UCLA
UCLA Previously Published Works

Title
Multimodal single-cell and whole-genome sequencing of small, frozen clinical specimens

Permalink
https://escholarship.org/uc/item/1sc8z1wh

Journal
Nature Genetics, 55(1)

ISSN
1061-4036

Authors
Wang, Yiping
Fan, Joy Linyue
Melms, Johannes C
et al.

Publication Date
2023

DOI
10.1038/s41588-022-01268-9

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Nat Genet. Author manuscript; available in PMC 2023 May 03.

Published in final edited form as:

Nat Genet. 2023 January ; 55(1): 19–25. doi:10.1038/s41588-022-01268-9.

Multi-modal single-cell and whole-genome sequencing of small, frozen clinical specimens

Yiping Wang[1,2,*], **Joy Linyue Fan**[3,*], **Johannes C. Melms**[1,4,*], **Amit Dipak Amin**[1,4], **Yohanna Georgis**[5], **Irving Barrera**[6], **Patricia Ho**[1,4], **Somnath Tagore**[1,7], **Gabriel Abril-Rodríguez**[8], **Siyu He**[3], **Yinuo Jin**[3], **Jana Biermann**[1,2], **Matan Hofree**[6], **Lindsay Caprio**[1,4], **Simon Berhe**[4], **Shaheer A. Khan**[1,5], **Brian S. Henick**[1,5], **Antoni Ribas**[8,9], **Evan Z. Macosko**[6,10], **Fei Chen**[6,11], **Alison M. Taylor**[5,12], **Gary K. Schwartz**[1,5], **Richard D. Carvajal**[1,5], **Elham Azizi**[3,5,13,#], **Benjamin Izar**[1,2,4,5,7,#]

[1]Department of Medicine, Division of Hematology/Oncology, Vagelos College of Physicians and Surgeons, Columbia University Irving Medical Center, New York, NY, USA

[2]Department of Systems Biology, Program for Mathematical Genomics, Columbia University, New York, NY, USA

[3]Department of Biomedical Engineering, Columbia University, New York, NY, USA

[4]Columbia Center for Translational Immunology, Columbia University Irving Medical Center, New York, NY, USA

[5]Vagelos College of Physicians and Surgeons, Herbert Irving Comprehensive Cancer Center, Columbia University, New York, NY, USA

[6]Broad Institute of Harvard and MIT, Cambridge, MA, USA

[7]Department of Systems Biology, Columbia University Irving Medical Center, New York, NY, USA

[#]correspondence: ea2690@columbia.edu, bi2175@cumc.columbia.edu.

[*]equal contribution

AUTHOR CONTRIBUTIONS

B.I. conceived of the study. B.I. and E.A. jointly provided overall supervision of the study. J.C.M., A.D.A., Y.G., I.B., P.H., L.C. and S.B. performed experiments. Y.W., J.L.F., J.C.M., Y.G., S.T., G.A-R., S.H., Y.J., J.B., M.H., A.M.T. performed analyses. S.A.K., B.S.H., A.R., G.K.S. and R.D.C. provided clinical specimens. R.D.C., E.Z.M., F.C., A.M.T. and G.K.S. provided additional supervision. Y.W., J.L.F, J.C.M., E.A. and B.I. wrote the manuscript. All authors reviewed, contributed to, and approved of the manuscript.

CONFLICT OF INTERESTS STATEMENTS

B.I. has received consulting fees from Volastra Therapeutics Inc, Merck, AstraZeneca and Janssen Pharmaceuticals and has received research funding to Columbia University from Alkermes, Arcus Biosciences, Checkmate Pharmaceuticals, Compugen, Immunocore, and Synthekine. G.A-R. has received honoraria from consulting with Arcus Biosciences. A.R. has received honoraria from consulting with Amgen, Bristol-Myers Squibb, Chugai, Genentech, Merck, Novartis, Roche and Sanofi, is or has been a member of the scientific advisory board and holds stock in Arcus, Compugen, CytomX, Highlight, ImaginAb, Isoplexis, Kite-Gilead, Lutris, Merus, PACT, RAPT, Synthekine and Tango Therapeutics. AMT receives research support from Ono Pharmaceuticals. B.S.H. participated in advisory boards for AstraZeneca and Ideaya. R.D.C. is a consultant for Alkermes, Bristol Myers Squibb, Castle Biosciences, Delcath, Eisai, Hengrui, Ideaya, Immunocore, InxMed, Iovance, Merck, Novartis, Oncosec, Pierre Fabre, PureTech Health, Regeneron, Sanofi Genzyme, Sorrento Therapeutics and Trisalus, serves on clinical/scientific advisory boards for Aura Biosciences, Chimeron, and Rgenix Research, and has received research funding to Columbia University from Amgen, Astellis, AstraZeneca, BioMed Valley, Bolt, Bristol-Myers Squibb, Corvus, Cstone, Foghorn, Ideaya, Immatics, Immunocore, InxMed, Iovance, Merck, Mirati, Novartis, Pfizer, Plexxikon, Regeneron, and Roche/Genentech. B.I. and J.C.M. filed a patent describing the generation of high quality single-cell genomics data from frozen tissues. The remaining authors declare no competing financial interests.

Code availability statement

Code is publicly available via: https://github.com/IzarLab/fresh_vs_frozen_comparison and https://github.com/azizilab/starfysh (10.5281/zenodo.6950760)

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

[8]Department of Medicine, Jonsson Comprehensive Cancer Center, University of California, Los Angeles (UCLA), Los Angeles, CA 90095, USA

[9]Parker Institute for Cancer Immunotherapy, San Francisco, CA 94129, USA

[10]Department of Psychiatry, Massachusetts General Hospital, Boston, USA

[11]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA, USA

[12]Department of Pathology and Cell Biology, Columbia University Irving Medical Center, New York, NY, USA

[13]Irving Institute for Cancer Dynamics, Columbia University, New York, NY, USA.

INTRODUCTION

Single-cell genomics are enabling technologies that are instrumental to dissecting tumor heterogeneity and molecular underpinnings of drug response at unprecedented resolution[1–11]. However, their broad clinical application remains challenging, due to several practical and pre-analytical challenges that are incompatible with typical clinical care workflows, namely the need for relatively large, fresh tissue inputs. Here, we show that multi-modal single-nucleus RNA/T cell receptor (TCR) sequencing, spatial transcriptomics and whole-genome sequencing (WGS) is feasible from small, frozen tissues that approximate routinely collected clinical specimens (e.g., core needle biopsies). In comparison to data from sample-matched fresh tissue, we find similar quality and biological outputs of snRNA/TCR-seq data, while substantially reducing artifactual signals and compositional biases introduced by fresh tissue processing. Profiling sequentially collected melanoma samples from a patient treated on the KEYNOTE-001 trial[12], we resolve cellular, genomic, spatial and clonotype dynamics that represents molecular patterns of heterogeneous intra-lesional evolution during anti-PD-1 therapy. To demonstrate applicability to banked biospecimens of rare diseases[13], we generated a uveal melanoma liver metastasis single-cell atlas with matched WGS data. These results show that single-cell genomics from archival, clinical specimens is feasible and provides a framework for translating these methods more broadly to the clinical arena.

MAIN

Single-cell genomics, namely single-cell RNA-sequencing (scRNA-seq), has enabled significant discoveries across all fields of biomedical research, including in several solid tumor, providing unique insights into tumor ecosystems[1–5], metastasis[5–8], and drug resistance[2,9]. Matched scRNA/TCR-seq is particularly informative for studies investigating the effect of transformative immunotherapies, such as anti-PD-1 therapies which lead to durable responses in patients with some cancers, such as metastatic melanoma[10,11]. While these methods have the potential to inform clinical studies, their broad application has been hampered by several challenges. In particular, the need for relatively large (milligrams), fresh tissue specimens and their immediate processing is incompatible with clinical workflows, thus, represents a major barrier challenges for standardized, multi-center analysis efforts. Furthermore, frequently used mechanical and enzymatic disaggregation of

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

solid tumor specimens to a single-cell suspension may introduce artifacts in resulting gene expression profiles. Consequently, single-cell studies to date have mostly been conducted in relatively small, heterogeneous patient populations (e.g. variable treatment exposures), which introduces challenging biases and may limit their generalizability. Building on recent developments that enable single-nucleus transcriptome sequencing (snRNA-seq) from frozen tissues[14,15], we show that multi-modal single-cell genomics and accompanying WGS can be performed in a rapid, scalable manner on very small, frozen clinical specimens, thus, overcoming several existing practical, pre-analytical and analytical challenges. Thus, we provide a framework for enabling harmonized single-cell genomics studies across multiple institutions and an opportunity to implement these methods for biomarker or target discovery. (Methods).

We first performed head-to-head comparisons of scRNA/snRNA-seq of matched fresh and frozen tissues from patients with different cancers, including non-small cell lung cancer (NSCLC), metastatic cutaneous melanoma, and uveal melanoma using different 10X Genomics chemistries (3', 5'v1, or 5' V2), cell sorting (Methods, Extended Data Fig. 1a) and RNase inhibitor protocols (Methods; Supplementary Table 1).

Overall, in snRNA-seq using 5' chemistries, the data quality (median number of genes detected per cell) was comparable to that of tissue-matched scRNA-seq and was consistently superior compared to 3' chemistries (Fig. 1a–c), while scRNA-seq protocols had a higher rate of mitochondrial reads (Extended Data Fig. 1b–d). In the NSCLC comparison, for example, we recovered a median of 3,117 genes/cell using 5'v1 (with addition of RNase inhibitor) compared to 2,521 in fresh scRNA-seq (Wilcoxon rank-sum p-value 1.67e-29) and 1,392 in the best 3' condition (Wilcoxon rank-sum p-value < 1e-30), respectively. Similar results were observed in cutaneous and uveal melanoma tissue comparisons (Fig. 1b–d Supplementary Table 2). Compared to previously published protocols (after correcting for sequencing saturation), our approach compared favorably with respect to required tissue input (~1000-fold lower), cell recovery, and data quality overall and on a per cell-type basis (Extended Data Fig. 1e and Supplementary Table 2).

Importantly, we noted substantial expression of an artifactual program[16] associated with tissue processing in scRNA-seq, while this artifact had lower expression in matched snRNA-seq data (Fig. 1d–f). This artifactual expression manifested particularly strongly in the NSCLC comparison, was increased by fluorescence-activated cell sorting (FACS), and was more prominent in immune cells. Notably, this stress signature captures important inflammatory pathways, among others, and may bias the interpretation of these pathways in scRNA-seq.

In 5' chemistry sc/snRNA-seq, we were able to robustly recover TCRs matched to single-cell transcriptomes (Extended Data Fig. 2). In both cutaneous and uveal melanoma comparisons, we recovered highly overlapping clonotypes (hypergeometric test, p=1.55e-62, and p=0.0018, between the single-cell 5p CD45+ and single-nuclei 5p with inhibitor samples, respectively) (Fig. 1g,h) and TCR diversity (Gini coefficient, Supplementary Table 3). Lastly, we inferred copy-number alterations (CNAs) and found strong agreement (Spearman R2=0.69) between scRNA-seq and snRNA-seq protocols (Fig. 1i).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Despite procedural and technical differences, matched sc/snRNA-seq showed good mixing of cell types (determined using the LISI score[17], Methods) following batch correction (Methods) indicating preservation of global transcriptional outputs (Extended Data Fig. 3,4). Accordingly, we identify comparable cell type diversity (estimated using the Shannon index, Extended Data Fig. 5a,b). Notably, cell type composition was highly consistent among different snRNA-seq runs within the same specimen, indicating high reproducibility (Extended Data Fig. 5a). Comparing sc/snRNA-seq, we noted one important outlier in cellular composition: there was low recovery of cancer cells in the NSCLC sample profiled using scRNA-seq (23% of all cells), while snRNA-seq robustly detected this population (87.9–92.8% of all cells) (Extended Data Fig. 5c). This is consistent with prior studies of NSCLC[7] using scRNA-seq that showed disproportionally low recovery of malignant cells, suggesting that these cells are vulnerable to sample processing, and emphasizes an advantage of snRNA-seq for this common cancer type.

To test the feasibility and strengths of these methods, we chose two application cases. First, we obtained sequentially collected (before, and two on-treatment) biopsy specimens from a patient treated on the first clinical trial (KEYNOTE-001)[12] using anti-PD-1 antibody MK-3475 (now known as pembrolizumab), and achieved a partial response (Extended Data Fig. 6a). Although these samples were ~10 years old, we achieved excellent technical quality (Extended Data Fig. 6b) with minimal artifactual gene expression (Extended Data Fig. 6c), while revealing cellular diversity (Fig. 2a, Extended Data Fig. 6d,e). We inferred CNAs and identified distinct cancer clones (Fig. 2b, Methods) and noted evolving CNA patterns during therapy (Fig. 2b) and corresponding transcriptional changes over time (Fig. 2c,d). Interestingly, immune-resistant clones (yellow 1 and purple 2) emerged from a small sub-population of pre-existing cancer cells (prior to receiving therapy) and enriched for expression of cancer cell intrinsic signatures of immunotherapy resistance[9] (Fig. 2e) that was not exclusively explained by gene dosage, but showed conserved gene regulatory networks (GRNs), Extended Data Fig. 7c,d; Methods) and de-differentiation (Fig. 2f). Importantly, despite the expression of antigen-presentation and IFNg pathway genes in clone 2 (Extended Data Fig. 7c), these cells were strongly enriched for pathways associated with immune evasion (Extended Data Fig. 7e,f) and reduced expression of *CD58* in (Fig. 2g), a recently identified mechanism of resistance to immune checkpoint blockade[18,19]. Among non-malignant cells, we observed increased infiltration with T cells and macrophages (Fig. 2a, Extended Data Fig. 7a,b). Integrated analysis of CD8+ T cells revealed tumor infiltration of both stem-like, precursor exhausted, and terminally differentiated cells[20] (Fig. 2h) with corresponding diversification of clonotypes and contraction of pre-existing T cell clones over time (Fig. 2i). On the same specimens, we also performed spatial transcriptomics using *Slide-seq V2*[21], and deconvolved and analyzed these data together with corresponding single-cell transcriptomes using *robust cell type deconvolution* (RCTD)[22] and *Starfysh*[23]. Spatial projection (Methods) of clone 2 revealed its distinct geographical location with strong expression of the immune resistance signature and lack of activated immune cell infiltration during anti-PD1 therapy, compared to tumor areas populated by other cancer clones, which showed dense infiltration and immune cell differentiation at the tumor/normal border, indicative of active immune editing (Fig. 3a,b and Extended Data Fig. 7g). Together, these findings suggest that similar to pre-existing cells with resistance mutations or cell

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

states that grow out under the pressure of oncogene-targeted therapies[2,24], pre-existing cancer cell populations defined by their underlying CNAs or conserved gene regulatory circuitry may emerge during cancer immunotherapy, *despite* apparently adequate T cell responses.

To demonstrate the scalability of performing such studies in larger cohorts from a multi-institutional clinical trial, we performed snRNA/TCR-seq of 169,015 cells from 20 core-needle biopsies collected from 7 patients with liver-metastatic uveal melanoma treated with MEK-inhibitor selumetinib[13] (Extended Data Fig. 8a,b). Liver metastases occur in most patients with uveal melanoma and are associated with immunotherapy resistance in different cancers, although the underlying mechanisms remain poorly understood[25]. We achieved excellent technical quality (Extended Data Fig. 8b) and recovered diverse cell types (Fig. 4a), thereby providing a large metastatic-niche-specific atlas of uveal melanoma. Among cancer cells, we find that in some patients, rapid changes in aneuploidy patterns (changes in large-scale CNAs) and distinct transcriptional outputs (indicated by distinct cancer cell clusters corresponding to different CNA-defined cancer clones) occur within days of therapy (Fig. 4b).

Given the importance of dynamic CNA/aneuploidy pattern changes in both case studies presented here, we next established population-matched low pass whole-genome sequencing (lp-WGS) of the same cell pool from which we performed snRNA-seq (Extended Data Fig. 9a). We predicted CNAs in single-cell data using *inferCNV*[2] (Extended Data Fig. 9b) and *ichorCNA*[26] in the lp-WGS data (Extended Data Fig. 9c,d), and found a strong correlation between both methods (R2=0.66) (Fig. 4c) (Methods). Matched snRNA-seq and lp-WGS performed on a melanoma sample also showed strong concordance across methods (Extended Data Fig. 9c). This data supports the use of single-cell inferred CNAs and corresponding gene expression changes to study genomic and dynamic transcriptional changes in clinical specimens. We next performed integrated analyses of T cell transcriptomes and TCRs across the uveal melanoma and treatment-naïve melanoma lesions presented here (Fig. 1b and Fig. 2), revealing a conserved trajectory of T cell differentiation (from naïve to terminal differentiation) and corresponding TCR clonality (Fig. 4d, e). Thus, these methods enable systematic comparison of immune environments of specific metastatic niches within and across different cancer lineages. Together, these analyses emphasize the feasibility of performing multi-modal single-cell studies on biospecimens collected on clinical trials and across institutions where immediate tissue preparation is both impractical and introduces several pre-analytical biases.

There are important considerations for future implementation of the approaches outlined here. snRNA-seq performs an unselected detection of cells from the tumor-ecosystem, including capture of cells that are poorly represented in scRNA-seq (e.g. lung cancer cells). Representation of TCRs will be dictated by the *in-situ* fraction of T cell abundances (or lack thereof). However, in frozen tissues, sorting of cell nuclei based on size and scatter patterns may be used to enrich T cells and enhance recovery of matching TCRs. Furthermore, reference atlases are increasingly becoming helpful in identifying cell types and states. As snRNA-seq is being more systematically implemented, it will be critical to build such references from frozen tissue specimens to enable rapid definition of cell types and states

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

relevant to specific tissue contexts and therapies[5,27]. Lastly, we expect that incorporating additional single-cell measurements from frozen tissue (e.g., chromatin accessibility or metabolomics) will be feasible, and will help determine which analytes may be best-suited to guide future research and clinical applications.

In summary, we show that performing multi-modal single-cell profiling and WGS is feasible from tissue inputs equivalent to small, routinely collected clinical specimens. This enables application of these methods to multi-institutional efforts through harmonized and scalable processes while reducing pre-analytical biases, thus representing an important step toward implementing these technologies in guiding clinical care.

METHODS

Patient tissue collections and Ethical approval

Fresh and frozen tissue specimens were collected under IRB approved protocols at New York Presbyterian Hospital/Columbia University Medical Center (AAAT7416, AAAT2278), Dana Farber Cancer Institute, and University of California, Los Angeles. All procedures performed on patient samples were in accordance with the ethical standards of the respective IRB and the Helsinki Declaration and its later amendments. For the comparison of single-cell RNA-sequencing (scRNA-seq) and single-nuclei RNA-sequencing (snRNA-seq) protocols, surgical specimens were reviewed by qualified pathologists according to institutional guidelines and immediately placed in ice-cold RPMI 1640 (Thermo Fisher, Waltham, MA; #21875034) without supplements and transported to the laboratory space for immediate processing (scRNA-seq) and parallel collection of matched flash frozen specimens (snRNA-seq). Frozen uveal melanoma liver metastases were collected as core-needle biopsies during a trial of targeted MEK inhibition[13] (NCT01143402). Frozen sequential biopsies of cutaneous melanoma prior and during treatment with anti-PD-1 therapy were collected during the KEYNOTE-001 trial [12].

Fresh tissue specimen processing

All steps apart from the digestion were carried out on wet ice with pre-cooled buffers. Tissue specimens were weighted and split in half, and pieces of ~5–8 mm edge length were placed in cryovials and snap frozen in liquid nitrogen and stored at −80°C. The remaining tissue was kept in ice cold RPMI in a petri dish and cut into 1 mm^3 cubes using two scalpels. The cubes and RPMI were transferred to a 50 ml Falcon tube (Corning) using a 10 ml serological pipette and collected by 5 min centrifugation at 300 x g at 4°C. The tissue was then digested using human tumor dissociation enzymes (Miltenyi Human Tumor Dissociation Kit; #30–095-929) according to manufacturer instructions based on the tissue weight. Briefly, tissue was resuspended in pre-warmed RPMI and human tumor dissociation enzymes were added. The sample was then placed in a 37°C water bath and agitated every 2 minutes. Every 5 minutes the tissue was further mechanically dissociated by pipetting using pipettes of decreasing calibers. This process was continued until most of the tissue had dissociated. After 10 minutes (melanoma) or 15 minutes (lung cancer) of dissociation the samples were filtered through a pre-wetted 70 μm cell strainer (Corning) into a separate 50 ml Falcon tube, collected by centrifugation for 5 min at 400 x g and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

4°C, and the supernatant was decanted. The cell pellet was resuspended in 3 ml ACK buffer (Thermo Fisher; #A1049201) to lyse red blood cells. After 1 min incubation the reaction was terminated by dilution with 30 ml ice cold sorting buffer (2% Fetal bovine serum/1mM EDTA in PBS). The cells were collected by centrifugation for 5 min at 400 x g and 4°C, resuspended in 1 ml ice cold sorting buffer and cell count and viability was assessed using trypan blue and disposable Neubauer counting chambers (Bulldog Bio, Inc. Portsmouth, NH; #DHC-N002). Cells were then allocated for direct loading (non-small cell lung cancer and primary uveal melanoma), or further processed for fluorescence-activated cell sorting (cutaneous melanoma and primary uveal melanoma).

Fluorescence-activated cell sorting

To enrich viable immune cells (primary uveal melanoma sample) or sort viable immune and non-immune cells (cutaneous melanoma) the samples were sorted using a FACS Influx instrument (BD Biosciences) gating on viable CD45+ (immune) or viable CD45- (non-immune) cells (Extended Data Fig. 1a). First, cells were stained for viability (Zombie NIR, 1:500 in PBS; Biolegend, San Diego, Ca; #423106) for 10 min at room temperature in the dark. Thereafter, cells were washed once with sorting buffer, collected by centrifugation and surface antigens were stained for 15 minutes on ice in the dark. The primary uveal melanoma sample was stained with Pacific-Blue-aCD45 (Biolegend, #304022). The cutaneous melanoma sample was stained with the following (all Biolegened): Human TruStain FcX (#422302), Pacific-Blue-aCD45 (#304022), PE-Dazzle594-aCD3 (#300450), PE-CY7-aCD66b (#305116), APC-aCD15 (#301908). After staining, the samples were washed twice with ice-cold sorting buffer and 1.5×10^3 cells per population of interest were sorted and immediately processed for scRNA-seq.

Single cell RNA library preparation

Sorted and unsorted single cell suspensions ($1.2–1.5\times10^3$ cells) were transferred into low-binding 1.5 ml Eppendorf tubes (Eppendorf, Hamburg, Germany), centrifuged and washed twice with 1 ml loading buffer (PBS with 0.05% RNase-free BSA; Thermo Fisher, #AM2616) using a swinging bucket centrifuge at 4°C with 400 x g for 5 min. After the final spin all but 31 µl buffer were removed and the samples were loaded on a Chromium controller with Chromium Single Cell V(D)J Reagents (10X Genomics, Pleasanton, CA; #1000006) for 5’ RNA capture. Gene expression libraries were then generated using Chromium Single Cell 5’ Library construction kit (#1000020) according to manufacturer instructions.

Tissue sparing extraction of single-nuclei from small frozen specimens

We adopted the previously described salt-tris (ST) based extraction method (protocol below)[14,15] and implemented changes to enable tissue sparing extraction of nuclei for single-nuclei RNA-sequencing from clinical-grade frozen tissue specimens and allow profiling of minute specimens such as core-needle biopsies. To this end, we used a Leica CM1950 cryostat (Leica, Wetzlar, Germany) for initial tissue processing and additional washing steps for OCT removal. Frozen tissue specimens with 2–10 mm edge length were embedded in optimal cutting temperature (OCT) compound (Tissue-Tek, Sankura; #4583) on dry ice. Samples were then mounted on sample holders and excess OCT was trimmed away using

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

a blade leaving ~5 mm OCT around each side of the tissue. Fine needle aspiration biopsies were directly mounted on sample holders of the cryostat using a small amount of OCT. Multiple 20 μm tissue curls were cut per tissue and collected in pre-cooled 5 ml tubes (Eppendorf, Hamburg, Germany) ensuring no thawing while transferring and stored on dry ice until processing. The number of curls required depends on the tissue size and ranges from 3–4 curls for large specimens (1 cm^2) to 10–15 for core needle biopsies, thus leaving most of the specimen intact for future investigation. All subsequent steps were performed on wet ice and all centrifuges were equipped with swinging buckets and cooled to 4°C. For extraction of nuclei, the tubes were moved from dry ice to wet ice and left to equilibrate briefly. After 30 seconds, 4 ml of ice-cold PBS without calcium or magnesium (Thermo Fisher, #10010023) were added and the tubes were inverted until all OCT had dissolved and the clean tissue could be collected by centrifugation at 300 x g for 2 min. The tissue was then resuspended in 1 ml ST buffer [146 mM NaCl, 10 mM Tris-HCL pH7.5, 1mM $CaCl_2$, and 21 mM $MgCl_2$ in ultrapure water] with 0.03% Tween-20 Sigma Aldrich, #p-7949 (=TST buffer) with 0.1% BSA (New England Biolabs, B9000S) and supplemented with or without 40 U/ml RNAse inhibitor (RNAse OUT, Thermo Fisher) (Supplementary Table 1). ST buffer was prepared from stock solutions listed in Supplementary Table 5. The suspension was thoroughly pipetted 15 x using a 1 ml pipette to mechanically dissociate the tissue and left to incubate for 5 min on ice. After 5 min the pipetting step was repeated, and the reaction was quenched using 4 ml ST buffer with or without 40 U/ml RNAse inhibitor (RNAse OUT, Thermo Fisher, #10777019) (Supplementary Table 1). The sample was filtered through a pre-wetted 70 μm nylon mesh filter (Fisher Scientific) into a 50 ml conical tube and the filter was washed with 5 ml ST buffer. The tube was then centrifuged at 500 x g for 5 min to collect the dissociated nuclei. After carefully decanting the supernatant, the nuclei were resuspended in 100–400 μl ST buffer without RNAse inhibitor and filtered with a 40 μm mesh filter attached to a FACS tube (Fisher Scientific). The nuclei concentration and dissociation quality was then determined in a 5 μl aliquot using Neubauer counting chambers (Bulldog Bio, Inc. Portsmouth, NH) and a fluorescent microscope (EVOS FL, Thermo Fisher) after staining of nuclear DNA with 50μg/ml Hoechst 33342 (Thermo Fisher, H3570).

Single-nuclei extraction using large tissue input

For comparison with established protocols, we extracted single nuclei from two specimens (melanoma and NSCLC) using the salt-tris (ST) based extraction method with mechanical dissociation using spring scissors[14]. Frozen tissue blocks (10 – 15 mg and from tissues measuring approximately 4×4×4 mm) from the same specimens as in the aforementioned tissue sparing extraction method were used to directly enable comparison of key quality control metrics. All steps were performed on wet ice and centrifuges were equipped with swinging bucket rotors and precooled to 4°C. To extract nuclei, frozen tissue was placed inside a 6 well plate and 1 ml of either TST buffer (see above) or ST buffer with 0.5 % CHAPS (Millipore, 220201) (CST buffer) were added. Tissues were then chopped for 10 minutes using Noyes spring scissors (Fine Science Tools, 15514–12). After chopping, the homogenized solution was filtered through a 40 μm cell strainer (Falcon, 08–771-1) and plate and filter were washed with 1 ml ST buffer. After addition of 3 ml ST buffer the suspension was transferred to a 5 ml tube (Eppendorf) and centrifuged for 5 min at 500g.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

After centrifugation the supernatant was carefully decanted. Depending on pellet size, the nuclei were resuspended in 100–200 µl ST buffer, filtered through a 35 µm cell strainer (Falcon, 352235). Finally, the nuclei concentration and dissociation quality was determined in a 5 µl aliquot using Neubauer counting chambers (Bulldog Bio) and a fluorescent microscope (EVOS FL, Thermo Fisher) after staining nuclear DNA with 50µg/ml Hoechst 33342.

Single nuclei RNA library preparation

$0.9–1.5\times10^3$ nuclei were loaded in ST buffer without RNAse inhibitor using a Chromium controller and chromium reagents (10X Genomics) for 3' (#1000075) or 5' capture (#1000006 and #1000263) as indicated (Supplementary Table 1). After reverse transcription and cleanup, cDNA libraries were generated according to manufacturer instructions with one additional cycle of cDNA amplification to account for the relatively lower amount of RNA in nuclei compared to whole cells.

Single cell and nuclei TCR library preparation

Single cell and single nuclei TCR libraries were prepared from amplified cDNA libraries according to manufacturer instructions using the following reagents (all 10X Genomics): Chromium Single Cell V(D)J Enrichment Kit for human T cells (#1000005) was used for cDNA generated with Chromium Single Cell V(D)J reagents (#1000006), and final sequencing libraries were prepared using Chromium i7 multiplexing kit (#120262)

Single Cell Human TCR Amplification Kit (#1000252) was used for cDNA generated with Chromium Next GEM Single Cell 5' V2 reagents (#1000263), and final sequencing libraries were prepared using Library construction kit (#1000190) and Dual Index Kit TT set A (#1000215).

Sequencing of single cell and single nuclei libraries

Final sequencing libraries were quantified using Tapestation D1000 and D5000 reagents (Agilent) and a 2200 TapeStation system. Samples were then mixed and sequenced to target >20,000 reads per cell for gene expression libraries and >5,000 reads per cell for TCR libraries using NovaSeq S4 or HiSeq 4000 (Illumina. San Diego, CA) with at least 2×100bp read length (Supplementary Table 1).

Genomic DNA extraction and library construction for for low-pass whole genome sequencing

Excess nuclei ($>1\times10^5$) from the sample preparations for sn-RNAseq were collected by centrifugation (500 x g, 5 min) and snap frozen after removing all but ~10 µL ST buffer and stored until further processing at −20°C. If insufficient numbers of nuclei were available after loading, additional curls were processed using the same methods as described above for single nuclei extraction. To extract genomic DNA from nuclei the nuclei were briefly thawed on wet ice and genomic DNA was extracted using DNeasy Blood and Tissue kit (Qiagen, Hilden, Germany; #69504) according to manufacturer instructions and eluted in RNAse and DNAse free water at 37°C for 5 minutes. The DNA concentration was then quantified using a Nanodrop. For library construction, see **Supplementary Information.**

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

lp-WGS sequencing and copy number assignment

Indexed WGS-libraries were mixed equimolarly and sequenced on an Illumina MiSeq instrument with 0.1–1X coverage using the V2–300 cycle kit (Illumina). Using Illumina pipelines, .bam and .bai files were generated from .fastq files which served as input for *ichorCNA*[26] generating .seg files for visualization. Finally, GISTIC 2.0[28] was used to assign a copy number to each gene.

Computational Methods

For routine raw data processing, alignment and initial quality control and filtering steps, see **Supplementary Information**.

CNA comparison between snRNA and scRNA samples

To measure correlation of inferred CNA profiles in snRNA and scRNA-seq samples (Fig. 1i), for datasets that included both types of samples, we calculated average *inferCNV* scores for each chromosome arm for the cancer cells in each individual sample. To exclude chromosome arms that did not exhibit any large-scale amplifications or deletions, we filtered out arms that had an average *inferCNV* score between −.01 and .01. We also did not consider the CD45+ sample from our melanoma dataset, as it mostly contained immune cells, or the NSCLC 3' sequencing protocols, as these appeared to be of lower quality in terms of gene counts and stress signature expression than the 5' protocols (Fig. 1d). For every possible pair of one fresh vs. one frozen sample in each of our datasets, we then calculated the Spearman correlation between average chromosome arm *inferCNV* scores in each of the two samples.

Comparison of TCR clonotype composition between fresh and frozen samples

For our primary uveal melanoma and cutaneous melanoma brain metastasis TCR datasets (Fig. 1g,h), we used a hypergeometric test to compare clonotype compositions between every pair of fresh vs. frozen samples. We used the clonotype composition of the fresh sample as a reference, and compared it with the composition of the frozen sample. Specifically, we used the clonotypes that were shared between the fresh sample and frozen sample as input to the *x* argument of the dhyper function in R v4.0.2, the clonotypes found in the fresh sample as the *m* argument, clonotypes found in the frozen sample but not the fresh sample as the *n* argument, and the total number of clonotypes in the frozen sample as the *k* argument. For each sample, we also calculated the Gini coefficient of the distribution of TCR clonotype frequencies, using the *gini* function downloaded from https://github.com/oliviaguest/gini/blob/master/gini.py

Comparison of arm-level CNAs between single-cell *inferCNV* predictions and lp-WGS measurements in uveal melanoma liver metastases

Using Illumina pipelines (automatic on miSeq machine), .bam and .bai files were generated from each sample from .fastq. *ichorCNA* analysis of these files generated .seg files for visualization. GISTIC 2.0 was used to assign a copy number to each gene. We calculated average *inferCNV* scores for each chromosome arm for the cancer cells in each sample of our uveal melanoma liver metastases dataset. We then compared this with the median

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

log relative copy number measured for each chromosome arm using lp-WGS (Fig. 4c). We calculated the Spearman correlation between these two values using two settings. First, we included all chromosome arms, and second, to exclude chromosome arms that did not exhibit any large-scale amplifications or deletions, we filtered out arms that had a lp-WGS median log relative copy number between −.1 and .1.

Evaluation of batch correction methods on fresh and frozen samples

To determine the extent of integration achievable between samples of fresh and frozen origin, we applied the following set of batch correction methods using python (v3.8.3) and R (v4.1.1): *STACAS*[29] (v1.1.0), *scVI*[30] (v0.8.1) and *Seurat* (v4.0.3), (Extended Data Fig. 3,4). For integration methods, genes were filtered using *scanpy*[31] by selecting only the top 8000 highly variable genes. Integration results were visualized using UMAP. For *scVI*, nearest neighbors were computed in the reconstructed gene space (with PCA preprocessing); for *STACAS* and *Seurat*, UMAP was computed in the integrated space with PCA preprocessing. The degree of integration achieved by each method was evaluated by computing the LISI score[17]. For all methods, LISI scores were computed on 20 principal components and visualized using UMAP. The mean LISI score was also computed for each method.

Analysis of tumor clonal dynamics in the KEYNOTE-001 patient

Preprocessing—Cells from all treatment time-points were normalized by library size together and log transformed using scanpy.pp.normalize_per_cell and scanpy.pp.log1p[32] (Fig. 2a). Batch correction was not performed, due to the observation that immune cells across samples showed more overlap in cluster assignments than tumor cells. We then selected tumor cells from the normalized anndata object from the KEYNOTE-001 patient only. For our data, we verified tumor cell identification using both inference of copy number alterations that were expected to be present in malignant cells using i*nferCNV* (v1.6.0) (Fig. 2b), as well as known lineage marker genes, including *MITF*, *MLANA*, as well as the MITF-high and AXL-high signature gene sets (Fig. 2e,f, Supplementary Table 6).

Identification of tumor clones—We utilized K-means clustering on copy number data generated from *inferCNV* to group cells into clones defined by shared patterns in copy number alterations across genes (Fig. 2b). We found through visual inspection of *inferCNV* results that, across all treatment time points, there appeared to be four distinct groups of cells, each having a unique *inferCNV* footprint. Thus, K-means clustering was performed with k=4. Visualization via UMAP was utilized to analyze clonal dynamics with treatment. Temporal analysis was performed by plotting the proportion of cells belonging to each clone at each treatment time-point. Expanding tumor clones were then defined as those that showed an increase in proportion at the “on_later” time point compared to the “pre” time point (Fig. 2d).

Characterization of differentially expressed genes in expanding clones—We selected 300 differentially expressed genes for each clone, via the procedures outlined above. The differentially expressed genes for expanding Clone 2 were analyzed using the pre-ranked option in GSEA[33] and the c5.goBP curated set of genes. The normalized

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

enrichment score and false discovery rate for the top 25 enriched gene sets was visualized (Extended Data Fig. 7d).

Analysis of patterns exhibited by known genes/gene signatures— Characterization of clonal and treatment-induced dynamics was performed by analyzing the expression of genes belonging to previously defined geneset signatures, including immune checkpoint inhibitor resistance (ICR) signature[9] and the AXL-high signature[2] (Supplementary Table 6). Normalized expression was averaged across all genes belonging to the signature set, and plotted using the UMAP representation (Fig. 2e,f). Individual genes of interest were manually selected and visualized using a heatmap, where cells were grouped on the x-axis according to treatment groups and clones within each (Extended Data Fig. 7c). For each gene, z-scoring was performed across all cells to normalize the data and show variability across all cells. Analysis of *CD58* expression (Fig. 2g) was performed by partitioning cells into clones. Normalized expression of *CD58* was summarized across cells in each group, and significance was assessed using a Mann-Whitney U test.

Diffusion component (DC) analysis of T cells—For CD8+ T-cells in our three sequential anti-PD1 therapy samples (Fig. 2h), we computed DCs using the 'DiffusionMap' function of the *Destiny* R-package[34]. The 'AddModuleScore' function in Seurat was applied to calculate average expression levels of several T-cell gene signatures on a single-cell level (Supplementary Table 4). We plotted expression of several of these signatures, as well as expression of the TOX and TCF7 genes, and TCR clonotype expansion, on the first three diffusion components.

Comparison of T cells across metastatic niches—To compare the dynamics of T-cells across varying sites of origin, we extracted cells bearing the T/NK annotation from the uveal samples, the cutaneous melanoma samples, and the pre-treatment sample from the KEYNOTE-001 patient. Cells were batch corrected using *scVI*. Granular annotation was performed in a hierarchical fashion through Louvain clustering and evaluation of expression of known marker genes as well as differentially expressed genes. CD4 and CD8 T-cells were then analyzed separately, with an emphasis on dynamics within CD8. Diffusion component analysis was performed in order to visualize cells in a three-dimensional diffusion component space (Fig. 4e). We found that visualization in the space of components 1, 2, and 3 highlighted the expected TCF7 to TOX progression. Louvain clustering was performed to further aid analysis. Lastly, diffusion pseudotime analysis (DPT)[34] was applied using the *Scanpy* implementation to confirm the trajectory observed from cells high in TCF7 to cells high in TOX.

Finally, we analyzed clonal expansion of TCRs among cells of fresh and frozen origin separately (Fig. 4d). We defined clonal expansion to be cells with a TCR clonality of >1 and <10, and clonal hyperexpansion to be cells with a clonality of >=10. Cells with no expansion were labeled as "singles," while cells for which no TCR data was available were labeled as "none."

Association of copy number alterations with immune resistance signature genes—We used Fisher's exact test to determine, for each individual cell in our data,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

whether immune resistance signature genes were more likely to be located in recurrently copy number-altered regions (Extended Data Fig. 7e). To do this, we defined an individual gene as being a recurrently copy number-altered in our *inferCNV* results if its copy number is significantly amplified or deleted in >10% of cells. We then defined a chromosomal region in an individual cell to be recurrently copy number altered if >10 individual genes were recurrently altered. Finally, we used Fisher's exact test to test whether, for each individual cell, there was an association between positive and negative marker genes in the immune resistance signature, compared to whether a marker gene was located in a recurrently amplified or deleted gene. For the vast majority of cells, we find no significant association ($p > 0.05$).

SCENIC analysis of sequential treatment samples

We performed analysis of putative gene regulatory networks using *SCENIC*[35] on tumor cells, using a list of 1390 curated TFs obtained from the *SCENIC* github site (https://github.com/aertslab/pySCENIC/hs_hgnc_curated_tfs.txt). We calculated AUCell scores for the strength of transcriptional programs mediated by these TFs. We then used the Wilcoxon rank-sum test to determine if the AUCell scores for any TF were significantly different in copy number clone 2 versus the other clones in each of our samples. We used a maximum threshold of .05 for Bonferroni-adjusted p-value, and a minimum threshold of 0.25 for log-fold change of AUCell score in clone 2 versus all other clones. Furthermore, we then determined TFs that passed these thresholds in all three timepoints (pre, on and on_later), resulting in a list of 26 TFs.

Analysis of spatial transcriptomic data

Preprocessing—We used the Slide-seq tools pipeline to process our Slide-seq data, and loaded the data into Seurat.

***Deconvolution of cell types* (RCTD)—**We used the *RCTD*[22] pipeline (v1.2.0), 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. The count matrices obtained after preprocessing of the *Slide-seq V2* assays of sequential treatment samples served as the first input. For the reference, we used previously published and annotated non-spatial single-nuclei sequencing data[5]. *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 (Extended Data Fig. 6g).

Mapping of signatures in spatial datasets—We used the FindMarkers function in Seurat, with default parameters, to find marker genes that differentiate clone 2 tumor cells in the sequential therapy dataset from non-clone 2 tumor cells. We then used the AddModuleScore function in Seurat to determine strength of the clone 2 signature as well as immune resistance signature in our sequential therapy *Slide-seq V2* samples.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

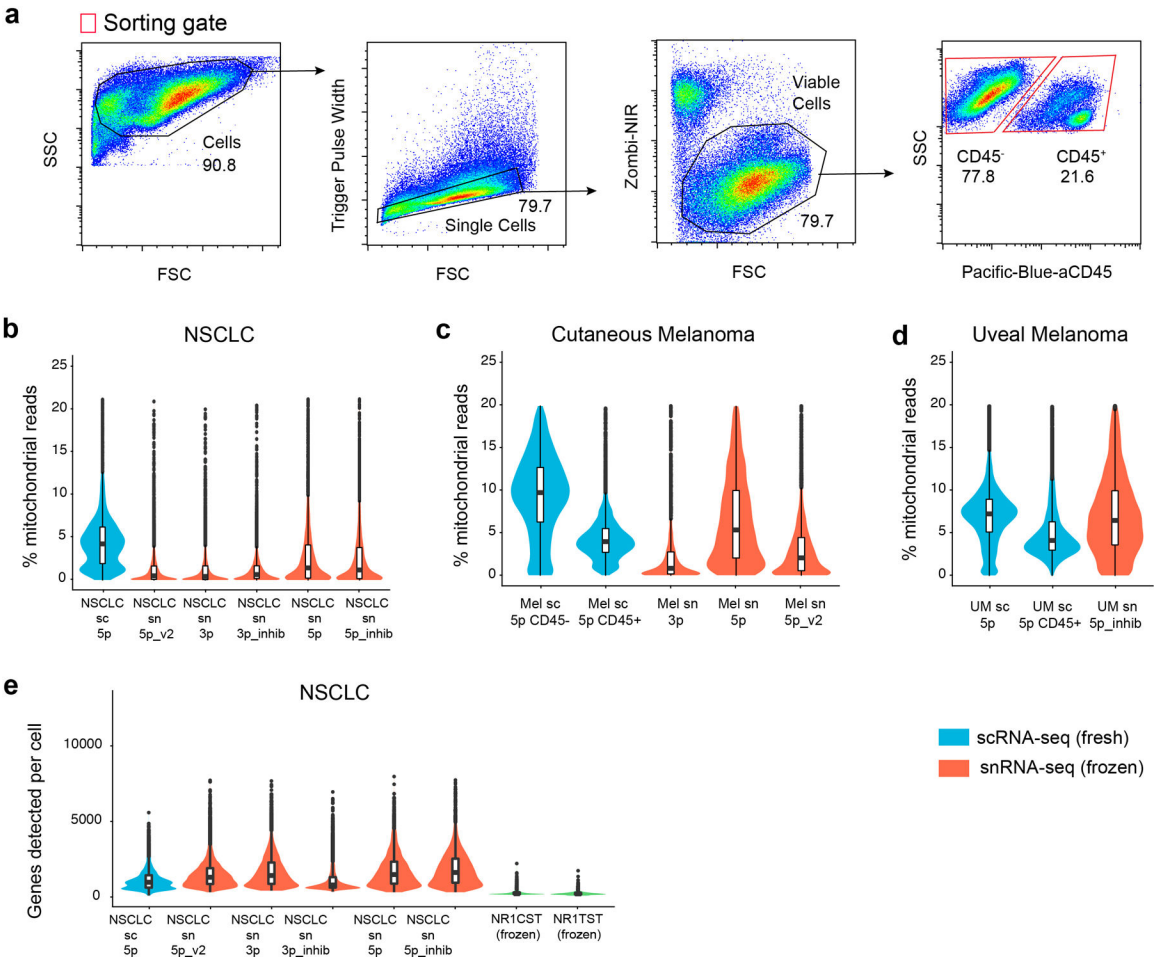
Spatial mapping of T cell states—To further deconvolve T cell states in the *Slide-seq V2* data, we applied our recently developed method *Starfysh* (Spatial Transcriptomic Analysis using Reference-Free auxiliarY deep generative modeling and Shared Histology) (He, et al, in preparation)[23]. In summary, Starfysh leverages a deep generative model with auxiliary variables representing cell state proportions in spots or barcoded areas. Importantly, Starfysh does not require paired single-cell RNA-seq data as reference and instead uses two sets of features to guide the deconvolution: (1) expression of known marker gene sets for cell states and (2) archetypal analysis which identifies the extreme vertices of a polytope encompassing all spots.We adapt *Starfysh* for use with *Slide-seq V2* data by reading the coordinates obtained from *Slide-Seq* into an AnnData object that can then be interfaced with downstream *Starfysh* methods. We then ran *Starfysh* with a set of immune and melanoma markers in order to obtain inferred proportions. The proportions of activated CD8 T cell and dysfunctional CD8 T cells (defined by published signatures[20], Supplementary Table 7) are shown in Fig. 3b and Extended Data Fig. 6f.

Statistics and reproducibility

Samples for comparison of scRNA-seq and snRNA-seq were based on available tissue with sufficient material for fresh and consecutive frozen profiling as collected in the Izar laboratory from 2019–2021. For clinical trial samples from the liver metastasis trial all patients from the original trial with tissue from pre, on, and post time points were included. For the case-study of a partial response lesion from KEYNOTE-001 a patient with samples obtained pre therapy and on two post therapy timepoints was selected based on sufficient material for matched snRNA-seq and spatial profiling. No statistical method was used to predetermine sample size. No data were excluded from the analysis. Upper and lower edges of boxplot indicate 75th and 25th percentiles respectively, and middle line indicates median. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

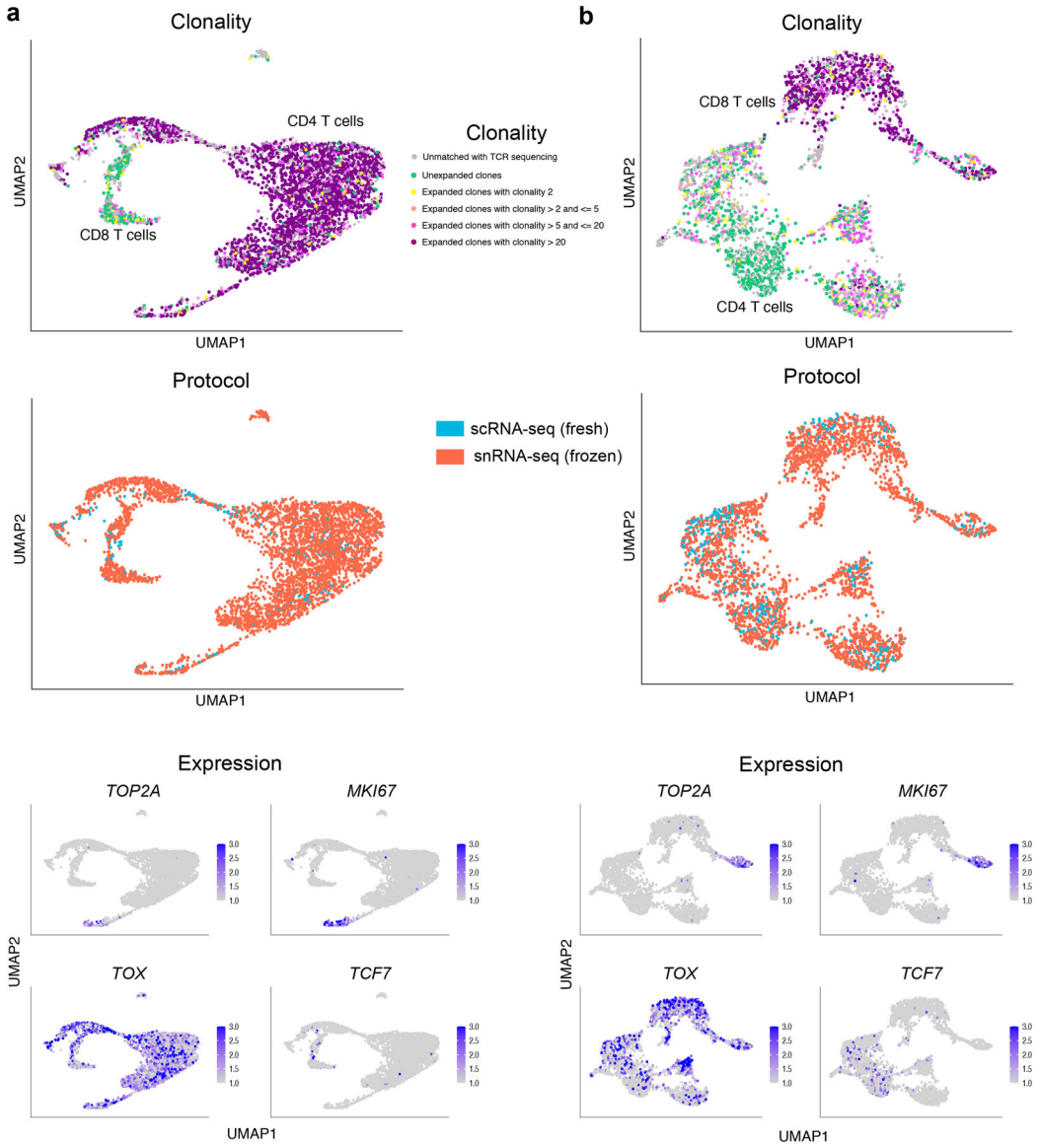
Extended Data



Extended Data Figure 1:

a, Gating strategy for sorting viable CD45+ and CD45- cells from freshly digested surgical specimens. Final sorting gates indicated in red. **b-d**, Violin plots and boxplots indicating percent of mitochondrial reads across samples and different experimental settings in (b) NSCLC, (c) cutaneous melanoma, (d) uveal melanoma. **e,** Genes detected per cell in the NSCLC sample using procedures presented here compared to two previously reported protocols (CST and TST, indicated in green) after adjustment for sequencing saturation (Methods). Upper and lower edges of boxplot indicate 75th and 25th percentiles respectively, and middle line indicates median.

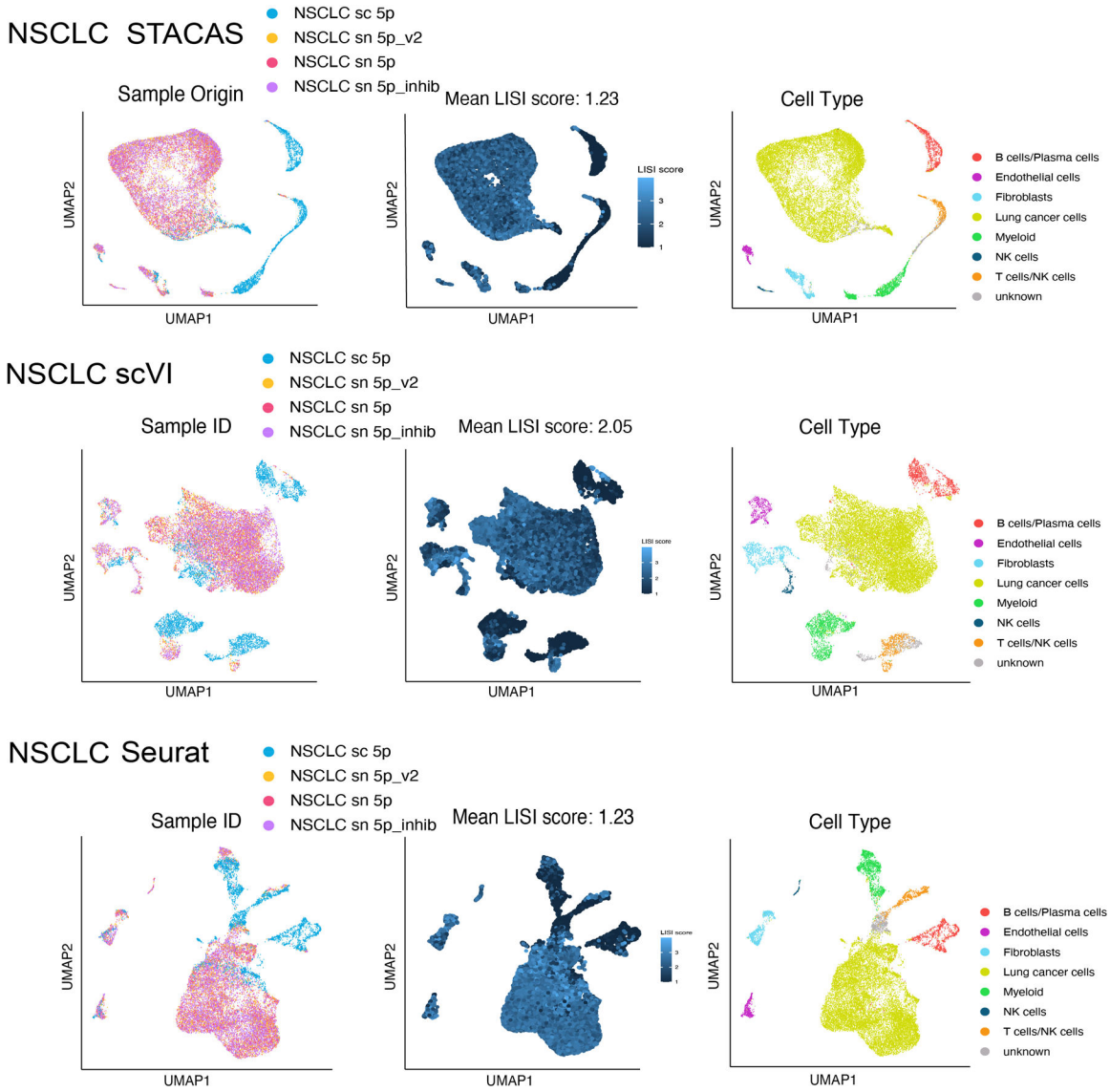
Author Manuscript Author Manuscript Author Manuscript Author Manuscript



Extended Data Figure 2:

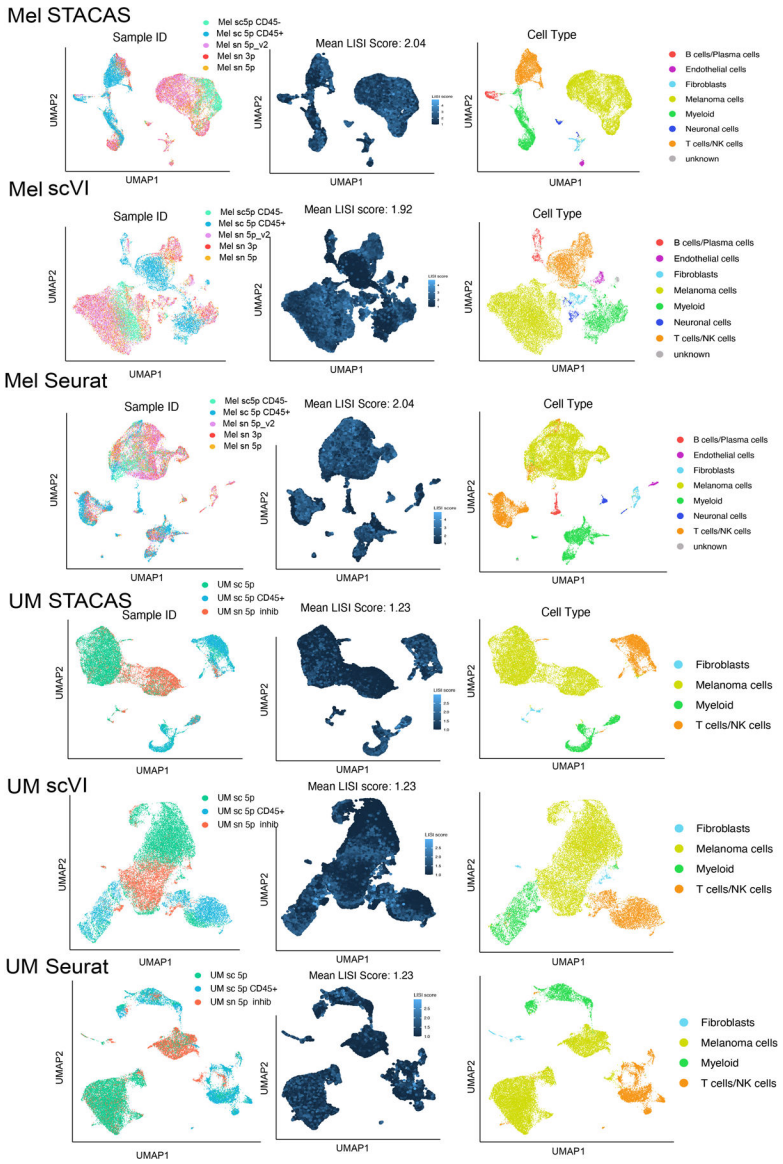
a, UMAP clustering of T cells, with projected clonality (top), sequencing source (middle) and cell cycle markers and T cell dysfunction and stemness markers (bottom) in primary uveal melanoma. **b**, same as (a) for cutaneous melanoma.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



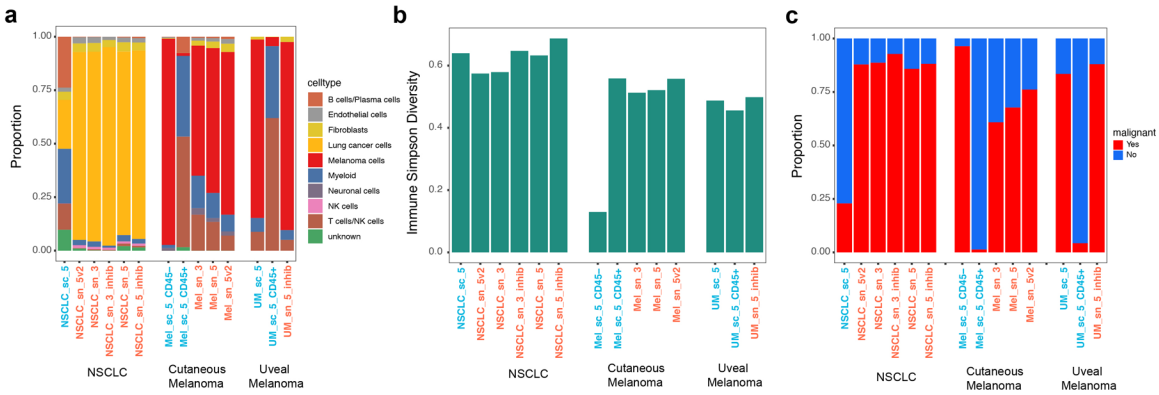
Extended Data Figure 3:
Application of three integration methods in NSCLC.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



Extended Data Figure 4:

Application of three integration methods across cutaneous and uveal melanoma.

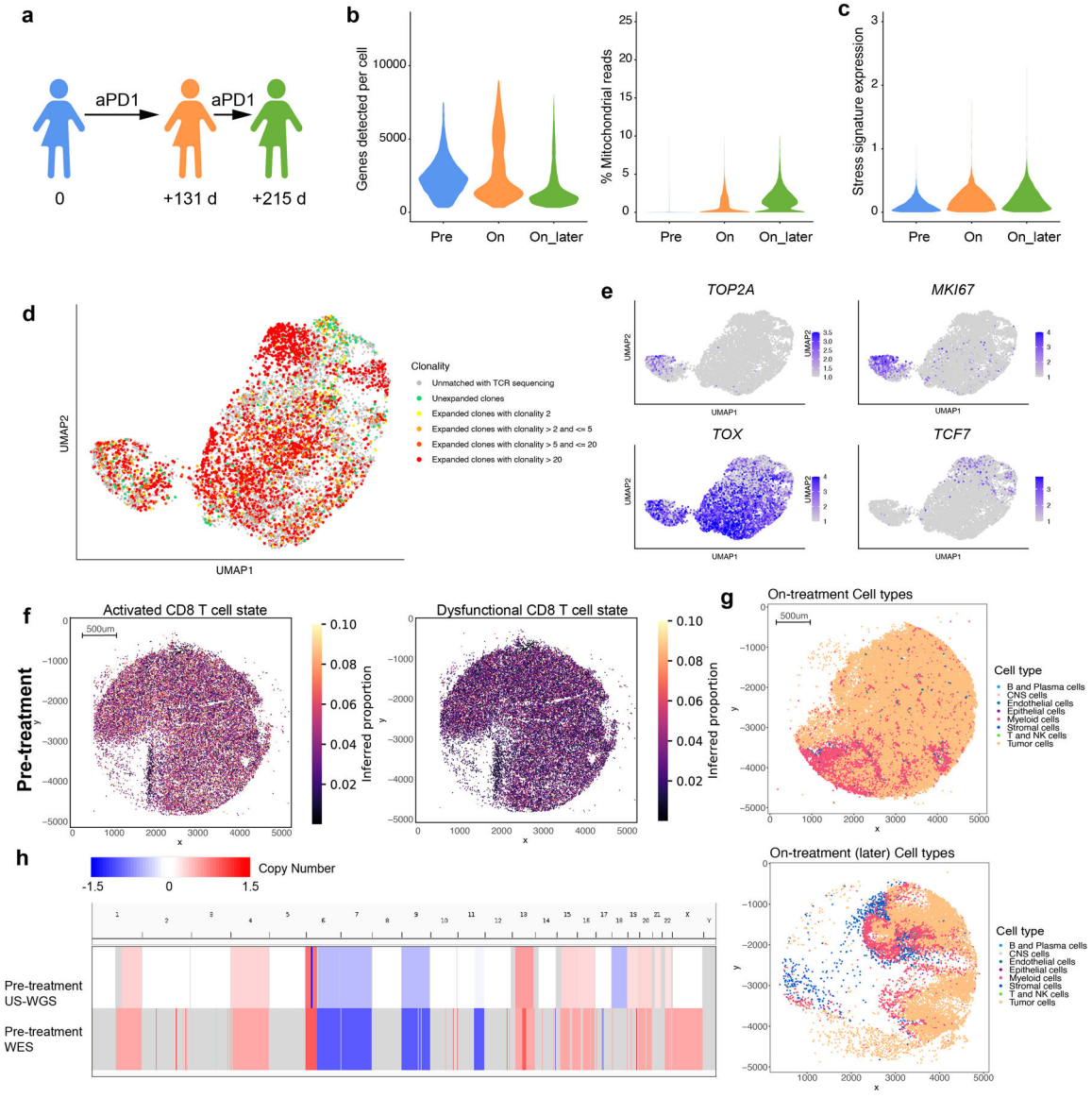


Extended Data Figure 5:

a, Stacked bar plots indicating proportions of all cell types in NSCLC, cutaneous and uveal melanoma across different methods. **b**, Simpson diversity index for immune cells in the same samples as (a). **c**, Stacked bar plots of malignant and non-malignant cell fractions in the same samples as (a).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

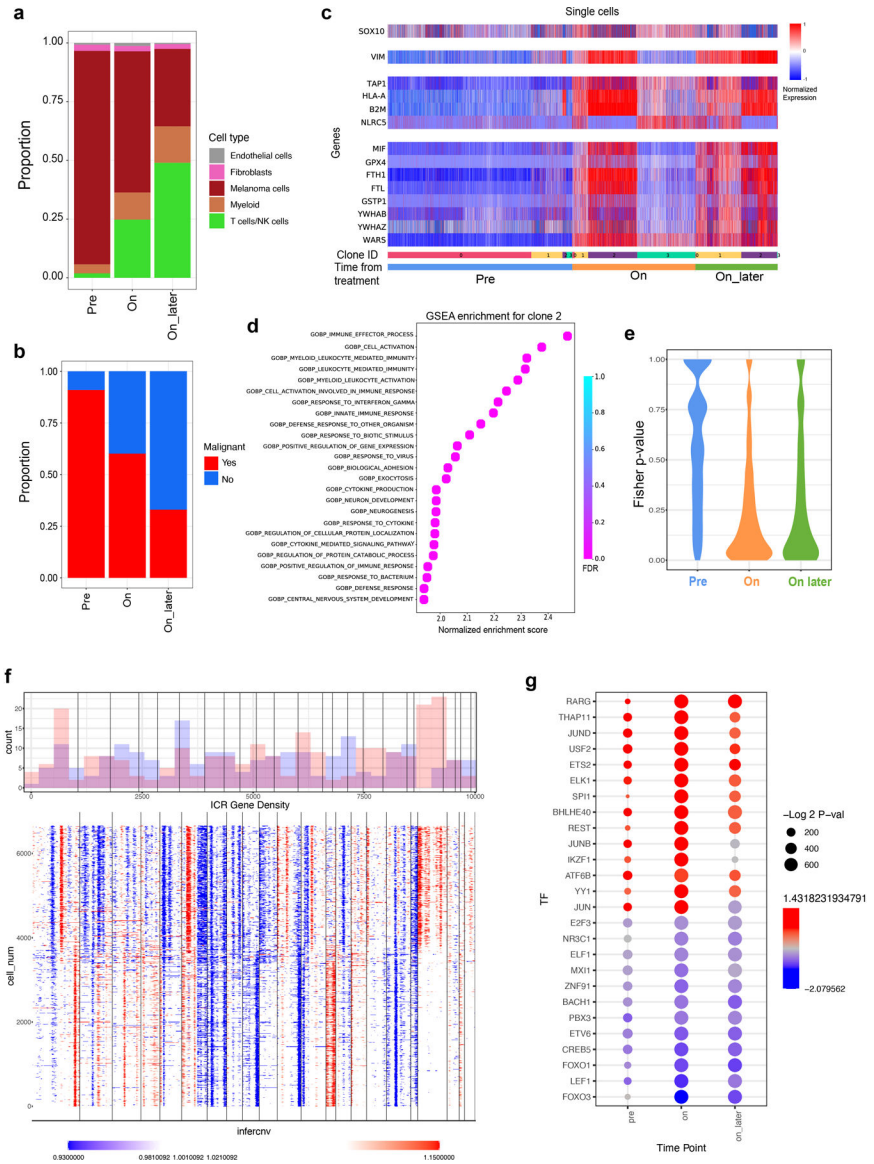
Author Manuscript Author Manuscript Author Manuscript Author Manuscript



Extended Data Figure 6:

a, Timing of sequentially collected specimens in a patient on anti-PD1 therapy. **b,c**, Violin plots and boxplots of (b) genes per cell detected (left) and percent of mitochondrial reads (right), and (c) expression of artifactual signature across samples collected over different time points. Upper and lower edges of boxplot indicate 75th and 25th percentiles respectively, and middle line indicates median. **d,e**, UMAP representation of CD8+ T cells across all time points and with projected TCR clonality, and (e) cell cycle markers (top) and T cell dysfunction and Stemness markers (bottom). **f,** Starfysh inferred proportions of cells corresponding to the CD8+ T cell activation and CD8+ T cell dysfunction signatures (columns), for data collected at the pre-treatment time point. **g**, Major cell types deconvolved using RCTD in on-treatment and on-treatment (later) timepoints. **h,** IGV plots of copy number changes measured in WES and US-WGS analysis of a pre-treatment melanoma specimen (presented in Fig. 2).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

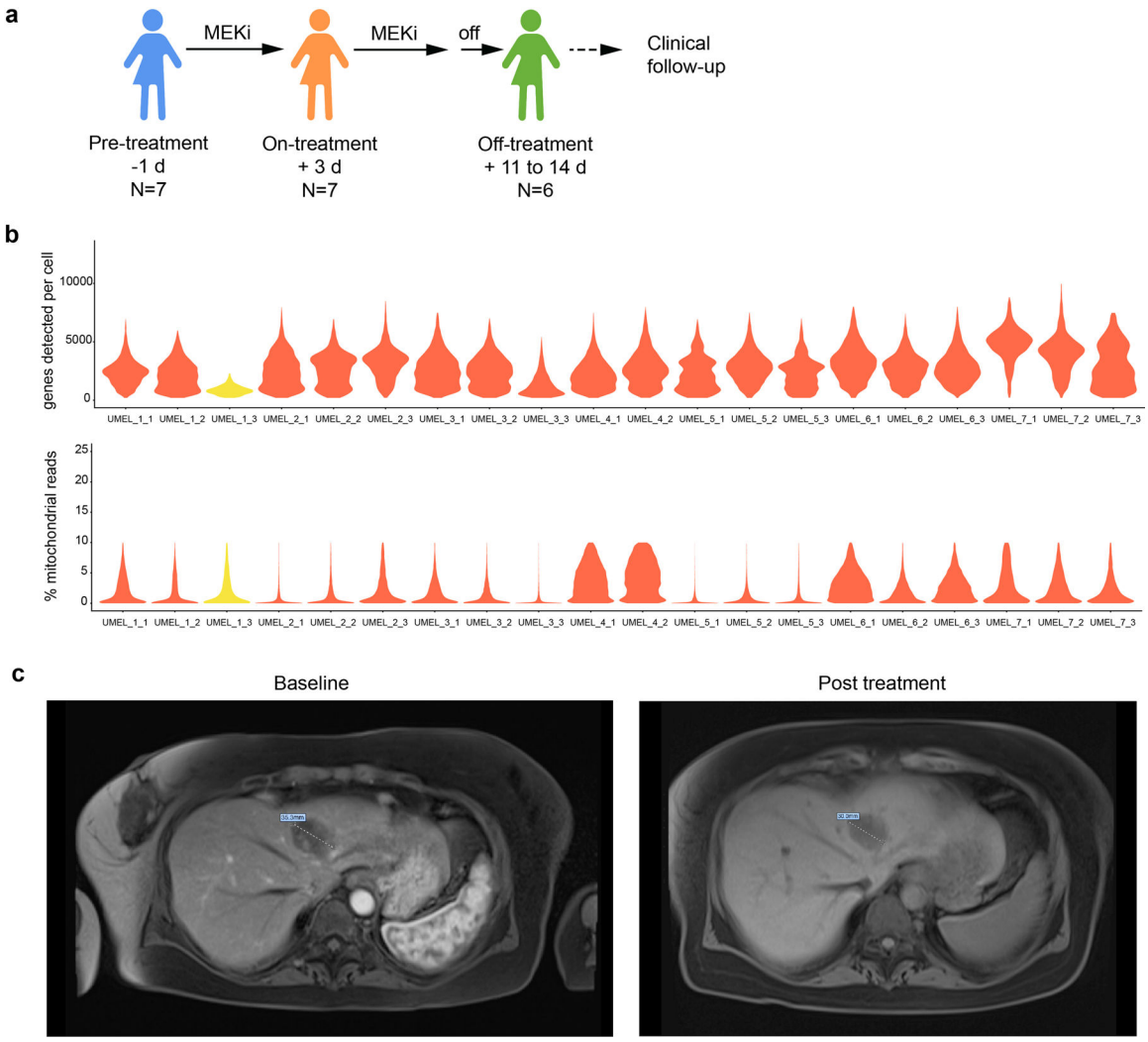


Extended Data Figure 7:

a,b, Stacked bar plots indicating proportion of (a) all cell types across sequentially collected anti-PD1 therapy tissue specimens and (b) malignant and non-malignant fractions. **c**, Heatmap of selected genes (rows) and their gene expression (normalized expression) in individual cells (column). Indicated on the bottom are time points of sample collection (pre, on, on_later) and clones (0–3) as defined in Fig. 2. **d**, Gene set enrichment analysis (GSEA) of genes differentially expressed in clone 2. **e**, Violin plots of the distribution of p-values from Fisher's exact test in pre, on and on_later timepoints (9969, 6651, and 4447 cells respectively), testing for association of recurrent copy number alterations in each cell with ICR signature gene location. **f,** Top: Histogram of gene density across genome for positive and negative genes in ICR signature. Bottom: Plot of *inferCNV* inferred copy number alterations in on timepoint sample. **g**, Transcription factors associated with Clone 2 in pre,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

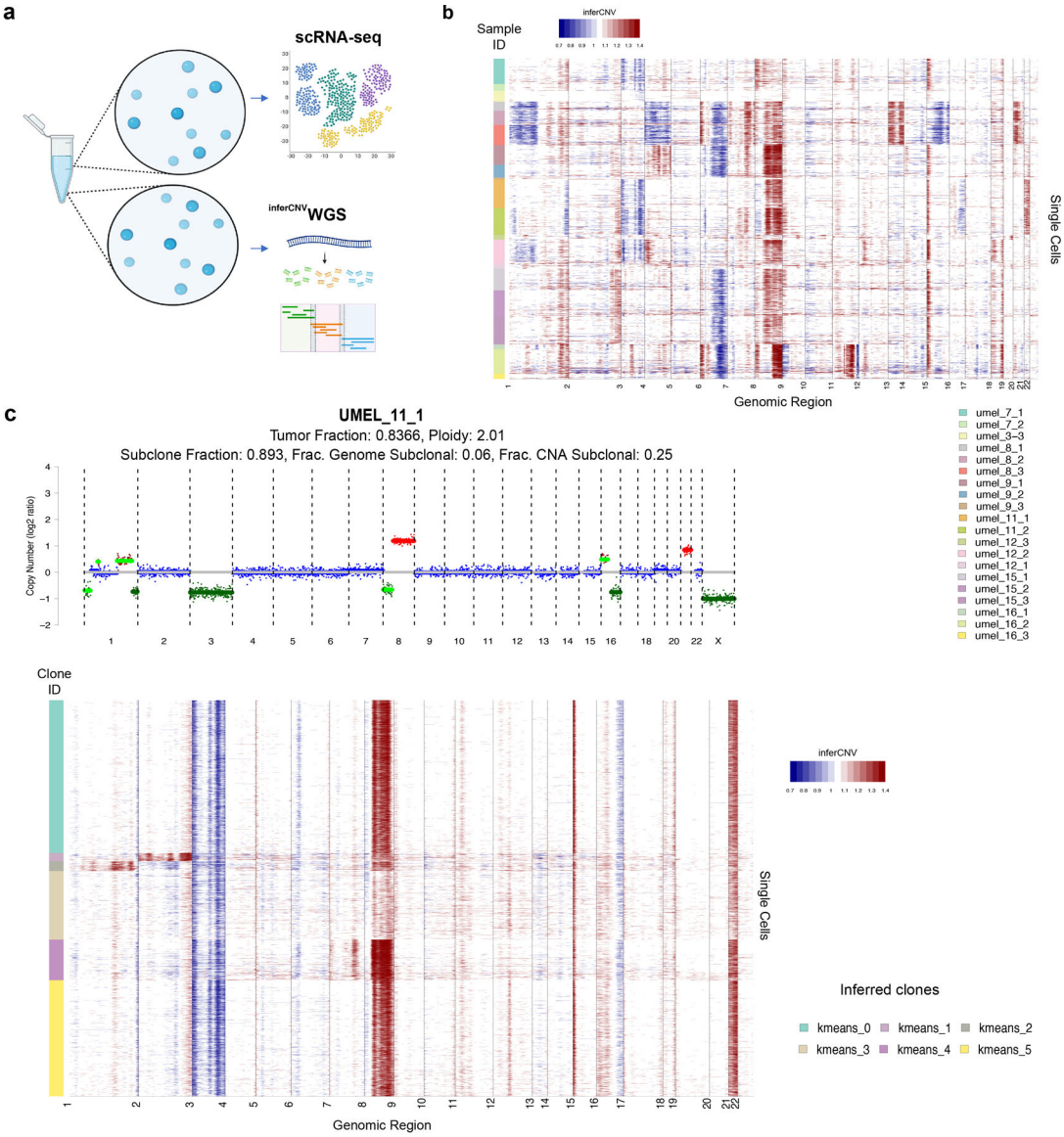
on-treatment and on-treatment later timepoints, based on Wilcoxon rank-sum test of AUCell scores in Clone 2 cells vs. non-Clone 2 cells () from SCENIC analysis.



Extended Data Figure 8:

a, Schematic of tissue collection and indicated number of specimens per time point. MEKi, MEK-inhibitor (Selumetinib). **b**, Violin plots and boxplots indicating number of genes detected per cell (top lane), percent of mitochondrial reads (middle lane) and expression of a stress signature (bottom lane) across 20 uveal melanoma specimens. Upper and lower edges of boxplot indicate 75th and 25th percentiles respectively, and middle line indicates median. Yellow color indicates a specimen sequenced with 10X 3' V3 chemistry, with lower quality, while data for the remainder of samples (indicated in red) were generated with 5' chemistry. **c,** Exemplary baseline and post-treatment MRI of the abdomen showing moderate response of a liver metastatic lesion in a uveal melanoma patient treated with selumetinib.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



Extended Data Figure 9:

a, Schematic design of generation of (low-pass) whole-genome sequencing from the same cell/nucleus pool that was also used for single-nucleus RNA and TCR sequencing. **b**, Inferred CNAs (columns) across samples (indicated by bar on the left) in the uveal melanoma cohort. **c**, Exemplary whole-genome sequencing result (top) showing copy number alterations (y axis, log2 ratio) with amplifications in red, deletions in green and unaltered chromosome regions in blue. Inference of CNAs of the using snRNA-seq that was generated from the same starting cell/nucleus pool as WGS.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

ACKNOWLEDGEMENTS

We thank Hanina Hibshoosh at Columbia University for fruitful discussions. We thank the members of the CUIMC Human Immune Monitoring Core for technical advice. Y.W. is supported by National Institute of Health (NIH), National Institute of Allergy and Infectious Disease (NIAID) training grant T32AI148099. B.I. is supported by NIH, National Cancer Institute (NCI) grants K08CA222663, R37CA258829, R01CA266446, U54CA225088, a Burroughs Wellcome Fund Career Award for Medical Scientists, a Velocity Fellows Award, the Louis V. Gerstner, Jr. Scholars Program and a Young Investigator Award by the Melanoma Research Alliance. R.D.C., E.A. and B.I. are supported by NCI grant R21CA263381 and a Columbia University Research Initiatives in Science & Engineering Award. E.A. was supported by NCI grant R00CA230195 and NSF grant CBET-2144542. J.L.F. acknowledges support from the Columbia University Van C. Mow fellowship. G.A-R. and A.R. are supported by Parker Institute for Cancer Immunotherapy and NIH grant P01CA168585. AMT is supported by NCI 5K22CA237733-03. This work was supported by NIH/NCI Cancer Center Support Grant P30CA013696, the Molecular Pathology Shared Resource and its Tissue Bank at Columbia University and the Flow-cytometry Core Facility supported by S10OD020056.

Data availability statement

Processed data is available on GEO accession number: GSE192402. Raw data are available on dbGAP: accession number phs003097.v1.p1.

REFERECES

1. Patel AP et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science 344, 1396–1401 (2014). [PubMed: 24925914]
2. Tirosh I et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science 352, 189–196 (2016). [PubMed: 27124452]
3. Azizi E et al. Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. Cell 174, 1293–1308.e36 (2018). [PubMed: 29961579]
4. Izar B et al. A single-cell landscape of high-grade serous ovarian cancer. Nat Med 26, 1271–1279 (2020). [PubMed: 32572264]
5. Biermann J et al. Dissecting the treatment-naive ecosystem of human melanoma brain metastasis. Cell 185, 2591–2608.e30 (2022). [PubMed: 35803246]
6. Puram SV et al. Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. Cell 171, 1611–1624.e24 (2017). [PubMed: 29198524]
7. Laughney AM et al. Regenerative lineages and immune-mediated pruning in lung cancer metastasis. Nat Med 26, 259–269 (2020). [PubMed: 32042191]
8. Gonzalez H et al. Cellular architecture of human brain metastases. Cell 185, 729–745.e20 (2022). [PubMed: 35063085]
9. Jerby-Arnon L et al. A Cancer Cell Program Promotes T Cell Exclusion and Resistance to Checkpoint Blockade. Cell 175, 984–997.e24 (2018). [PubMed: 30388455]
10. Yost KE et al. Clonal replacement of tumor-specific T cells following PD-1 blockade. Nat Med 25, 1251–1259 (2019). [PubMed: 31359002]
11. Luoma AM et al. Tissue-resident memory and circulating T cells are early responders to pre-surgical cancer immunotherapy. Cell 185, 2918–2935.e29 (2022). [PubMed: 35803260]
12. Hamid O et al. Safety and tumor responses with lambrolizumab (anti-PD-1) in melanoma. N. Engl. J. Med 369, 134–144 (2013). [PubMed: 23724846]
13. Carvajal RD et al. Effect of selumetinib vs chemotherapy on progression-free survival in uveal melanoma: a randomized clinical trial. JAMA 311, 2397–2405 (2014). [PubMed: 24938562]
14. Slyper M et al. A single-cell and single-nucleus RNA-Seq toolbox for fresh and frozen human tumors. Nat Med 26, 792–802 (2020). [PubMed: 32405060]
15. Melms JC et al. A molecular single-cell lung atlas of lethal COVID-19. Nature (2021) doi:10.1038/s41586-021-03569-1.
16. van den Brink SC et al. Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. Nat Methods 14, 935–936 (2017). [PubMed: 28960196]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

17. Korsunsky I et al. Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods 16, 1289–1296 (2019). [PubMed: 31740819]
18. Frangieh CJ et al. Multimodal pooled Perturb-CITE-seq screens in patient models define mechanisms of cancer immune evasion. Nat Genet 53, 332–341 (2021). [PubMed: 33649592]
19. Ho P et al. The CD58:CD2 axis is co-regulated with PD-L1 via CMTM6 and governs anti-tumor immunity 2022.03.21.485049 Preprint at 10.1101/2022.03.21.485049 (2022).
20. Bachireddy P et al. Mapping the evolution of T cell states during response and resistance to adoptive cellular therapy. Cell Rep 37, 109992 (2021). [PubMed: 34758319]
21. Stickels RR et al. Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. Nat Biotechnol 39, 313–319 (2021). [PubMed: 33288904]
22. Cable DM et al. Robust decomposition of cell type mixtures in spatial transcriptomics. Nat Biotechnol 40, 517–526 (2022). [PubMed: 33603203]
23. https://zenodo.org/record/6950761#.YveolOzMLeB https://github.com/azizilab/starfysh.
24. Turke AB et al. Pre-existence and clonal selection of MET amplification in EGFR mutant NSCLC. Cancer Cell 17, 77–88 (2010). [PubMed: 20129249]
25. Yang J, Manson DK, Marr BP & Carvajal RD Treatment of uveal melanoma: where are we now? Ther Adv Med Oncol 10, 1758834018757175 (2018). [PubMed: 29497459]
26. Adalsteinsson VA et al. Scalable whole-exome sequencing of cell-free DNA reveals high concordance with metastatic tumors. Nat Commun 8, 1324 (2017). [PubMed: 29109393]
27. Hwang WL et al. Single-nucleus and spatial transcriptome profiling of pancreatic cancer identifies multicellular dynamics associated with neoadjuvant treatment. Nat Genet 54, 1178–1191 (2022). [PubMed: 35902743]
28. Mermel CH et al. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biology 12, R41 (2011). [PubMed: 21527027]
29. STACAS: Sub-Type Anchor Correction for Alignment in Seurat to integrate single-cell RNA-seq data | Bioinformatics | Oxford Academic https://academic.oup.com/bioinformatics/article/37/6/882/5897412.
30. Gayoso A et al. scvi-tools: a library for deep probabilistic analysis of single-cell omics data 2021.04.28.441833 10.1101/2021.04.28.441833v1 (2021) doi:10.1101/2021.04.28.441833.
31. Wolf FA, Angerer P & Theis FJ SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. 19, 15 (2018). [PubMed: 29409532]
32. Virtanen P et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods 17, 261–272 (2020). [PubMed: 32015543]
33. Subramanian A et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U.S.A 102, 15545–15550 (2005). [PubMed: 16199517]
34. Angerer P et al. destiny: diffusion maps for large-scale single-cell data in R. Bioinformatics 32, 1241–1243 (2016). [PubMed: 26668002]
35. Van de Sande B et al. A scalable SCENIC workflow for single-cell gene regulatory network analysis. Nat Protoc 15, 2247–2276 (2020). [PubMed: 32561888]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

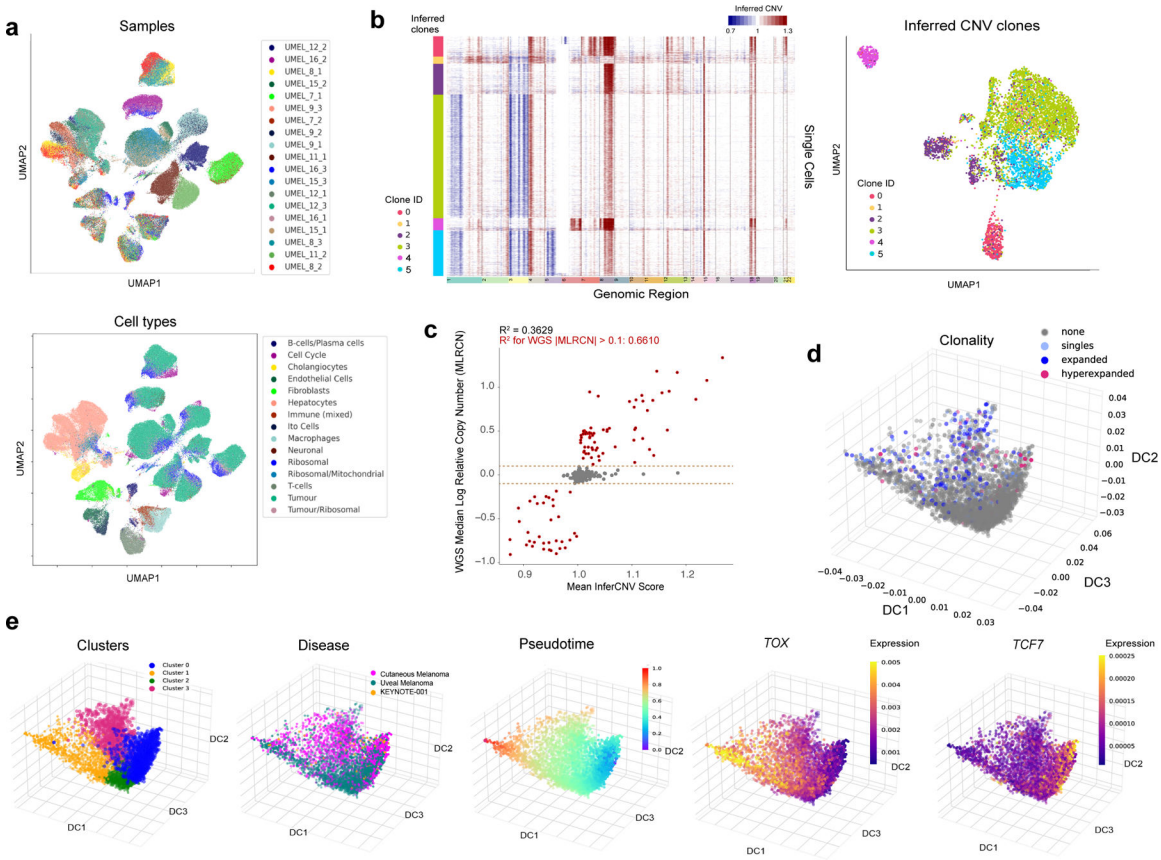


Figure 1.

a-c, Violin plots and boxplots of genes detected per cell in (a) non-small cell lung cancer, (b) cutaneous melanoma and (c) uveal melanoma samples. Upper and lower edges of boxplot indicate 75th and 25th percentiles respectively, and middle line indicates median. Blue plots indicate scRNA-seq from fresh tissue, and red indicates snRNA-seq from frozen tissue. 10X chemistry type and presence/absence of RNAse inhibitor is indicated on labels beneath each violin. **d-f**, Violin plots and boxplots of relative expression of an artifactual stress-associated gene expression signature. Samples and experimental settings corresponding to panels (a-c) above. **g,h,** Circos plots of T cell receptor clonotypes in (g) cutaneous and (h) uveal melanoma, respectively. Connections indicate overlap of identical TCRs found in both fresh scRNA-seq and frozen snRNA-seq samples (hypergeometric test, (g) p=1.55e-62, using the dhyper function in R with x=66, m=1204, n=55, and k=121, and (h) p=0.0018, between the single-cell 5p CD45+ and single-nuclei 5p with inhibitor samples, using the dhyper function in R with x=19, m=488, n=2, and k=21). **i,** Spearman correlation of average cutaneous melanoma cell arm-level CNAs predicted by *inferCNV*, in a fresh CD45- scRNA-seq sample (x-axis) vs. two frozen 5p snRNA-seq protocols (y-axis).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

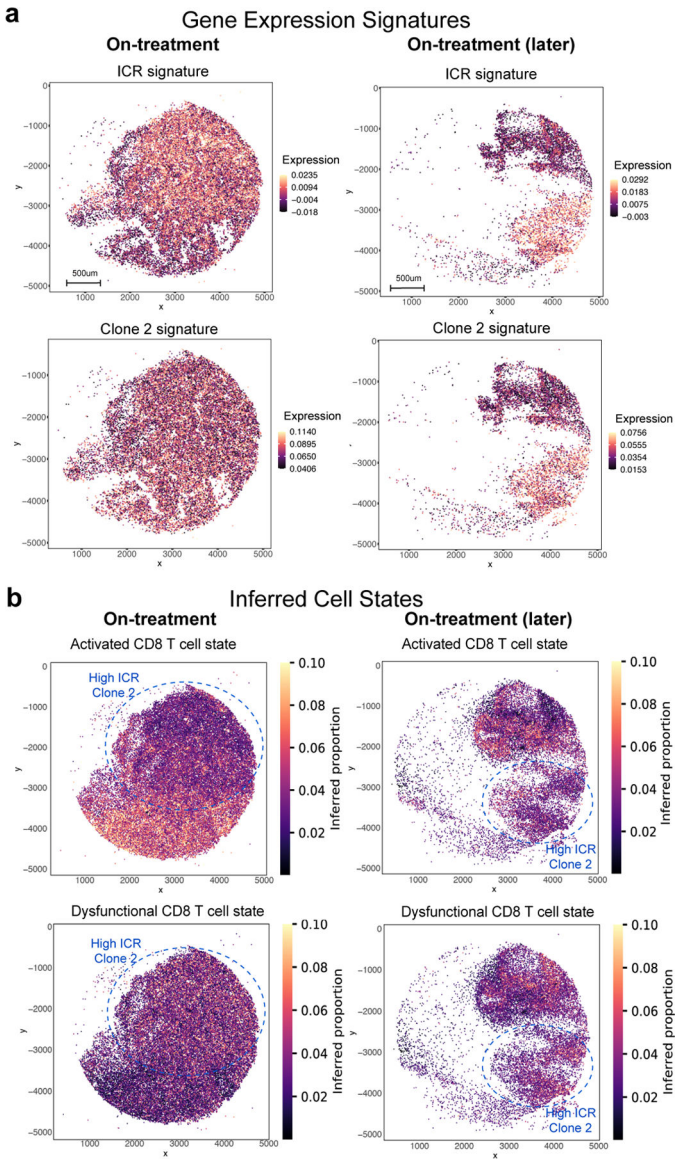


Figure 2.

a, Merged, unintegrated UMAP and annotation of clusters of cells from sequentially collected specimens during anti-PD-1 therapy. **b**, Inferred copy number alterations across the chromosomal landscape for melanoma cells, in pre- and sequential on-treatment biopsies (from left to right). Genomic location is indicated across the x axis, with chromosomes delineated by vertical lines. Individual cells are plotted along the y axis, with each row representing the CNA profile of one cell, and amplifications in red and deletions in blue. Colored bar to the left indicates clones identified by k-means clustering. **c**, Merged, unintegrated UMAP embedding of cancer cells from three time points indicated by different colors. **d**, (left) Same projection as (c) indicating cancer clones defined by aneuploidy patterns in (b) and their proportion (right) across pre- and sequential on-treatment biopsies. **e,f,** Same projection as in (c) showing expression of (e) immunotherapy resistance program[9] and (f) AXL-signature[2]. **g**, Violin plots of expression of *CD58* in emerging clone 2

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

compared to other cancer clones based on aneuploidy patterns. (n=7921, 4858, 4808, 3486; box denotes Q1-Q3 and whiskers correspond to farthest data point within 1.5*IQR). **h**, Diffusion component (DC) analysis of CD8+ T cells with projections of cells in first 3 DCs colored by indicated genes, signatures and clonotypes. **i**, Circos plots of T cell receptor clonotypes across different time points (indicated on different aspects of the circle). Connections indicate overlap of identical TCRs between time points.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

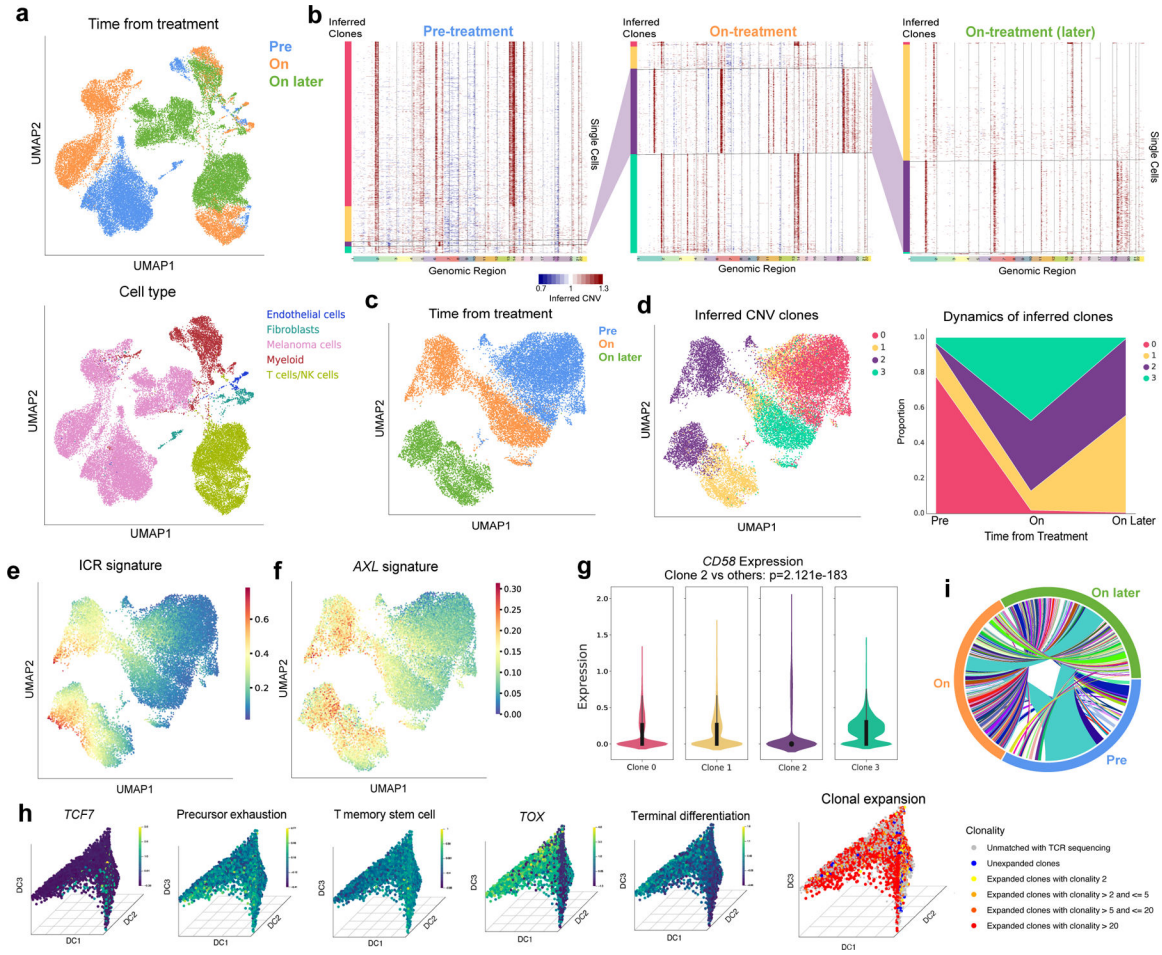


Figure 3.

a-b, Representation of spatial transcriptomics generated with *Slide-seq V2* with indicated expression of the immune resistance signature (ICR) (first row), clone projection (second row) (a) and distribution of deconvolved T cell states (b) in the on-treatment (first column) and on-treatment later (second column) specimens. ICR and clone 2 signatures were defined by average expression of marker genesets, and differentially expressed genes, respectively. Deconvolution of cell states was achieved with *Starfysh*. Major cell types inferred with RCTD are presented in Extended Data Fig. 6g.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

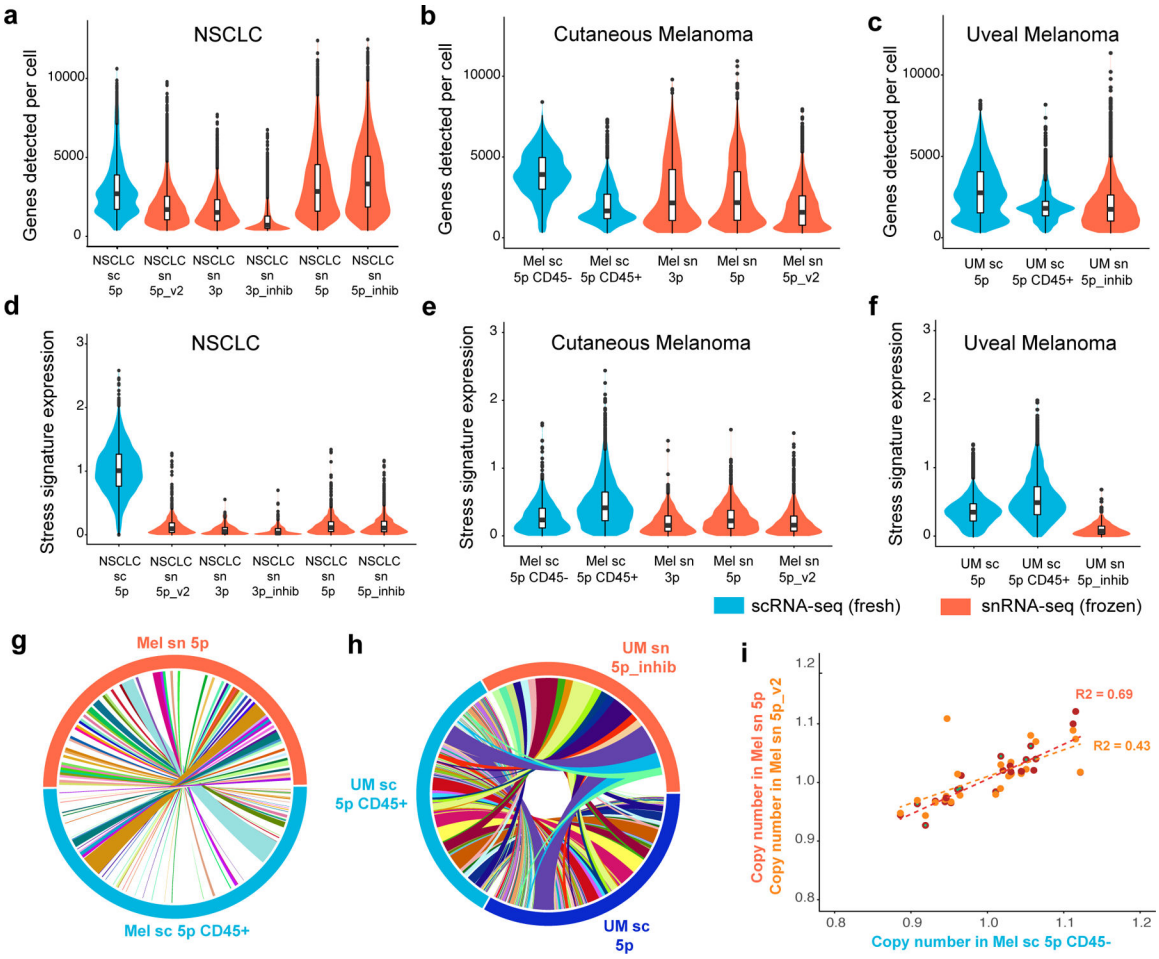


Figure 4.

a, Merged, unintegrated UMAP and annotated clusters of snRNA-seq transcriptomes across 20 uveal melanoma liver metastasis samples colored by specimen of origin (left) and cell type (right). **b**, Exemplary representation of inferred copy number alterations across the chromosomal landscape of an uveal melanoma liver metastasis specimen (left; color bar on left delineating CNA clones identified by k-means clustering), and corresponding UMAP embedding and clustering (right) colored by respective CNA clones, demonstrating impact of CNA heterogeneity on Transcriptional output. **c**, Spearman correlation of chromosome arm copy number alterations predicted by *inferCNV* from uveal melanoma liver metastases snRNA-seq data vs. population-matched low-pass whole-genome sequencing. **d**, TCR clonotypes of CD8+ T cells from the uveal melanoma samples, the cutaneous melanoma samples, and the pre-treatment sample from the KEYNOTE-001 patient projected onto the first 3 diffusion components. **e**, CD8+ T cells projected into diffusion component space (same as d), colored by cluster, tissue origin, inferred pseudotime, and T cell differentiation in the *TCF7+* to *TOX+* continuum.