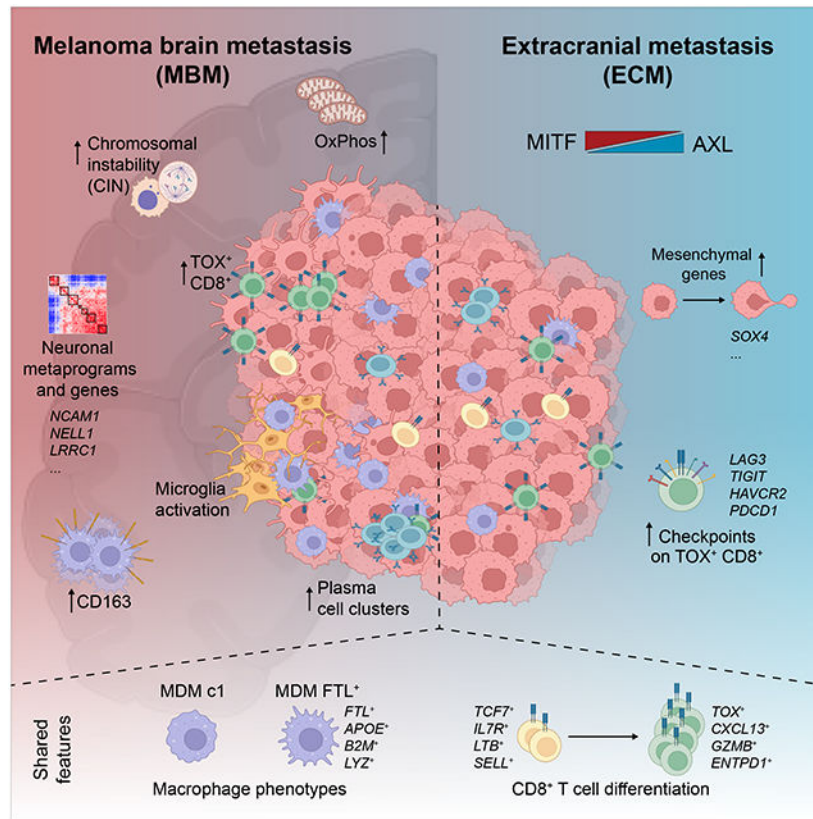
\*Correspondence: bi2175@cumc.columbia.edu.

#### AUTHOR CONTRIBUTIONS

B.I. provided overall supervision of the study. J.C.M., W.L.B., and B.I. conceptualized the study. J.C.M., A.D.A., L.A.C., A.K., I.B., B.T.F., T.T.T.N., M.Ro., P.H., M.B., N.H., A.M., Z.H.W., S.B.S., D.H.V., M.M., S.B., A.M.L., A.A. and P.A. performed experiments. M.I., S.A.K., J.D., C.L.S., T.E., M.Rö., R.C., M.B.A., M.A.D., H.N.B., W.L.B., A.R., P.C., J.N.B., G.K.S. and B.I. took care of patients and provided critical tissue resources. J.B., J.C.M., A.D.A., Y.W., A.K., S.T., M.A.I.A., M.A., K.H.G., V.S.M., V.S., A.N., P.D.K., F.W., M.C., B.C., E.A., S.J.C., and B.I. performed analyses. S.F.B., M.S., C.L., S.J.C., P.C., J.N.B., W.L.B., D.S., E.H., E.Z.M. and F.C. provided additional supervision. H.H. and P.C. performed pathological review of tissue specimens. J.C.M., J.B., A.D.A., Y.W., S.T., and B.I. wrote the manuscript.

#### DECLARATION OF INTERESTS

B.I. has received honoraria from consulting with Merck, J&J/Janssen Pharmaceuticals, AstraZeneca and Volastra Therapeutics. MAD has been a consultant to Roche/Genentech, Array, Pfizer, Novartis, BMS, GSK, Sanofi-Aventis, Vaccinex, Apexigen, Eisai, and ABM Therapeutics, and he has been the PI of research grants to MD Anderson by Roche/Genentech, GSK, Sanofi-Aventis, Merck, Myriad, and Oncothyreon. A.R. has received honoraria from consulting with CStone, Merck, and Vedanta, is or has been a member of the scientific advisory board and holds stock in Advaxis, Appia, Apricity, Arcus, Compugen, CytomX, Highlight, ImaginAb, ImmPact, ImmuneSensor, Inspira, Isoplexis, Kite-Gilead, Lutris, MapKure, Merus, PACT, Pluto, RAPT, SyntheKine and Tango, has received research funding from Agilent and from Bristol-Myers Squibb through Stand Up to Cancer (SU2C), and patent royalties from Arsenal Bio. T.E. has acted as a consultant for Almiral Hermal, Bristol-Myers Squibb, MSD, Novartis, Pierre Fabre and Sanofi. E.Z.M. is a consultant for Curio Bioscience. The other authors do not declare any conflict of interest.



### “In Brief”:

Multi-modal single-cell analysis reveals genomic, transcriptional, and tumor-microenvironmental features of treatment-naïve human melanoma brain metastases compared to extracranial metastases.

## INTRODUCTION

Melanoma brain metastases (MBM) are the third most common cause of brain metastases after carcinomas of the lung and breast (Eichler et al., 2011) and lead to significant morbidity and mortality (Davies et al., 2011). While treatment with combination immune checkpoint blockade can be effective in patients with MBM (Tawbi et al., 2018), many patients do not respond, and MBM frequently develop or evolve in situations where extracranial disease is controlled, thus posing a particular therapeutic challenge (Brastianos et al., 2013). Prior genomic profiling studies identified divergent somatic mutations (and putative drivers) in brain metastasis across carcinomas (Brastianos et al., 2015), but not in MBM (Fischer et al., 2019). This suggests that other genomic, phenotypic or tumor-microenvironmental features play a role in MBM (Fischer et al., 2019; Fukumura et al., 2021), yet a precise understanding of the cellular composition, as well as molecular and spatial underpinnings of the MBM ecosystem remains rudimentary.

Single-cell RNA-sequencing (scRNA-seq) has provided important biological insights in multiple cancer types, including metastatic melanoma, revealing cell states associated with

resistance to targeted/immune-therapies (Jerby-Arnon et al., 2018; Tirosh et al., 2016). To date, scRNA-seq studies (including a study of brain metastasis (Gonzalez et al., 2022)) have been conducted on fresh tissue specimens, the use of which poses significant practical challenges, and frequently result in inclusion of patients with variable prior treatment exposures, thus, making it challenging to discern underlying biology from potential therapy effects.

Here, we built a multi-omic single-cell atlas of treatment-naïve MBM and extracranial melanoma metastases (ECM). Integrated analyses of this atlas coupled with validation in pre-clinical models and additional patient cohorts identified genomic, adaptive and tumor-microenvironmental features enriched in MBM ecosystems. This study offers important insights into salient MBM biology, identifies putative targets for therapeutic modulation, and a rich reference for the brain metastasis research community.

## RESULTS

### Building a single-cell atlas from small archival specimens

We performed single-cell(sc)/nuclei(sn) transcriptome profiling of 22 treatment-naïve MBM (from 21 patients) and 10 ECM (Figure 1A, Table S1), totaling 114,455 cell transcriptomes (**Methods**), 7,575 matched T cell clonotypes from five patients and spatial transcriptomics using SlideSeqV2 (Stickels et al., 2021) of 16 matched specimens (Table S1). Performing high-quality snRNA-seq from very small frozen tissue specimens, which approximates routinely collected clinical biopsy specimens (~1–10 µg) (**Methods**), enabled access to tissue banks, including specimens collected more than 15 years ago, and permitted profiling on well-curated patient cohorts (Slyper et al., 2020; Wang et al., 2022). To determine whether comparable technical and biological outputs can be gleaned, we split a metastatic lesion (MBM05) immediately following resection for scRNA-seq (MBM05\_sc), and snap-froze the other half, on which we performed snRNA-seq (MBM05\_sn) approximately 12 months later. This comparison yielded several important insights: at approximately the same sequencing saturation and following streamlined quality control filters (**Methods**), the median number of genes detected per cell was 2,137 in MBM05\_sc and 2,504 in MBM05\_sn and remained comparable within the CD45<sup>+</sup> and CD45<sup>-</sup> compartments (3,817 and 1,698 in MBM05\_sc and 3,551 and 1,457 in MBM05\_sn, respectively) (Figures S1A and S1B). We evaluated the expression levels of general “stress” signatures and a recent microglia specific stress signature (Table S1) that are artifactually introduced during fresh tissue processing (Li et al., 2019; Marsh et al., 2022), and found these more strongly expressed in MBM05\_sc compared to MBM05\_sn (Wilcoxon rank-sum test;  $p < 2.2e-16$ ) (Figures 1B and S1C). Nonetheless, following batch correction (Butler et al., 2018), transcriptomes of MBM05\_sc and MBM05\_sn samples clustered by respective cell types in Uniform Manifold Approximation and Projection (UMAP) embedding rather than by method (Figures 1C and 1D), indicating that data from these two modalities yield comparable global biological outputs and recovery of major cell types. Finally, we inferred copy-number alterations (CNAs) from single-cell transcriptomes (Tirosh et al., 2016) (**Methods**) and find that the canonical CNAs (e.g. Chr. 7 amplification) were detected MBM05\_sc and MBM05\_sn (Figures 1E and S1D). Overall, this comparison

demonstrates that biological and inferred genomic insights can be gleaned from snRNA-seq data generated from minute frozen samples.

### The cellular landscape of melanoma brain and extracranial metastases

To enable integrated analyses across samples, where possible, we used a standardized quality-control pipeline (Figure S1E; **Methods**) and tested the performance of different integration methods (Figure S1F; **Methods**). UMAP embedding and clustering of all transcriptomes yielded clusters determined by cell types and independent from profiling method, including transcriptomes from multiple individuals (Figures 2A-D and S2A-D). Malignant cells (cycling and non-cycling) (Figure 2E) were identified by CNA inference (Tirosh et al., 2016) (Figure 2F). We defined nine major cell types, including malignant (n=92,733 cells), myeloid (n=20,449), T and NK lymphocytes (n=18,650), B lymphocyte and plasma (n=3,695), low-quality (n=3,804), stromal (n=3,344), central nervous system (CNS) (n=1,708), endothelial (n=896), and epithelial cells (n=276) (Tables S1 and S2). Re-clustering of main clusters of non-malignant cells, identification of cluster markers, manual annotation, and cross-referencing with external signatures yielded cell type labels for sub-populations (**Methods**, Table S1). At the most granular level, we identified a total of 26 different cell types (or major cell states) (Figures S2A-H, Table S2). Important compositional differences included a higher fraction of “dysfunctional” CD8<sup>+</sup> T cells and myeloid cells in MBM compared to ECM (see below) (Figures S2G and S2H).

### MBM is associated with increased chromosomal instability

Among malignant cells, we find significant heterogeneity of underlying aneuploidy patterns described previously (Akbari et al., 2015) suggestive of underlying chromosomal instability (CIN) (Figure 2F). Analysis of whole-exome sequencing (WES) data (**Methods**) (Fischer et al., 2019), revealed that MBM<sub>WES</sub> (n=21) indeed show a significantly larger fraction of genome altered compared to ECM<sub>WES</sub> (n=23) (Figures 2G and S2I-J). Because CIN is a dynamic process (Bakhoun and Cantley, 2018), we sought to validate our observation in isogenic MBM and ECM cell culture models. Using WM239A-derived 131/4-5B1 cells (hereafter 5B1) and 113/6-4L cells, (hereafter 4L) (Cruz-Munoz et al., 2008); and 12-273 BM and 12-273 LN (Kleffman et al., 2022; de Miera et al., 2012), which were extracted from two metastases in the same patient, we performed high-throughput microscopy and quantified the frequency of micronuclei, a commonly used measure for CIN (Bakhoun et al., 2018). In both matched pairs, MBM-derived cancer cells had a significantly higher rate of micronuclei compared to their respective ECM counterparts (Figure 2H). Importantly, in addition to increased migratory capacity of MBM-derived models *in vitro* (Figure 2I), 5B1 (and additional patient-derived MBM models from our cohort) maintained their brain-metastatic organotropism following intracardiac injection in immunocompromised mice (Figure 2J), while 4L metastasized nearly exclusively to extracranial organs (Figures 2K and 2L, Figure S2K-M). Together, our data suggest that CIN is not only associated with increased migratory capacity as previously described (Bakhoun et al., 2018), but preferentially with brain metastasis. While further work is necessary to elucidate underlying mechanisms, our results are consistent with a recent large-scale clinical genomic study (Nguyen et al., 2021), suggesting that CIN associates with brain metastasis in melanoma and lung cancer.

## A cancer cell signature enriched in MBM

We next determined whether tumor cells from treatment-naïve MBM differ from those from ECM in expression of the previously described AXL-high program, which is associated with increased invasiveness and drug resistance, and the MITF-high program, which is defined by expression of melanocyte-lineage markers including the master regulator MITF (Garraway et al., 2005; Tirosh et al., 2016). MBM scored more strongly for the MITF-high program, while ECM scored stronger for the AXL-high signature in ECM (Figures 3A-B) suggesting that the AXL-high program is not a defining feature of MBM.

To further explore site-specific differences, we performed differential gene expression (DGE) and found that cancer cells from MBM had significantly higher expression of several genes implicated in tumorigenesis and tumor maintenance (e.g. *MET*), as well as neural differentiation, morphogenesis, and adhesion (e.g., *NRG3*, *NCAM1*, *NELL1*, *LRRRC1*, *AKAP12* and *CADPS2*) (Figure 3C; Table S3), while ECM strongly expressed genes involved in epithelial-to-mesenchymal transition (EMT), including *CTNNA2*. Gene set enrichment analysis (**Methods**) revealed several pathways overrepresented in MBM, including oxidative phosphorylation (OxPhos), PI3K (Chen et al., 2014; Fischer et al., 2021; Fukumura et al., 2021), insulin, platelet-derived growth factor receptor beta (PDGFRB), ERBB, and KIT signaling, and genes involved in resistance to tyrosine kinase inhibitors (TKI), while ECM enriched for EMT, cell adhesion molecules, and MTORC1 signaling (Figure 3D; Table S3). Notably, there was significant heterogeneity in the expression of these pathways. For example, while OxPhos overall enriched in MBM, there was a broad distribution of cells within MBM samples, and we also detected cell populations in ECM with strong expression of OxPhos (Figures 3D and S3A). This heterogeneity also manifested on a gene level related to these pathways, when probing for recurrent drivers of variability across metastatic sites (**Methods**; Figures S3B-E). In another approach, we used network-based inferences (VIPER) (Alvarez et al., 2016) to infer activity of proteins and transcription factors (TFs) based on the expression of their downstream regulons (**Methods**). VIPER revealed increased activity of MITF and RELB, a key mediator of the non-canonical NFkB pathway that is activated in CIN<sup>high</sup> cancer cells (Bakhroum et al., 2018), in MBM (Figure 3E; Table S3). The top differentially regulated proteins in ECM cells included CREB1; BACH1, a regulator of several genes involved in metabolism and metastasis (Kaur et al., 2021); and SOX4, a key EMT TF (Cheng et al., 2017). Using an orthogonal method for TF activity inference (SCENIC) (Aibar et al., 2017) (**Methods**) revealed consistently top-enriched TFs, including MITF and RELB in MBM, and SOX4 in ECM (Figures S3F and S3G; Table S3).

Next, we asked whether the MBM cancer cell intrinsic signature identified in our single-cell analysis was conserved in other human datasets. We scored an external cohort of MBM (n=88) and ECM (n=50) tumors profiled using bulk-RNA-seq (MBM<sub>bulk</sub> and ECM<sub>bulk</sub>, respectively) (Fischer et al., 2019) (**Methods**) and found that our MBM signature significantly enriched in MBM<sub>bulk</sub> compared to ECM<sub>bulk</sub> (Figure 3F).

Furthermore, using the 5B1 and 4L cell lines, we generated MBM and ECM *in vivo*, and performed RNA-seq of 17 metastases (**Methods**). Transcriptomes from MBM clustered

distinctly from ECM and scored significantly stronger for the single-cell-derived MBM signature (Figures 3G and S4A). Additionally, for two patients (MBM03 and MBM05) from our single-cell profiling cohort, we established cell lines and performed bulk RNA-seq which correlated strongly with our *in-situ* profiling (Spearman's Rho of 0.83 and 0.80 in MBM03\_sc and MBM05\_sc, respectively). After intracardiac injections we found that 45% and 90% of mice injected with MBM03 and MBM05 developed macroscopic MBM, respectively (Figures S2L and S2M). This high rate of MBM development is notable in light of recent work showing that brain metastasis is a highly inefficient process (Kleffman et al., 2022), and suggests that MBM-derived models maintain their brain-metastatic organotropism. We reasoned that MBM and ECM may differ in their chromatin accessibility. ATAC-seq (and matched RNA-seq) of four MBM- and four ECM-derived patient cell lines (including two matched pairs) indeed showed increased chromatin-accessibility of lineage genes (e.g., *TYRP1*, *DCT*) in MBM, and EMT-related genes (e.g., *COL5A1*, *COL4A5*, *ITGA3/4*) in ECM (Figures S4B-D; Table S3). Lastly, we identified enriched motifs and binding TFs using foot printing (**Methods**), which revealed five TFs enriched in MBM that were also shared with SCENIC and VIPER analyses, all of which are involved in melanocytic lineage and neuronal differentiation (MITF, SOX10, NPAS2, ZNF317 and CUX1) (Figures S4C-E; Table S3).

Overall, these results suggest that cancer cells show significant transcriptional heterogeneity and variation between MBM and ECM. Our MBM signature, which includes several neuronal genes, enriches in both human and murine brain metastasis profiling, and may be involved in promoting brain-metastatic organotropism.

### Neuronal-like differentiation is a distinct feature of MBM

We next sought to discover coherently regulated programs of tumor heterogeneity among MBM and ECM cancer cells. For this purpose, we used KINOMO, a semi-supervised non-negative matrix factorization (NMF) approach (**Methods**) (Tagore et al., 2022) to first discover programs (factors) of heterogeneity within each individual patient, followed by identification of metaprograms (MPs) defining shared programs across patients (**Methods**, Table S4). Across MBM and ECM, we identified 87 factors that converged on seven MPs (Figures 3H-K, Table S4). Five MPs contained factors from MBM and ECM samples, while MP7 and MP4 were specific to MBM and ECM, respectively (Figures 3J and 3K). Shared MPs represented *MYC* targets; mTORC1 signaling and antigen presentation (MP1); lineage-defining genes and antigen presentation (MP2); cell-cycle and mitosis genes (MP3); apoptosis and stress response (MP5); and NOTCH signaling and *IL2/STAT5* signaling (MP6) (Figure 3I). MP4 contained meta-genes of glucose metabolism, hypoxia response, unfolded protein response, matrix proteins, negative regulators of cell cycle progression, and TNF-alpha signaling (Figures S4F).

MP7 was exclusively representative of MBM (Figures 3J and 3K). While preserving meta-genes of melanocytic lineage, MP7 contained a wide range of genes involved in neuronal development and differentiation (e.g., *NGFR*, *NLGN3*, *NRXN*), and synapse function and formation (e.g., *SNCA*, *SYT11*, *GPHN*) (Figure 3I), and concordantly scored significantly for signatures of neuronal differentiation and lineage (Figures S4G). Notably, neuronal

growth factor receptor (*NGFR*) is associated with melanoma plasticity and drug resistance (Fallahi-Sichani et al., 2017; Landsberg et al., 2012; Liu et al., 2021).

Together, multiple analyses suggest that neuronal-like differentiation may represent a feature of cancer cells in MBM. To further interrogate this observation, we performed an integrated analysis of our single-cell data, public RNA-seq data of MBM and ECM, and proteomics data from patient-derived MBM/ECM short-term cell cultures (Kleffman et al., 2022) (**Methods**). Across datasets, platforms, and analytes, six genes/proteins were consistently enriched/upregulated in MBM compared to ECM, including *NCAM1*, *NELL1*, and *LRRCL1*, involved in neuronal adhesion and differentiation, and *ST6GALNAC3*, *TBXAS1*, and *WNK4* (Figure 3L; Table S3). We assembled 42 additional melanoma (cutaneous and uveal) patient samples, including 19 from MBM and 23 from ECM, and performed multiplexed immunofluorescence staining, including markers for neural adhesion (NCAM1), melanocyte lineage (SOX10 and HMB-45), and lymphoid and myeloid immune infiltrates (CD8, CD68 and CD138) (**Methods**). NCAM1 was strongly enriched in cancer cells in MBM (Figure 3M), but only sporadically found in non-malignant cells in ECM (Figure 3N) (likely representing NK cells that express NCAM1, which is also known as CD56). Single-cell segmentation and quantification revealed a significantly higher fraction of cancer cells expressing NCAM1 and higher NCAM1 intensity in melanoma cells from MBM (Figures 3O and P and S4H-K). Together, these data suggest that malignant cells from MBM express a neuronal-like phenotype, which may be important for brain-metastatic organotropism and represent a potential opportunity for therapeutic development.

### Compositional and phenotypic differences among monocyte-derived macrophages in MBM

We recovered a total of 17,562 myeloid cell transcriptomes, including 12,579 from MBM and 4,983 from ECM, encompassing monocyte-derived macrophages (MDM; n=13,049), monocytes (n=2,119), microglia (n=1,658), three types of dendritic cells (DCs; n=663; including cDC1, cDC2, and DC3), and mast cells (n=73) (Figures 4A-D and S5A-H, Table S2). Diffusion component (D.C.) analysis (**Methods**) revealed two major clusters (MDM-c1 and MDM-c2) among macrophages arising from a classical monocyte pool (Figures 4A and 4E), that expressed a gradient of putative pro-tumorigenic MDM features (e.g., *CD163L1*, *SELENOP*, *F13A1*, *DAB2*, *SIGLECI*) (Adamson et al., 2016; Cassetta et al., 2019; Gonzalez-Dominguez et al., 2015; Solinas et al., 2010; Wang et al., 2021) (Figures 4D, Table S5). Importantly, recovered MDMs were not clearly representative of the previously reported M1-like/M2-like phenotypes (Figures S5G and S5H) (Yunna et al., 2020). Among the most strongly differentially expressed genes (DEGs) between the two MDM clusters was *FTL* (in MDM-c2); we thus labeled MDM-c2 as *FTL*<sup>+</sup> MDM, and the remainder MDM-c1. *FTL* encodes for the light chain of ferritin (which is further composed of *FTH1* encoding for the heavy chain), an intracellular iron storage protein complex critical in regulating iron homeostasis and sequestration (Winn et al., 2020) that is suppressed in pro-tumorigenic macrophages (Recalcati et al., 2010). Compared to MDM-c1, *FTL*<sup>+</sup> MDM showed increased expression of mitochondrial genes encoding proteins of the electron transport chain (*MT-CO1/2/3*, *MT-CYB*, mitochondrial NADH dehydrogenase subunits, and ATP-synthase subunits), reactive oxygen species detoxification



(*GPX4*), interferon gamma response genes, and antigen presentation genes, among others, while MDM-c1 showed higher expression of efferocytosis receptors (*AXL*, *MERTK*) and several marker genes associated with a pro-tumorigenic macrophage program, such as *CD163* (Skytthe et al., 2020) (Figures S5I and 5J, Table S5). To glean insights into the potential temporal relationship among MDMs, we used RNA velocity (La Manno et al., 2018) and scVelo (Bergen et al., 2020) (**Methods**), and identified FTL<sup>+</sup> MDM cells as an early cell state from which MDM-c1 may arise (Figures 4F-G and S5K-L). Comparison of macrophages among metastatic sites revealed that those from MBM showed lower expression of antigen presentation genes (*HLA-A/B/C/E*, *B2M*, *CIITA*, *CD74*, *HLA-DP/DQ/DR*) and matrix proteases (*MMP2/14*) and higher expression of *CD163*, *TLR2*, *MERTK*, *AXL*, and *IL2RA* (Figures 4H, Table S5). Differences between metastatic sites were also preserved when stratifying by FTL<sup>+</sup> MDM and MDM-c1 (Figures S5M and S5N), suggesting that macrophages in the MBM ecosystem showed generally higher expression of genes associated with a pro-tumorigenic phenotype. While our cell fraction analysis suggested that myeloid cells, and in particular MDM-c1 were more frequent in MBM ( $p=0.061$ ) (Figure S2H) we next analyzed the frequency in an independent patient cohort of 19 MBM and 23 ECM and found that MBM indeed showed a significantly higher fraction of macrophages (Figure 4I). Furthermore, macrophages in MBM had a higher expression intensity of CD163 protein in MBM, corroborating our transcriptional findings (Figure 4J). Together, these data suggest the presence of two major MDM populations, including FTL<sup>+</sup> MDM, which express genes associated with anti-tumor immunity, and MDM-c1, which express several genes linked with pro-tumorigenic polarization. MBM enriched for tumor-permissive MDM, defining an important feature of their ecosystem.

### Inflammatory *SPP1*<sup>+</sup> microglia

Recovery and interpretation of microglia (MG) single-cell transcriptomes from human tissues is challenging as these cells are particularly sensitive to fresh-tissue processing (Marsh et al., 2022) and strongly express artifactual stress signatures (Figure S1C). Nonetheless, we recovered key MG marker genes (e.g., *P2RY12*, *SLC2A5*, *NAV3*) in both snRNA-seq and scRNA-seq (Figure 4D, Table S2). Using D.C. analysis, we identified two major clusters (MG1 and MG2), wherein MG1 was characterized by significantly higher expression of *SPP1* (encoding for osteopontin; an indicator for a pro-tumor phenotype (Wei et al., 2019)), chemokines (e.g., *CCL2/3/4*), and inflammatory cytokines (e.g., *TNFA* and *IL1B*), irrespective of profiling modality (Figures 4K, 4L, S5O and S5P). Accordingly, MG1 enriched strongly for inflammatory and activation pathways, among others (Table S5). Together, this data suggests that profiling MG from frozen tissues is feasible, avoids cell-specific artifacts, and in MBM identifies a *SPP1*<sup>+</sup> sub-population with an activated phenotype.

### Integrated analyses of T cells define common axis of variation in MBM and ECM

The T/NK cell compartment comprised 17,149 transcriptomes with six major clusters, including NK cells ( $n=759$ ), T-regulatory cells (T<sub>regs</sub>;  $n=2,284$ ), conventional CD4<sup>+</sup> T cells ( $n=3,657$ ), T follicular helper (T<sub>FH</sub>)-like cells ( $n=1,300$ ), and CD8<sup>+</sup> T cells ( $n=9,149$ ) (Figures 5A-D and S6A-E; Table S2). Among CD8<sup>+</sup> T cells, we distinguished two major populations (Figure 5D): CD8<sup>+</sup> T cells, with high expression of *TOX* ( $n=5,474$ )

(TOX<sup>+</sup>CD8<sup>+</sup> T cells), a marker associated with and required for T cell dysfunction (Khan et al., 2019), and effector memory T cells characterized by expression of *TCF7* (n=3,675) (TCF7<sup>+</sup>CD8<sup>+</sup> T cells), a marker of progenitor-like function (Zhao et al., 2021). TOX<sup>+</sup>CD8<sup>+</sup> T cells showed strong expression of multiple immune checkpoints, effector genes, and chemokines (Figure 5D). TCF7<sup>+</sup>CD8<sup>+</sup> T cells expressed markers consistent with a naïve or memory phenotype as well as effector molecules, and had low expression of immune checkpoints, such as *PDCD1*/PD1 and *HAVCR2*/TIM3 (Figure 5D). Because we noted differential capture of important T cell hallmark genes in scRNA-seq and snRNA-seq data (Figure S6F, Table S5), we validated our annotation with an orthogonal approach using ProjecTILs (Andreata et al., 2021) (**Methods**), an automated method for reference-based analysis of T cell states. This confirmed our manual annotation, yielding robust identification of T cell subtypes irrespective of the queried sequencing technology (Figures S6G-M).

Compared to ECM, MBM showed a higher abundance of TOX<sup>+</sup>CD8<sup>+</sup> T cells and CXCL13<sup>+</sup> CD4<sup>+</sup> T<sub>FH</sub>-like cells (Cohen et al., 2022) (Figures 5D and S2H). While TOX<sup>+</sup>CD8<sup>+</sup> tumor-infiltrating T cells have previously been defined as “exhausted”, recent work shows that this is a dynamic pool of cells with proliferative capacity (Li et al., 2019) and potential anti-tumor activity (Ho et al., 2022). Furthermore, T<sub>FH</sub>-like cells resemble T helper tumor cells, which were recently defined in non-small cell lung cancer, breast cancer, and melanoma (Cohen et al., 2022), and may have an important role in tumor-antigen recognition and response to PD-1 blockade. We next aimed to delineate differences among T cells from MBM and ECM. Among others, we found significantly lower expression of *PDCD1*/PD1, *LAG3*, *TIGIT*, and *HAVCR2*/TIM-3 in TOX<sup>+</sup>CD8<sup>+</sup> T cells from MBM compared to ECM (Figure 5E; Table S5).

### T cell expansion and differentiation in MBM

We generated matched TCR-seq data for 7,575 cells from five MBM patients, which enabled annotation of clonally expanded and non-expanded cells in T cell compartments, including mucosal associated invariant T (MAIT, n=1,396) cells and invariant NKT (iNKT, n=59) cells that have limited anti-tumor activity (Godfrey et al., 2019) (Figure S7A). CD4<sup>+</sup> T cells accounted for 1,893 unique clonotypes with a clone size up to 157, wherein clonal expansion was most frequent among T<sub>regs</sub> and T<sub>FH</sub>-like cells (Figure S7B). Among CD8<sup>+</sup> T cells, we identified 1,470 unique clonotypes, with a clone size of up to 336 (Figure S7B).

D.C. analysis integrating clonality and gene expression showed clonal expansion varied along the TCF7<sup>+</sup>CD8<sup>+</sup> to TOX<sup>+</sup>CD8<sup>+</sup> T cell trajectory (Figures 5F and 5G), which is consistent with prior work in ECM and indicates progressive loss of progenitor function (Li et al., 2019) and increased differentiation (Azizi et al., 2018) (Figure S7B). Within individual samples, shared clonotypes were captured among TCF7<sup>+</sup> and TOX<sup>+</sup>CD8<sup>+</sup> T cell pools, directly supporting the notion that detected clonotypes arise from a specific progenitor pool (Figure S7C). Comparison of clonally expanded (n=2,121) and non-expanded (n=901) CD8<sup>+</sup> T cells showed significant enrichment of *IL7R*, *LTB*, *SELL*, and *TCF7* in non-expanded T cells, and increased expression of genes associated with cytokine production, cytotoxicity (*IFNG*, *GZMB*), co-stimulation (*TNFRSF94*-1BB),

immune checkpoints (e.g., *CTLA4*, *HAVCR2*), and terminal differentiation (*MIR155UG*) (Dudda et al., 2013) in expanded T cells (Figure 5H; Table S5). Overall, this analysis suggests that while the differentiation trajectory of T cells in MBM is consistent with that described in ECM, important differences may exist with respect to expression of immune checkpoints.

### Intra-tumoral plasma cell differentiation and lymphoid aggregates

We recovered 3,175 single-cell transcriptomes spanning a functional differentiation trajectory from naïve (n=166 cells; *LAIR1*, *TCL1A*) to activated B cells (n=1,032; *BANK1*, *BLK*) and plasma cells (n=1,977; *SDCI*, *CD38*, *PRDMI*) across 28 patients (Figures 5I, S7D and S7E; Table S2). We reconstructed immunoglobulins by determining mRNA co-expression of the variable heavy (IGHV) and light (IGLV) chains and isotypes on a per cell basis (Melms et al., 2021) and identified recurrent combinations (**Methods**; Table S5). Notably, we recovered the entire differentiation arch from naïve to activated B cells to plasma cells, and in some specimens, cells along this developmental trajectory shared the same variable chain expression (Figure S7F), suggesting that differentiation of B to plasma cells occurs intratumorally. Next, we analyzed the spatial distribution of plasma cells (IF staining of CD138) in 42 additional melanoma samples (19 MBM and 23 ECM, as above) and found that plasma cells were present in a spatially restricted manner in both MBM and ECM (Figure 5J), and surprisingly, plasma cell aggregates were significantly more abundant in MBM (Figure 5K). Recent studies in kidney cancer showed that intra-tumoral plasma cell differentiation occurs in tertiary lymphoid structure (TLS) (Meylan et al., 2022), and that lymphoid aggregative or TLS are predictive of response to anti-PD-L1 therapy in non-small cell lung cancer (Patil et al., 2022). Thus, our results raise the possibility that intra-tumoral B to plasma cell differentiation occurs in a spatially restricted manner and indicates the presence of a subset of metastatic melanomas that are more likely to respond to immunotherapies.

### Spatial single-cell transcriptomics identify distinct cell clusters and geographically variable transcriptional features

We performed spatial transcriptomics using SlideSeqV2 (Stickels et al., 2021) of 16 tissue sections (11 MBM and five ECM) from 11 matching tumors also profiled by snRNA-seq (Figure 6A, Tables S1 and S6) (**Methods**). Due to the fine resolution, the same fixed pixel may represent transcripts captured from different cells. Thus, we first leveraged the matched snRNA-seq/SlideSeqV2 data and used *robust cell type decomposition* (RCTD) (Cable et al., 2021) to deconvolve cell mixtures and assign the spatial configuration of discrete cell types (Figure 6B) (**Methods**). This approach resulted in recovery of fractions of malignant and non-malignant cells that was highly correlated with that of matching snRNA-seq cell type composition (e.g., for malignant cells,  $R^2=0.75$ ;  $p=1.6e^{-5}$ ) (Figures 6C, Table S6), and showed expected compositional variability with high technical reproducibility (Figure S7G). To discover spatially regulated genes, we determined genes with the highest spatial auto-correlation by Moran's  $I$  ( $>0.05$  and  $p<0.001$ ) for each individual tissue slide (**Methods**, Table S6) and integrated identified spatial expression with corresponding cell type assignments (Figure 6B). Orthogonally, we used *C-SIDE* (Cable et al., 2022), to identify spatially variable gene expression, which yielded highly

overlapping results (Table S6). We then categorized these genes by biological function, such as metabolism (oxidative phosphorylation, glycolysis), antigen presentation, interferon and chemokine responses, hemorrhagic regions, and immunoglobulins (indicating clustering of plasma cells) (Table S6) and determined their spatial expression across the entire cohort and assessed cell type co-localization (Table S6). Consistent with analyses and imaging of B/plasma cells (Figure 5K), we found strongly spatially restricted clusters of lymphoid aggregates that are dominated by plasma cells (with enrichment for specific variable *IGHV/IGHL* combinations) in multiple specimens (Figures 6D-G). Another spatial pattern involved large regional expression of cancer cell intrinsic antigen presentation genes that was reciprocal to expression of *TIMPI*, a matrix metalloproteinase with a potential role in promoting melanoma immune evasion (Zurac et al., 2016) (Figure 6H). Notably, while antigen presentation genes were more broadly expressed, there was a spatially restricted expression of type I interferon response (e.g., *ISG15*, *IFI27*, *CXCL9*) in specific areas, indicating that antigen presentation was not sufficient to promote anti-tumor responses (Figure 6H). Furthermore, cancer cells show spatially dichotomous expression of oxidative phosphorylation and glycolysis metabolic pathways (Figure 6I). While our snRNA-seq analyses indicated heterogeneous expression of oxidative phosphorylation (Figure 3D), these findings suggest that this variability was also spatially defined. Together, these analyses provide important examples for spatial contexts of cell states, responses, and interactions in a preserved tumor architecture.

## DISCUSSION

Despite significant therapeutic advances in the treatment of metastatic cancers, MBM remains a challenging problem. Here, we begin addressing this challenge by interrogating MBM and ECM ecosystems using multi-modal single-cell transcriptomics coupled with functional approaches and human and murine models. Focusing on treatment-naïve MBM enabled us to examine the salient biology of this metastatic site, unaffected by the potential effects of prior treatments. Among cancer cells, increased chromosomal instability (CIN) was associated with MBM. In matched cell line cultures (MBM and ECM from the same individual) we demonstrate that CIN was enriched in MBM-derived models, which maintain an increased invasive capacity and brain-metastatic organotropism *in vivo*. While CIN is associated with metastasis (Bakhroum et al., 2018), recent large-scale genomics studies suggest enrichment of CIN in brain metastases in melanoma and non-small cell lung cancer (NSCLC) (Nguyen et al., 2021). CIN and associated cellular adaptation may represent promising therapeutic vulnerabilities, that are particularly relevant in MBM. Furthermore, using KINOMO, we detect a *relative* increase (beyond potential underlying neural crest lineage gene expression (Opdecamp et al., 1997)) in neuronal differentiation genes exclusively in MBM. Phenotypic plasticity is a hallmark of melanoma biology, and can be associated with invasive melanoma behavior and drug resistance (Fallahi-Sichani et al., 2017; Hoek et al., 2008; Konieczkowski et al., 2014; Liu et al., 2021; Sun et al., 2019; Tirosh et al., 2016; Tsoi et al., 2018). Thus, the interplay of the brain metastatic niche and the neuronal-like cell state implicated by our study, which is further supported by recent xenograft work (Wingrove et al., 2019), may be important for niche-specific metastasis and drug resistance and warrants further mechanistic investigation. Furthermore, our integrated

snRNA-seq and spatial transcriptomics analyses reveal increased yet variable expression of metabolic pathways in MBM (e.g. OxPhos), and spatially restricted cancer cell expression of type I interferon responses (indicating ensuing T cell responses), despite broad expression of antigen presentation. This suggests that antigen presentation alone is not sufficient to promote anti-tumor immunity, as recent work in patient models suggests (Ho et al., 2022), and may indicate areas prone to cancer immune evasion, possibly through expression of *TIMPI* (Zurac et al., 2016).

Collectively, this study represents the most comprehensive single-cell profiling effort of human brain metastases to date, providing foundational insights into the molecular underpinnings of cancer cells and cells in the tumor microenvironment that are distinct or shared in different metastatic sites. By focusing on treatment-naïve MBM, this study sets the stage for further mechanistic studies and serves as an important reference for understanding how different perturbations (e.g., various systemic or local treatments) influence brain metastasis ecosystems. Building on this work, future experimental and analytical innovations that capture additional layers of single-cell biology and resolve the role of genomic, cellular, and transcriptional features (e.g., CIN, neuronal-like differentiation, and varying immune checkpoint expression) enriched in MBM may facilitate therapeutic discovery.

### Limitations of the study

Although we validated key observations gleaned from single-cell analyses, our study did not include matching MBM and ECM collected from the same patients which would allow us to directly study differences between metastatic niches in genomically similar specimens. Concurrent resection or biopsy of an MBM and ECM (and possibly primary tumor) from the same patient is rarely indicated, however, the ability to perform single-cell profiling from snap-frozen tissue removes important barriers to enable multi-institutional efforts to collect and analyze larger sample sizes in such instances. This may also enable investigation of differences among matched primary tumors and multiple different metastatic sites. Although we identify important features associated with MBM (e.g., CIN, neuronal-like differentiation, distinct features of the microenvironment), further work is necessary to determine their mechanistic interactions and roles along the metastatic cascade. Lastly, expanded studies including other common cancers associated with brain metastasis (e.g., lung cancer) will further inform whether these mechanisms are relevant in other lineages.

## STAR METHODS

### Resource availability

**Lead contact**—Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Benjamin Izar (bi2175@cumc.columbia.edu).

**Materials availability**—Presented materials are made available upon reasonable request to the lead contact.

**Data and code availability**—Processed data generated using sc/snRNA-seq, TCR-seq, SlideSeqV2, RNA-seq and ATAC-seq have been deposited at GEO and are publicly available. Accession numbers are listed in the key resources table. Raw data derived from human samples will be made available through dbGaP and are listed in the key resources table. All original code has been deposited at GitHub and is publicly available as of the date of publication. The repository is listed in the key resources table. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

## Experimental models and subject details

**Patient tissue collections**—Frozen tissue specimens were collected under institutional review board (IRB) approved protocols at New York Presbyterian Hospital/Columbia University Medical Center (from 2001 to 2017), Dana Farber Cancer Institute (from 2017 to 2018), Georgetown University (2014 to 2015) and UCLA (from 2012 to 2013). Treatment-naïve melanoma brain metastasis specimens were collected during surgery and banked at  $-80^{\circ}\text{C}$  after allocation by qualified pathologists according to institutional guidelines. Extracranial melanoma metastases were collected as biopsies and banked at  $-80^{\circ}\text{C}$  after allocation by qualified pathologists according to institutional guidelines. Formalin-Fixed Paraffin-Embedded (FFPE) tissue specimens of melanoma brain and liver metastases were collected under IRB approved protocols at Universität Tübingen (2007-2020). Samples were selected based on treatment status and metastatic site and allocated to respective groups. Sample size was based on prior studies of similar scope. All available patient information is summarized in Table S1. All procedures performed on patient samples were in accordance with the ethical standards of the IRB and the Helsinki Declaration and its later amendments.

**Cell lines and cell culture**—Human melanoma cell lines were derived under IRB approved protocols. Cell line 2686 was originally derived and provided by MD Anderson Cancer Center (IRB 2004-0069). Cell line MaMel-134 was kindly provided by UK-Essen. MBM03 and MBM05 cell lines were derived from surgical resections of melanoma brain metastases at Dana Farber Cancer Institute (IRB 10-417). 113/6-4L (“4L”) cells and 131/4-5B1 (“5B1”) cells were originally generated by Cruz-Munoz *et al.* (Cruz-Munoz *et al.*, 2008). 12-273-BM and LN were originally derived in the laboratory of Iman Osman at NYULH Melanoma Program (Kleffman *et al.*, 2022; de Miera *et al.*, 2012).

2686, MaMel-134 and MBM cell lines were cultured in melanoma growth medium (RPMI 1640 supplemented with 10% heat-inactivated fetal bovine serum (FBS), GlutaMax, 10 mM HEPES, 10 mg/L insulin, 5.5 mg/L transferrin, and 6.7  $\mu\text{g/L}$  sodium selenite, and 55  $\mu\text{M}$  2-mercaptoethanol). 4L/5B1 and 12-273 cell lines were cultured in DMEM supplemented with 10% FBS, and 1% (v/v) penicillin/streptomycin. For 12-273 the medium further contained 1% non-essential amino acids. All cell lines were maintained in a 5%  $\text{CO}_2$  incubator at  $37^{\circ}\text{C}$ . Cell lines were routinely tested to exclude Mycoplasma contamination using Plasmotest (InvivoGen).

**Animal Studies**—All mice experiments performed at Columbia university were following the Institutional animal care and use committee (IACUC) protocol #AC-AABE6570 using

6–8 week-old, male NOD.Cg-Prkdc<sup>scid</sup> Il2rg<sup>tm1Wjl</sup>/SzJ mice (NSG, Jackson labs, #5557). All mice experiments at NYULH were performed in compliance with a referenced protocol (IA16-00051) approved by the NYULH IACUC using 4–6 week-old NSG mice (male and female, Jackson labs, #5557) All mice were maintained under standard pathogen free conditions.

## Method details

**Fresh tissue specimen processing**—After allocation of specimens by pathology, samples were immediately placed in ice cold RPMI 1640 (Thermo Fisher, #21875034) without supplements and transported to the laboratory space for processing. Samples were cut to 0.5-1 mm<sup>3</sup> cubes and transferred to a 50 mL Falcon tube (Corning) filled with ice cold RPMI. The tissue was collected by centrifugation. All centrifugation steps were conducted in a swinging bucket rotor for 5 minutes at 400 x g at 4°C. The tissue was resuspended in 4.7 mL RPMI (prewarmed to 37°C) supplemented with human tumor dissociation enzymes (Miltenyi Human Tumor Dissociation Kit, #30-095-929) prepared according to manufacturer's guidelines. The sample was digested in a 37°C water bath and swirled every 2 minutes. After 5- and 10-minutes incubation the sample was mechanically dissociated using pipettes with decreasing orifice size. After the final dissociation step the sample was filtered through a pre-wetted 70 µm cell strainer (Corning) into a fresh 50 mL Falcon tube. The cells were collected by centrifugation and the supernatant was carefully discarded. The cell pellet was resuspended in 3 mL ACK buffer (Thermo Fisher; #A1049201) and incubated for 1 minute at room temperature to lyse red blood cells. After dilution with 30 mL ice cold sorting buffer (2% FBS/1 mM EDTA in PBS) the cells were collected by centrifugation. After this step the cells were resuspended in ice cold PBS, filtered through a 40 µm cell strainer attached to a FACS tube, counted and 1e<sup>6</sup> cells were processed for fluorescence-activated cell sorting.

**Fluorescence-activated cell sorting**—To sort viable immune and non-immune cells the cell suspension was stained with Zombie NIR viability dye (Biolegend, San Diego, CA; #423106) in PBS (1:500) for 10 minutes at room temperature in the dark. The reaction was terminated with sorting buffer and the cells were collected by centrifugation. After collection, the cells were stained with the following antibodies for 15 minutes on ice in the dark (all Biolegend): Human TruStain FcX (#422302), Pacific-Blue-aCD45 (#304022), PE-Dazzle594-aCD3 (#300450), PE-CY7-aCD66b (#305116), APC-aCD15 (#301908). After staining, cells were washed twice with sorting buffer and filtered through a 40 µm cell strainer attached to a FACS tube and sorted into ice-cold sorting buffer using a FACS Aria II (BD Biosciences). For each specimen we sorted 12–15x10<sup>3</sup> viable cells for the following populations: 1.) CD45<sup>-</sup> cells including tumor and stromal cells. 2.) CD45<sup>+</sup>/CD66b<sup>-</sup> cells including all immune cells but CD66b<sup>+</sup> granulocytes which are reported to interfere with scRNA-seq using 10x genomics. After sorting, the cells were placed on ice and processed immediately for scRNA-sequencing.

**Single cell RNA library preparation and TCR dial out**—Sorted single cell suspensions were centrifuged and washed two times in PBS with 0.05% RNase-free BSA (Thermo Fisher, #AM2616). After the final wash the supernatant was carefully

removed to leave a final volume of 31  $\mu$ L. Each sorted single cell suspension was loaded in separate channels on a Chromium controller using Chromium Single Cell V(D)J Reagents (10x genomics, Pleasanton, CA; #1000006) for 5' RNA capture. Gene expression libraries and TCR expression libraries (from CD45<sup>+</sup> channels) were generated according to manufacturer's instructions using the Chromium Single Cell 5' Library construction kit (#1000020) and Chromium Single Cell V(D)J Enrichment Kit for human T cells (#1000005) respectively.

**Sequencing of single-cell RNA libraries**—Final libraries were quantified using TapeStation D5000 and D1000 reagents on a 2200 TapeStation system (Agilent). Each donor sample was mixed individually based on the molarity of the libraries (45% for each gene expression library, 10% TCR library) and sequenced with 2x150 bp paired-end sequencing on an Illumina HiSeq 4000.

**Frozen tissue processing**—Frozen archival tissue specimens were embedded in optimal cutting temperature (OCT) compound (Tissue-Tek, Sankura) on dry ice. Tissue sections were collected in pre-cooled 5 mL tubes (Eppendorf, Hamburg, Germany) and stored on dry ice until processing. Single nuclei were isolated using salt tris (ST) buffer (146 mM NaCl, 10 mM Tris-HCL pH7.5, 1 mM CaCl<sub>2</sub>, and 21 mM MgCl<sub>2</sub> in ultrapure water) with Tween-20 as previously described (Slyper et al., 2020) with the following modifications. All centrifugation steps were carried out in a swinging bucket centrifuge cooled to 4°C. Tubes were placed on wet ice and left to equilibrate for 30 seconds. Then 4 mL of ice-cold PBS without calcium or magnesium (Thermo Fisher) was added and the tube was inverted until all OCT had dissolved. The tissue pellet was then collected by centrifugation at 300 x g for 2 minutes. The supernatant was discarded, and the tissue pellet was resuspended in 1 mL salt-tris buffer with 0.03% Tween-20 (Sigma Aldrich, p-7949) and 0.1% BSA (New England Biolabs, B9000S) (TST buffer) supplemented with 40 U/mL RNase OUT (Thermo Fisher). The suspension was vigorously pipetted and then incubated for 5 minutes on ice. After 5 minutes the tissue suspension was again pipetted and the reaction was quenched by adding 4 mL of ST buffer supplemented with 40 U/mL RNaseOUT. The suspension was filtered through a 70  $\mu$ m nylon mesh filter (Fisher Scientific) into a precooled 50 mL conical on ice. The filter was washed with additional 5 mL ST buffer and the tube was centrifuged at 500 x g for 5 minutes. The supernatant was discarded, and the nuclei were resuspended in 100–400  $\mu$ L ST buffer without RNase inhibitor and filtered through a 40  $\mu$ m mesh filter attached to a FACS tube (Fisher Scientific). A 5  $\mu$ L aliquot of the filtered nuclei suspension was mixed at equal volume with PBS with a final concentration of 50  $\mu$ g/mL Hoechst 33342 (Thermo Fisher) to stain nuclear DNA and the nuclei concentration was determined using a fluorescent microscope (EVOS FL, Thermo Fisher) and disposable Neubauer counting chambers (Bulldog Bio, Inc. Portsmouth, NH).

**Single nuclei RNA library preparation**—14,000 nuclei were diluted in ST buffer without RNase inhibitor and loaded on a Chromium controller using Chromium Single Cell 3' reagents v3.0 (10x genomics, Pleasanton, CA; #1000006) for 3' RNA capture. After reverse transcription and cleanup, gene expression libraries were generated according to manufacturer instructions using the Chromium Single Cell 3' Library construction kit. To



increase cDNA yield from nuclei the initial amplification was carried out with 1 additional cycle compared to the recommendations for single cells.

**Sequencing of single nuclei RNA libraries**—Gene expression libraries were quantified as described above and equimolarly mixed before sequencing. Libraries were sequenced targeting a minimum of  $2e^8$  reads per sample using paired-end sequencing with 2x150 bp on a NovaSeq 6000 with S4 flow cell (Illumina, San Diego, CA).

**Generation of cell lines from melanoma brain metastases**—Cell lines from MBM03 and MBM05 were established by placing excess melanoma tissue pieces (~0.5 mm<sup>3</sup>) not used for scRNA-sequencing into individual wells of a 24 well cell culture plate filled with melanoma growth medium (RPMI 1640 medium supplemented with 10% heat-inactivated fetal bovine serum, GlutaMax, 10 mM HEPES, 10 mg/L insulin, 5.5 mg/L transferrin, and 6.7 µg/L sodium selenite, and 55 µM 2-mercaptoethanol (all from Thermo Fisher Scientific)). For the first two weeks the medium was supplemented with 100 IU/mL penicillin and 100 µg/mL streptomycin (Thermo). Tissue pieces were cultured at 37°C and 5% CO<sub>2</sub>, and ambient O<sub>2</sub> in a humidified incubator until adherent melanoma cells had crawled out of the tissue. The tissue was then removed, and the cells were allowed to grow to 80% confluence before pooling of individual wells for further expansion. Cells were then tested for mycoplasma contamination using Plasmotest (InvivoGen) and cryopreserved in Bambanker (Bulldog Bio) and stored in liquid nitrogen vapor phase until future use.

**In vitro culture of melanoma cell lines**—Human MBM, 2686 and MaMel-134 melanoma cell lines were grown in melanoma growth media (see above). MBM03 and MBM05 were generated from MBM as described above. ECM melanoma cell lines 2686 and MaMel-134 were previously derived and provided by MD Anderson Cancer Center and UK-Essen, respectively. 113/6-4L (“4L”) cells and 131/4-5B1 (“5B1”) cells were originally generated by Cruz-Munoz *et al.* (Cruz-Munoz et al., 2008). 12-273-BM and LN were originally derived in the laboratory of Iman Osman at the NYULH Melanoma Program (Kleffman et al., 2022; de Miera et al., 2012). Culture media for these cell lines is comprised of DMEM, 10% fetal bovine serum, and 1% (v/v) penicillin/streptomycin. All cell lines were grown under adherent conditions at 37°C and 5% CO<sub>2</sub> in a humidified incubator.

**Migration assays**—Falcon FluoroBlok 96-well HTS plates were used to evaluate *in vitro* cell migration. Cells were resuspended at a density of 10,000 cells per well in 1%FBS supplemented growth media and migrated towards 10% FBS growth media. The transwell migration plate was incubated for 22 h at 37°C in 5% CO<sub>2</sub>, followed by staining with a 4 µg/mL solution of Calcein AM in Hanks Balanced Salt Solution for 1 h at 37°C. Fluorescence intensity was measured from the bottom of the migration plate using Synergy H1 Hybrid Reader.

**Micronuclei quantification**—Cells were seeded at a density of 1,000 cells per well in a 96 well plate, fixed, and stained after 24 h as previously described (Melms et al., 2020). In brief, cells were fixed with 4% paraformaldehyde for 30 minutes and incubated with 1:5000 Hoechst in Odyssey blocking buffer overnight at 4°C before imaging on the Zeiss Celldiscoverer 7. Cells were counted using ImageJ software. Each individual data point is

the percentage of micronuclei out of total nuclei per high power field. 28–30 high-power fields were imaged per cell line, encompassing 750–1250 cells per cell line.

***In vitro* culture for RNA-seq/ATAC-seq**—For bulk RNA-seq and matching ATAC-seq melanoma cells were cultured as described above split in triplicates for 10 days. Throughout this time cells were split and propagated to obtain ~80% cell density on day 10. On day 10 cells were then detached using 0.025% trypsin, quenched with full melanoma media, washed once with PBS and immediately processed for ATAC-seq or stored as frozen pellets for RNA extraction.

**RNA extraction and bulk RNA-seq of *in vitro* cultured cell lines**—Total RNA was extracted from frozen pellets using RNeasy Plus (Qiagen) with 5e5 cells per column. The RNA samples were quantified using Qubit 2.0 Fluorometer (ThermoFisher Scientific) and RNA integrity was checked using TapeStation (Agilent Technologies). The RNA sequencing libraries were prepared using the NEBNext Ultra II RNA Library Prep Kit for Illumina using manufacturer's instructions (New England Biolabs). Briefly, mRNAs were initially enriched with Oligod(T) beads. Enriched mRNAs were fragmented for 15 minutes at 94°C. First strand and second strand cDNA were subsequently synthesized. cDNA fragments were end repaired and adenylated at 3'ends, and universal adapters were ligated to cDNA fragments, followed by index addition and library enrichment by PCR with limited cycles. The sequencing libraries were validated on the Agilent TapeStation (Agilent Technologies), and quantified by using Qubit 2.0 Fluorometer (ThermoFisher Scientific) as well as by quantitative PCR (KAPA Biosystems, Wilmington, MA, USA). Final libraries were sequenced on a NovaSeq 6000 with S4 flow cell with 2x150 bp paired-end sequencing.

**ATAC-seq processing**—ATAC-seq was carried out as previously described (Ackermann et al., 2016; Buenrostro et al., 2013) with some minor modifications. Briefly, 50,000 cells were washed once with ice-cold PBS and the pellet subsequently resuspended in 50 µL of ice-cold lysis buffer (10 mM Tris-HCl, 10 mM NaCl, 3 mM MgCl<sub>2</sub>, 0.1% w/v NP-40). After centrifugation at 500 x g for 10 minutes at 4°C, the supernatant (comprising cytoplasm) was discarded and the pellet (nuclei) was resuspended in 50 µL of 2x TD Buffer (25 µL), pre-assembled Tn5 (2.5 µL) and distilled water (22.5 µL). TD Buffer and Tn5 were prepared as described previously (Ma et al., 2020; Wang et al., 2013). The transposition mix was incubated at 37°C for 30 minutes, after which DNA was isolated using the Qiagen MiniElute Reaction Cleanup Kit, eluting in 10 µL of Buffer EB.

A 5 cycle PCR was carried out using all 10 µL of isolated DNA, 10 µL nuclease-free water, 2.5 µL 25 µM Ad1\_noMX primer, 2.5 µL 25 µM Ad2.\* indexing primer and 25 µL NEBNext High-Fidelity 2x PCR Master Mix under the following conditions: 72°C for 5 minutes, and then for 5 cycles: 98°C for 30 seconds, 98°C for 10 seconds, 63°C for 30 seconds and 72°C for 1 minute. 5 µL of partially amplified DNA was removed and combined with 4.41 µL nuclease-free water, 0.25 µL 25 µM Ad1\_noMX primer, 0.25 µL 25 µM Ad2.\* indexing primer, 0.09 µL 100x SYBR Green 1 and 5 µL High-Fidelity 2x PCR Master Mix under the following conditions: 98°C for 30 seconds, and then for 20 cycles: 98°C for 10 seconds, 63°C for 30 seconds and 72°C for 1 minute. By plotting the relative fluorescence against cycle number, the number of additional PCR cycles required for each

sample was determined by assessing the cycle number needed to reach 1/3 of the relative fluorescence. The PCR reaction was then continued using the remaining 45  $\mu$ L of partially amplified DNA under the following conditions: 98°C for 30 seconds, and then for required number of cycles determined by qPCR: 98°C for 10 seconds, 63°C for 30 seconds and 72°C for 1 minute. A double-sided bead purification using AMPure XP Beads was then carried out using a 0.5x cleanup followed by 1.3x cleanup. Libraries were quantified using a 2200 TapeStation system (Agilent) and D5000 tapes. Final libraries were sequenced on a NovaSeq 6000 with S4 flow cell with 2x150 bp paired-end sequencing.

### ***In vivo* metastasis assay**

**Intracardiac injection:** 4L and 5B1 cells were first transduced with Green Fluorescent Protein/luciferase reporter lentivirus. MBM03 and MBM05 cell lines were used without additional modification. 4L and 5B1 were resuspended in sterile PBS at a concentration of  $1 \times 10^5$  cells per 100  $\mu$ L aliquoted into Eppendorf tubes and maintained on ice until injection. Male NOD.Cg-Prkdc<sup>scid</sup> Il2rg<sup>tm1Wjl</sup>/SzJ (NSG) mice were anesthetized by exposure to isoflurane. On day 0, anesthetized mice were injected with  $1 \times 10^5$  4L or 5B1 cells resuspended in 100  $\mu$ L sterile PBS into the left ventricle of the heart using Visualsonics Vevo 770 Ultrasound Imaging System. For MBM03 and MBM05  $5 \times 10^4$  cells were injected.

**Harvesting of melanoma cells:** Animals were sacrificed between 3–8 weeks post-intracardiac injection, once the animal exhibited 20% weight loss or symptoms meeting the IACUC protocol standard (e.g., limb paralysis). Animals were anesthetized with a ketamine (100 mg/kg) and xylazine (10 mg/kg) cocktail. Organs of interest were dissected and placed in a plate containing HBSS (Hank's buffered salt solution) on ice. GFP+ metastases were visualized using a Leica M205 FA fluorescence stereo (dissecting) scope. GFP+ areas were dissected away from the organ parenchyma. Tissue fragments were then minced and dissociated using type I collagenase (Worthington, LS004196) and DNase I (Worthington, LS002139) at 37°C for 45 minutes, with vortexing every 5 minutes. Samples were then strained successively through a 70  $\mu$ M and 40  $\mu$ M strainer and centrifuged. Cells were resuspended in 1 mL of RBC lysis buffer, incubated for 1 minute at RT, followed by quenching in 20 mL of dPBS. For brain samples, cells were resuspended in 38% (v/v) Percoll solution and centrifuged at 800 x g for 20 minutes, after which the top layer of myelin was removed. All cells were resuspended in 1 mL of RBC lysis buffer, incubated for 1 minute at RT, followed by quenching in 20 mL of dPBS. After pelleting and resuspending in FACS buffer (5% FBS in dPBS), cells were sorted for GFP+ status, pelleted, and snap-frozen. Cells were stored at –80°C until RNA extraction using the miRNEasy kit (Thermo Fischer, 217004). RNA quality was verified by Bioanalyzer for all samples. All samples had a RIN score of >7.8, with an average RIN score of 9.53. After RiboZero Plus library preparation, RNA sequencing took place using the Illumina NovaSeq 6000 platform and an S1 100 Cycle Flow Cell v1.5.

**Quantification of brain metastasis:** MBM03 and MBM05 brain metastasis were quantified macroscopically, and brain metastatic burden was calculated. To compare the organotropism of 4L and 5B1 mice were sacrificed after 4–5 weeks and livers and brains were dissected. The organs were formalin-fixed and paraffin embedded and three levels of each organ were

stained with hematoxylin and eosin. Whole slides were then scanned at 40x magnification using a Leica AT2 (Leica) slide scanner and metastatic burden was quantified by a blinded investigator in QuPath v.0.3.2.

**Multiplexed immunofluorescence**—Multiplexed immunofluorescence (mIF) staining of 19 metastatic brain and 23 liver lesions from patients with metastatic melanoma was performed using antibodies targeting CD8 (Leica, #4B11; 1:100), CD68 (Biogenex, #KP1; 1:1), CD138 (Leica, #MI15; 1:2), HMB45 (Cell Marque, #HMB-45; 1:100), NCAM1 (Invitrogen, #123C3; 1:200), and SOX10 (BioCar, #BC34; 1:100) using the Opal 7-color automation IHC kit (Akoya Bioscience) on a Leica Bond RX automated stainer (Leica Biosystems). 10 matching specimens from each group were also stained with CD163 (Leica, #10D6; 1:500). FFPE tissue sections (5  $\mu$ m) were baked for 2 h at 60°C, followed by automatic deparaffinization, rehydration, and antigen retrieval in BOND Epitope Retrieval Solution 2, pH 9 (Leica Biosystems) for 30 minutes at 95°C. Immunofluorescence staining with Opal and tyramide signal amplification (TSA) were performed in six cycles. In each cycle, the tissue was incubated sequentially with a primary antibody for 30 minutes at room temperature, the secondary antibody conjugated to polymeric horseradish peroxidase (HRP), an Opal fluorophore in TSA buffer, and BOND Epitope Retrieval Solution 1, pH 6 (Leica Biosystems) for 20 minutes at 95°C to strip the tissue-bound primary–secondary antibody complexes before the next staining cycle. For the full panel, the cycles contained NCAM1, no antibody (antibody diluent in lieu of primary antibody and 1X Bond Wash solution in lieu of HRP and Opal fluorophore), an HMB45 and SOX10 mixture (defined as lineage, LIN), CD8, CD138, and CD68, respectively. For CD163 quantification the antibody was added in position 2 of the multiplexed panel. After nuclear counterstaining with DAPI, slides were coverslipped with Vectashield HardSet Antifade mounting medium (Vector Laboratories). Whole slide scans (WSS) were captured at 10 $\times$  magnification using the Vectra Polaris automated multispectral microscope (Akoya/PerkinElmer) with Vectra Polaris 1.0.13 software. Regions of interest were chosen by the investigator for multispectral imaging (MSI) at 20 $\times$  magnification using the Phenochart 1.1.0 software (Akoya). Spectral unmixing was performed using the InForm v2.5.1 software (Akoya). Demultiplexed images were exported as 32-bit component TIFF files for further analysis.

## Quantification and statistical analysis

### **Generating single-cell and single-nucleus gene expression matrices**—

Demultiplexed FASTQ files from raw 3'-single-nuclei and 5'-single-cell RNA sequencing reads were aligned to the human GRCh38 genome and gene counts were quantified using Cell Ranger 'count' with introns included (v6.1.1; 10x Genomics).

**Removal of background noise in gene expression matrices**—We used the 'remove-background' function of CellBender v0.2.0 to remove technical ambient-RNA counts and empty droplets from the gene expression matrices (Fleming et al., 2019). Cell Ranger generated 'raw\_feature\_bc\_matrix.h5' files, which served as input for CellBender. The parameter 'expected-cells' was obtained from the Cell Ranger metric 'Estimated Number of Cells', while the parameter 'total-droplets-included' was set to a value between 10,000–40,000 representing a point within the plateau of the barcode-rank plot.

**Quality control and filtering**—The resulting expression matrices were processed individually in R v4.1.1 using Seurat v4.1.0 (Stuart et al., 2019). Filters were applied to keep cells with 500–10,000 genes, 1,000–60,000 UMIs, and <10% of mitochondrial reads in each individual sample. Additionally, Scrublet v0.2.1 (Wolock et al., 2019) and DoubletFinder v2.0.3 (McGinnis et al., 2019) were applied to identify and remove doublets with an expected doublet rate ranging from 2.8–6.4% based on the loading rate. Filtered gene-barcode matrices were normalized with the ‘NormalizeData’ function using the ‘LogNormalize’ method. The top 2,000 variable genes were identified using the ‘vst’ method in the ‘FindVariableFeatures’ function. Gene expression matrices were scaled and centered using ‘ScaleData’. Next, we performed principal component analysis (PCA) as well as uniform manifold approximation and projection (UMAP) dimension reduction using the top 30 principal components. UMAPs of individual samples were inspected prior to integration. In scRNA-seq samples, the two batches per patient derived from CD45<sup>+</sup> and CD45<sup>-</sup> cells were merged. For patient MBM01\_sc the CD45<sup>-</sup> batch failed and was excluded from subsequent analyses. Cell types were preliminarily annotated using SingleR v1.8.0 (Aran et al., 2019) against its built-in reference ‘BlueprintEncodeData’.

**Identification of malignant cells**—Chromosomal CNA profiles of individual cells were inferred from transcriptional data using inferCNV v1.10.1 (Tirosh et al., 2016). For each sample, we used cells that were identified as immune cells by SingleR (Aran et al., 2019) using ‘BlueprintEncodeData’ as a diploid reference to estimate CNAs in the non-immune cells. We applied a cutoff of 0.1 for the minimum average read counts per gene among reference cells/nuclei, set the clustering to ‘subcluster’, denoised the output using the default ‘sd\_amplifier’ of 1.5, and ran Hidden Markov Models (HMM) to predict the CNA level. The proportion of scaled CNAs was then averaged over all chromosomes in each cell individually. We identified malignant cells using sample-specific thresholds for the average proportion of inferred copy number alterations (CNAs) per cell, which is largely defined by the cell-type composition of the sample.

**Integration of individual samples**—All individual samples were integrated in Seurat using the canonical correlation analysis (CCA) pipeline to remove batch effects. The ‘SelectIntegrationFeatures’ function was applied to choose the features ranked by the number of datasets they were detected in. Next, the ‘FindIntegrationAnchors’ function selected 2,000 anchors between different samples using the top 50 dimensions from the CCA to specify the neighbor search space. ‘IntegrateData’ was then applied to integrate the datasets using the pre-computed anchors and the integrated dataset was scaled using ‘ScaleData’. PCA and UMAP dimension reduction based on the top 50 principal components were performed. In addition, we also integrated three separate batches of the dataset for cell type annotation: MBM with scRNA-seq, MBM with snRNA-seq and ECM snRNA-seq. These batches were subset to non-malignant cells, rescaled and PCA, UMAP, finding-neighbor and cluster analyses were applied to create the dataset for manual cell-type annotation.

**Benchmarking of integration methods**—To identify the best approach for batch correction of our dataset consisting of samples from different patients, different tissues

and different sequencing methods, we performed a comprehensive comparison of different integration methods (CCA (Butler et al., 2018), Harmony (Korsunsky et al., 2019), Conos (Barkas et al., 2019), STACAS (Andreatta and Carmona, 2021)) in comparison with the non-integrated baseline. We performed the different integration methods on four batches consisting of all profiled cells (MBM+ECM scn, i.e. single cells and nuclei), snRNA-seq alone (MBM+ECM sn), and also specific comparisons for T cells as STACAS was largely benchmarked on T cells (MBM+ECM scn T cell; MBM+ECM sn T cell).

**Harmony:** To integrate the different batches with Harmony v0.1.0 (Korsunsky et al., 2019), we used the top 30 principal components as input for Harmony and removed the bias resulting from the ‘patient’ variable. Subsequently, the top 30 Harmony dimensions were selected for quantifying UMAP embeddings.

**Conos:** We generated a Conos object containing the individual, Seurat-preprocessed samples using Conos v1.4.4 (Barkas et al., 2019). Next, we built a joint graph with default settings that encompasses all samples using PCA space including the top 50 principal components. Then, we generated an embedding of the joint graph using the function ‘embedGraph’ with UMAP setting selected.

**STACAS:** Sub-Type Anchoring Correction for Alignment in Seurat (STACAS) v1.1.0 (Andreatta and Carmona, 2021) is a method for anchor identification and was performed on T-cell subsets of our dataset. We excluded samples with <50 T cells and ran the Seurat preprocessing pipeline on each sample individually with 20 principal components and 500 anchor features excluding mitochondrial, RPL, RPS and cell cycle genes. Next, we computed an integration tree with STACAS anchors that was passed as a parameter to the Seurat function ‘IntegrateData’.

**Local Inverse Simpson Index (LISI) score:** To quantify how different integration methods perform, we used the recently developed Local Inverse Simpson Index (LISI) score (Korsunsky et al., 2019). Specifically, we determined the integration LISI score (iLISI) which measures how well cells from different batches (patients) mix, and where higher scores indicate better mixing and therefore integration. We also used the cell-type LISI (cLISI), which measures how well specific cell-type clusters are separated from other cell-type clusters, and where a lower score is desirable.

**Differential gene expression (DGE) and geneset enrichment—**For cell type annotation, DGE was performed using the Seurat function ‘FindAllMarkers’ on normalized count data to identify positive (overexpressed) markers in each population (Table S2). The MAST algorithm was used to identify differentially expressed genes (DEGs) between two groups of cells unless stated otherwise and the log-fold change was set to 0.25. The parameter ‘min.pct’ was set to 0.25 to assure that genes are detected at a minimum fraction of 25% of cells in either of the populations.

To identify differences between MBM and ECM for each cell type, we first downsampled the snRNA-seq dataset to ensure that the average per-cell total count is equal between MBM and ECM using the function ‘downsampleBatches’ from the scuttle R-package v1.4.0

(McCarthy et al., 2017). Additionally, the number of cells for DGE was downsampled to the number of cells in the smaller batch. The resulting DEGs were analyzed for geneset enrichment using hypeR v1.10.0 (Federico and Monti, 2020). The background population of genes was set to include all detected genes and geneset over-representation was determined using the hypergeometric test. In the MBM vs. ECM comparisons, we scored the following pathways of MSigDB v7.4.1 (Dolgalev, 2021): GO BP, GO MF, Canonical Pathways, Hallmarks, Kegg, Reactome, PID, and WikiPathways. Pathways with an FDR <0.0001 were applied as module scores to the Seurat object. We then used a modified version of the GmAMisc R-package v1.2.0 (Alberti, 2021) to perform a permutation-based *t*-test to compare the module scores identified as significantly different between MBM and ECM. Outliers (defined as >90<sup>th</sup> percentile or <10<sup>th</sup> percentile) were excluded from the input dataset and 10,000 permutations were performed to identify the most robust differential pathways between MBM and ECM. Violin plots of selected, permutation-significant pathway module scores were reported including p-values from Wilcoxon rank-sum tests.

To identify differences between the single-cell and single-nuclei transcriptomes in TOX<sup>+</sup> CD8<sup>+</sup> T cells and microglia, we removed all RPL, RPS and MT genes to find differences aside from known stressors based on sample handling. Next, we downsampled using the function 'downsampleBatches' to adjust for differences in quality between the two sequencing techniques followed by DGE using the 'FindMarkers' function with 'logFC' and 'min.pct' set to zero and 'max.cells.per.ident' set to the number of cells in the smaller batch. DEGs with an adjusted p-value <0.05 and log<sub>2</sub>FC >2 or <-2 were considered as differentially regulated. The average expression of each sequencing group was calculated by taking the log<sub>2</sub> (with a pseudocount of 1) of the averaged exponent (minus 1) of the normalized output from the downsampled Seurat assay. The resulting two vectors were plotted in a scatter plot and correlation was assessed using Pearson correlation.

Volcano plots are based on DEGs using the 'FindMarkers' function with 'logFC' and 'min.pct set' to zero. In DGE of T cell subsets, *TRAV/TRBV* genes were removed before applying the 'FindMarkers' function.

**Cell type annotation**—We integrated profiles based on their sequencing method and tissue origin to enable the most concise cell-type annotation. The main cell types were identified by manual annotation of differential gene expression (DGE) between clusters in the non-malignant data subset based on known markers and signatures (Table S1) (Azizi et al., 2018; Cahoy et al., 2008; Lein et al., 2007; Olah et al., 2020; Zilionis et al., 2019). The initial labeling resulted in the identification of endothelial, epithelial, stromal, CNS, myeloid, low-quality, T/NK and B/plasma cell populations. Next, we split the Seurat object into subsets of the main labels and reran scaling, PCA, UMAP dimension reduction, clustering and DGE analysis on each subset. The resulting clusters were annotated manually or using cell-type-specific single-cell signatures from respective cell atlases (Table S2), and the cell-type labels were added to the main integrated object. The 'AddModuleScore' function was applied to calculate average expression levels of gene signatures on single-cell level. Additionally, cell-cycle phases were scored in the subsets using the 'CellCycleScoring' function, adjusted for individual cutoffs and added to the main

object. Mouse-derived signatures were converted to human homolog genes using biomaRt v2.46.3 (Durinck et al., 2009).

**Cell type frequency comparison**—We calculated frequencies of cell types in all snRNA-seq samples from MBM and ECM groups and compared medians of the two groups to determine differences in frequency. For MBM samples, cell type frequencies between scRNA-seq and snRNA-seq samples were calculated as well. Significance was assessed using Wilcoxon rank-sum test.

**Diffusion component (D.C.) analysis**—We applied diffusion maps as a nonlinear dimensionality reduction technique to examine the major components of variation across different cell types. We computed D.C.s using the ‘DiffusionMap’ function of the Destiny R-package v3.9.0 (Angerer et al., 2016). To identify drivers of tumor-cell variability, non-cycling snRNA-seq samples of MBM and ECM tumor cells were analyzed individually. The Destiny-derived eigenvectors were added to the respective Seurat objects as embedding for a dimension reduction object. For each sample, we considered the top 50 features with the largest absolute value of projected loadings for each of the first three D.C.s and computed the overlap of recurring genes among samples. Genes that were detected at a frequency of >0.25 were included in geneset enrichment analysis using hypeR v1.10.0 (Federico and Monti, 2020).

**ProjecTILs analysis of T cell states**—Without batch effect correction, T cell profiles clustered mainly by technology (scRNA-seq vs. snRNA-seq) and by patient/batch. To interpret T cell states irrespective of batch and technology, we applied a reference-based approach (Andreatta et al., 2021). Briefly, pure T cell transcriptomes from large samples (>200 cells) were integrated using Seurat-STACAS (Andreatta and Carmona, 2021; Stuart and Satija, 2019), and unsupervised cell clusters were manually annotated to generate a “reference atlas”. This hybrid sc/snRNA-seq reference atlas defines five TIL subtypes that are consistently observed across sequencing technologies and samples: TOX<sup>+</sup> CD8<sup>+</sup> T cells co-expressing *TOX* and multiple co-inhibitory receptors including *HAVCR2* and *ENTPDI*; TCF7<sup>+</sup> CD8<sup>+</sup> T cells with low levels of inhibitory receptors and high levels of *IL7R*, *CD69* and *CCL5*; helper/memory CD4<sup>+</sup> T cells expressing high levels of *TCF7*, *SELL*, *IL7R*; T<sub>regs</sub> expressing *FOXP3* and *IL2RA*; and follicular helper-like CD4<sup>+</sup> T cells (T<sub>fh</sub>-like), co-expressing *TOX*, *TOX2* and *TCF7*. ProjecTILs v1.0 using default parameters was then used to project new samples (either sc- or snRNA-seq datasets) into the reference hybrid sc/snRNA-seq atlas.

**TCR data processing and integration**—TCR FASTQ files were aligned using Cell Ranger ‘vdj’ (v6.1.1; 10x Genomics). The resulting filtered contig annotations were consolidated and added to the Seurat object. TCRs with sequencing information for both chains were considered for downstream analyses. Clones with more than one detected copy were labelled as expanded.

**B cell chain analysis**—To analyze the distribution of heavy and light chains in B and plasma cells, the dataset was subset to include only B and plasma cells. For the identification of variable chain regions, we selected the highest expressed heavy and light chain gene of



each cell that expressed both, heavy (starting with *IGHV*) and light (starting with '*IGLV*' or '*IGKV*') chain encoding genes. Next, we identified the highest expressed constant chain region among expressed genes following the pattern '*IGH*[*G, M, A, or E*][number]'. The resulting pairs of heavy and light chains were visualized as a heatmap using average linkage for hierarchical clustering analysis.

**Non-negative matrix factorization (NMF) using KINOMO**—Non-negative matrix factorization (NMF) is a widely used method for performing dimensional reduction and feature extraction on 'non-negative' data. The conventional NMF (Lee and Seung, 1999) decomposes matrix  $A$  into two matrices with non-negative entries with smaller ranks,  $A \approx WH$ , where  $A \in \mathbb{R}^{(n \times m)}$ ,  $W \in \mathbb{R}^{(n \times k)}$ ,  $H \in \mathbb{R}^{(k \times m)}$ . Without loss of generalization, rows of  $A$  represent features (e.g., genes) and columns of  $A$  represent samples. Depending on context,  $W$  can be interpreted as a feature mapping, rows of  $W$  represent disease profiles or meta-genes (Brunet et al., 2004) and columns of  $H$  are compact representations of samples, i.e., sample profiles. In this manuscript we rely on using Kernel differentiability correlation-based NON-negative Matrix factorization algorithm using Kullback-Leibler divergence loss Optimization (KINOMO), a semi-supervised NMF model that is robust to noises and uses 'prior' biological knowledge for better refinement (Tagore et al., 2022). KINOMO has three major steps, namely, i) subsetting the dataset to snRNA-seq-derived tumor cells, ii) NMF core module, and iii) meta-gene and meta-program (MP) estimation. The filtration and cleaning steps begin with a scRNA-seq sample, followed by tumor cell estimation, normalization and scaling. The second step consists of two sub-steps a) Factorization Error analysis, using  $L_{2,1}$  norm loss to handle outliers, add prior knowledge by introducing graph regularization parameters, sequential quadratic approximation for Kullback-Leibler divergence loss, local geometrical structure preservation, optimizing the update rules for approximation matrix  $WH$  and clustering using Kernel differentiability correlation; and b) factor rank survey analysis. The third step consists of meta-gene and factor block estimation by performing co-correlation analysis of estimated factors (sample-wise) using Spearman's correlation. The consensus factors are selected using significance testing (p-value) and/or correlation value among all factors and using a correlation threshold of 0.4–0.9. This is done iteratively, until consensus MPs (metaprograms) are obtained (Tagore et al., 2022).

**Master regulator analysis**—The regulatory network in this study was reverse-engineered from scRNA-seq data using the ARACNe-AP algorithm (Basso et al., 2005; Lachmann et al., 2016). We generated networks for each patients' tumor cells and integrated the networks by taking a union of the predictions of all networks. p-values of Master regulator (MR)-target interactions predicted by the networks were integrated using Fisher's method (Fisher, 1992). The final tumor network contained predictions for regulators regulating target genes through interactions, respectively. The final tumor network contained predictions for 3,112 regulators regulating 8,210 target genes through 145,065 interactions, respectively. The relative activity of each protein represented in the tumor network was inferred using the VIPER algorithm v1.26.0 (Alvarez et al., 2016; Ding et al., 2018). Conceptually, the VIPER algorithm is similar to the master regulator inference algorithm (MARINA) (Lachmann et al., 2016; Lefebvre et al., 2010), which uses the MR targets inferred by the ARACNe algorithm to predict drivers of changes in cellular phenotypes.

In addition to calculating the enrichment of ARACNe-predicted targets in the signature of interest, VIPER also considers the regulator mode of action, regulator–target gene interaction confidence, and the pleiotropic nature of each target gene’s regulation. Statistical significance, including p-value and normalized enrichment score (NES), was estimated by comparison to a null model generated by permuting the samples uniformly at random 1,000 times.

**SCENIC (single-cell regulatory network inference and clustering)**—SCENIC/pyscenic v0.11.2 (Aibar et al., 2017) is a computational method for gene regulatory network (GRN) reconstruction and cell-state identification that predicts transcription factor (TF) activities. Co-expression modules between TFs and target genes were inferred using GRNBoost based on correlations between expression of genes across snRNA-seq-derived tumor cells. RcisTarget then refined the selection of modules by keeping only modules with significantly enriched TF binding motif. Subsequently, AUCell scored the activity of each regulon in each cell.

**RNA velocity**—We used velocity v0.17 (La Manno et al., 2018) to generate loom files from bam files, which were generated by Cell Ranger. Next, the individual loom files were merged and Harmony-integrated using scanpy v1.8.2 (Wolf et al., 2018) in python v3.9.9. We then ran scVelo v0.2.4 (Bergen et al., 2020) using default settings to obtain RNA velocity and pseudotime.

**Processing of cell line bulk RNA-seq (*in vitro*)**—We quantified abundances of transcripts from bulk RNA-sequencing of cell lines using Kallisto v0.46.1 (Bray et al., 2016). Next, we imported the abundance.h5 files into R using tximport v1.22.0 (Soneson et al., 2016) and summarized the counts to gene-level. Differential gene expression was performed using DESeq2 v1.34.0 (Love et al., 2014) on raw counts with an additive model including the cell line and tissue origin (i.e., MBM vs. ECM).

**Processing of cell line bulk ATAC-seq data (*in vitro*)**—Raw sequencing FASTQ files were evaluated for quality and adapter content using FastQC v0.11.9 followed by adapter removal (-a CTGTCTCTTATA -A CTGTCTCTTATA) and quality trimming using cutadapt v3.6 (Martin, 2011) and finally aligned with hisat2 v2.2.1 (Kim et al., 2019) (default, -X 1000) to the human genome (GRCh38). Using samtools the BAM files were sorted, filtered for extra-chromosomal DNA, mitochondrial DNA and for reads overlapping blacklisted regions (Amemiya et al., 2019). Duplicate reads were removed using sambamba v0.6.8 (Tarasov et al., 2015) before downstream analyses. Accessible peaks were identified for each sample using MACS2 (v2.2.7.1, parameters: --nomodel --shift -100 --extsize 200 --broad) (Feng et al., 2012). BigWig files were generated using bamcoverage from deepTools2 (v3.5.1; --normalizeUsing CPM --binSize 25 --smoothLength 100) (Ramírez et al., 2016). The BAM, BED, and BigWig files were used as input in CoBRA v2.0 (Containerized Bioinformatics Workflow for Reproducible ChIP/ATAC-seq Analysis) (Qiu et al., 2021) to perform differential peak calling, generate differential site heatmaps and motif discovery. Footprinting analysis was done using the TOBIAS pipeline v0.13.2 (Bentsen et al., 2020) starting with transcription factor motifs from the JASPAR CORE 2020 (Castro-Mondragon

et al., 2022). When comparing differentially expressed genes with ATAC-seq, we used a q-value of 0.1 and a two-fold-change in expression as thresholds. The negative log of p-adjusted values was plotted against  $\log_2FC$ .

**Scoring signatures on bulk RNA-seq data**—We used the top 100 genes of our snRNA-seq-derived DEG comparisons for MBM and ECM tumor cells as signatures on the bulk-RNA-seq of the Fischer dataset (Fischer et al., 2019). The singscore R-package v4.0.3 (Foroutan et al., 2018) was used to apply bidirectional gene signatures (with the top 100 snRNA-seq derived MBM markers as up-set and ECM as down-set in  $MBM_{bulk}$ ; and vice versa for  $ECM_{bulk}$  samples) on bulk-RNA-seq data. The bulk-RNA-seq genes were ranked based on their transcript abundance in increasing order for the up-set and in decreasing order for the down-set. Mean ranks were normalized separately, centered and then summed to provide the score which ranges between  $-1$  and  $1$ . High scores show concordance to the specified signature. The resulting scores reflect the relative mean percentile rank of the provided markers within each sample.

**Analysis of bulk RNA-seq data from xenograft models (*in vivo*)**—Alignment of FASTQ files was done using the seq-n-slide pipeline route rna-star (Dolgalev, 2022). Briefly, the Trimmomatic package v0.36 (Bolger et al., 2014) was used to trim adapters and low-quality bases. STAR v2.7.3a (Dobin et al., 2013) was used to align to human genome hg38, in addition to screening for alignment to other species and common contaminants. The featureCounts package v1.6.3 (Liao et al., 2014) was used to generate a gene-samples counts matrix. Read quality was assessed using FASTQC v0.11.7, picard v2.18.20, and FastQ Screen v0.13.0 (Wingett and Andrews, 2018). Differential gene expression was performed using DESeq2 v1.34.0 (Love et al., 2014) on raw counts with an additive model including the cell line and tissue origin (i.e., MBM vs. ECM). The sample distance matrix was generated based on the Euclidean distances of the ‘vst’ (varianceStabilizingTransformation) function output and clustered using ‘complete’ linkage.

**Differential protein expression in proteomics data of short-term cultures**—The proteomics dataset was obtained from Dr. Eva Hernando’s Lab and is part of a recently published study (Kleffman et al., 2022) and deposited at MassIVE (accessible at <https://massive.ucsd.edu/>; ID: MassIVE MSV000088814). In brief, the MS/MS spectra were searched against the UniProt human reference proteome with the Andromeda search engine (Cox et al., 2011) integrated into the MaxQuant environment v1.5.2.8 (Cox and Mann, 2008) using the following settings: oxidized methionine (M), TMT-labeled N-term and lysine, acetylation (protein N-term) and deamidation (asparagine and glutamine) were selected as variable modifications, and carbamidomethyl (C) as fixed modifications; precursor mass tolerance was set to 10 ppm; fragment mass tolerance was set to 0.01 Th. The identifications were filtered using a false-discovery rate (FDR) of 0.01 using a target-decoy approach at the protein and peptide level. Only unique peptides were used for quantification and only proteins with at least two unique peptides were reported. Data analysis was performed using Perseus (Tyanova et al., 2016). Protein levels were median centered and  $\log_2$ -normalized. To identify differentially expressed proteins between the MBM and ECM cohorts, a paired *t*-test was performed on three sample pairs.

**MBM signature overlap**—To identify overlapping genes differentially expressed in MBM across techniques in snRNA-seq (this study), RNA-seq (Fischer et al., 2019), and proteomics (Kleffman et al., 2022) datasets, we applied a  $\log_2$ FC cutoff of 0.4 for all cohorts and an adjusted p-value of 0.05 in snRNA-seq and RNA-seq.

**Copy number alteration (CNA) analysis**—For each MBM and ECM sample in our dataset, we used inferCNV results to calculate the median tumor-cell copy number of each gene. We then defined a gene as amplified if its median copy number was higher than the median inferred copy number of all genes in each sample, and conversely as deleted if its median copy number was lower. We obtained WES seg files from the authors of the Fischer *et al.* dataset (Fischer et al., 2019) and processed them using the Bioconductor package DNACopy v1.66.0 (Venkatraman and Olshen, 2007) with an interval of  $[-0.4, 0.4]$  as lower and upper thresholds to define a segment as deleted or amplified, respectively.

**Fraction of the genome altered**—For the MBM<sub>WES</sub> and ECM<sub>WES</sub> samples obtained from the authors of the Fischer *et al.* (Fischer et al., 2019) dataset, we calculated the fraction of the genome for each sample that was either amplified or deleted according to the above criterion. We then used the ‘plotFreq’ function from the Bioconductor package copynumber v1.32.0 (Nilsen et al., 2012) to visualize the altered genome fraction across each chromosome.

#### **Initial processing and quality control of spatial data generated using**

**SlideSeqV2**—We used the Slide-seq tools pipeline (Stickels et al., 2021) to process our Slide-seq data, and loaded the data into Seurat. For all pucks besides MBM05 replicate #3 and MBM11 replicate #3, we did not apply any additional quality control thresholds. For MBM05 replicate #3 and MBM11 replicate #3p, due to their lower quality compared to the other samples, we filtered out cells with fewer than 20 unique genes.

#### **Assignment of discrete cell types with Robust Cell Type Decomposition**

**(RCTD)**—We used the RCTD pipeline v1.2.0 (Cable et al., 2021), as part of the R package spacexr, which accepts two inputs: 1. count matrices for a spatial single-cell sequencing dataset, and 2. a non-spatial single-cell sequencing dataset with cell type annotations. We used count matrices from our SlideSeqV2 dataset for MBM and ECM samples. For our reference in this case, we used non-spatial single-nuclei sequencing data from all MBM and ECM samples, with the main cell types manually annotated as described previously. RCTD then uses the reference dataset to learn expression profiles for each annotated cell type, and uses these profiles to deconvolve cell type proportions at each location in the spatial single-cell data. Based on these inferred cell type proportions, discrete cell types were assigned by taking the cell type with the highest inferred proportion at a particular location.

#### **Analysis of spatial differential gene expression (DGE) in SlideSeqV2 samples**

—We used the ‘SCtransform’ function in Seurat (Hafemeister and Satija, 2019) to normalize and stabilize the variance of the spatial data. We then used the function ‘FindVariableFeatures’, with ‘selection.method’ parameter set to ‘vst’, to determine the top 1,000 genes whose expression was variable in each puck. Finally, we used the function

‘SpatiallyVariableFeatures’, with ‘selection.method’ parameter set to ‘moransi’, to rank genes whose expression was spatially variable in each puck.

We then manually examined patterns of gene expression in our spatial data in order to determine shared patterns of spatial co-expression. We observed that genes with a Moran's *I* value greater than .05, roughly corresponding to a p-value of  $<.01$ , often showed high levels of spatial DGE, with expression being clustered in distinct regions of the puck. On the other hand, genes below this threshold usually did not show any distinct pattern of spatial DGE. Therefore, we extracted all genes with Moran's *I* above .05, and manually classified those with shared spatial DGE patterns into groups (Table S6). We then used these groups as signatures to score spatial and non-spatial data using the Seurat function ‘AddModuleScore’, as described previously.

Finally, focusing on ECM\_01 replicate #2, we observed that *TIMP1* and *HLA-A* appeared to be spatially anti-correlated, with *TIMP1* expression highest on the left side of the puck, and *HLA-A* higher on the right side. We therefore calculated the Spearman correlation of expression of each gene in our data with *TIMP1*, and selected those with a Bonferroni-corrected p-value  $<0.05$ . We then manually examined plots of gene expression for each gene with significant correlation with *TIMP1*, to confirm that they were spatially correlated or anti-correlated.

**Cell type co-occurrence analysis**—We wrote custom R and Python scripts to transfer spatial cell type annotations from RCTD into AnnData objects within the squidpy framework (Palla et al., 2022). We then used the `spatial_neighbors` and `nhood_enrichment` functions in squidpy to calculate the significance of co-occurrence between every pair of cell types in each sample. The significance of each co-occurrence interaction is given as a Z-score value after permutation testing within the `nhood_enrichment` function. We gathered all interactions within a Z-score  $>3.09$ , corresponding to a p-value of .001.

**Spatial DEG detection with CSIDE, and comparison with spatial DEG from Moran's *I***—We used the `run.CSIDE.nonparam` function from the `spacexr` package to calculate cell type-specific spatial DEG, using as input the RCTD output objects that were generated by RCTD, also from the `spacexr` package. Following the online tutorial for CSIDE, we decided to use very permissive parameters of `gene_threshold=0.01`, `fdr=0.25`, and `cell_type_threshold=15` for the `run.CSIDE.nonparam` function. However, with these parameters, we encountered a problem during our analysis in which some of the parameter estimates for some genes failed to converge in the course of the algorithm. The algorithm thus halted before reporting any results. After correspondence with the author of the `spacexr` package, Dylan Cable, we decided to work around this problem by slightly modifying the code of the `spacexr` package, so that it no longer checks for gene convergence before reporting results.

For each sample, we thus obtained a list of spatial DEG from CSIDE for each cell type in the sample. We combined the lists of cell-type specific spatial DEG into one gene list for each sample, and then checked for overlap with the list of spatial DEG obtained by Moran's *I* on the same sample. We used two statistical tests for this comparison. First, we used the

hypergeometric test with the R function `phyper(q, m, n, k, lower.tail=FALSE)`, with  $q$  being the number of overlapping spatial DEG between the two methods,  $m$  the number of Moran's  $I$  spatial DEG,  $n$  the number of genes in the sample – the number of Moran's  $I$  spatial DEG, and  $k$  the number of CSIDE spatial DEG. Second, we also used a simple permutation test. Given a sample with the number of Moran's  $I$  spatial DEG= $x$ , we randomly selected 10,000 subsets of size  $x$  from all genes in that sample, and calculated the overlap of each random subset with the set of CSIDE spatial DEG for that sample. We then compared the resulting distribution of overlaps with the overlap of the true Moran's  $I$  and CSIDE spatial DEG, and computed the probability of getting an overlap as large or greater than the true overlap. Both the hypergeometric and permutation tests returned very significant p-value results, which are reported in Table S6.

**Multiplexed image analysis**—Raw images were processed with `bfconvert` from `bftools` v5.8.2 for performing cell segmentation (Linkert et al., 2010). Whole-cell segmentation was performed on raw images using `MESMER` v0.3.0 with the containerized version using `singularity` v3.9.2 (Greenwald et al., 2021) on a computational cluster. Mesmer is a deep-learning cell segmentation algorithm trained on a large number of images to segment nuclei and cells in histological images with human level performance. We used DAPI staining to identify nuclei and CD138 staining for the membrane marker as input for Mesmer with whole-cell segmentation and `--mpp 0.49` settings to specify the micron per pixel image resolution. Cell segmentation masks were then used with the `MCMICRO` (Schapiro et al., 2021) quantification module using the `nextflow` (v21.10.6.5660) (Di Tommaso et al., 2017) implementation of `MCMICRO`, to quantify marker intensities for all identified cells for each image. Additional cell segmentation by using DAPI staining, and marker classifiers were created with `QuPath` v0.3.2 (Bankhead et al., 2017), used to cross check thresholds utilized in image data analysis.

Quantified marker intensities were analyzed and processed using a local instance of Python and Jupyter notebook v 3.9.0 (Kluyver et al., 2016). A metadata file was created for each image file, containing tissue and sample information as well as the number of cells detected by `MCMICRO`. For the CD68 marker the raw intensities were log-scaled, all the raw intensities equal to 0 were removed from the log-scaled data. The data distribution for both raw and log-scaled marker intensities were visualized. A raw intensity value of 1.7564 was chosen as the cutoff to call CD68+ cells on all images. This cutoff value was visually reviewed by an expert by overlaying the original marker intensity with the Mesmer predicted mask (Figure S4H,I,J).. The marker intensity distribution of CD68+ cells was calculated and visualized by tissue of origin and by sample. We calculated the distribution of the number of CD68+ cells (per image) and the proportion of CD68+ cells to the total number of identified cells (per image). The proportion of CD68+ cells was then visualized by tissue and sample. To test the null hypothesis that two given populations are equal, we performed a two-sided Wilcoxon rank-sum test to compare both the intensities and the proportion of CD68+ cells between MBM and ECM.

The raw intensities for the CD138 marker were log-scaled, any raw intensity equal 0 was removed from the log-scaled data. The intensity distribution was visualized for both raw and log-scaled intensities and a cutoff value of 1.6312 was chosen to label CD138+ cells.

Expert visual inspection of the cutoff value was performed by overlaying the original marker intensity with the predicted mask (Figure S4H,I,J). The average CD138+ cells / total cells ratio per image was obtained by applying the CD138 cutoff value to all cells in an image and dividing it by the total number of detected cells. All images without CD138+ cells were removed from further analysis. The CD138+/Total ratio was visualized by sample and tissue of origin. A two-sided Wilcoxon rank-sum test was performed to test the CD138+ ratio between MBM and ECM samples. For all images with at least two CD138+ cells a neighborhood analysis was performed. The analysis consisted in the identification of all CD138+ cells in the close neighborhood of all other CD138+ cells per image. To obtain the CD138+ cell neighborhood the pairwise euclidean distances between all CD138+ cells was calculated. From these distances we counted all the CD138+ cells below a given distance threshold and averaged the neighborhood value for each cell in the image. For this study a threshold value of 50 micrometers was used. After the CD138+ neighborhood value was calculated for all images, two-sided Wilcoxon rank-sum tests were performed at all threshold values to compare the neighborhood values between MBM and ECM samples. For NCAM1 and melanoma lineage markers (LIN; SOX10 and HMB-45) visualization of raw and log-scaled intensity values were performed and gates were visualized (Figure S4H,I,J). Cutoff values of 1 for NCAM1 and 1 for LIN were chosen and validated by visual inspection by overlaying original image values with the predicted mask. Data was gated for LIN and the NCAM1 intensity values for the gated data were then visualized by tissue of origin and compared between MBM and ECM samples using a two-sided Wilcoxon rank-sum test. We also visualized and calculated the proportion of double positive NCAM1+ LIN+ cells to all the LIN+ cells (NCAM1+ LIN+/ LIN+). This was performed by first gating on LIN and subsequently on NCAM1. The number of double positive cells was then divided by the number of LIN+ cells by image. Finally, a two-sided Wilcoxon rank-sum test was performed comparing the double positive to LIN+ ratio between MBM and ECM samples.

**Quantification of CD163 protein expression**—Images for the DAPI and CD163 stains were imported into CellProfiler-3.1.9 (Broad Institute). For image analysis, nuclei were first segmented using minimum cross entropy method, and then, using the segmented nuclei as seeds, total cellular areas were segmented using the propagation method. Size restrictions were entered in the primary nuclei segmentation to improve accuracy; objects with diameters either below or above a pre-specified pixel length were excluded. After detection of cells, a local background area surrounding each cell was determined by expanding the cells by a set number of pixels, and then subsequently subtracting out the initial cellular area. For fluorescence calculation, the mean fluorescent intensities inside the cellular area (MFI\_cell) and in the background area (MFI\_sig) were measured, and the final intensity was calculated as  $MFI_{CD163} = MFI_{cell} - MFI_{sig}$ . To restrict analysis to CD163+ cells and lower the number of false positives due to autofluorescence, we ultimately performed calculation on cells whose MFI\_CD163 was above a threshold set at the mean of the total distribution.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Authors

Jana Biermann<sup>1,2,32</sup>, Johannes C. Melms<sup>1,3,32</sup>, Amit Dipak Amin<sup>1,3</sup>, Yiping Wang<sup>1,2</sup>, Lindsay A. Caprio<sup>1,3</sup>, Alcida Karz<sup>4</sup>, Somnath Tagore<sup>5</sup>, Irving Barrera<sup>6</sup>, Miguel A. Ibarra-Arellano<sup>7</sup>, Massimo Andreatta<sup>8,9</sup>, Benjamin T. Fullerton<sup>1</sup>, Kristjan H. Gretarsson<sup>10</sup>, Varun Sahu<sup>10</sup>, Vaibhav S. Mangipudy<sup>10</sup>, Trang T.T. Nguyen<sup>11</sup>, Ajay Nair<sup>1</sup>, Meri Rogava<sup>1,3</sup>, Patricia Ho<sup>1,3</sup>, Peter D. Koch<sup>12</sup>, Matei Banu<sup>13</sup>, Nelson Humala<sup>13</sup>, Aayushi Mahajan<sup>13</sup>, Zachary H. Walsh<sup>12</sup>, Shivem B. Shah<sup>12</sup>, Daniel H. Vaccaro<sup>12</sup>, Blake Caldwell<sup>12</sup>, Michael Mu<sup>12</sup>, Florian Wünnemann<sup>7</sup>, Margot Chazotte<sup>7</sup>, Simon Berhe<sup>1</sup>, Adrienne M. Luoma<sup>14</sup>, Joseph Driver<sup>15</sup>, Matthew Ingham<sup>1</sup>, Shaheer A. Khan<sup>1</sup>, Suthee Rapisuwon<sup>16</sup>, Craig L. Slingluff Jr.<sup>17</sup>, Thomas Eigentler<sup>18,19</sup>, Martin Röcken<sup>18</sup>, Richard Carvajal<sup>1</sup>, Michael B. Atkins<sup>16</sup>, Michael A. Davies<sup>20</sup>, Albert Agustinus<sup>21,22</sup>, Samuel F. Bakhoun<sup>21,23</sup>, Elham Azizi<sup>24,25</sup>, Markus Siegelin<sup>11</sup>, Chao Lu<sup>10</sup>, Santiago J. Carmona<sup>8,9</sup>, Hanina Hibshoosh<sup>11</sup>, Antoni Ribas<sup>26</sup>, Peter Canoll<sup>11</sup>, Jeffrey N. Bruce<sup>13</sup>, Wenya Linda Bi<sup>15,27</sup>, Praveen Agrawal<sup>28</sup>, Denis Schapiro<sup>7,29</sup>, Eva Hernando<sup>4</sup>, Evan Z. Macosko<sup>6,30</sup>, Fei Chen<sup>6,31</sup>, Gary K. Schwartz<sup>1</sup>, Benjamin Izar<sup>1,2,3,33,\*</sup>

## Affiliations

<sup>1</sup>Department of Medicine, Division of Hematology/Oncology, and Herbert Irving Comprehensive Cancer Center, Columbia University Irving Medical Center, New York, NY 10032, USA

<sup>2</sup>Program for Mathematical Genomics, Columbia University, New York, NY 10032, USA

<sup>3</sup>Columbia Center for Translational Immunology, Columbia University Irving Medical Center, New York, NY 10032, USA

<sup>4</sup>Department of Pathology, NYU Grossman School of Medicine, New York, NY 10016, USA

<sup>5</sup>Department of Systems Biology, Columbia University Irving Medical Center, New York, NY 10032, USA

<sup>6</sup>Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA

<sup>7</sup>Heidelberg University, Faculty of Medicine, and Heidelberg University Hospital, Institute for Computational Biomedicine, Bioquant, 69120 Heidelberg, Germany

<sup>8</sup>Department of Oncology UNIL CHUV, Lausanne Branch, Ludwig Institute for Cancer Research Lausanne, CHUV and University of Lausanne, Lausanne, 1066 Épalinges, Switzerland

<sup>9</sup>Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland

<sup>10</sup>Department of Genetics and Development, Columbia University Medical Center, New York, NY 10032, USA

<sup>11</sup>Department of Pathology and Cell Biology, Columbia University Medical Center, New York, NY 10032, USA



<sup>12</sup>Columbia University Vagelos College of Physicians and Surgeons, New York, NY 10032, USA

<sup>13</sup>Department of Neurological Surgery, New York Presbyterian/Columbia University Irving Medical Center, New York, NY 10032, USA

<sup>14</sup>Department of Cancer Immunology and Virology, Dana-Farber Cancer Center, Boston, MA 02215, USA

<sup>15</sup>Department of Neurosurgery, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA

<sup>16</sup>Division of Hematology/Oncology, Medstar Washington Cancer Institute, Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, Washington, DC 20007, USA

<sup>17</sup>Department of Surgery, University of Virginia, Charlottesville, VA, USA

<sup>18</sup>Department of Dermatology, Eberhard Karls University Tübingen, 72076 Tübingen, Germany

<sup>19</sup>Charité–Universitätsmedizin Berlin, Freie Universität Berlin and Humboldt-Universität zu Berlin, Department of Dermatology, Venereology and Allergology, 10117, Berlin, Germany

<sup>20</sup>Department of Melanoma Medical Oncology, MD Anderson Cancer Center, Houston, TX 77030, USA

<sup>21</sup>Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA

<sup>22</sup>Department of Pharmacology, Weill Cornell Graduate School, New York, NY 10065, USA

<sup>23</sup>Department of Radiation Oncology, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA

<sup>24</sup>Department of Biomedical Engineering, Columbia University, New York, NY 10027, USA

<sup>25</sup>Irving Institute for Cancer Dynamics, Columbia University, New York, NY 10027, USA

<sup>26</sup>Department of Medicine, Jonsson Comprehensive Cancer Center, University of California, Los Angeles (UCLA), Los Angeles, CA 90024, USA

<sup>27</sup>Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02215, USA

<sup>28</sup>Department of Molecular Pharmacology, Albert Einstein College of Medicine, Bronx, New York, NY 10461, USA

<sup>29</sup>Institute of Pathology, Heidelberg University Hospital, 69120 Heidelberg, Germany

<sup>30</sup>Department of Psychiatry, Massachusetts General Hospital, Boston, MA 02114, USA

<sup>31</sup>Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138, USA

<sup>32</sup>These authors contributed equally

<sup>33</sup>Lead contact

## ACKNOWLEDGEMENTS

We thank Robert Schwabe (Columbia University) for fruitful discussions, Iman Osman (Melanoma Program, NYU Langone Health) for providing the 12-273BM and 12-273LN melanoma short-term cultures, Robert S. Kerbel (University of Toronto) for providing the 4L and 5B1 melanoma cell lines. B.I. is supported by National Institute of Health (NIH) National Cancer Institute (NCI) grants K08CA222663, R37CA258829, U54CA225088 and R21CA263381 (with E.A. and R.C.), an American Cancer Society Research Scholar Grant, a Burroughs Wellcome Fund Career Award for Medical Scientists, a Velocity Fellows Award, the Louis V. Gerstner, Jr. Scholars Program, a V Foundation Scholar Award, a Columbia University RISE award (with E.A. and R.C.), and the Tara Miller Young Investigator Award by the Melanoma Research Alliance. L.A.C., P.H., Z.H.W. and M.M. are supported by a NIH grant T32GM007367. M.A.D. is supported by the Dr. Miriam and Sheldon G. Adelson Medical Research Foundation, the AIM at Melanoma Foundation, the NIH/NCI (P50CA221703 and U54CA224070), the American Cancer Society and the Melanoma Research Alliance, Cancer Fighters of Houston, the Anne and John Mendelsohn Chair for Cancer Research, and philanthropic contributions to the Melanoma Moon Shots Program of MD Anderson. A.A. is supported by a PhRMA Foundation Predoctoral Fellowship. M. Rö and T.E. are supported by Wilhelm Sander-Stiftung (2020.100.1), DFG RO 764/15-2, Cluster of Excellence iFIT (EXC 2180) “Image-Guided and Functionally Instructed Tumor Therapies”, University of Tübingen, Germany, funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy - EXC 2180 – 390900677 and acknowledge project support from S. Forchhammer and F. Fend. M.B.A. is supported by the NIH/NCI (P30CA051008), a Melanoma Research Alliance Senior Scientist Award and the William M. Scholl Chair for Cancer Research. D.S. is supported by the German Federal Ministry of Education and Research (BMBF 01ZZ2004). E.H. contributions were supported by a Melanoma Research Alliance (MRA)/American Cancer Society (ACS) Team award and the NYU Melanoma SPORE P50 CA225450. A.K. is supported by the NYU Medical Scientist Training program (MSTP; 5T32GM136573). We thank the NYULH Genomics Technology Center, Preclinical Imaging and Experimental Pathology Cores, which are partially supported by the Cancer Center Support Grant NIH/NCI P30 CA016087 to the Perlmutter Cancer Center at NYULH. Illustrations were created with [BioRender.com](https://BioRender.com). This work was supported by NIH/NCI Cancer Center Support Grant P30CA013696, including the Oncology Precision Therapeutics and Imaging Core and the Molecular Pathology Shared Resource and its Tissue Bank at Columbia University.

## REFERENCES

- Ackermann AM, Wang Z, Schug J, Naji A, and Kaestner KH (2016). Integration of ATAC-seq and RNA-seq identifies human alpha cell and beta cell signature genes. *Mol. Metab* 5, 233–244. [PubMed: 26977395]
- Adamson SE, Griffiths R, Moravec R, Senthivinayagam S, Montgomery G, Chen W, Han J, Sharma PR, Mullins GR, Gorski SA, et al. (2016). Disabled homolog 2 controls macrophage phenotypic polarization and adipose tissue inflammation. *J. Clin. Invest* 126, 1311–1322. [PubMed: 26927671]
- Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, Rambow F, Marine JC, Geurts P, Aerts J, et al. (2017). SCENIC: Single-cell regulatory network inference and clustering. *Nat. Methods* 14, 1083–1086. [PubMed: 28991892]
- Akbani R, Akdemir KC, Aksoy BA, Albert M, Ally A, Amin SB, Arachchi H, Arora A, Auman JT, Ayala B, et al. (2015). Genomic Classification of Cutaneous Melanoma. *Cell* 161, 1681–1696. [PubMed: 26091043]
- Alberti G (2021). GmAMisc: “Gianmarco Alberti” Miscellaneous. R package version 1.2.0.
- Alvarez MJ, Shen Y, Giorgi FM, Lachmann A, Ding BB, Hilda Ye B, and Califano A (2016). Functional characterization of somatic mutations in cancer using network-based inference of protein activity. *Nat. Genet* 48, 838–847. [PubMed: 27322546]

- Amemiya HM, Kundaje A, and Boyle AP (2019). The ENCODE Blacklist: Identification of Problematic Regions of the Genome. *Sci. Reports* 2019 91 9, 1–5.
- Andreatta M, and Carmona S (2021). STACAS: Sub-Type Anchor Correction for Alignment in Seurat to integrate single-cell RNA-seq data. *Bioinformatics* 37, 882–884. [PubMed: 32845323]
- Andreatta M, Corria-Osorio J, Müller S, Cubas R, Coukos G, and Carmona SJ (2021). Interpretation of T cell states from single-cell transcriptomics data using reference atlases. *Nat. Commun* 12, 2965. [PubMed: 34017005]
- Angerer P, Haghverdi L, Büttner M, Theis FJ, Marr C, and Buettner F (2016). Destiny: Diffusion maps for large-scale single-cell data in R. *Bioinformatics* 32, 1241–1243. [PubMed: 26668002]
- Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, Chak S, Naikawadi RP, Wolters PJ, Abate AR, et al. (2019). Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat. Immunol* 20, 163–172. [PubMed: 30643263]
- Azizi E, Carr AJ, Plitas G, Cornish AE, Konopacki C, Prabhakaran S, Nainys J, Wu K, Kisieliovas V, Setty M, et al. (2018). Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. *Cell* 174, 1293–1308.e36. [PubMed: 29961579]
- Bakhoun SF, and Cantley LC (2018). The Multifaceted Role of Chromosomal Instability in Cancer and Its Microenvironment. *Cell* 174, 1347–1360. [PubMed: 30193109]
- Bakhoun SF, Ngo B, Laughney AM, Cavallo J-A, Murphy CJ, Ly P, Shah P, Sriram RK, Watkins TBK, Taunk NK, et al. (2018). Chromosomal instability drives metastasis through a cytosolic DNA response. *Nat.* 2018 5537689 553, 467–472.
- Bankhead P, Loughrey MB, Fernández JA, Dombrowski Y, McArt DG, Dunne PD, McQuaid S, Gray RT, Murray LJ, Coleman HG, et al. (2017). QuPath: Open source software for digital pathology image analysis. *Sci. Rep* 7.
- Barkas N, Petukhov V, Nikolaeva D, Lozinsky Y, Demharter S, Khodosevich K, and Kharchenko PV (2019). Joint analysis of heterogeneous single-cell RNA-seq dataset collections. *Nat. Methods* 2019 168 16, 695–698. [PubMed: 31308548]
- Basso K, Margolin AA, Stolovitzky G, Klein U, Dalla-Favera R, and Califano A (2005). Reverse engineering of regulatory networks in human B cells. *Nat. Genet* 2005 374 37, 382–390. [PubMed: 15778709]
- Bentsen M, Goymann P, Schultheis H, Klee K, Petrova A, Wiegandt R, Fust A, Preussner J, Kuenne C, Braun T, et al. (2020). ATAC-seq footprinting unravels kinetics of transcription factor binding during zygotic genome activation. *Nat. Commun* 2020 111 11, 1–11. [PubMed: 31911652]
- Bergen V, Lange M, Peidli S, Wolf FA, and Theis FJ (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat. Biotechnol* 2020 3812 38, 1408–1414. [PubMed: 32747759]
- Bolger AM, Lohse M, and Usadel B (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114. [PubMed: 24695404]
- Brastianos PK, Curry WT, and Oh KS (2013). Clinical Discussion and Review of the Management of Brain Metastases. *J. Natl. Compr. Cancer Netw* 11, 1153–1164.
- Brastianos PK, Carter SL, Santagata S, Cahill DP, Taylor-Weiner A, Jones RT, Van Allen EM, Lawrence MS, Horowitz PM, Cibulskis K, et al. (2015). Genomic Characterization of Brain Metastases Reveals Branched Evolution and Potential Therapeutic Targets. *Cancer Discov.* 5, 1164–1177. [PubMed: 26410082]
- Bray NL, Pimentel H, Melsted P, and Pachter L (2016). Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol* 2016 345 34, 525–527. [PubMed: 27043002]
- Brunet JP, Tamayo P, Golub TR, and Mesirov JP (2004). Metagenes and molecular pattern discovery using matrix factorization. *Proc. Natl. Acad. Sci* 101, 4164–4169. [PubMed: 15016911]
- Buenrostro JD, Giresi PG, Zaba LC, Chang HY, and Greenleaf WJ (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* 2013 1012 10, 1213–1218. [PubMed: 24097267]
- Butler A, Hoffman P, Smibert P, Papalexi E, and Satija R (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol* 36, 411–420. [PubMed: 29608179]

- Cable DM, Murray E, Zou LS, Goeva A, Macosko EZ, Chen F, and Irizarry RA (2021). Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat. Biotechnol*
- Cable DM, Murray E, Shanmugam V, Zhang S, Diao M, Chen H, Macosko EZ, Irizarry RA, and Chen F (2022). Cell type-specific inference of differential expression in spatial transcriptomics. *BioRxiv* 2021.12.26.474183.
- Cahoy JD, Emery B, Kaushal A, Foo LC, Zamanian JL, Christopherson KS, Xing Y, Lubischer JL, Krieg PA, Krupenko SA, et al. (2008). A transcriptome database for astrocytes, neurons, and oligodendrocytes: A new resource for understanding brain development and function. *J. Neurosci* 28, 264–278. [PubMed: 18171944]
- Cassetta L, Fragkogianni S, Sims AH, Swierczak A, Forrester LM, Zhang H, Soong DYH, Cotechini T, Anur P, Lin EY, et al. (2019). Human Tumor-Associated Macrophage and Monocyte Transcriptional Landscapes Reveal Cancer-Specific Reprogramming, Biomarkers, and Therapeutic Targets. *Cancer Cell* 35, 588–602.e10. [PubMed: 30930117]
- Castro-Mondragon JA, Riudavets-Puig R, Rauluseviciute I, Berhanu Lemma R, Turchi L, Blanc-Mathieu R, Lucas J, Boddie P, Khan A, Manosalva Pérez N, et al. (2022). JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* 50, D165–D173. [PubMed: 34850907]
- Chen G, Chakravarti N, Aardalen K, Lazar AJ, Tetzlaff MT, Wubbenhorst B, Kim SB, Kopetz S, Ledoux AA, Vashisht Gopal YN, et al. (2014). Molecular profiling of patient-matched brain and extracranial melanoma metastases implicates the PI3K pathway as a therapeutic target. *Clin. Cancer Res* 20, 5537–5546. [PubMed: 24803579]
- Cheng Q, Wu J, Zhang Y, Liu X, Xu N, Zuo F, and Xu J (2017). SOX4 promotes melanoma cell migration and invasion through the activation of the NF- $\kappa$ B signaling pathway. *Int. J. Mol. Med* 40, 447–453. [PubMed: 28627651]
- Cohen M, Giladi A, Barboy O, Hamon P, Li B, Zada M, Gurevich-Shapiro A, Beccaria CG, David E, Maier BB, et al. (2022). The interaction of CD4 + helper T cells with dendritic cells shapes the tumor microenvironment and immune checkpoint blockade response. *Nat. Cancer*
- Cox J, and Mann M (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol* 2008 2612 26, 1367–1372. [PubMed: 19029910]
- Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, and Mann M (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res* 10, 1794–1805. [PubMed: 21254760]
- Cruz-Munoz W, Man S, Xu P, and Kerbel RS (2008). Development of a preclinical model of spontaneous human melanoma central nervous system metastasis. *Cancer Res.* 68, 4500–4505. [PubMed: 18559492]
- Davies MA, Liu P, McIntyre S, Kim KB, Papadopoulos N, Hwu W-J, Hwu P, and Bedikian A (2011). Prognostic factors for survival in melanoma patients with brain metastases. *Cancer* 117, 1687–1696. [PubMed: 20960525]
- Ding H, Douglass EF, Sonabend AM, Mela A, Bose S, Gonzalez C, Canoll PD, Sims PA, Alvarez MJ, and Califano A (2018). Quantitative assessment of protein activity in orphan tissues and single cells using the metaVIPER algorithm. *Nat. Commun* 9.
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. [PubMed: 23104886]
- Dolgalev I (2021). msigdb: MSigDB Gene Sets for Multiple Organisms in a Tidy Data Format.
- Dolgalev I (2022). Seq-N-Slide.
- Dudda JC, Salaun B, Ji Y, Palmer DC, Monnot GC, Merck E, Boudousquie C, Utschneider DT, Escobar TM, Perret R, et al. (2013). MicroRNA-155 is required for effector CD8+ T cell responses to virus infection and cancer. *Immunity* 38, 742–753. [PubMed: 23601686]
- Durinck S, Spellman PT, Birney E, and Huber W (2009). Mapping identifiers for the integration of genomic datasets with the R/ Bioconductor package biomaRt. *Nat. Protoc* 4, 1184–1191. [PubMed: 19617889]

- Eichler AF, Chung E, Kodack DP, Loeffler JS, Fukumura D, and Jain RK (2011). The biology of brain metastases—translation to new therapies. *Nat. Rev. Clin. Oncol* 2011 86 8, 344–356. [PubMed: 21487419]
- Fallahi-Sichani M, Becker V, Izar B, Baker GJ, Lin J, Boswell SA, Shah P, Rotem A, Garraway LA, and Sorger PK (2017). Adaptive resistance of melanoma cells to RAF inhibition via reversible induction of a slowly dividing de-differentiated state. *Mol. Syst. Biol* 13, 905. [PubMed: 28069687]
- Federico A, and Monti S (2020). HypeR: An R package for geneset enrichment workflows. *Bioinformatics* 36, 1307–1308. [PubMed: 31498385]
- Feng J, Liu T, Qin B, Zhang Y, and Liu XS (2012). Identifying ChIP-seq enrichment using MACS. *Nat. Protoc* 2012 79 7, 1728–1740. [PubMed: 22936215]
- Fischer GM, Jalali A, Kircher DA, Lee WC, McQuade JL, Haydu LE, Joon AY, Reuben A, de Macedo MP, Carapeto FCL, et al. (2019). Molecular profiling reveals unique immune and metabolic features of melanoma brain metastases. *Cancer Discov.* 9, 628–645. [PubMed: 30787016]
- Fischer GM, Guerrieri RA, Hu Q, Joon AY, Kumar S, Haydu LE, McQuade JL, Gopal YNV, Knighton B, Deng W, et al. (2021). Clinical, molecular, metabolic, and immune features associated with oxidative phosphorylation in melanoma brain metastases. *Neuro-Oncology Adv.* 3.
- Fisher RA (1992). *Statistical Methods for Research Workers.* 66–70.
- Fleming S, Marioni J, and Babadi M (2019). CellBender remove-background: a deep generative model for unsupervised removal of background noise from scRNA-seq datasets. *BioRxiv* 791699.
- Foroutan M, Bhuva DD, Lyu R, Horan K, Cursons J, and Davis MJ (2018). Single sample scoring of molecular phenotypes. *BMC Bioinforma.* 2018 191 19, 1–10.
- Fukumura K, Malgulwar PB, Fischer GM, Hu X, Mao X, Song X, Hernandez SD, Zhang XH-F, Zhang J, Parra ER, et al. (2021). Multi-omic molecular profiling reveals potentially targetable abnormalities shared across multiple histologies of brain metastasis. *Acta Neuropathol.* 141, 303–321. [PubMed: 33394124]
- Garraway LA, Widlund HR, Rubin MA, Getz G, Berger AJ, Ramaswamy S, Beroukhim R, Milner DA, Granter SR, Du J, et al. (2005). Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma. *Nat.* 2005 4367047 436, 117–122.
- Godfrey DI, Koay H-F, McCluskey J, and Gherardin NA (2019). The biology and functional importance of MAIT cells. *Nat. Immunol* 2019 209 20, 1110–1128. [PubMed: 31406380]
- Gonzalez-Dominguez E, Samaniego R, Flores-Sevilla JL, Campos-Campos SF, Gomez-Campos G, Salas A, Campos-Pena V, Corbi AL, Sanchez-Mateos P, and Sanchez-Torres C (2015). CD163L1 and CLEC5A discriminate subsets of human resident and inflammatory macrophages in vivo. *J. Leukoc. Biol* 98, 453–466. [PubMed: 25877931]
- Gonzalez H, Mei W, Robles I, Spitzer MH, Roose JP, Werb Z, Hagerling C, Allen BM, Line T, Okholm H, et al. (2022). Cellular architecture of human brain metastases. *Cell* 0, 1–17.
- Greenwald NF, Miller G, Moen E, Kong A, Kagel A, Dougherty T, Fullaway CC, McIntosh BJ, Leow KX, Schwartz MS, et al. (2021). Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. *Nat. Biotechnol*
- Hafemeister C, and Satija R (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* 20.
- Ho P, Melms JC, Rogava M, Frangieh CJ, Shah SB, Walsh Z, Kyrasyuk O, Amin AD, Caprio L, Fullerton BT, et al. (2022). The CD58:CD2 axis is co-regulated with PD-L1 via CMTM6 and governs anti-tumor immunity. *BioRxiv* 2022.03.21.485049.
- Hoek KS, Eichhoff OM, Schlegel NC, Döbbling U, Kobert N, Schaerer L, Hemmi S, and Dummer R (2008). In vivo switching of human melanoma cells between proliferative and invasive states. *Cancer Res.* 68, 650–656. [PubMed: 18245463]
- Jerby-Arnon L, Shah P, Cuoco MS, Rodman C, Su MJ, Melms JC, Leeson R, Kanodia A, Mei S, Lin JR, et al. (2018). A Cancer Cell Program Promotes T Cell Exclusion and Resistance to Checkpoint Blockade. *Cell* 175, 984–997.e24. [PubMed: 30388455]
- Kaur P, Nagar S, Bhagwat M, Uddin M, Zhu Y, Vancurova I, and Vancura A (2021). Activated heme synthesis regulates glycolysis and oxidative metabolism in breast and ovarian cancer cells. *PLoS One* 16, e0260400. [PubMed: 34807950]

- Khan O, Giles JR, McDonald S, Manne S, Ngoi SF, Patel KP, Werner MT, Huang AC, Alexander KA, Wu JE, et al. (2019). TOX transcriptionally and epigenetically programs CD8+ T cell exhaustion. *Nature* 571, 211–218. [PubMed: 31207603]
- Kim D, Paggi JM, Park C, Bennett C, and Salzberg SL (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol* 2019 378 37, 907–915. [PubMed: 31375807]
- Kleffman K, Levinson G, Rose IVL, Blumenberg LM, Shadaloey SAA, Dhabaria A, Wong E, Galan-Echevarria F, Karz A, Argibay D, et al. (2022). Melanoma-secreted Amyloid Beta Suppresses Neuroinflammation and Promotes Brain Metastasis. *Amyloid Beta Promotes Brain Metastasis. Cancer Discov.* 4, 10016.
- Kluyver T, Ragan-Kelley B, Pérez F, Granger B, Bussonnier M, Frederic J, Kelley K, Hamrick J, Grout J, Corlay S, et al. (2016). Jupyter Notebooks—a publishing format for reproducible computational workflows. *Position. Power Acad. Publ. Play Agents Agendas - Proc. 20th Int. Conf. Electron. Publ. ELPUB* 2016 87–90.
- Konieczkowski DJ, Johannessen CM, Abudayyeh O, Kim JW, Cooper ZA, Piris A, Frederick DT, Barzily-Rokni M, Straussman R, Haq R, et al. (2014). A melanoma cell state distinction influences sensitivity to MAPK pathway inhibitors. *Cancer Discov.* 4, 816–827. [PubMed: 24771846]
- Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, Baglaenko Y, Brenner M, Loh P. ru, and Raychaudhuri S (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* 16, 1289–1296. [PubMed: 31740819]
- Lachmann A, Giorgi FM, Lopez G, and Califano A (2016). ARACNe-AP: gene network reverse engineering through adaptive partitioning inference of mutual information. *Bioinformatics* 32, 2233–2235. [PubMed: 27153652]
- Landsberg J, Kohlmeyer J, Renn M, Bald T, Rogava M, Cron M, Fatho M, Lennerz V, Wölfel T, Hölzel M, et al. (2012). Melanomas resist T-cell therapy through inflammation-induced reversible dedifferentiation. *Nature* 490, 412–416. [PubMed: 23051752]
- Lee DD, and Seung HS (1999). Learning the parts of objects by non-negative matrix factorization. *Nature* 401, 788–791. [PubMed: 10548103]
- Lefebvre C, Rajbhandari P, Alvarez MJ, Bandaru P, Lim WK, Sato M, Wang K, Sumazin P, Kustagi M, Bisikirska BC, et al. (2010). A human B-cell interactome identifies MYB and FOXM1 as master regulators of proliferation in germinal centers. *Mol. Syst. Biol* 6.
- Lein ES, Hawrylycz MJ, Ao N, Ayres M, Bensinger A, Bernard A, Boe AF, Boguski MS, Brockway KS, Byrnes EJ, et al. (2007). Genome-wide atlas of gene expression in the adult mouse brain. *Nature* 445, 168–176. [PubMed: 17151600]
- Li H, van der Leun AM, Yofe I, Lubling Y, Gelbard-Solodkin D, van Akkooi ACJ, van den Braber M, Rozeman EA, Haanen JBAG, Blank CU, et al. (2019). Dysfunctional CD8 T Cells Form a Proliferative, Dynamically Regulated Compartment within Human Melanoma. *Cell* 176, 775–789.e18. [PubMed: 30595452]
- Liao Y, Smyth GK, and Shi W (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. [PubMed: 24227677]
- Linkert M, Rueden CT, Allan C, Burel JM, Moore W, Patterson A, Loranger B, Moore J, Neves C, MacDonald D, et al. (2010). Metadata matters: access to image data in the real world. *J. Cell Biol* 189, 777–782. [PubMed: 20513764]
- Liu D, Lin J-R, Robitschek EJ, Kasumova GG, Heyde A, Shi A, Kraya A, Zhang G, Moll T, Frederick DT, et al. (2021). Evolution of delayed resistance to immunotherapy in a melanoma responder. *Nat. Med* 27, 985–992. [PubMed: 33941922]
- Love MI, Huber W, and Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014 1512 15, 1–21.
- Ma S, Zhang B, LaFave LM, Earl AS, Chiang Z, Hu Y, Ding J, Brack A, Kartha VK, Tay T, et al. (2020). Chromatin Potential Identified by Shared Single-Cell Profiling of RNA and Chromatin. *Cell* 183, 1103–1116.e20. [PubMed: 33098772]
- La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, Lidschreiber K, Kastri ME, Lönnerberg P, Furlan A, et al. (2018). RNA velocity of single cells. *Nat.* 2018 5607719 560, 494–498.

- Marsh SE, Walker AJ, Kamath T, Dissing-Olesen L, Hammond TR, de Soysa TY, Young AMH, Murphy S, Abdulraouf A, Nadaf N, et al. (2022). Dissection of artifactual and confounding glial signatures by single-cell sequencing of mouse and human brain. *Nat. Neurosci* 2022 253 25, 306–316. [PubMed: 35260865]
- Martin M (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal* 17, 10–12.
- McCarthy DJ, Campbell KR, Lun ATL, and Wills QF (2017). Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* 33, 1179–1186. [PubMed: 28088763]
- McGinnis CS, Murrow LM, and Gartner ZJ (2019). DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst.* 8, 329–337.e4. [PubMed: 30954475]
- Melms JC, Vallabhaneni S, Mills CE, Yapp C, Chen JY, Morelli E, Waszyk P, Kumar S, Deming D, Moret N, et al. (2020). Inhibition of Haspin Kinase Promotes Cell-Intrinsic and Extrinsic Antitumor Activity. *Cancer Res.* 80, 798–810. [PubMed: 31882401]
- Melms JC, Biermann J, Huang H, Wang Y, Nair A, Tagore S, Katsyv I, Rendeiro AF, Amin AD, Schapiro D, et al. (2021). A molecular single-cell lung atlas of lethal COVID-19. *Nat.* 2021 5957865 595, 114–119.
- Meylan M, Petitprez F, Becht E, Bougouin A, Pupier G, Calvez A, Giglioli I, Verkarre V, Lacroix G, Verneau J, et al. (2022). Tertiary lymphoid structures generate and propagate anti-tumor antibody-producing plasma cells in renal cell cancer. *Immunity* 55.
- de Miera EVS, Friedman EB, Greenwald HS, Perle MA, and Osman I (2012). Development of five new melanoma low passage cell lines representing the clinical and genetic profile of their tumors of origin. *Pigment Cell Melanoma Res.* 25, 395–397. [PubMed: 22404973]
- Nguyen B, Fong C, Luthra A, Smith SA, DiNatale RG, Nandakumar S, Walch H, Chatila WK, Madupuri R, Kundra R, et al. (2021). Genomic characterization of metastatic patterns from prospective clinical sequencing of 25,000 patients. *BioRxiv* 2021.06.28.450217.
- Nilsen G, Liestøl K, Van Loo P, Moen Vollan HK, Eide MB, Rueda OM, Chin S-F, Russell R, Baumbusch LO, Caldas C, et al. (2012). Copynumber: Efficient algorithms for single- and multi-track copy number segmentation. *BMC Genomics* 13.
- Olah M, Menon V, Habib N, Taga MF, Ma Y, Yung CJ, Cimpean M, Khairallah A, Coronas-Samano G, Sankowski R, et al. (2020). Single cell RNA sequencing of human microglia uncovers a subset associated with Alzheimer’s disease. *Nat. Commun* 11.
- Opdecamp K, Nakayama A, Nguyen MT, Hodgkinson CA, Pavan WJ, and Arnheiter H (1997). Melanocyte development in vivo and in neural crest cell cultures: crucial dependence on the Mitf basic-helix-loop-helix-zipper transcription factor. *Development* 124, 2377–2386. [PubMed: 9199364]
- Palla G, Spitzer H, Klein M, Fischer D, Schaar AC, Kuemmerle LB, Rybakov S, Ibarra IL, Holmberg O, Virshup I, et al. (2022). Squidpy: a scalable framework for spatial omics analysis. *Nat. Methods* 19, 171–178. [PubMed: 35102346]
- Patil NS, Nabet BY, Müller S, Koepfen H, Zou W, Giltane J, Au-Yeung A, Srivats S, Cheng JH, Takahashi C, et al. (2022). Intratumoral plasma cells predict outcomes to PD-L1 blockade in non-small cell lung cancer. *Cancer Cell* 40, 289–300.e4. [PubMed: 35216676]
- Qiu X, Feit AS, Feiglin A, Xie Y, Kesten N, Taing L, Perkins J, Gu S, Li Y, Cejas P, et al. (2021). CoBRA: Containerized Bioinformatics Workflow for Reproducible ChIP/ATAC-seq Analysis. *Genomics. Proteomics Bioinformatics*
- Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F, and Manke T (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* 44, W160–W165. [PubMed: 27079975]
- Recalcati S, Locati M, Marini A, Santambrogio P, Zaninotto F, De Pizzol M, Zambataro L, Girelli D, and Cairo G (2010). Differential regulation of iron homeostasis during human macrophage polarized activation. *Eur. J. Immunol* 40, 824–835. [PubMed: 20039303]

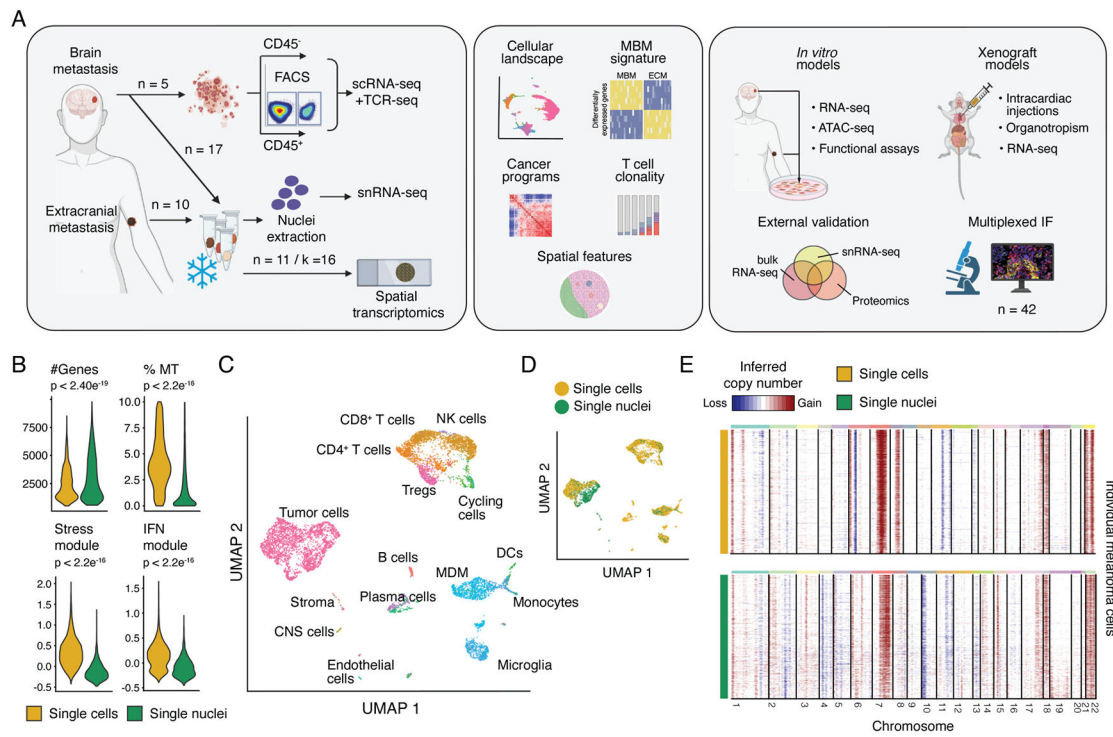
- Schapiro D, Sokolov A, Yapp C, Chen YA, Mühlich JL, Hess J, Creason AL, Nirmal AJ, Baker GJ, Nariya MK, et al. (2021). MCMICRO: a scalable, modular image-processing pipeline for multiplexed tissue imaging. *Nat. Methods* 2021 193 19, 311–315. [PubMed: 34824477]
- Skytte MK, Graversen JH, and Moestrup SK (2020). Targeting of CD163+ Macrophages in Inflammatory and Malignant Diseases. *Int. J. Mol. Sci* 2020, Vol. 21, Page 5497 21, 5497.
- Slyper M, Porter CBM, Ashenberg O, Waldman J, Drokhlyansky E, Wakiro I, Smillie C, Smith-Rosario G, Wu J, Dionne D, et al. (2020). A single-cell and single-nucleus RNA-Seq toolbox for fresh and frozen human tumors. *Nat. Med* 26, 792–802. [PubMed: 32405060]
- Solinas G, Schiarea S, Liguori M, Fabbri M, Pesce S, Zammataro L, Pasqualini F, Nebuloni M, Chiabrando C, Mantovani A, et al. (2010). Tumor-conditioned macrophages secrete migration-stimulating factor: a new marker for M2-polarization, influencing tumor cell motility. *J. Immunol* 185, 642–652. [PubMed: 20530259]
- Soneson C, Love MI, and Robinson MD (2016). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Research* 2016 41521 4, 1521.
- Stickels RR, Murray E, Kumar P, Li J, Marshall JL, Di Bella DJ, Arlotta P, Macosko EZ, and Chen F (2021). Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. *Nat. Biotechnol* 39, 313–319. [PubMed: 33288904]
- Stuart T, and Satija R (2019). Integrative single-cell analysis. *Nat. Rev. Genet* 20, 257–272. [PubMed: 30696980]
- Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, Hao Y, Stoeckius M, Smibert P, and Satija R (2019). Comprehensive Integration of Single-Cell Data. *Cell* 177, 1888–1902.e21. [PubMed: 31178118]
- Sun Q, Lee W, Mohri Y, Takeo M, Lim CH, Xu X, Myung P, Atit RP, Taketo MM, Moubarak RS, et al. (2019). A novel mouse model demonstrates that oncogenic melanocyte stem cells engender melanoma resembling human disease. *Nat. Commun* 10.
- Tagore S, Wang Y, Biermann J, Rabadan R, Azizi E, and Izar B (2022). KINOMO: A non-negative matrix factorization framework for recovering intra- and inter-tumoral heterogeneity from single-cell RNA-seq data. *BioRxiv* 2022.05.02.490362.
- Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, and Prins P (2015). Sambamba: fast processing of NGS alignment formats. *Bioinformatics* 31, 2032–2034. [PubMed: 25697820]
- Tawbi HA, Forsyth PA, Algazi A, Hamid O, Hodi FS, Moschos SJ, Khushalani NI, Lewis K, Lao CD, Postow MA, et al. (2018). Combined Nivolumab and Ipilimumab in Melanoma Metastatic to the Brain. *N. Engl. J. Med* 379, 722. [PubMed: 30134131]
- Tirosh I, Izar B, Prakadan SM, Wadsworth MH, Treacy D, Trombetta JJ, Rotem A, Rodman C, Lian C, Murphy G, et al. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* (80-. ). 352, 189–196.
- Di Tommaso P, Chatzou M, Floden EW, Barja PP, Palumbo E, and Notredame C (2017). Nextflow enables reproducible computational workflows. *Nat. Biotechnol* 2017 354 35, 316–319. [PubMed: 28398311]
- Tsoi J, Robert L, Paraiso K, Galvan C, Sheu KM, Lay J, Wong DJL, Atefi M, Shirazi R, Wang X, et al. (2018). Multi-stage Differentiation Defines Melanoma Subtypes with Differential Vulnerability to Drug-Induced Iron-Dependent Oxidative Stress. *Cancer Cell* 33, 890–904.e5. [PubMed: 29657129]
- Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, Mann M, and Cox J (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* 2016 139 13, 731–740. [PubMed: 27348712]
- Venkatraman E, and Olshen A (2007). A faster circular binary segmentation algorithm for the analysis of array CGH data. *Bioinformatics* 23, 657–663. [PubMed: 17234643]
- Wang Q, Gu L, Adey A, Radlwimmer B, Wang W, Hovestadt V, Bähr M, Wolf S, Shendure J, Eils R, et al. (2013). Tagmentation-based whole-genome bisulfite sequencing. *Nat. Protoc* 2013 810 8, 2022–2032. [PubMed: 24071908]
- Wang Y, Yan K, Lin J, Li J, and Bi J (2021). Macrophage M2 Co-expression Factors Correlate With the Immune Microenvironment and Predict Outcome of Renal Clear Cell Carcinoma. *Front. Genet* 12.



- Wang Y, Fan JL, Melms JC, Amin AD, Georgis Y, Ho P, Tagore S, Abril-Rodríguez G, Biermann J, Hofree M, et al. (2022). Multi-modal single-cell and whole-genome sequencing of minute, frozen specimens to propel clinical applications. *BioRxiv* 2022.02.13.480272.
- Wei J, Marisetty A, Schrand B, Gabrusiewicz K, Hashimoto Y, Ott M, Grami Z, Kong L-Y, Ling X, Caruso H, et al. (2019). Osteopontin mediates glioblastoma-associated macrophage infiltration and is a potential therapeutic target. *J. Clin. Invest* 129, 137–149. [PubMed: 30307407]
- Wingett SW, and Andrews S (2018). FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Research* 7, 1338. [PubMed: 30254741]
- Wingrove E, Liu ZZ, Patel KD, Arnal-Estapé A, Cai WL, Melnick MA, Politi K, Monteiro C, Zhu L, Valiente M, et al. (2019). Transcriptomic Hallmarks of Tumor Plasticity and Stromal Interactions in Brain Metastasis. *Cell Rep.* 27, 1277–1292.e7. [PubMed: 31018140]
- Winn NC, Volk KM, and Hasty AH (2020). Regulation of tissue iron homeostasis: the macrophage “ferrostat”. *JCI Insight* 5, 132964. [PubMed: 31996481]
- Wolf FA, Angerer P, and Theis FJ (2018). SCANPY: Large-scale single-cell gene expression data analysis. *Genome Biol.* 19, 1–5. [PubMed: 29301551]
- Wolock SL, Lopez R, and Klein AM (2019). Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst.* 8, 281–291.e9. [PubMed: 30954476]
- Yunna C, Mengru H, Lei W, and Weidong C (2020). Macrophage M1/M2 polarization. *Eur. J. Pharmacol* 877.
- Zhao X, Shan Q, and Xue H-H (2021). TCF1 in T cell immunity: a broadened frontier. *Nat. Rev. Immunol*
- Zilionis R, Engblom C, Pfirschke C, Savova V, Zemmour D, Saaticioglu HD, Krishnan I, Maroni G, Meyerovitz CV, Kerwin CM, et al. (2019). Single-Cell Transcriptomics of Human and Mouse Lung Cancers Reveals Conserved Myeloid Populations across Individuals and Species. *Immunity* 50.
- Zurac S, Neagu M, Constantin C, Cioplea M, Nedelcu R, Bastian A, Popp C, Nichita L, Andrei R, Tebeica T, et al. (2016). Variations in the expression of TIMP1, TIMP2 and TIMP3 in cutaneous melanoma with regression and their possible function as prognostic predictors. *Oncol. Lett* 11, 3354–3360. [PubMed: 27123116]

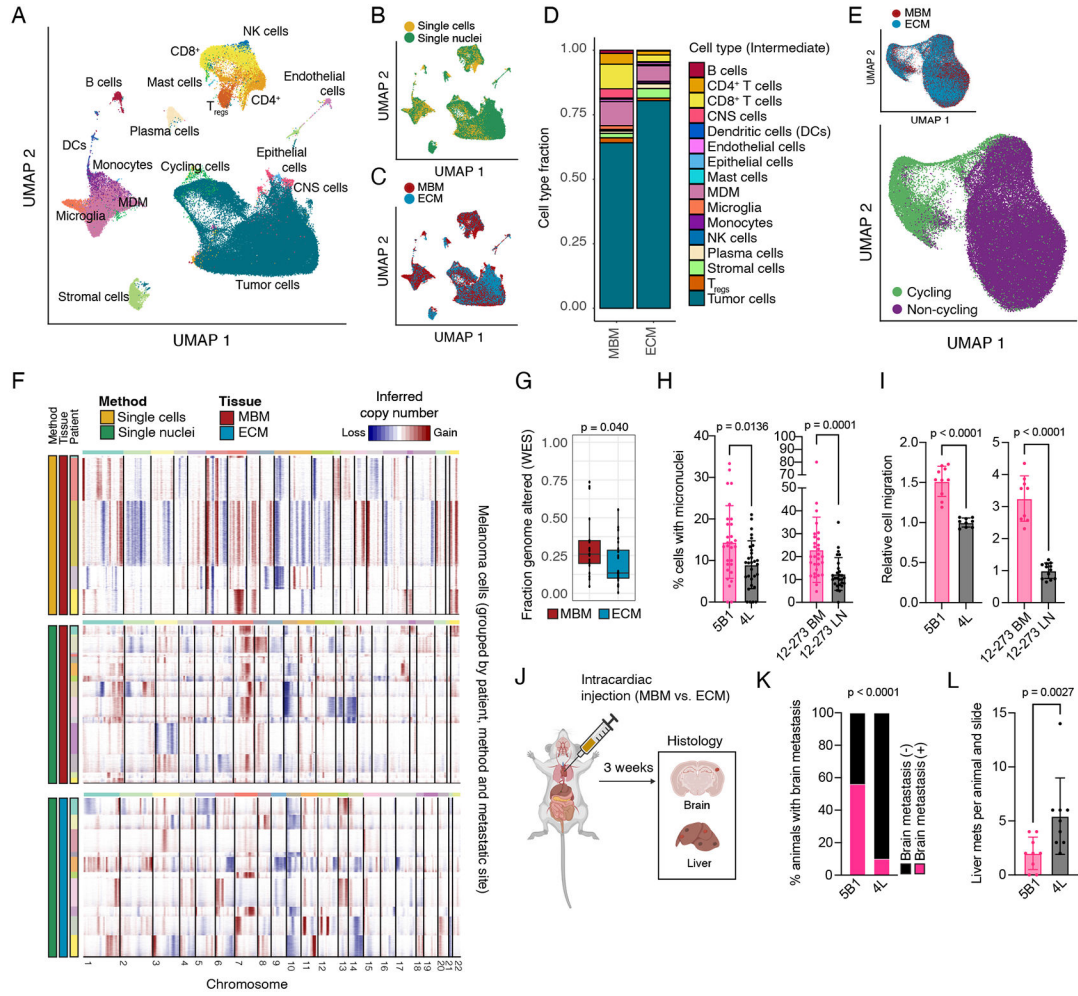
**Highlights:**

- Single-cell/spatial transcriptomics atlas of melanoma brain and extracranial metastases
- Chromosomal instability is associated with brain metastasis
- Cancer cells in brain metastases enrich for a neuronal-like metaprogram
- Macrophages have a pro-tumorigenic phenotype in brain metastases

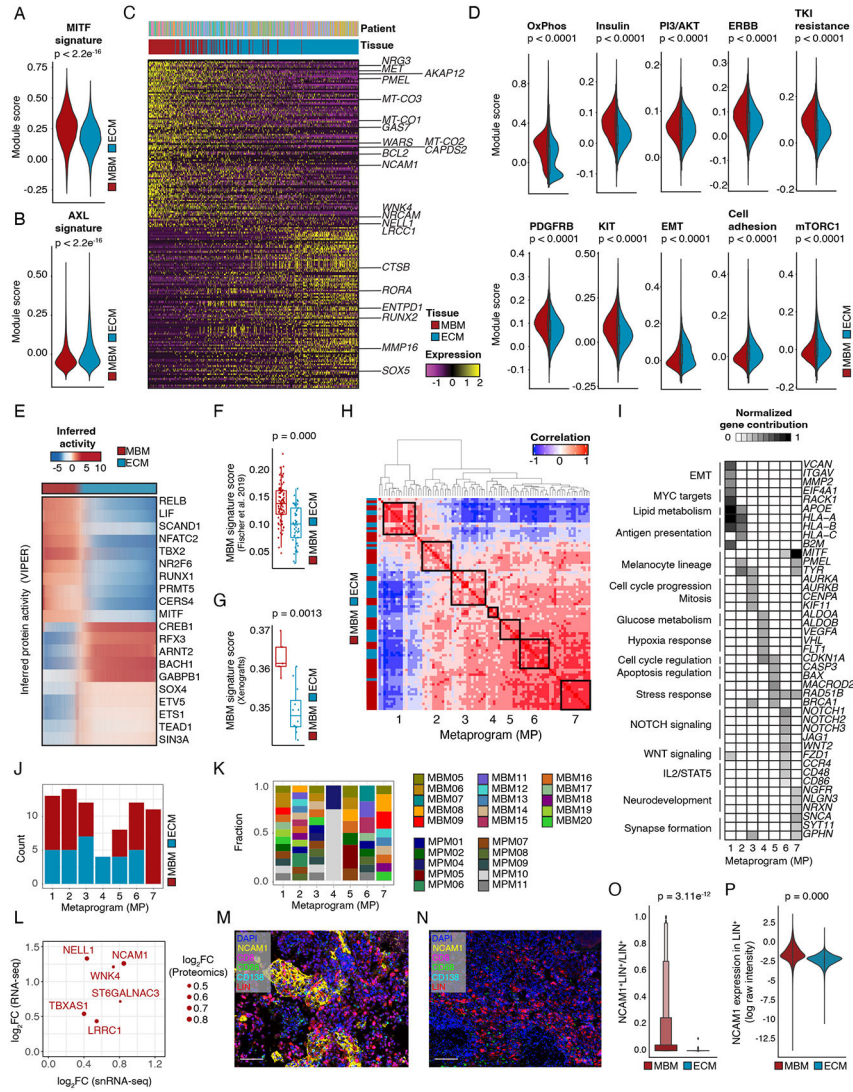


**Figure 1. Study design and comparison of fresh vs. frozen profiling**

(A) Study design including specimens, analyses and validation approaches. (B) Quality-control parameters of a matched MBM specimen profiled with sc/snRNA-seq. (C,D) UMAP embedding with (C) cell type assignment and (D) profiling method. (E) Inferred copy number alterations (CNAs) from a matched MBM specimen processed with scRNA-seq (top) and snRNA-seq (bottom).



**Figure 2. Cellular and genomic landscape of MBM**  
 (A) UMAP of integrated transcriptomes of 22 MBM and 10 ECM samples showing cell type assignment, (B) sequencing, and (C) metastatic site. (D) Cell type distribution of MBM/ECM profiled by snRNA-seq. (E) UMAP of integrated malignant cell transcriptomes profiled by snRNA-seq data and indicated cycling status and tissue origin (inset). (F) CNA inference in cancer cells (rows) across chromosomes (columns) with amplifications (red) and deletions (blue) based on method (far left bar), tissue site (middle bar) and patient (inner bar). (G) Fraction of genome altered in MBM<sub>WES</sub> and ECM<sub>WES</sub> from Fischer et al. (H) Frequency of micronuclei per visual field in patient-derived MBM (5B1 and 12-273 BM) or ECM (4L and 12-273 LN, respectively) cell line cultures. (I) Relative *in vitro* migration of MBM/ECM models as in (H). (J) Design for *in vivo* experiments using 5B1 and 4L models. (K) Frequency of animals with/without MBM and (L) ECM burden from experiment in 5B1 and 4L injection as in (J). Bars, mean±SD. Wilcoxon rank-sum test.



**Figure 3. Cancer cell specific programs of MBM profiled by snRNA-seq**  
 (A,B) Module scores of (A) MITF and (B) AXL programs in melanoma cells of MBM/ECM. Wilcoxon rank-sum test. (C) DEGs between MBM and ECM cancer cells. Patient and tissue origin of individual cells (ticks) indicated on top bar. Selected genes (rows) indicated with gene name. (D) Pathway module scores of selected, differentially enriched pathways (adjusted  $p < 0.0001$ ) in MBM/ECM melanoma cells. (E) Differential inferred protein activity in melanoma cells from MBM/ECM (indicated by ticks on top bar) with selected rows (proteins) highlighted. (F,G) Scoring of MBM signature on MBM<sub>bulk</sub>/ECM<sub>bulk</sub> from Fischer et al. and (G) mouse xenograft-derived MBM/ECM transcriptomes from 5B1/4L and 12-273 BM/LN; Wilcoxon rank-sum test. (H) Spearman correlation of individual programs from cancer cells and metaprograms (MP) (black boxes) using KINOMO; left bar indicating tissue origin. (I) Normalized gene contribution of MPs (columns) identified in (H) and biological function (left). (J,K) Count of programs from tissue sites (J) and individual samples (K) in metaprograms from (H). (L) Log<sub>2</sub>FC of DEGs in MBM with recurrent identification in snRNA-seq (this study; x-axis), bulk RNA-seq (Fischer et al.; y-axis)

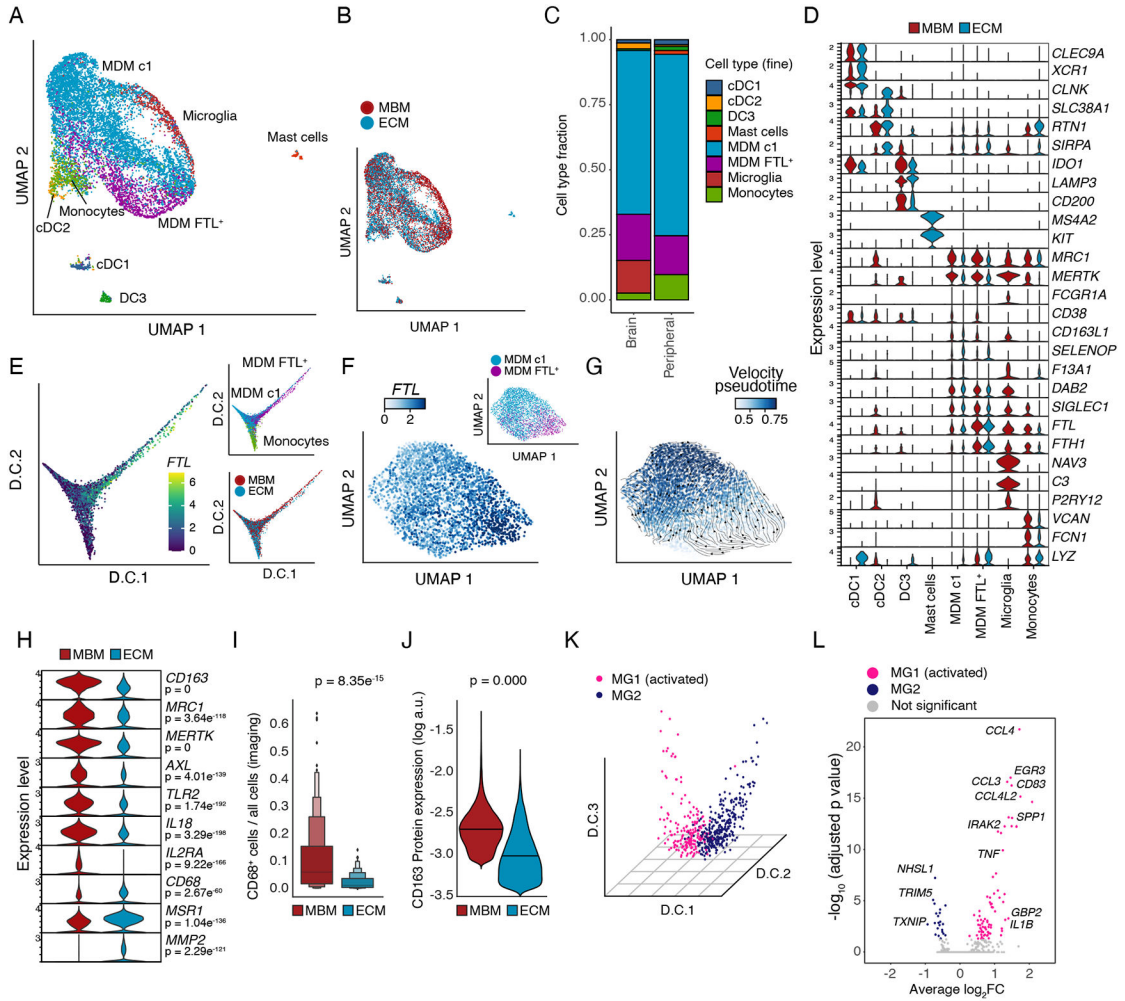
and proteomics datasets (Kleffman et al.; dot size). (M,N) Exemplary immunofluorescence micrograph showing nuclei (DAPI), CD8, NCAM1, CD68, CD138, and SOX10/HMB-45 (=LIN) expression in MBM (M) and ECM (N). Scale bar=100  $\mu$ m. (O) Fraction of NCAM1+ melanoma (LIN+) cells in MBM/ECM (boxplot indicating mean + quartiles) and (P) intensity of NCAM1 in melanoma cells. Wilcoxon rank-sum test.

Author Manuscript

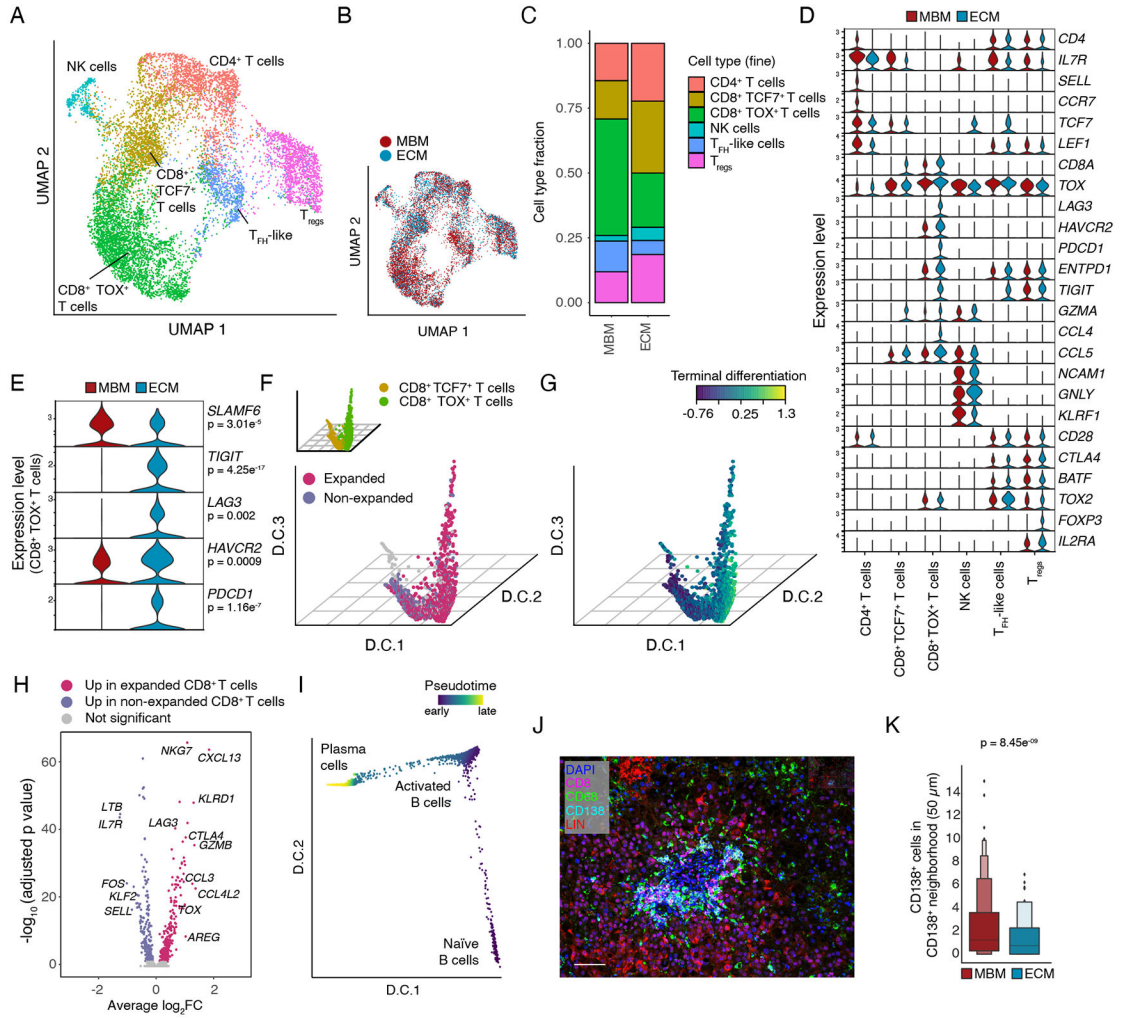
Author Manuscript

Author Manuscript

Author Manuscript

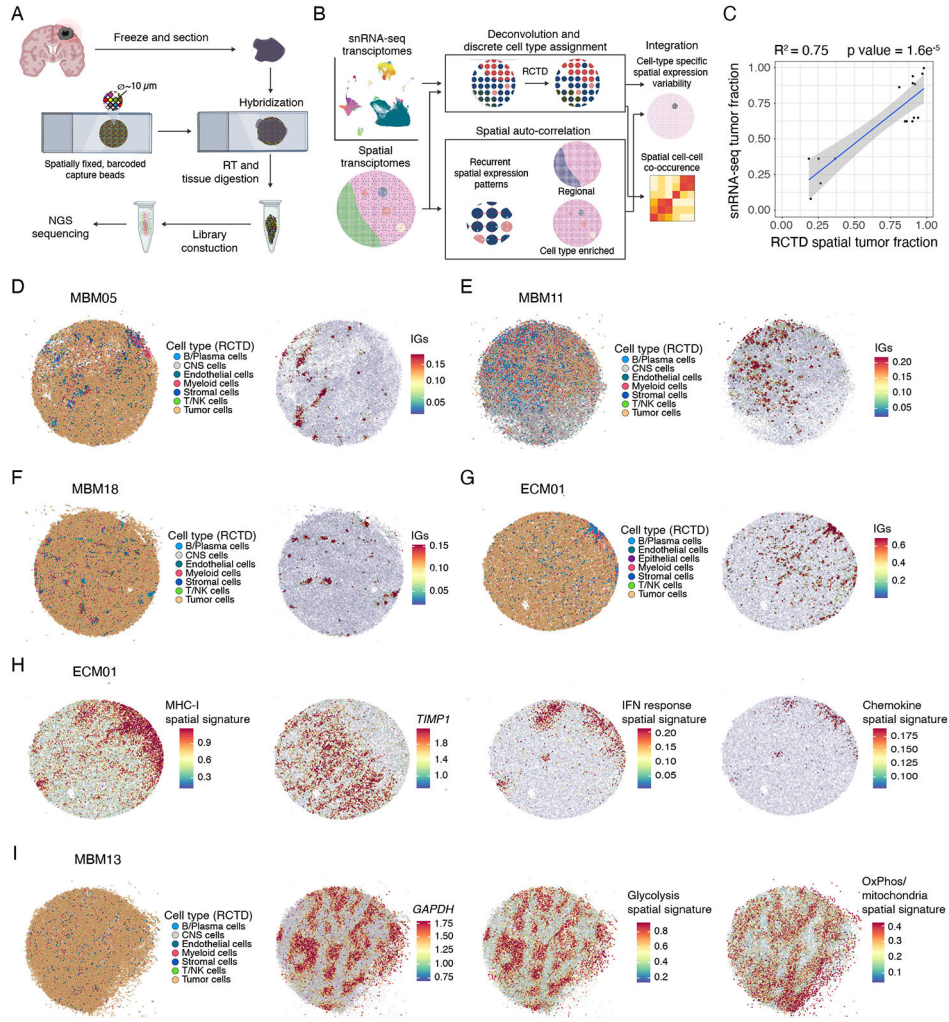


**Figure 4. Landscape and differences of myeloid cells in MBM/ECM profiled by snRNA-seq.** (A) UMAP of integrated myeloid cells from 17 MBM and 10 ECM samples, indicating cell type assignment. (B) UMAP as in (A) with tissue origin. (C) Fractions of cell types in the myeloid compartment of MBM and ECM. (D) Violin plots of marker genes in myeloid cells separated by tissue origin. Columns indicate cell type assignment. Rows represent selected marker genes. (E) Diffusion component (D.C.) analysis of monocytes/macrophages colored by *FTL* expression (left), cell type (top right) and tissue origin (bottom right). (F,G) RNA velocity-based UMAP of MDM-c1 and *FTL*<sup>+</sup> MDM in MBM showing *FTL* expression (F) with cell type assignment (inset) and pseudotime (G). (H) Violin plots of selected genes across all MDMs and separated by tissue site; MAST, adj. p-value as indicated. (I,J) Fraction of macrophages (CD68<sup>+</sup> cells, boxplot indicating mean and quartiles) (I) and intensity of CD163 protein (J) measured by IF in an independent patient cohort of MBM/ECM. Wilcoxon rank-sum test. (K) D.C. 1–3 of microglia profiled showing two major populations, microglia (MG)-1 and MG-2. (L) Volcano plot of DEGs between MG-1 (activated) and MG-2 subpopulations.



**Figure 5. Transcriptional and clonal T and B cell landscape in MBM.** (A,B) UMAP embedding of T/NK cells profiled by snRNA-seq showing cell type assignment (A) and tissue origin (B). (C) Fraction of T/NK cell subsets shown in (A). (D) Violin plots of selected genes (rows) in T/NK cell subsets (columns) and separated by tissue origin. (E) Violin plots of CD8<sup>+</sup> TOX<sup>+</sup> T cells profiled by snRNA-seq showing differentially expressed immune checkpoints (rows) by tissue origin. MAST, adj. p-value as indicated. (F) D.C. 1-3 of CD8<sup>+</sup> T cells (profiled by scRNA-seq) indicating subsets (TCF7<sup>+</sup> or TOX<sup>+</sup>, inset) and clonal expansion. (G) T cell terminal differentiation signature score on in D.C. embedding shown in (F). (H) Volcano plot depicting DGE of expanded and non-expanded CD8<sup>+</sup> T cells in MBM profiled by scRNA-seq. (I) D.C. embedding of B cell differentiation showing pseudotime projection in B and plasma cells profiled by sc/snRNA-seq. (J) Exemplary IF micrograph showing plasma cell aggregates in an MBM. Scale bar = 100 μm. (K) Box plot showing local neighborhood in MBM and ECM quantifying number of CD138<sup>+</sup> plasma cells in direct vicinity of CD138<sup>+</sup> plasma cells as a metric for plasma cell clustering in tissue. Boxes display mean and quartiles, Wilcoxon rank-sum test.





**Figure 6. Spatial features of metastatic melanoma.** (A) Illustration of SlideSeqV2 experimental protocol. (B) Illustration of computational approach to analyze spatially auto-correlated genes (using Moran’s  $J$ ) in specific cell types and tumor regions, and measurement of cellular co-occurrence. RCTD, robust cell-type decomposition. (C) Correlation of malignant cell fraction in spatial transcriptomics and matched snRNA-seq. Error bands, 95% s.e. interval on the Spearman’s correlation coefficient. (D-G) RCTD-based cell type assignment (left puck) and spatial expression pattern of immunoglobulins (IG signature, right puck) in MBM05 (D), MBM11 (E), MBM18 (F) and ECM01 (G), identifying plasma cell clusters. (H) Spatial plots of ECM01 showing expression of MHC-I genes, *TIMP1*, type I interferon response genes, and chemokines (pucks from left to right). (I) Malignant cell rich MBM13 with spatial plots showing expression of *GAPDH*, a glycolysis signature, and antithetical expression of OxPhos signature (pucks from left to right).

## Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Fluorescence-activated cell sorting		
anti-human CD15, APC	Biolegend	Cat# 301908
anti-human CD3, PE-Dazzle594	Biolegend	Cat# 300450
anti-human CD45, Pacific blue	Biolegend	Cat# 304022
anti-human CD66b, PE-CY7	Biolegend	Cat# 305116
anti-human TruStain FcX	Biolegend	Cat# 422302
Multiplexed Immunofluorescence		
anti-human CD138 (MI15)	Leica	Cat# PA0088
anti-human CD163 (10D6)	Leica	Cat# CD163-L-CE
anti-human CD68 (KP1)	Biogenex	Cat# AM416-5M
anti-human CD8 (4B11)	Leica	Cat# CD8-4B11-L-CE
anti-human HMB45 9HMB-45)	Cell Marque	Cat# 282M-95
anti-human NCAM1 (123C3)	Invitrogen	Cat# MA1-06801
anti-human SOX10 (BC34)	BioCare	Cat# ACI3099C
Bacterial and virus strains		
C3013 Competent Cells	New England Biolabs	Cat# C3013I
Biological samples		
Fresh human melanoma brain metastases	This study	Table S1
Frozen human melanoma brain and extracranial metastases	This study	Table S1
Formalin-fixed paraffin-embedded melanoma brain and liver metastases	This study	Table S1
Chemicals, peptides, and recombinant proteins		
10% NP-40 Surfact-Amps Detergent Solution	Thermo Fisher Scientific	Cat# 28324
2-mercaptoethanol	Thermo Fisher Scientific	Cat# 21-985-023
4',6-diamidino-2-phenylindole (DAPI)	Akoya Bioscience	Cat# NEL821001KT
70 $\mu$ m cell strainer	Corning	Cat# 431751
Acetic Acid	Thermo Fisher Scientific	Cat# A38S-500
ACK buffer	Thermo Fisher Scientific	Cat# A1049201
Amicon Ultra-4 Centrifugal Filters, 30kDa	EMD Millipore	Cat# UFC803024
AMPure XP Beads	Beckman Coulter	Cat# A63880
Bambanker	Bulldog Bio	Cat# BB01
Benzonase Nuclease	Millipore Sigma	Cat# E1014-25KU

REAGENT or RESOURCE	SOURCE	IDENTIFIER
BOND Dewax Solution	Leica	Cat# AR9222
BOND Epitope Retrieval Solution 1, pH 6	Leica	Cat# AR9961
BOND Epitope Retrieval Solution 2, pH 9	Leica	Cat# AR9640
Bovine serum albumin (BSA)	New England Biolabs	Cat# B9000S
CaCl <sub>2</sub> 1M	VWR	Cat# 97062-820
Calcein AM	Invitrogen	Cat# C1430
Chitin Resin	New England Biolabs	Cat# S6651L
Collagenase I	Worthington	Cat# LS004196
cOmplete Protease Inhibitors	Millipore Sigma	Cat# 1167498001
D1000 Reagents	Agilent Technologies	Cat# 5067-5583
D1000 ScreenTape	Agilent Technologies	Cat# 5067-5582
D5000 Reagents	Agilent Technologies	Cat# 5067-5589
D5000 ScreenTape	Agilent Technologies	Cat# 5067-5588
Dimethylformamide	Thermo Fisher Scientific	Cat# D119-500
DMEM	Gibco	Cat# 11965092
DNAse I	Worthington	Cat# LS002139
DTT 1M	Millipore Sigma	Cat# 646563-10X.5ML
Econo Pac	BioRad	Cat# 7321010
Econo-Column Chromatography Columns	BioRad	Cat# 7372512
EDTA 0.5M	Thermo Fisher Scientific	Cat# AM9262
Ethanol 99.5% 200 proof	Thermo Fisher Scientific	Cat# 615090010
Falcon® Round-Bottom Tubes with Cell Strainer Cap	Thermo Fisher Scientific	Cat# 08-771-23
Fetal Bovine Serum	Gibco	Cat# 10-437-028
GlutaMax	Thermo Fisher Scientific	Cat# 35050061
Glycerol	Thermo Fisher Scientific	Cat# BP229-1
H <sub>2</sub> O (RNA/DNA clean)	Thermo Fisher Scientific	Cat# 10-977-015
Hank's Balanced Salt Solution	Thermo Scientific	Cat# J67763.K2
HEPES	Thermo Fisher Scientific	Cat# 15-630-106
HEPES	Thermo Fisher Scientific	Cat# BP310-100
Hoechst 33342	Invitrogen	Cat# H3570
Human Tumor Dissociation Kit	Miltenyi	Cat# 30-095-929
Insulin-Transferrin-Selenium	Thermo Fisher Scientific	Cat# 41400-045
IPTG	Thermo Fisher Scientific	Cat# BP1755-1
Luria Broth	Thermo Fisher Scientific	Cat# BP9723-2
MEM Non-essential Amino Acids	Gibco	Cat# 11140050
MgCl <sub>2</sub> 1M	Thermo Fisher Scientific	Cat# AM9530G
MinElute Reaction Cleanup Kit	Qiagen	Cat# 28206
NaCl 5M	Thermo Fisher Scientific	Cat# AM9759
NEBNext High-Fidelity 2X PCR Master Mix	New England Biolabs	Cat# M0541S

REAGENT or RESOURCE	SOURCE	IDENTIFIER
NEBNext Ultra II RNA Library Prep Kit	New England Biolabs	Cat# E7420L
Odyssey Blocking Buffer	Li-Cor	Cat# 927-40000
Opal 7-Color Automated IHC Staining Kit	Akoya Bioscience	Cat# NEL821001KT
Optimal cutting temperature compound (OCT)	Tissue-Tek, Sankura	Cat# 94-4583
Optiprep Density Gradient Medium	Sigma Aldrich	Cat# D1556
PBS	Thermo Fisher Scientific	Cat# 20-012-050
PBS (10X)	Thermo Fisher Scientific	Cat# BP399-500
Penicillin Streptomycin (10,000 U/mL)	Thermo Fisher Scientific	Cat# 15-140-122
Plasmotest	InvivoGen	Cat# rep-pt1
Qubit 1X dsDNA HS Assay Kit	Thermo Fisher Scientific	Cat# Q33230
RNAeasy Plus	Qiagen	Cat# 74134
RNAse OUT	Thermo Fisher Scientific	Cat# 10-777-019
RPMI 1640	Thermo Fisher Scientific	Cat# 21875034
Slide-A-Lyzer Dialysis Cassete Kit, 10K MWCO	Thermo Fisher Scientific	Cat# 66807
SPRIselect beads	Beckman Coulter	Cat# B23318
SYBR Green I Nucleic Acid Gel Stain 10000x	Thermo Fisher Scientific	Cat# S7563
Tris HCL pH7.5 1M	Thermo Fisher Scientific	Cat# 5567027
Tris pH8.0 1M	Thermo Fisher Scientific	Cat# 15568-025
Tris(hydroxymethyl)aminomethane	Millipore Sigma	Cat# 252859-100G
Triton X100	Millipore Sigma	Cat# X100-500ML
Tween-20	Sigma Aldrich	Cat# p-7949
UltraPure Distilled Water	Thermo Fisher Scientific	Cat# 10977-023
Vectashield HardSet Antifade mounting medium	Vector Laboratories	Cat# H-1400
Zombie NIR viability dye	Biolegend	Cat# 423106
Critical commercial assays		
Chromium Single Cell 3' reagents v3.0	10x genomics	Cat# 1000006
Chromium Single Cell 5' Library construction kit	10x genomics	Cat# 1000020
Chromium Single Cell V(D)J Enrichment Kit for human T cells	10x genomics	Cat# 1000005
Chromium Single Cell V(D)J Reagents	10x genomics	Cat# 1000006
Falcon™ FluoroBlok™ HTS 96-Well Insert Systems, 8.0 um pore	Thermo Fisher Scientific	Cat# 08-771-007
Deposited data		
sc/snRNA-seq data of brain and extracranial melanoma metastases	This study	GSE185386
TCR-seq of melanoma brain metastases	This study	GSE185386
SlideSeqV2 of brain and extracranial melanoma metastases	This study	GSE185386
RNA-seq (patient-derived cell lines of brain and extracranial melanoma metastases)	This study	GSE185386
ATAC-seq (patient-derived cell lines of brain and extracranial melanoma metastases)	This study	GSE185386

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Raw data generated using sc/snRNA-seq, TCR-seq, SlideSeqV2, RNA-seq, and ATAC-seq	This study	dbGaP ( <a href="https://www.ncbi.nlm.nih.gov/gap">https://www.ncbi.nlm.nih.gov/gap</a> )
RNA-seq (patient samples of brain and extracranial melanoma metastases)	Fischer et al., 2019	DOI: 10.1158/2159-8290.CD-18-1489
Whole-exome sequencing (patient samples of brain and extracranial melanoma metastases)	Fischer et al., 2019	DOI: 10.1158/2159-8290.CD-18-1489
Proteomics data (brain and extracranial melanoma metastases)	Kleffman et al., 2022	<a href="https://massive.ucsd.edu/">https://massive.ucsd.edu/</a> (ID: Massive MSV000088814)
RNA-seq data (xenografts)	This study	Available upon request
Experimental models: Cell lines		
131/4-5B1 "5B1"	Cruz-Munoz et al, 2008	Authenticated by STR analysis by ATCC
113/6-4L "4L"	Cruz-Munoz et al, 2008	Authenticated by STR analysis by ATCC
12-273-BM "273-BM"	Eva Hernando, NYU	N/A
12-273-LN "273-LN"	Eva Hernando, NYU	N/A
2686	MD Anderson	N/A
MaMel-134	UK-Essen	N/A
MBM03	This study	N/A
MBM05	This study	N/A
Experimental models: Organisms/strains		
Mouse: NOD.Cg-Prkdc <sup>scid</sup> Il2rg <sup>tm1Wjl</sup> /SzJ (NSG)	The Jackson Laboratory	Strain# 005557
Oligonucleotides		
Ad1_noMX	Buenrostro et al., 2013	AATGATACGGCGAC CACCGAGATCTACA CTCGTCGGCAGCG TCAGATGTG
Ad2.1_NNNNNNNN	Buenrostro et al., 2013	CAAGCAGAAGACG GCATACGAGATNNN NNNNNGTCTCGTG GGCTCGGAGATGT
Recombinant DNA		
CMV-Luciferase-EF1 $\alpha$ -copGFP	BD Biosciences	Cat# BLIV511PA-1
pTXB1-Tn5	Addgene	Cat# 60240
Software and algorithms		

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Andromeda	Cox et al., 2011	<a href="https://bioinformaticshome.com/tools/proteomics/descriptions/Andromeda.html">https://bioinformaticshome.com/tools/proteomics/descriptions/Andromeda.html</a>
ARACNe-AP algorithm	Basso et al., 2005; Lachmann et al., 2016	<a href="http://califano.c2b2.columbia.edu/aracne">http://califano.c2b2.columbia.edu/aracne</a>
biomaRt v2.46.3	Durinck et al., 2009	<a href="https://bioconductor.org/packages/biomaRt/">https://bioconductor.org/packages/biomaRt/</a>
Cell Ranger v6.1.1	10x Genomics	<a href="https://support.10xgenomics.com/single-cell-gene-expression/software/overview/welcome">https://support.10xgenomics.com/single-cell-gene-expression/software/overview/welcome</a>
CellBender v0.2.0	Fleming et al., 2019	<a href="https://github.com/broadinstitute/CellBender">https://github.com/broadinstitute/CellBender</a>
CoBRA v2.0	Qiu et al., 2021	<a href="https://doi.org/10.1016/j.gpb.2020.11.007">https://doi.org/10.1016/j.gpb.2020.11.007</a>
Conos v1.4.4	Barkas et al., 2019	<a href="https://github.com/kharchenkolab/conos">https://github.com/kharchenkolab/conos</a>
cutadapt v3.6	Martin, 2011	<a href="http://journal.embnet.org/index.php/embnetjournal/article/view/200">http://journal.embnet.org/index.php/embnetjournal/article/view/200</a>
deepTools2 v3.5.1	Ramírez et al., 2016	<a href="https://academic.oup.com/nar/article/44/W1/W160/2499308">https://academic.oup.com/nar/article/44/W1/W160/2499308</a>
DESeq2 v1.34.0	Love et al., 2014	<a href="https://bioconductor.org/packages/DESeq2/">https://bioconductor.org/packages/DESeq2/</a>
Destiny v3.9.0	Angerer et al., 2016	<a href="https://github.com/theislab/destiny">https://github.com/theislab/destiny</a>
DoubletFinder v2.0.3	McGinnis et al., 2019	<a href="https://github.com/chris-mcginnis-ucsf/DoubletFinder">https://github.com/chris-mcginnis-ucsf/DoubletFinder</a>
FastQC v0.11.9	Babraham Institute	<a href="https://www.bioinformatics.babraham.ac.uk/projects/fastqc/">https://www.bioinformatics.babraham.ac.uk/projects/fastqc/</a>
Fastqscreen v0.13.0	Wingett and Andrews, 2018	<a href="https://www.bioinformatics.babraham.ac.uk/projects/fastq_screen/">https://www.bioinformatics.babraham.ac.uk/projects/fastq_screen/</a>
featureCounts v1.6.3	Liao et al., 2014	<a href="https://doi.org/10.1093/bioinformatics/btt656">https://doi.org/10.1093/bioinformatics/btt656</a>
GmAMisc v1.2.0	Alberti, 2021	<a href="https://CRAN.R-project.org/package=GmAMisc">https://CRAN.R-project.org/package=GmAMisc</a>
Harmony v0.1.0	Korsunsky et al., 2019	<a href="https://github.com/immunogenomics/harmony">https://github.com/immunogenomics/harmony</a>

REAGENT or RESOURCE	SOURCE	IDENTIFIER
hisat2 v2.2.1	Kim et al., 2019	<a href="https://www.nature.com/articles/s41587-019-0201-4">https://www.nature.com/articles/s41587-019-0201-4</a>
hypeR v1.10.0	Federico and Monti, 2020	<a href="https://github.com/montilab/hypeR">https://github.com/montilab/hypeR</a>
ImageJ	ImageJ	<a href="https://imagej.nih.gov/ij/">https://imagej.nih.gov/ij/</a>
inferCNV v1.10.1	Tirosh et al., 2016	<a href="https://github.com/broadinstitute/inferCNV">https://github.com/broadinstitute/inferCNV</a>
inForm v2.5.1	Akoya Biosciences	<a href="https://www.akoyabio.com/support/software/">https://www.akoyabio.com/support/software/</a>
Kallisto v0.46.1	Bray et al., 2016	<a href="https://pachterlab.github.io/kallisto/">https://pachterlab.github.io/kallisto/</a>
LivingImage software	Xenogen Corp., Alameda, CA	<a href="http://www.perkinelmer.com/product/spectrum-200-living-image-v4series-1-128113">http://www.perkinelmer.com/product/spectrum-200-living-image-v4series-1-128113</a>
MACS2 v2.2.7.1	Feng et al., 2012	<a href="https://www.nature.com/articles/nprot.2012.101">https://www.nature.com/articles/nprot.2012.101</a>
MaxQuant v1.5.2.8	Cox et al., 2008	<a href="https://maxquant.org">https://maxquant.org</a>
MCMICRO	Schapiro et al., 2021	<a href="https://mcmicro.org">https://mcmicro.org</a>
Mesmer 0.3.0	Greenwald et al., 2021	<a href="https://hub.docker.com/r/vanvalenlab/deepcell-applications">https://hub.docker.com/r/vanvalenlab/deepcell-applications</a>
MSigDB v7.4.1	Dolgalev, 2021	<a href="https://CRAN.R-project.org/package=msigdb">https://CRAN.R-project.org/package=msigdb</a>
Perseus	Tyanova et al., 2016	<a href="http://www.perseus-framework.org">http://www.perseus-framework.org</a>
Phenochart v1.1.0	Akoya Biosciences	<a href="https://www.akoyabio.com/support/software/">https://www.akoyabio.com/support/software/</a>
Picard v2.18.20	Broad Institute	<a href="https://broadinstitute.github.io/picard/">https://broadinstitute.github.io/picard/</a>
ProjecTILs v1.0	Andreatta et al., 2021	<a href="https://github.com/carmonalab/ProjecTILs">https://github.com/carmonalab/ProjecTILs</a>
Python v3.9.9	Python Software Foundation	<a href="https://www.python.org">https://www.python.org</a>
R v4.1.1	The R Foundation	<a href="https://www.r-project.org/">https://www.r-project.org/</a>
RCTD v1.2.0	Cable et al., 2021	<a href="https://github.com/dmcable/spacexr">https://github.com/dmcable/spacexr</a>
sambamba v0.6.8	Tarasov et al., 2015	<a href="https://lomereiter.github.io/sambamba/index.html">https://lomereiter.github.io/sambamba/index.html</a>
SCANPY v1.8.2	Wolf et al., 2018	<a href="https://github.com/theislab/Scanpy">https://github.com/theislab/Scanpy</a>

REAGENT or RESOURCE	SOURCE	IDENTIFIER
SCENIC/pyscenic v0.11.2	Aibar et al., 2017	<a href="https://github.com/aertslab/pySCENIC">https://github.com/aertslab/pySCENIC</a>
Scrublet v0.2.1	Wolock et al., 2019	<a href="https://github.com/swolock/scrublet">https://github.com/swolock/scrublet</a>
scuttle v1.4.0	McCarthy et al., 2017	<a href="https://bioconductor.org/packages/scuttle/">https://bioconductor.org/packages/scuttle/</a>
scVelo v0.2.4	Bergen et al., 2020	<a href="https://scvelo.readthedocs.io/">https://scvelo.readthedocs.io/</a>
Seq-n-slide	Dolgalev, 2022	<a href="https://github.com/igordot/sns">https://github.com/igordot/sns</a>
Seurat v4.1.0	Stuart et al., 2019	<a href="https://github.com/satijalab/seurat/">https://github.com/satijalab/seurat/</a>
SingleR v1.8.0	Aran et al., 2019	<a href="https://github.com/dviraran/SingleR">https://github.com/dviraran/SingleR</a>
singscore v4.0.3	Foroutan et al., 2018	<a href="https://CRAN.R-project.org/package=devtools">https://CRAN.R-project.org/package=devtools</a>
squidpy v1.1.2	Palla et al., 2022	<a href="https://github.com/theislab/squidpy">https://github.com/theislab/squidpy</a>
STACAS v1.1.0	Andreatta and Carmona, 2021	<a href="https://github.com/carmonalab/STACAS">https://github.com/carmonalab/STACAS</a>
STAR v2.7.3a	Dobin et al., 2013	<a href="https://github.com/alexdobin/STAR">https://github.com/alexdobin/STAR</a>
TOBIAS v0.13.2	Bentsen et al., 2020	<a href="https://doi.org/10.1038/s41467-020-18035-1">https://doi.org/10.1038/s41467-020-18035-1</a>
Trimmomatic v0.36	Bolger et al., 2014	<a href="http://www.usadellab.org/cms/?page=trimmomatic">http://www.usadellab.org/cms/?page=trimmomatic</a>
tximport v1.22.0	Soneson et al., 2015	<a href="https://github.com/mikelove/tximport">https://github.com/mikelove/tximport</a>
velocity v0.17	La Manno et al., 2018	<a href="http://velocityto.org/">http://velocityto.org/</a>
VIPER v1.26.0	Alvarez et al., 2016; Ding et al., 2018	<a href="https://doi.org/10.1038/ng.3593">https://doi.org/10.1038/ng.3593</a>
Other		
Original code generated in this study	This study	<a href="https://github.com/IzarLab/Melanoma_Brain_Metastasis">https://github.com/IzarLab/Melanoma_Brain_Metastasis</a>
Bond RX Fully Automated Research Stainer	Leica	Cat# 21.2821
Chromium Controller	10x genomics	Cat# 120270
Evos FL Auto 2 microscope	Thermo Fisher Scientific	Cat# AMAFD2000
Facs Aria II	BD Biosciences	N/A
HiSeq 4000	Illumina	N/A
Leica AT2	Leica	N/A
Leica CM1950	Leica	N/A
Neubauer counting chambers	Bulldog Bio	Cat# DHC-N01
NovaSeq 6000	Illumina	N/A



REAGENT or RESOURCE	SOURCE	IDENTIFIER
PhenoImager HT	Akoya Biosciences	<a href="https://www.akoyabio.com/phenoimager/instruments/phenoimager-ht/">https:// www.akoyabio.com/ phenoimager/ instruments/ phenoimager-ht/</a>
Qubit Flex Fluorometer	Thermo Fisher Scientific	Q33327
Synergy H1 Hybrid Reader	Agilent Biotek	N/A
TapeStation 2200	Agilent	N/A
Vevo 770 Ultrasound Imaging System	Visualsonics	N/A
Zeiss Celldiscoverer 7	Zeiss	N/A

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript