Cellular Dynamics Upon Immune Checkpoint Inhibition

Zhongyang Lin[1], Ekaterina Petrenko[1], Mitalee Chandra[2,3], Nalini Srinivas[2,3,4], Selma Ugurel[3,4], Jürgen C. Becker[2,3,4*], & Dvir Aran[1,6*]

[1] Faculty of Biology, Technion-Israel Institute of Technology, Haifa, Israel

[2] Translational Skin Cancer Research, German Cancer Consortium (DKTK), Partner Site Essen, Medical Faculty, University of Duisburg-Essen, Essen, Germany

[3] German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany

[4] Department of Dermatology, University Hospital Essen, Essen, Germany

The Taub Faculty of Computer Science, Technion-Israel Institute of Technology, Haifa, Israel.

* Equal last authors

Corresponding author: Dvir Aran

21 Sderot Rose, Technion-Israel Institute of Technology, Haifa, Israel, 3095230

dviraran@technion.ac.il

Keywords: Immunotherapy; Single-cell RNA-seq; Tumor microenvironment; Treatment Response

Abstract

Cellular dynamics in the tumor microenvironment upon immune checkpoint inhibition (ICIs) treatment remain largely known. To elucidate these dynamics and associated mechanisms of response and resistance, we analyzed a large-scale single-cell RNA-sequencing dataset of 326 samples from 163 patients across multiple cancer types, with longitudinally site-matched samples. We developed a robust, integration-free deep phenotyping framework, and annotated 1,022,526 cells into 83 granular cell states. Our analysis revealed consistent compositional changes in 19 cell subtypes during ICI treatment, including enhanced adaptive immune responses (e.g., CD8+ IL7R+ central memory T cells, GZMK+ effector memory T cells, germinal center B cells), reduced interferon-responsive immune cells, quiescent CD4+CREM- memory T cells, and angiogenesis tip cells. Co-regulatory network analysis revealed a continuum between two patterns resembling tertiary lymphoid structure maturity. Accumulation of interferon-responsive subtypes was associated with T cell expansion. We further identified a B cell-centric response-specific pattern indicating an enhanced early-phase adaptive immune response favored treatment efficacy, which was subsequently validated through supervised subgroup analysis. Our study presents a comprehensive landscape of TME cellular dynamics during ICI treatment, showcasing the importance of a systems-level understanding in improving patient stratification and advancing personalized cancer immunotherapy.

Introduction

The tumor microenvironment (TME) is pivotal in shaping the dynamics of tumor progression and immune responses[1-3]. Thus, this intricate network comprising malignant cells, immune cells, and non-immune stromal components, has become a focal point of contemporary therapeutic strategies[4]. These strategies have evolved from targeting cancer cells directly to modulating the immune system, inhibiting angiogenesis, and disrupting the supportive roles of cancer-associated fibroblasts (CAFs). Among these approaches, the emergence of immune checkpoint inhibitors (ICIs) – particularly those targeting the programmed cell death protein 1 (PD1), cytotoxic T-lymphocyte-associated protein 4 (CTLA4) has marked a significant advancement, especially in the treatment of melanoma[5-7].

Despite the notable success of ICIs in increasing survival rates across several cancer types, a significant proportion of patients remain unresponsive to this therapy or develop secondary resistance. For instance, while approximately half of the metastatic melanoma patients experience a long-lasting response to ICIs[8,9], the response rates in cancers like non-small lung cell cancer and renal cell cancer are below 30%[6,10]. Moreover, the treatment is often accompanied by severe adverse effects, with about 60% of melanoma patients experiencing high-grade toxicities, leading to the discontinuation of therapy[6,11,12]. These challenges, coupled with the high costs of ICI treatment, underscore the urgent need for enhanced research tools and methodologies to better predict and understand the varied responses to immunotherapy.

While the clinical success of ICIs validates the potential of immunotherapy, our understanding of the immunological changes they induce remains limited. The current predictive models for ICI response are based on the characteristics of 'hot' tumors, which are rich in T-cell infiltration and pro-inflammatory cytokines, often associated with a high tumor mutational burden (TMB)[13,14]. As the primary target of ICI therapy, CD8+ T cell dynamics are critical factors affecting treatment outcomes. However, these predictors have limited efficacy. With accumulating RNA sequencing data, a pan-cancer analysis revealed that a bulk RNA-seq-derived immune-enriched, non-fibrotic subtype significantly correlates with better ICI response[15]. At the single-cell level, higher abundance of naïve T cells and exhaustion precursor T cells (Texp) have been observed in responsive tumors[16,17]. Despite these findings, the therapy-induced changes in TME cell composition and their potential as predictive biomarkers remain incompletely understood.

Recent studies utilizing site-matched samples from different time points have provided unprecedented insights into the evolving landscape of TME under the influence of ICI. Clonal replacement following PD1 blockade in cutaneous cancer demonstrated limited reinvigoration capacity of T cells[18]. Liu et al. further confirmed the local Texp state after treatment was derived from clonal revival or pre-existing

Texp[19]. The emerging 'window of opportunity' for neoadjuvant ICI treatment has facilitated access beyond easy-to-biopsy cancer types. With a sample collection interval of approximately 9 days, the expanded cytotoxic CD8+T, T helper-1 (Th1), and T follicular helpers (Tfh) cells potentially contribute to response in breast cancer[20]. In triple negative breast cancer, *CXCL13*+ T cells were concurrently expanded alongside pro-inflammatory macrophages and follicular B cells, and this expansion was linked to a positive response during neoadjuvant chemo-ICI combination therapy[21]. A reduction in regulatory T cells (Tregs) was observed in both colorectal and head and neck cancer tumors that achieved pathological complete response[22,23]. Increasing evidence suggest that pre-existing tertiary lymphoid structure (TLS) components consisting of T cells, B cells, and dendritic cells (DC) contribute to response[21,24,25]. Yet, the global TME dynamics upon ICI treatment remain unknown.

The limited number of patients profiled in individual studies and inconsistent cell annotation depth among them necessitates a broader, more inclusive analysis. Furthermore, scRNA-seq studies typically consider cells from all patients as replications, hindering the examination of variations across patients. To accurately portray the complex landscape of diverse cellular states in TME at an individual patient level, it is imperative to integrate multiple datasets from various cancer types and leverage pan-cancer TME atlases for unified annotation. This approach, coupled with an in-depth analysis of the biological relevance and clinical significance of different cell subtypes within larger patient cohorts, can enhance statistical power, improve the robustness of findings, and pave the way to identify signals that are critical for ICI treatment response.

To accurately portray the complex landscape of diverse cellular states in TME at an individual patient level, we developed a computational framework to integrate multiple datasets from various cancer types and leverage pan-cancer TME atlases for unified annotation. By focusing on cross-tissue analyses and individual variability in a cohort of 163 patients with pre- and on-ICI treatment scRNA-seq samples, we unraveled the intricate cellular changes that underpin the response to ICI treatment, ultimately contributing to the advancement of personalized cancer immunotherapy.

Results

Large-scale pan-cancer atlas of longitudinally paired ICI-treated tumors

We generated scRNA-seq data from twelve skin cutaneous melanoma patients at day 0 of anti-PD-1 or dual anti-PD1 plus anti-CTLA4 treatment and 7-9 days later (SKCM; Becker). To expand our analysis, we curated a broad range of datasets from 11 recent studies that profiled scRNA-seq from patients at

baseline and at a subsequent time point after initiating ICI treatment. This aggregated dataset incorporated profiles from a diverse range of cancer types, including head and neck squamous cell carcinoma [HNSC; Luoma[26], van der Leun (IMCISION)[22], and Franken[24])], breast cancer (BRCA; Zhang[21], Bassez[20], and Shiao[25]), colorectal cancer (CRC; Li[23], Chen[27]), prostate cancer (PCa; Hawley[28]), non-small cell lung cancer (NSCLC; Liu[29]), and skin cutaneous basal and squamous cell carcinoma (BCC and SCC; Yost[18]). After rigorous filtering, our assembled cohort comprised 141 patients, each represented by two scRNA-seq time points (**Figure 1A-B and Table S1**).

For 89 patients, the second time point was categorized as 'early on' treatment, i.e., less than 21 days from treatment initiation, while for the remaining 74 patients, it was designated as 'later on' treatment, up to an interval of 121 days. We collected and curated metadata for all patients, including information on prior treatments, the specifics of the ICI regimen, and various metrics of response to treatment. Patients in the cohort were administered different agents and combinations, which we categorized into monotherapy (anti-PD1/PDL1 with or without chemotherapy; N = 115) and combination therapy (anti-PD1/PDL1 + anti-CTLA4; N = 26). Response metrics varied across studies, including Response Evaluation Criteria in Solid Tumors (RECIST)/iRECIST criteria, prognosis (early death after treatment), pathology, and, in some cases, only T-cell expansion levels with or without tumor size changes. We harmonized these diverse response metrics into a binary classification of responders and non-responders (**Methods**). After rigorous quality control at both cell and sample levels, the dataset represented 1,022,526 cells across 163 samples and 10 cancer types (**Figure 1A**).

To identify granular cell states, we developed a framework for deep phenotyping of the scRNA-seq data. This annotation pipeline circumvents the need for manual intervention in parameter fine-tuning procedures, such as highly variable genes selection, integration, and clustering. The pipeline employs a two-step hierarchical annotation approach, primarily utilizing the reference-based method SingleR. In the first step, we identified the major non-cancer TME components using a complementary combination of clustering by scGate[30] and SingleR[31] (**Figure 1B-D**). The second step, particularly advantageous for resolving continuous cell states, involved re-applying SingleR with multiple curated granular-level pan-cancer references. We enhanced this process by dividing the reference data and combining scores to mitigate potential batch effects across reference studies in a hierarchical framework (**Methods**). This approach allowed us to achieve high-resolution annotation for each cell, avoiding the pitfalls of clusters derived from integrated embeddings. We initially discerned 11 main non-cancer TME components across the 326 scRNA-seq samples (**Figure 1B, D-F**). The reliability of these annotations was validated by marker gene expression analysis (**Figure 1C**), ensuring the robustness of our subsequent analyses.

Patient-level Insights into Cellular Dynamics during ICI Therapy

To investigate the impact of ICI treatment on the cellular composition of the TME, we applied linear mixed-effects models (LMMs), which accounts for the paired nature of the data and allows for the inclusion of the fixed effects (cancer type, timepoint, and treatment modality) and for patients as the random effect for each of the 11 main TME components identified in our analysis. Surprisingly, none of these components showed significant changes in their proportions after treatment (**Figure 1F, G**).

To characterize the functional dynamics, we utilized curated gene signatures and assessed the functional shifts in CD4+ T cells, CD8+ T cells, and macrophages within pseudo-bulk samples, and applied the LMMs to identify patterns across patients. In both CD4+ and CD8+ T cells, we observed significant changes in several functional programs in response to ICI treatment (**Figure 1H, I**). Notably, we observed significant reduction in interferon response signatures (Effect size (ES)=-0.43, p-value=0.009 and ES=-0.43, p-value=0.027, in CD4+ and CD8+ T-cells respectively) and cytokine/cytokine receptor signatures (ES=-0.61, p-value=0.002 and ES=-0.7, p<0.001, respectively), which may indicate a dampening of the initial inflammatory response following ICI treatment (**Figure 1H**). Interestingly, CD8+ T cells exhibited additional functional changes that were not observed in CD4+ T cells (**Figure 1I**). These included an increase in anergy (ES=0.75, p-value<0.001), pro-apoptosis (ES=0.66, p-value<0.001), chemokine/chemokine receptor (ES=0.37, p-value=0.026), and exhaustion signatures (ES=0.33, p-value=0.007), suggesting a complex interplay between T cell activation and dysfunction in response to ICI treatment. Macrophages also underwent significant functional changes in response to ICI treatment, further highlighting the complex dynamics within the TME (**Figure 1J**). We identified a significant increase in phagocytosis (ES=0.91, p<0.001), M2 (ES=0.78, p-value<0.001), MHC class I antigen processing (ES=0.49, p-value=0.013), and angiogenesis (ES=0.45, p-value=0.043), indicating a potential shift towards a more immunosuppressive and tissue-remodeling phenotype. The changes in macrophage function may contribute to the modulation of the TME and influence the overall response to ICI treatment.

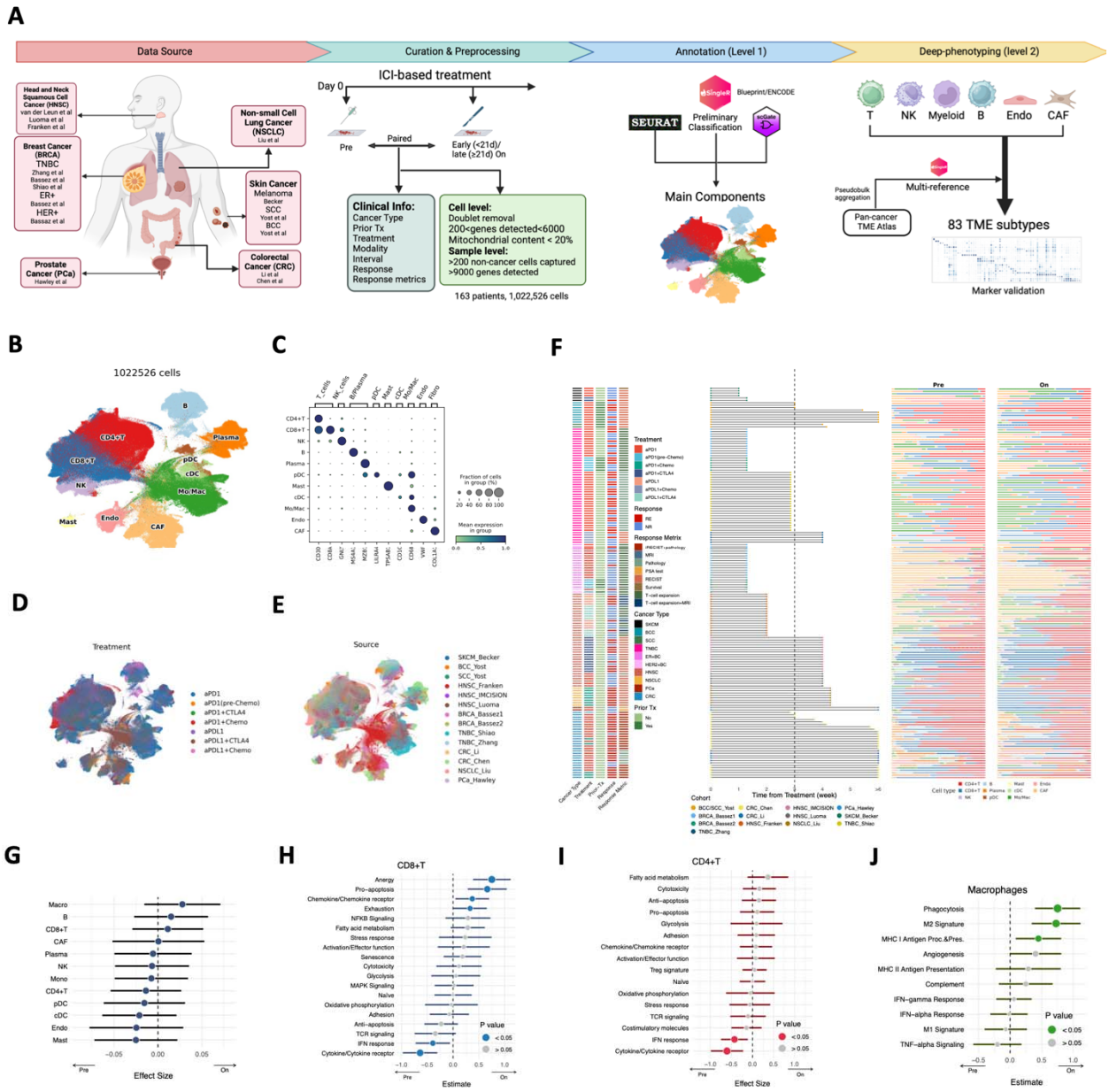
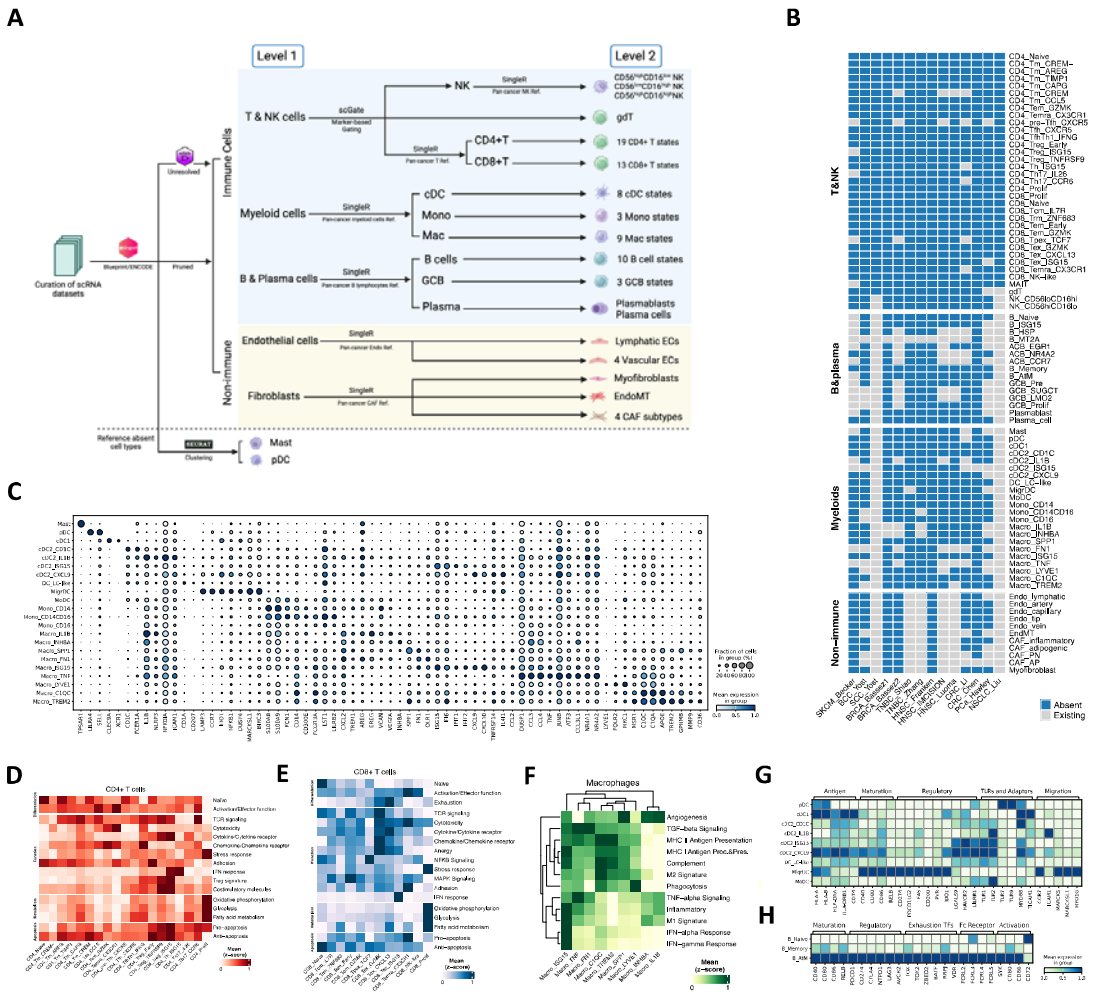


Figure 1. *Large-scale pan-cancer atlas of longitudinally paired ICI-treated tumors.* ***A.*** *Summary of study design and analysis workflow.* ***B-D.*** *UMAP plots showing main components, source, treatment of integrated dataset.* ***E.*** *Dotplot showing marker expression for main components.* ***F.*** *Clinical information and composition of main components by patient.* ***G-J.*** *Forest plots showing the association between main components, pathways (CD4+T, CD8+T, and macrophages) and time points during ICI treatment, the effect sizes (error bars indicate 95% confidence intervals) were based on the coefficients from linear mixed-effects models.*

To gain a more detailed view of the cellular dynamics upon ICI therapy, we applied the hierarchical deep-phenotyping framework for each distinct lineage utilizing a set of specialized references (**Figure 2A and Methods**). This process yielded 83 cell subtypes/states, most of which were represented across datasets (**Figure 2B and Table S2**). It is noteworthy that some datasets were sorted for specific subsets of TME cells, such as CD3+ cells (Liu and Yost) or CD45+ cells (Hawley, Shiao, Zhang, van der Leun, and

Luoma). To validate our deep-phenotyping performance, we employed comprehensive gene markers and curated gene signatures to validate these cell types and states (**Figure 2C, S1A-D**). Regulatory T cells (Treg) showed a gradient in both regulatory (e.g. *IL2RA, CTLA4, and LAYN*) and costimulatory (*TNFRSF9, TNFRSF18, TNFRSF1B*) signatures, transitioning from an early-phase state towards the activated state (**Figure 2C, S1A**). The recently reported cellular stress signature was predominantly enriched in both CD4+ and CD8+ T cells with high proliferation marker expression (*MKI67* and *TOP2A*), which also exhibited a moderate pro-apoptosis activation (**Figure 2D-E**). Comparatively, T cells in the early phase displayed relatively higher anti-apoptosis activity (e.g. *BIN1, BIN2,* and *BIRC3*) than mature T cells. Macrophages characterized by *C1QC*, *SPP1*, *TREM2*, and *LYVE1* primarily exhibited in anti-inflammatory, M2-like immunosuppressive phenotype, including TGF-beta signaling, MHC I&II antigen processing and presentation, complement, and phagocytosis. Conversely, *INHBA*+ and *IL1B*+ macrophages instead exhibited high expression of genes associated with angiogenesis and inflammatory pathways. *TNF*+ and *FN1*+ macrophages showed activation in both immunosuppression and inflammation, but lacked involvement in angiogenesis and phagocytosis, respectively, demonstrated broad activation across multiple pathways, with particularly strong expression in MHC I antigen presentation, complement pathways, and inflammatory response (**Figure 2C, 2F**). The activated *LAMP3*+ dendritic cells (MigrDC) showed high expression of maturation (e.g. *CD40, CD80, CD86*), regulatory (e.g. *CD274, PDCD1LG2, FAS*), and migratory (e.g. *CCR7, MARCKS, MYO1G*) markers (**Figure 2G**). The atypical memory B cells displayed upregulated maturation, regulatory, exhaustion translation factors (TF), Fc receptors, and activation signatures, compared to naive B cells and memory B cells (**Figure 2H**).

Regarding cellular composition, in line with previous pan-cancer studies, activated regulatory T cells were the most abundant in the CD4+ T-cell compartment, followed by naive and memory CD4+ T cells (**Figure S1E**). Resident memory cells dominated the CD8+ T-cell compartment, followed by early effector memory, naive, and exhausted CD8+ T cells (**Figure S1F**). In B cells, plasma cells were the most abundant, followed by memory B cells (**Figure S1G**). The *C1QC*+ macrophages dominated the myeloid compartment (**Figure S1H**). Among non-immune cells, the inflammatory CAF and myofibroblasts were most frequently detected (**Figure S1I**). These validations confirmed that our approach can robustly differentiate closely related cell states without the need for data integration and clustering — methods plagued by batch effects and the generation of impure clusters that, while enriched for certain states, are often contaminated with others. It should be noted that attempts to identify these cell states using conventional clustering-based methods were unsuccessful and provided noisy associations with gene markers.



IFN-responsive immune cells was observed, primarily affecting T lymphocytes, gamma-delta T cells, and

cDC (**Figure 3F**). These cell states, which showed significant increases or decreases on-treatment, exhibited consistent dynamics across most datasets. However, due to low statistical power, they often did not reach statistical significance (**Figure 3B-E**). The concordance amongst datasets bolsters our confidence in these findings and emphasizes the enhanced statistical power afforded by our meta-analytical approach. These results were further validated by paired student t-test (**Figure 3F and S2E**).

Interestingly, *CREM*- CD4+ memory T cells, with lowest T cell receptor (TCR) signaling among CD4+T memory cells, which showed significant decrease upon ICI treatment (ES=-0.13; p-value<0.001), exhibited enhanced gene expression in oxidative stress responses (*TXNIP*), immune signaling regulation (*TRAF3IP3*), gene expression (*MALAT1*), vesicle trafficking (*TBC1D10C*), and cell survival (*GIMAP4*). Alternatively, they show reduced expression in cell proliferation (*RGCC*, *SRGN*, *YPEL5*), iron metabolism (*FTH1*), and altered gene regulation (*CREM*) (**Figure S2C**). These cells are characterized by upregulated pathways related to interferon-alpha response and downregulated pathways involved in inflammation, cell proliferation, and metabolism (TNF-alpha Signaling via NF-kB, Myc Targets V1, G2-M Checkpoint, mTORC1 Signaling, E2F Targets, **Figure S2D**). This quiescent yet primed state consistently decreases during ICI treatment, suggesting a shift towards a more active anti-tumor immune response. Notably, as primary target of anti-angiogenesis therapy, tip-like endothelial cells (EC) showed the most significant reduction during the treatment (ES=-0.14; p-value<0.001). This endothelial subtype exhibited the highest score in angiogenesis and collagen adhesion activities compared to other subtypes (**Figure S2E**).

Next, to investigate divergent dynamics between ICI regimens, we conducted a subgroup analysis using LMMs to represent general patterns in each treatment modality (**Figure 3G and Table S3**). Additionally, we performed paired t-tests within treatment-cohort subgroups as an orthogonal validation method (**Figure 3H**). Overall, the paired t-tests revealed consistent trends in cell type dynamics across treatment-cohort subgroups, aligning with the general patterns observed from the LMMs. Although some did not reach statistical significance, this consistency suggests that dynamic changes in cell states may reflect a general mechanism of response to ICI therapy, rather than being specific to cancer type or treatment regimen. Collectively, while the subtypes exhibited similar trends in cell state changes across the two treatment modalities, the magnitude of these changes varied. Patients receiving combination therapy showed stronger evolutionary trends, underscoring the synergistic effect of adding CTLA4 inhibition. The pro-angiogenesis non-classical *CD16*+ monocytes decreased in monotherapy but tended to increase in combination therapy. Conversely, immune-suppressive TREM2+ macrophages accumulated in monotherapy cohorts but were reduced in those undergoing combination therapy. Restricted by the imbalanced distribution of patient numbers between treatment modality groups, further studies are needed

to better understand the mechanisms underlying the different modalities, ultimately allowing for more precise tailoring of immunotherapeutic strategies based on these dynamics.

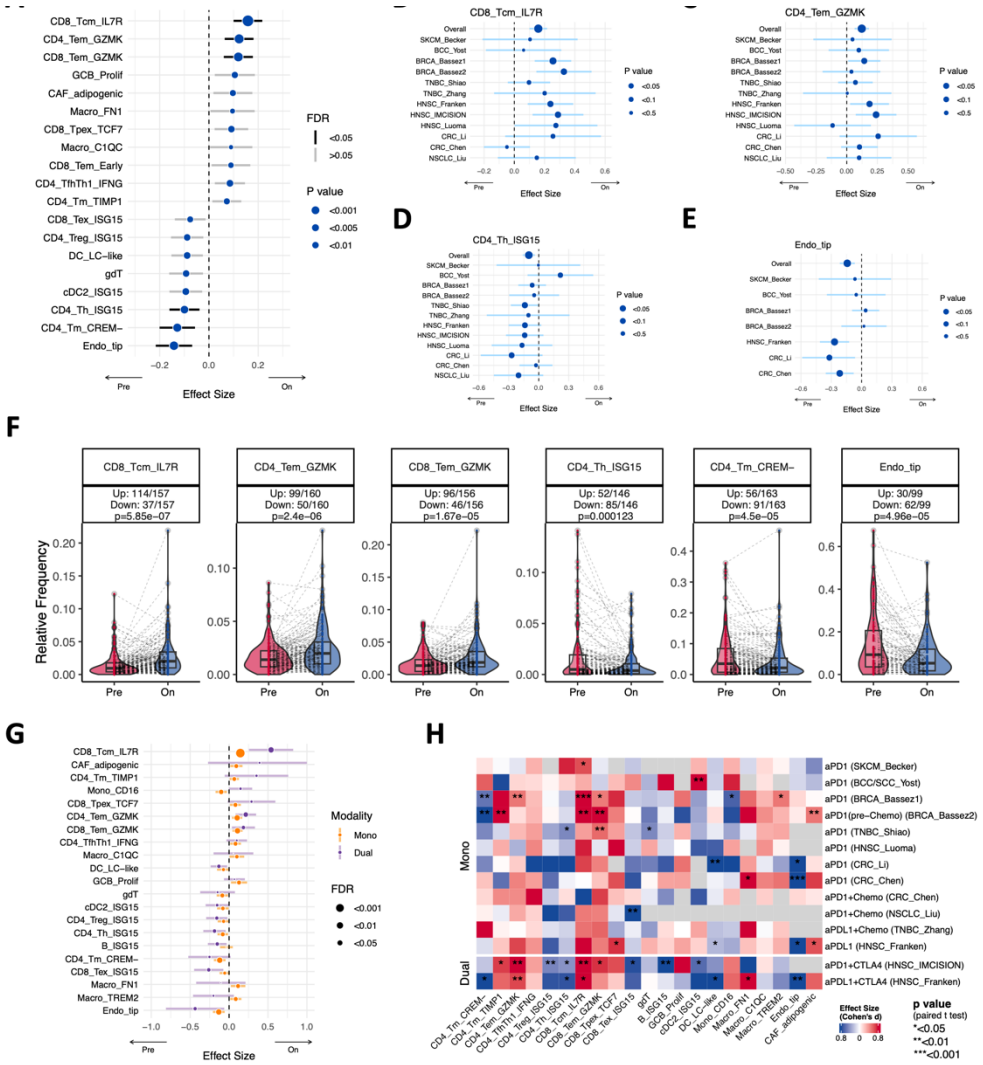


Figure 3. *Patient-level compositional analysis.* ***A.*** *Forest plot for the associations between cellular abundance (relative frequency, RF) and treatment time points.* ***B-E.*** *Forest plots for the associations between cellular abundance (RF) and time points by cohort in selected cell types.* ***F.*** *Paired box and violin plots comparing RF between time points for the top-6 cell states exhibiting the strongest changes. Boxes show 25th, median, and 75th percentiles, the RF changes were classified as "Up" (Δ > 10%), "Down" (Δ < -10%) during the treatment.* ***G.*** *Effect sizes from subgroups analysis of monotherapy (Mono) treated patients and combination therapy (Dual) treated patients.* ***H.*** *Heatmap showing the effect sizes (Cohen's d) of compositional change for each cell type (column) in modality-cohort subgroups (row), *p<0.05, **p<0.01, ***p<0.001, based on paired Student's t-test. Gray color indicates that the cell type wasn't detected in < 3 patients in the subgroup.*

Tissue-Specific Cellular Programs and Co-Regulatory Networks Upon ICI Treatment

To further elucidate coordinated behaviors within the TME and to discern patterns of cellular dynamics in response to ICI therapy, we utilized non-negative matrix factorization (NMF) based on cellular

compositions across all samples (**Methods**). Three co-occurrence cellular programs were identified (**Figure 4A and Table S4**). Distinct from recent pan-cancer studies[32,33], which reported an IFN-responsive community on tumor samples, *ISG15*+ B and cDC2 cells were enriched in NMF3, while other *ISG15*+ immune subtypes were primarily enriched in NMF2. NMF1 was enriched with activated B cells and memory T lymphocytes, indicating an adaptive immune response characterized by antigen-experienced cells ready to engage in immune defense. NMF2 exhibited a mature immune status, featuring a mixture of proliferating, terminally differentiated effector, and exhausted T lymphocytes, along with antigen-secreting cells such as plasmablasts and plasma cells. This profile also included distinct macrophage subsets (*SPP1*+, *C1QC*+, and *TREM2*+), suggesting a complex interplay of immune activation and suppression within the tumor microenvironment. In contrast, NMF3 reflected an early immune status, consisting of naïve and memory B cells, pre-germinal center B cells, and proliferating GCB cells. This program was also marked by the presence of CD4+ naïve T cells, early regulatory T cells, and cDCs, indicating an initial phase of immune activation. Notably, *TNFRSF9*+ Tregs were most abundant in NMF2, with moderate levels in NMF1, highlighting the potential regulatory influence in both experienced and mature immune contexts. Correlation analysis between immune cellular programs and non-immune cells revealed distinct interaction patterns for NMF2 and NMF3 with endothelial cells and fibroblasts (**Figure 4B**). NMF2 and NMF3 exhibited disparate correlation with endothelial cells and fibroblasts. Interestingly, NMF2 demonstrated a significant positive correlation with inflammatory CAFs, indicative of an inflamed ecotype that may support tumor-promoting inflammation.

We then investigated whether these cellular program changes were associated with treatment efficacy. The changes in activity for each program were assessed by calculating the difference in activity scores between baseline and on-treatment tumor samples. Notably, distinct patterns emerged across patients with varying response statuses (**Figure 4C**). Specifically, responders exhibited a preferential accumulation of NMF3 following treatment, while the NMF2 program increased in non-responders.

To further delineate the co-regulated pattern across TME cell types during the treatment, we next performed pairwise correlation analysis of compositional changes in cell types and states across patients (**Figure 4D, S3B**). This analysis was conducted in two iterations: the first encompassed all patients with available immune cell data (**Figure 4D**), and the second was restricted to patients for whom all TME data was accessible (**Figure S3E**). The correlation analysis unveiled several intriguing groups of cells that exhibited similar dynamics across the patient cohort. These clusters represent groups of cell types and states that, potentially, are co-regulated or are part of a shared response pathway to ICI treatment.

The unsupervised Leiden clustering[34] highlighted inter-lineage cellular communities (**Figure 4E**). A *ISG15*+ cellular network was identified, reflecting a concomitant interferon response (**Figure S3D**).

Notably, the overall change of *ISG15+* cellular abundance significantly correlated with T cell expansion in breast cancer and head and neck cancer (**Figure 4F**). We also identified a large immune network, comprising both adaptive and innate immune response. This community includes naïve lymphocytes, memory T cells, activated/GC B cells, cytotoxic NK cells, terminally differentiated effector T cells, gdT cells, cDCs, and monocytes. Furthermore, we identified correlations between *TNFRS9+* Tregs and inflammatory CAFs (**Figures 4F**) as well as between *TNFRS9+* Tregs and tip-like endothelial cells (**Figures S3F**), underscoring Tregs as potential node between immune and non-immune cells.

In the negative correlation graph, a strong negative correlation was observed between plasma cells and naive/memory B cells. Additionally, *C1QC+* macrophages and CD4+ Tregs, which showed a positive correlation to each other, negatively correlated with CD4+ naive T cells (**Figure 4A, S3E**). Jointly, two patterns were observed (**Figures 4D and S3F**): the first TME pattern encompasses cell types pivotal in the early stages of immune response activation, including naive T and B cells, memory B cells, and cytotoxic effector cells (NK_CD56loCD16hi, Temra_CX3CR1), i.e., indicating a readiness for antigen engagement and initiation of adaptive immune responses; the second TME pattern comprises cell types integral to the modulation of immune activation and the establishment of an immunosuppressive environment (*TNFRSF9+* Tregs, *C1QC+* macrophages, plasma cells). The correlation between these two patterns illustrates a continuum of immune modulation within the TME (**Figure 4G**), from an initial state of readiness through active engagement to a more complex interplay of activation and regulatory suppression, revealing the dynamics in response to ICB therapy.

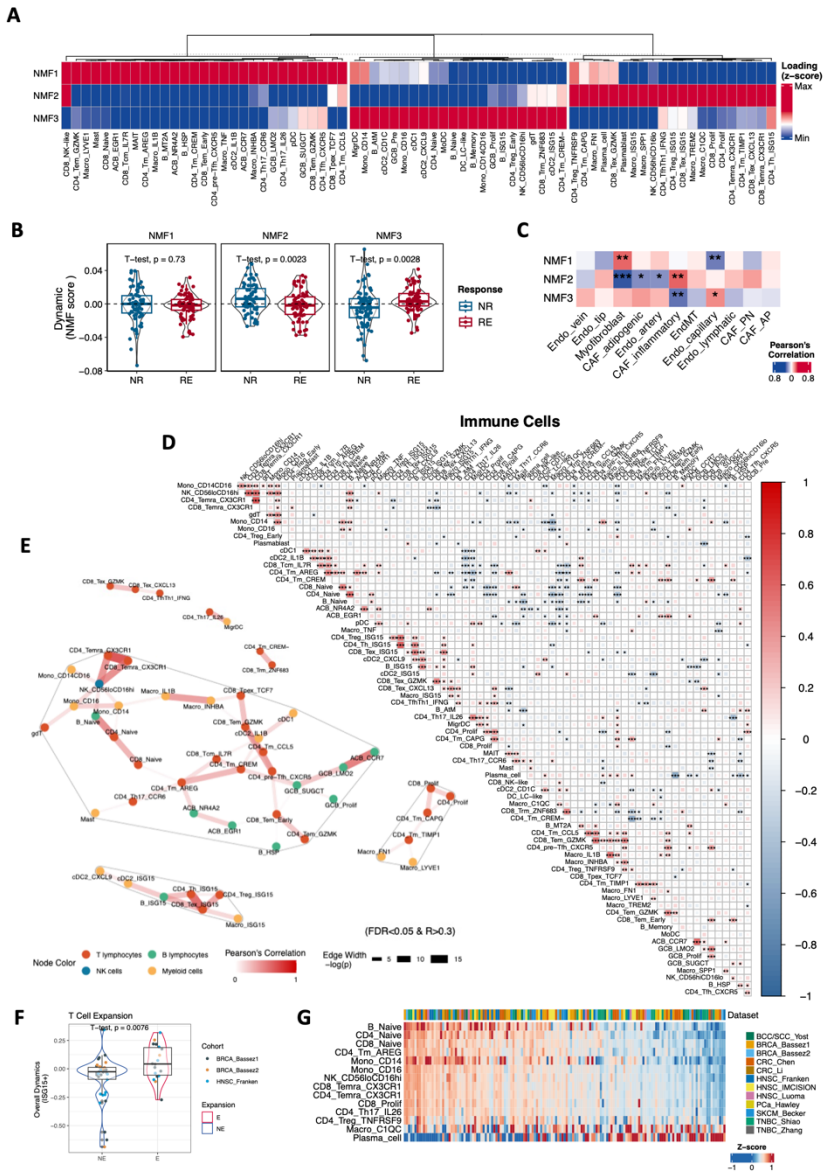


Figure 4. *Tumor ecosystem and Co-regulating patterns upon ICI treatment.* ***A.*** *NMF-based TME immune programs. The color represents the loading (column-wise scaled) of every cell type in each program.* ***B.*** *Box plot and violin plot showing the dynamic changes of three immune programs in tumors. Boxes show 25th, median, and 75th percentiles, each dot represents a patient. P values were based on two-side Student's t-test.* ***C.*** *Heatmap showing the correlation between immune programs and non-immune compartments.* ***D.*** *Heatmap of correlation matrix showing the pairwise Pearson correlations between dynamic changes (ΔRF) of immune cell types per patient, *p < 0.05, **p < 0.01, ***p < 0.001, based on the correlation.* ***E.*** *Undirected graph showing significant positive (FDR<0.05, Pearson's rho>0.3) correlations (edges) between cell types (nodes).* ***F.*** *Box and violin plot comparing overall dynamics of ISG15+ cell subtypes between expanders and non-expanders. Boxes show 25th, median, and 75th percentiles.* ***G.*** *Heatmap showing a continuum from early activation/innate immune response to immune suppression. Patients (column) were sorted according to sum of corresponding cell type (row) dynamic changes (column-wise scaled).*

Association of Immune Cell Dynamics with Treatment Response

To examine whether these cell type compositional changes displayed associations with treatment efficacy, we performed a pairwise correlation analysis across patients to identify groups exhibiting similar cellular

dynamics in response to ICI treatment (**Figure 5A**). The resulting correlation matrix facilitated the identification of three distinct patient clusters, with clusters 1 and 2 demonstrating particularly strong intra-group correlations and encompassing 48 and 52 patients, respectively. Cluster 1 was enriched for responders (RE; 30 out of 48 patients), whereas cluster 2 consisted more of non-responders (NR; 30 out of 50 patients) (**Figure 5B**). This finding emphasizes the significance of cellular dynamics, rather than static proportions at individual time points, as a determinant of patient response to ICI therapy.

The delineation of patient clusters based on treatment dynamics offers a compelling insight into the potential of cellular dynamics as a biomarker for ICI responsiveness. This dynamic-based patient stratification could prove to be a powerful tool in predicting treatment outcomes and tailoring therapy to individual patients. By examining the cell states compositional changes between the clusters, we observed that cluster 1 (referred to as RE-enriched) was enriched with CD4+ naïve T cells, CD4+Tfh cells, and non-antibody secreting B cells (**Figure 5C**). Cluster 2 (referred to as NR-enriched) on the other hand was enriched with plasma cells, plasmablasts, EndMT. The retention of Langerhans cell-like follicular DCs, and CD4+CREM+ memory T cells and the accumulation of cytotoxic CD56lowCD16high NK cells, NK-like CD8+ T cells, and CD8+GZMK+ T cells were exclusively found while not significant in cluster 2 (**Figure 5D**).

To further identify the cellular determinants of treatment response, we performed a supervised subgroup analysis, dividing the aggregated cohort into responders and non-responders (**Figure 5E**). When subsequently subdivided the group by datasets, we observed similar trends, while most were not significant due to the limited sample sizes (**Figure 5F**). Memory and proliferating GC B cells were specifically accumulated in RE group compared to the inconsistent tendencies in the NR group. For cell states that showed consistent dynamics between the two groups, *IL7R*+ CD8+ Tcm increased significantly in both groups, with a more pronounced increase in the NR group. Conversely, *GZMK*+ CD4+ Tem cells showed greater increments in the RE group. For cell states that exhibited divergent trends, *TNFRSF*+ Tregs displayed significant reductions in the RE group particularly in CRC, HNSC and NSCLC, while their increases in the NR group were not significant. Among non-immune cells, tip-like endothelial cells decreased in both response groups, with a more marked reduction in the CRC and HNSC RE group. EndMT cells were significantly reduced in the RE group, while lymphatic endothelial cells showed a significant increase in the NR group.

Highlighted by both unsupervised and supervised analysis, these cell subtypes may serve as potential biomarkers for patient stratification and could be investigated further for their mechanistic roles in mediating the efficacy of ICI therapy.

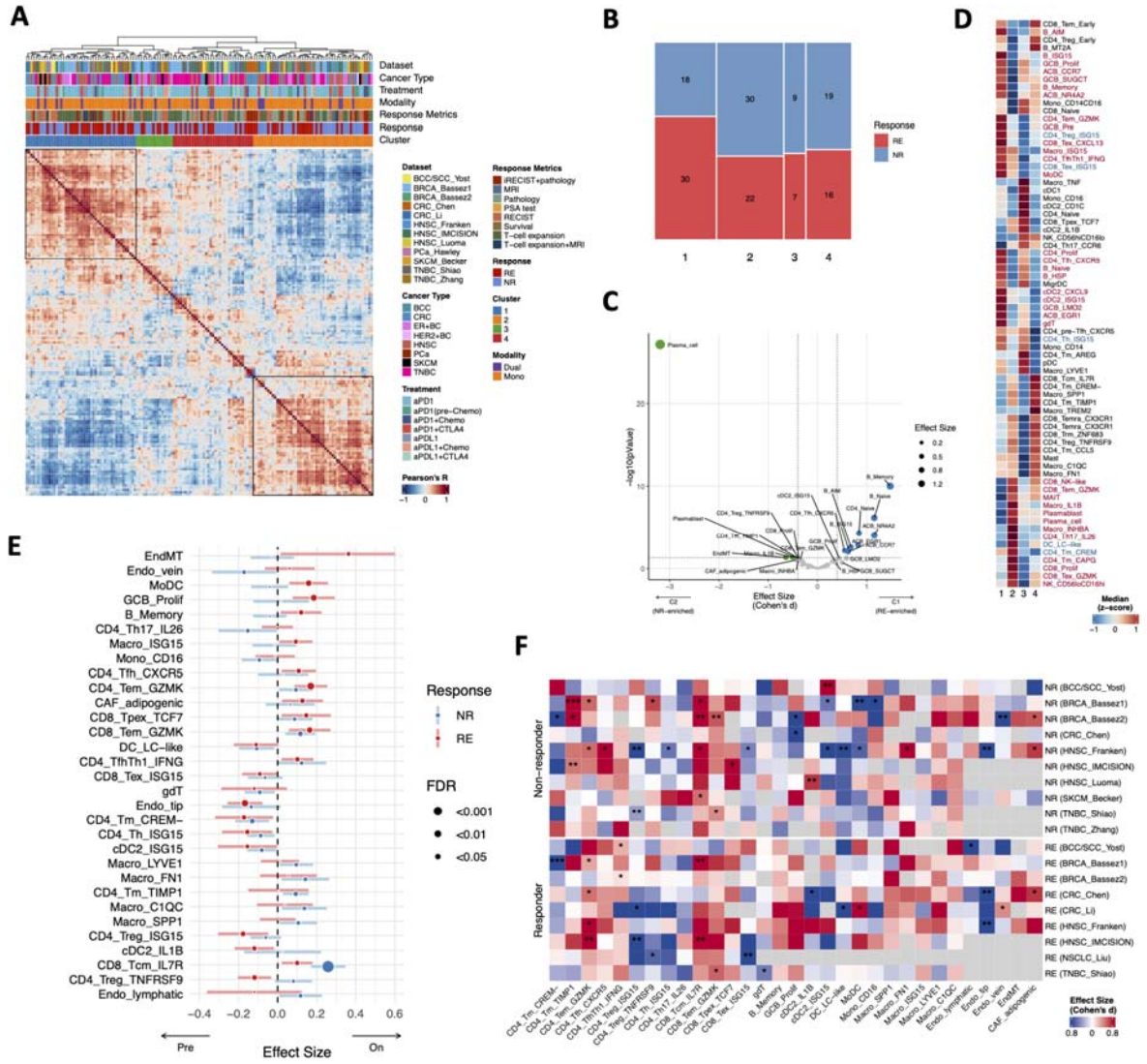


Figure 5. *Association of Immune Cell Dynamics with Treatment Response.* ***A.*** *Heatmap showing the pairwise Pearson correlation matrix across the patients according to the dynamics of immune cells.* ***B.*** *Mosaic plot showing numbers of responders and non-responders in dynamic-derived clusters.* ***C.*** *Volcano plot showing differential dynamics of immune cell dynamics between cluster 1 (responder-enriched) and cluster 3 (non-responder enriched), effect sizes were based on Cohen's d and p values were based on paired t-test.* ***D.*** *Heatmap showing the median dynamics (Δ relative frequency, row-wise scaled) of immune cell types (row) in the 3 clusters.* ***E.*** *Association between cell types and time points in responders and non-responders, the effect sizes (error bars indicate 95% confidence intervals) were based on the coefficients from linear mixed-effects models.* ***F.*** *Heatmap showing the effect sizes (Cohen's d) of compositional change for each cell types (column) in response-cohort subgroups (row), *p<0.05, **p<0.01, ***p<0.001, based on paired Student's t-test. Gray color indicates that the cell type detected in <3 patients from the subgroup.*

Discussion

We present a comprehensive analysis of cellular dynamics within the TME in response to ICI therapy across multiple cancer types. By leveraging a large-scale pan-cancer atlas of longitudinally paired ICI-

treated tumors studied by scRNA-seq, we provide novel insights into the compositional dynamics during ICI treatment and their relationship to treatment response.

One of the key technical contributions of our study is the development of a robust deep-phenotyping framework for accurate identification of cell types, subtypes, and functional states within the TME. Traditional clustering-based approaches often struggle with batch effects and generate impure clusters, leading to suboptimal cell type annotations. Annotating within individual studies can be challenging due to inconsistent annotation depth across studies. To overcome these limitations, we employed a two-step annotation process that minimizes reliance on data integration and clustering. By utilizing multiple reference datasets and carefully reviewing the annotations, our approach enables the identification of granular cell states with high purity and interpretability. Another important technical aspect of our study is the shift in the statistical unit of analysis from single cells to samples. While most single-cell studies focus on individual cells as the primary unit of analysis, connecting cellular dynamics to clinical outcomes, such as treatment response, requires a sample-level approach. However, as highlighted in our subgroup analysis, the small sample sizes in individual studies often limit the ability to detect relevant associations between cellular dynamics and clinical variables. By aggregating data from multiple studies with the pan-cancer TME annotation, our meta-analysis substantially increases the statistical power, enabling the identification of consistent evolution across datasets that may not be apparent in individual studies alone. This highlights the importance of collaborative efforts and data sharing in the single-cell research community to uncover biologically meaningful and clinically relevant patterns.

Several interesting findings were uncovered regarding the cellular dynamics within the TME upon ICI treatment. We observed a significant increase in immune subtypes, such as *IL7R*+ Tcm and *GZMK*+ Tem, CD4+ ThfTh1. However, these changes did not differ between patients under different treatments or between responders and non-responders. The addition of anti-CTLA4 intensified cellular dynamics and significantly promoted the accumulation of pro-angiogenesis non-classical monocytes in head and neck cancer.

A global decrease in IFN-responsive states across immune components was observed in our study, but better responses when these states were retained. It should be noted that IFN-responsive immune subtypes were not consistently characterized across studies due to variations in T cell annotation strategies, and some studies even regressed out the IFN signal during data scaling[16,20]. By comparing tissue-specific cellular programs with co-regulatory networks, we found that inter-lineage IFN-responsive cells from different cellular programs had coordinated regulation and was further confirmed to correlate with favorable treatment efficacy in terms of T cell expansion. Although there remains a gap between T cell expansion and actual clinical benefit, the mechanism by which this impacts treatment response warrants

further investigation. This finding aligns with recent pan-cancer studies that identified the IFN-enriched network in correlation with a favorable response[2,32]. However, it contrasts with studies in mice that blocking IFN signaling could resensitize resistant tumors[35,36] and a recent phase 2 trial for metastatic NSCLC that combined JAK inhibition with PD-1 immunotherapy to inhibit chronic IFN stimulation[37], which showed high response rates. These analyses were based on sorted CD8+ T cells from pooled blood and tumor tissue samples without specific ISG+ subset characterization. The differences in sample sources and annotation strategies highlight the complexity of understanding the dynamics and impact of the IFN response in the TME upon ICI treatment.

The co-regulatory network illustrated a continuum between two patterns partially resembling TLS maturity[38], from an initial immune status for antigen engagement and initiation of adaptive immune responses towards the modulation of immune activation and the establishment of an immunosuppressive environment. This finding was corroborated by the subsequent response-specific analysis across patients, which revealed a B cell-centric pattern, characterized by a significant accumulation of memory and proliferating GC B cells, along with a reduction in TNFRSF9+ Tregs in responders.

While our deep-phenotyping approach represents an improvement over existing methods, it is important to acknowledge the inherent challenges in delineating cell states within the TME. The continuum of cell states and the subtle differences between them make it difficult to establish clear boundaries, even with advanced analytical techniques. This notion underscores the complexity of the biological system and the need for further refinement of computational methods to better capture the nuances of cellular heterogeneity. Another limitation of our study is the heterogeneity in response metrics across studies, which may introduce some variability in the associations between cellular dynamics and treatment outcomes. Future efforts to standardize response criteria and harmonize data collection across institutions should further improve the reliability and robustness of such meta-analyses.

In conclusion, our analysis presents a distinct perspective in addressing inter-sample variation of large-scale longitudinally paired scRNA-seq cohort, by treating samples as statistical replication. Through compositional comparison, it provides a panoramic view of the cellular composition dynamics within the TME during ICI therapy across multiple cancer types. The granular annotation by the proposed robust deep-phenotyping framework facilitated the identification of tissue-specific cellular programs and co-regulatory networks that may govern the efficacy of ICI therapy. These findings highlight the importance of a systematic approach to understanding the complex interplay within the TME and pave the way for the development of more effective immunotherapeutic strategies. Future studies building upon these results will further elucidate the mechanisms underlying treatment response and resistance, ultimately leading to improved patient stratification and personalized cancer immunotherapy.

Methods

Single-cell RNA-seq data generation

Patients with advanced melanoma who were to receive immunotherapy with anti-PD-1 antibodies alone or in combination with anti-CTLA4 antibodies from Departments of Translational Skin Cancer Research and Dermatology were included in the study upon written informed consent (ethics vote numbers 16-6784-BO und 16-7097-BO). Tumor tissue biopsies from melanoma lesions were obtained on the day of therapy initiation (day 0) and 7 days thereafter (day 7). Fresh biopsies were immediately processed after excision. The generation of the single cell suspension was performed according to the tumor dissociation protocol from Miltenyi Biotec (Bergisch Gladbach, Germany). Briefly, the tissue biopsy was cut into small cubes of 2 mm and then minced with the gentleMACS program “h_tumor_01” and two times “h_tumor_02” with 30 min incubation time after each mince. During the incubation, the sample was digested with an enzyme cocktail mix consisting of 200 µl Enzyme H, 100 µl Enzyme R and 25 µl Enzyme A (catalogue #130-095-929 Miltenyi Biotec). The cell suspension was reconstituted in PBS with 0.05% BSA and washed two to three times, before passing through a 100 µm cell strainer. Subsequently, the cells were processed with the Single Cell 5’ Library & Gel Bead Kit v1.1 (catalogue #1000165, 10xGenomics, Pleasanton, CA, USA) according to the manufacturer’s protocol. As input for the 10x protocol, 10,000 cells from the single cell suspension, at a concentration of 1000 cells/µl. The resulting gene expression library was sequenced at the Cologne Centre for Genomics (CCG), University of Cologne. The sequencing data were aligned with the GRCh38 human reference genome and quantified using Cell Ranger (version 3.0, 10x Genomics Inc).

Single-cell RNA-seq data curation and standardization

We supplemented our data with public datasets, identified through an extensive literature review conducted up to August 2024. After quality control, we curated and standardized longitudinal paired transcriptomic data from 163 patients undergoing immunotherapy with immune checkpoint inhibitors (ICIs), encompassing 11 cancer subtypes across 12 cohorts (**Table S1**).

The interval between the initiation of treatment and the collection of the second sample was categorized into two groups: less than 21 days (early on) and 21 days or more (late on). Treatment regimens were classified based on their composition into several categories: anti-PD-1 (aPD1), anti-PD-1 before chemotherapy [aPD1(pre-Chemo)], anti-PD-1 combined with chemotherapy (aPD1+Chemo), anti-PD-1

combined with CTLA-4 inhibition (aPD1+CTLA4), anti-PD-L1 (aPDL1), anti-PD-L1 combined with chemotherapy (aPDL1+Chemo), and anti-PD-L1 combined with CTLA-4 inhibition (aPDL1+CTLA4). These were further grouped into monotherapy and dual therapy based on the checkpoint antibodies used.

Due to the heterogeneity in study designs across datasets, response metrics varied, incorporating methods like early/standard radiological assessment (MRI/RECIST/iRECIST), blood tests (PSA), T-cell expansion (scTCR-seq), and clinical outcomes (survival). The classification of patients as responders in our analysis adopts the criteria used in their respective original studies. An exception is made for the HNSC_Franken cohort, where only those patients who did not exhibit tumor growth were categorized as responders in our study.

Single-cell RNA-seq Data preprocessing

Newly generated single-cell sequencing data were aligned with the GRCh38 human reference genome and quantified using Cell Ranger (version 3.0, 10x Genomics Inc). The quality of cells was assessed based on two metrics: (1) The number of detected genes per cell; and (2) The proportion of mitochondrial gene counts per cell. Specifically, cells with detected genes fewer than 200 or greater than 6000, and mitochondrial unique molecular identifier (UMI) count percentage larger than 20% were filtered out. The genes detected in more than 3 cells were retained. To remove the potential doublets, scDblFinder[39] and DoubletFinder[40] were used for each sequencing library with the expected doublet rate set to be 0.08, and cells identified as doublets by both methods were further filtered out. For each dataset, gene symbols were converted into a common dictionary with STACAS:::standardizeGeneSymbols()[41].

After the removal of low-quality cells and deep phenotyping, sample-level quality control was applied to remove samples with fewer than 200 non-cancer cells or less than 9000 detected genes.

Single-cell RNA-seq Annotation

The annotation of single-cell RNA-seq data is performed at 2 levels, built on the reference-based SingleR annotation tool. At Level 1, cells are divided into discrete lineage populations. Level 2 focuses on the annotation of cells to the most granular cell types/states.

Level 1 starts with the assignment of major lineage components utilizing SingleR[31] with the BlueprintEncodeData[42,43] as a reference and using the built-in pruning approach in SingleR. The pruning will cause low-quality assignments to not be labeled. To assign those low-quality cells, we utilized the scGate[30]. First, we integrated all the data with Harmony[44] and clustered with Seurat[45] with a resolution of 0.5. Unknown cell types were then annotated using scGate with a pre-defined multi-classifier

(TME_HiRes). Plasmacytoid dendritic cells (pDC) were isolated by the positive expression of *LILRA4* and *JCHAIN*. Mixed cell types, including non-immune cells present in the CD45+ sorted data and cells identified as multi-label by scGate, were filtered out. Additionally, due to the technical limitations of 10X sequencing, the neutrophils were also excluded from further analysis. At the end of this step, cells are classified into major groups, including T and NK cells, myeloid cells, B lymphocytes, endothelial cells, and fibroblasts.

Level 2 annotations were designed to cater to various cellular lineages, and consists of two primary steps - preparation of reference sets and subsequent deep phenotyping:

Reference preparation. The reference data for T-cells, NK cells, myeloid cells, and CAFs were retrieved from the pan-cancer atlas studies[17,46-49], loaded in R as *SingleCellExperiment* objects, and processed individually. The objects were split by cancer types as lists, normalized with *logNormCounts*() function, and further aggregated into pseudo-bulk samples *aggregateReference*() function by the granular level labels from original studies. Additionally, to capture the diversity of main components, data for T cells, myeloid cells, and B lymphocytes were aggregated at the main level (e.g., CD4+ T cells, CD8+ T cells, cDC, macrophages, B cells, plasma cells) to ensure accurate intra-lineage reclassification. Due to the relatively lower abundance, the reference data of NK cells was aggregated only at the main level ($CD56^{high}CD16^{low}$, $CD56^{high}CD16^{high}$, and $CD56^{low}CD16^{high}$).

Deep phenotyping. Immune cells were first classified into main cell types within each lineage by SingleR based on the pan-cancer reference at the main level. For the T and NK cell group, this included CD4+ T cells, CD8+ T cells, and NK cells. For the myeloid cells group, this included cDCs, monocytes, and macrophages. For the B lymphocytes, this included B cells, GC B cells, and plasma cells.

Due to the distinct marker expression, the scGate was applied to separate cells from the T and NK cells group, with *KLRD1* and *NCAM1* served as positive markers, and CD3D as a negative marker for NK cells, and *CD3D*, *TRGC2*, *TRDV2*, *TRGV9*, *TRGV10*, and *TRDC* as positive markers, and CD8A and CD4 as negative markers for the gdT cells. After separating the NK and gd T cells, the remaining CD4+/CD8+ T cells underwent further ‘gating’ by the generic CD4T and CD8T models from scGate. Cells with overlapped labels were identified by SingleR, based on an aggregated pan-cancer reference at the main level. Following the main-level annotation, the immune cells were further annotated within each main cell type by SingleR with pan-cancer reference at the deep level.

Non-immune cells including endothelial cells and fibroblasts cells were annotated directly to the deep level by using SingleR with corresponding reference. After prediction, the labels were manually adjusted according to the extensively updated markers.

Cell Type Composition

Based on the annotation, TME cellular composition (referred to as the relative frequency, RF) was calculated within in their corresponding major groups, including T and NK cells, B and plasma cells, myeloid cells, and non-immune cells. The change in RF was calculated by subtracting the relative frequency at baseline from the RF at the second time point.

Linear Mixed Effects Models

Dynamic changes in functional pathways and composition during the treatment were evaluated using mixed effects regression models with the lmerTest package[50]. These models were adjusted for factors such as dataset, sample collection interval, and treatment modality, incorporating individual patients as random effects to account for the paired sample design. P values were adjusted for the false discovery rate (FDR) to control for multiple comparisons in the analysis.

Cell Type-specific Gene Set Enrichment

Gene set enrichment scoring was conducted using the UCell[51], which is based on the relative ranking of genes for individual cells, robust to dataset composition. T cell-specific signatures, encompassing differentiation, function, metabolism, and apoptosis for both CD4+ and CD8+ T cells, were derived from an earlier pan-cancer T cell analysis. For macrophages, gene sets including KEGG's antigen processing and presentation, along with hallmark pathways such as interferon-gamma response, complement, and TNFα signaling via NFκB, were retrieved via the msigdbr package[52]. Additionally, the M1, M2, angiogenesis, and phagocytosis signatures from a previous pan-cancer myeloid cell study were incorporated.

Differential Expression Analysis and Over-representation Enrichment Analysis

The differential expression analysis was performed by using *rank_genes_groups*() function from the scanpy package based on the Wilcoxon rank-sum test. P-values were corrected by Benjamini–Hochberg method. Genes with adjusted p-values < 0.05 were included in functional enrichment analysis. The functional enrichment analysis was performed by using the GSEApy package[53] with the MSigDB_Hallmark library[54].[54]. The enriched pathways with overlapping genes $\leq 50\%$ were filtered out.

NMF Analysis for Identifying Tissue-specific Cellular Programs

With the compositional data as matrix V (cell type × sample), we performed NMF by using R package NMF across samples from both pre- and -on treatment time points to identify tissue-specific programs within the TME. With n samples and m cell types, V is a non-negative matrix in $V \in R^{\mathrm{m}\times\mathrm{n}}$. Then V could be decomposed into 2 non-negative matrices: $W \in R^{\mathrm{m}\times\mathrm{k}}$ and $H \in R^{\mathrm{k}\times\mathrm{n}}$, where k represents the pre-selected factor number. Each element in W matrix indicates the loading of a cell type on a corresponding factor, while elements of H capture activity levels of these factor across tumor samples.

To assess the influence of the k parameter on NMF outcomes, we ran the NMF analysis on the immune matrix using a range of values for k (2 to 8), with 40 iterations for each value and a fixed seed of 10000. Following the guidelines suggested in the NMF vignette[55], we selected a factorization rank of 3 based on two criteria: choosing the first value of k where the cophenetic coefficient begins to decrease, and identifying the first point where the RSS curve shows an inflection.

We considered each NMF factor to represent a distinct cellular program. To identify the representative cell types associated with each program, we applied hierarchical clustering to the scaled W matrix using single linkage and Euclidean distance. The dynamics of these cellular programs were then assessed for each patient by calculating the difference in activity scores between on-treatment and baseline samples, with positive changes indicating an increase in program activity following treatment.

Unsupervised Clustering

The correlation analysis was performed using Pearson correlation test. The unsupervised hierarchical clustering and heatmap visualization were performed based on the Euclidean distance of correlation by the corrplot[56] and ComplexHeatmap[57] R packages.

Statistical Analysis

All statistical analyses were performed in R software 4.3.0 and Python 3.9. The evenness of relative frequencies for cell subtypes at a single time-point was calculated with the Shannon equitability index, which ranges from 0 to 1 (complete evenness). Comparisons between two groups were performed using either a two-tailed paired Student's t-test (unless otherwise indicated), and the effect size was calculated with Cohen's d. The statistical significance was set to p-value < 0.05, The other statistical methods for specific analyses are present in the sections above.

Data availability

All single-cell datasets used in this study are accessible from their respective original publications, as detailed in **Table S1**. The Becker dataset has been deposited in the European Genome-phenome Archive under accession number EGAS50000000459.

Code availability

The code for reproducing all the analyses in the study is available at https://github.com/zlin-bioinfo/Pan-cancer_ICI.

Acknowledgments

This work was funded by the Israel Science Foundation (1543/21) and Azrieli Faculty Fellowship (to DA), by Deutsches Konsortium für Translationale Krebsforschung (ED03 to JCB) and Else-Kröner-Fresenius Stiftung (EFKS), University Medicine Essen Medical Scientist Academy (UMESciA) (to NS).

Declaration of interests

JCB is receiving speaker's bureau honoraria from Amgen, Pfizer, Recordati and Sanofi, is a paid consultant/advisory board member/DSMB member for Almirall, Boehringer Ingelheim, InProTher, ICON, MerckSerono, Pfizer, Regeneron, 4SC, and Sanofi. His group receives research grants from Merck Serono, HTG, IQVIA, and Alcedis. DA is a consultant to Carelon Digital Platforms and Link Cell Therapies. All other authors declare no competing financial interests.

Author contributions

DA and JCB conceived the study and supervised the analyses. ZL performed all bioinformatics analyses. EP provided bioinformatics support. MC and NS performed the scRNA-seq. SU collected the samples from the patients. DA and ZL wrote the manuscript with inputs from JCB.

References

1. Shaked, Y. (2016). Balancing efficacy of and host immune responses to cancer therapy: the yin and yang effects. Nat Rev Clin Oncol *13*, 611-626. 10.1038/nrclinonc.2016.57.
2. Thorsson, V., Gibbs, D.L., Brown, S.D., Wolf, D., Bortone, D.S., Ou Yang, T.H., Porta-Pardo, E., Gao, G.F., Plaisier, C.L., Eddy, J.A., et al. (2018). The Immune Landscape of Cancer. Immunity *48*, 812-830 e814. 10.1016/j.immuni.2018.03.023.

3. Ren, X., Zhang, L., Zhang, Y., Li, Z., Siemers, N., and Zhang, Z. (2021). Insights Gained from Single-Cell Analysis of Immune Cells in the Tumor Microenvironment. Annu Rev Immunol *39*, 583-609. 10.1146/annurev-immunol-110519-071134.
4. Rood, J.E., Maartens, A., Hupalowska, A., Teichmann, S.A., and Regev, A. (2022). Impact of the Human Cell Atlas on medicine. Nat Med *28*, 2486-2496. 10.1038/s41591-022-02104-7.
5. Hodi, F.S., O'Day, S.J., McDermott, D.F., Weber, R.W., Sosman, J.A., Haanen, J.B., Gonzalez, R., Robert, C., Schadendorf, D., Hassel, J.C., et al. (2010). Improved Survival with Ipilimumab in Patients with Metastatic Melanoma. New England Journal of Medicine *363*, 711-723. doi:10.1056/NEJMoa1003466.
6. Brahmer, J.R., Tykodi, S.S., Chow, L.Q.M., Hwu, W.-J., Topalian, S.L., Hwu, P., Drake, C.G., Camacho, L.H., Kauh, J., Odunsi, K., et al. (2012). Safety and Activity of Anti–PD-L1 Antibody in Patients with Advanced Cancer. New England Journal of Medicine *366*, 2455-2465. doi:10.1056/NEJMoa1200694.
7. Ribas, A., and Wolchok, J.D. (2018). Cancer immunotherapy using checkpoint blockade. Science *359*, 1350-1355. doi:10.1126/science.aar4060.
8. Wolchok, J.D., Chiarion-Sileni, V., Gonzalez, R., Grob, J.-J., Rutkowski, P., Lao, C.D., Cowey, C.L., Schadendorf, D., Wagstaff, J., Dummer, R., et al. (2022). Long-Term Outcomes With Nivolumab Plus Ipilimumab or Nivolumab Alone Versus Ipilimumab in Patients With Advanced Melanoma. Journal of Clinical Oncology *40*, 127-137. 10.1200/jco.21.02229.
9. Huang, A.C., and Zappasodi, R. (2022). A decade of checkpoint blockade immunotherapy in melanoma: understanding the molecular basis for immune sensitivity and resistance. Nature Immunology *23*, 660-670. 10.1038/s41590-022-01141-1.
10. Motzer, R.J., Rini, B.I., McDermott, D.F., Redman, B.G., Kuzel, T.M., Harrison, M.R., Vaishampayan, U.N., Drabkin, H.A., George, S., Logan, T.F., et al. (2015). Nivolumab for Metastatic Renal Cell Carcinoma: Results of a Randomized Phase II Trial. Journal of Clinical Oncology *33*, 1430-1437. 10.1200/jco.2014.59.0703.
11. Topalian, S.L., Hodi, F.S., Brahmer, J.R., Gettinger, S.N., Smith, D.C., McDermott, D.F., Powderly, J.D., Carvajal, R.D., Sosman, J.A., Atkins, M.B., et al. (2012). Safety, Activity, and Immune Correlates of Anti–PD-1 Antibody in Cancer. New England Journal of Medicine *366*, 2443-2454. doi:10.1056/NEJMoa1200690.
12. Boutros, C., Tarhini, A., Routier, E., Lambotte, O., Ladurie, F.L., Carbonnel, F., Izzeddine, H., Marabelle, A., Champiat, S., Berdelou, A., et al. (2016). Safety profiles of anti-CTLA-4 and anti-PD-1 antibodies alone and in combination. Nat Rev Clin Oncol *13*, 473-486. 10.1038/nrclinonc.2016.58.
13. Lauss, M., Donia, M., Harbst, K., Andersen, R., Mitra, S., Rosengren, F., Salim, M., Vallon-Christersson, J., Törngren, T., Kvist, A., et al. (2017). Mutational and putative neoantigen load predict clinical benefit of adoptive T cell therapy in melanoma. Nature Communications *8*, 1738. 10.1038/s41467-017-01460-0.
14. Galon, J., and Bruni, D. (2019). Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. Nature Reviews Drug Discovery *18*, 197-218. 10.1038/s41573-018-0007-y.
15. Bagaev, A., Kotlov, N., Nomie, K., Svekolkin, V., Gafurov, A., Isaeva, O., Osokin, N., Kozlov, I., Frenkel, F., Gancharova, O., et al. (2021). Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. Cancer Cell *39*, 845-865 e847. 10.1016/j.ccell.2021.04.014.
16. Qian, J., Olbrecht, S., Boeckx, B., Vos, H., Laoui, D., Etlioglu, E., Wauters, E., Pomella, V., Verbandt, S., Busschaert, P., et al. (2020). A pan-cancer blueprint of the heterogeneous tumor microenvironment revealed by single-cell profiling. Cell Res *30*, 745-762. 10.1038/s41422-020-0355-0.

17. Zheng, L., Qin, S., Si, W., Wang, A., Xing, B., Gao, R., Ren, X., Wang, L., Wu, X., Zhang, J., et al. (2021). Pan-cancer single-cell landscape of tumor-infiltrating T cells. Science *374*, abe6474. 10.1126/science.abe6474.
18. Yost, K.E., Satpathy, A.T., Wells, D.K., Qi, Y., Wang, C., Kageyama, R., McNamara, K.L., Granja, J.M., Sarin, K.Y., Brown, R.A., et al. (2019). Clonal replacement of tumor-specific T cells following PD-1 blockade. Nat Med *25*, 1251-1259. 10.1038/s41591-019-0522-3.
19. Liu, B., Hu, X., Feng, K., Gao, R., Xue, Z., Zhang, S., Zhang, Y., Corse, E., Hu, Y., Han, W., and Zhang, Z. (2022). Temporal single-cell tracing reveals clonal revival and expansion of precursor exhausted T cells during anti-PD-1 therapy in lung cancer. Nat Cancer *3*, 108-121. 10.1038/s43018-021-00292-8.
20. Bassez, A., Vos, H., Van Dyck, L., Floris, G., Arijs, I., Desmedt, C., Boeckx, B., Vanden Bempt, M., Nevelsteen, I., Lambein, K., et al. (2021). A single-cell map of intratumoral changes during anti-PD1 treatment of patients with breast cancer. Nat Med *27*, 820-832. 10.1038/s41591-021-01323-8.
21. Zhang, Y., Chen, H., Mo, H., Hu, X., Gao, R., Zhao, Y., Liu, B., Niu, L., Sun, X., Yu, X., et al. (2021). Single-cell analyses reveal key immune cell subsets associated with response to PD-L1 blockade in triple-negative breast cancer. Cancer Cell *39*, 1578-1593 e1578. 10.1016/j.ccell.2021.09.010.
22. van der Leun, A.M., Traets, J.J.H., Vos, J.L., Elbers, J.B.W., Patiwael, S., Qiao, X., Machuca-Ostos, M., Thommen, D.S., Haanen, J., Schumacher, T.N.M., and Zuur, C.L. (2023). Dual Immune Checkpoint Blockade Induces Analogous Alterations in the Dysfunctional CD8+ T-cell and Activated Treg Compartment. Cancer Discov *13*, 2212-2227. 10.1158/2159-8290.CD-22-0851.
23. Li, J., Wu, C., Hu, H., Qin, G., Wu, X., Bai, F., Zhang, J., Cai, Y., Huang, Y., Wang, C., et al. (2023). Remodeling of the immune and stromal cell compartment by PD-1 blockade in mismatch repair-deficient colorectal cancer. Cancer Cell. 10.1016/j.ccell.2023.04.011.
24. Franken, A., Bila, M., Mechels, A., Kint, S., Van Dessel, J., Pomella, V., Vanuytven, S., Philips, G., Bricard, O., Xiong, J., et al. (2024). CD4+ T cell activation distinguishes response to anti-PD-L1+anti-CTLA4 therapy from anti-PD-L1 monotherapy. Immunity. 10.1016/j.immuni.2024.02.007.
25. Shiao, S.L., Gouin, K.H., 3rd, Ing, N., Ho, A., Basho, R., Shah, A., Mebane, R.H., Zitser, D., Martinez, A., Mevises, N.Y., et al. (2024). Single-cell and spatial profiling identify three response trajectories to pembrolizumab and radiation therapy in triple negative breast cancer. Cancer Cell *42*, 70-84 e78. 10.1016/j.ccell.2023.12.012.
26. Luoma, A.M., Suo, S., Wang, Y., Gunasti, L., Porter, C.B.M., Nabilsi, N., Tadros, J., Ferretti, A.P., Liao, S., Gurer, C., et al. (2022). Tissue-resident memory and circulating T cells are early responders to pre-surgical cancer immunotherapy. Cell *185*, 2918-2935 e2929. 10.1016/j.cell.2022.06.018.
27. Chen, Y., Wang, D., Li, Y., Qi, L., Si, W., Bo, Y., Chen, X., Ye, Z., Fan, H., Liu, B., et al. (2024). Spatiotemporal single-cell analysis decodes cellular dynamics underlying different responses to immunotherapy in colorectal cancer. Cancer Cell *42*, 1268-1285.e1267. 10.1016/j.ccell.2024.06.009.
28. Hawley, J.E., Obradovic, A.Z., Dallos, M.C., Lim, E.A., Runcie, K., Ager, C.R., McKiernan, J., Anderson, C.B., Decastro, G.J., Weintraub, J., et al. (2023). Anti-PD-1 immunotherapy with androgen deprivation therapy induces robust immune infiltration in metastatic castration-sensitive prostate cancer. Cancer Cell. 10.1016/j.ccell.2023.10.006.
29. Li, B., Cui, Y., Diehn, M., and Li, R. (2017). Development and Validation of an Individualized Immune Prognostic Signature in Early-Stage Nonsquamous Non-Small Cell Lung Cancer. JAMA Oncol *3*, 1529-1537. 10.1001/jamaoncol.2017.1609.

30. Andreatta, M., Berenstein, A.J., and Carmona, S.J. (2022). scGate: marker-based purification of cell types from heterogeneous single-cell RNA-seq datasets. Bioinformatics *38*, 2642-2644. 10.1093/bioinformatics/btac141.
31. Aran, D., Looney, A.P., Liu, L., Wu, E., Fong, V., Hsu, A., Chak, S., Naikawadi, R.P., Wolters, P.J., Abate, A.R., et al. (2019). Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. Nat Immunol *20*, 163-172. 10.1038/s41590-018-0276-y.
32. Gavish, A., Tyler, M., Greenwald, A.C., Hoefflin, R., Simkin, D., Tschernichovsky, R., Galili Darnell, N., Somech, E., Barbolin, C., Antman, T., et al. (2023). Hallmarks of transcriptional intratumour heterogeneity across a thousand tumours. Nature. 10.1038/s41586-023-06130-4.
33. Kang, J., Lee, J.H., Cha, H., An, J., Kwon, J., Lee, S., Kim, S., Baykan, M.Y., Kim, S.Y., An, D., et al. (2024). Systematic dissection of tumor-normal single-cell ecosystems across a thousand tumors of 30 cancer types. Nat Commun *15*, 4067. 10.1038/s41467-024-48310-4.
34. Traag, V.A., Waltman, L., and van Eck, N.J. (2019). From Louvain to Leiden: guaranteeing well-connected communities. Scientific Reports *9*, 5233. 10.1038/s41598-019-41695-z.
35. Chen, W., Teo, J.M.N., Yau, S.W., Wong, M.Y.-M., Lok, C.-N., Che, C.-M., Javed, A., Huang, Y., Ma, S., and Ling, G.S. (2022). Chronic type I interferon signaling promotes lipid-peroxidation-driven terminal CD8+ T cell exhaustion and curtails anti-PD-1 efficacy. Cell Reports *41*, 111647. https://doi.org/10.1016/j.celrep.2022.111647.
36. Benci, J.L., Johnson, L.R., Choa, R., Xu, Y., Qiu, J., Zhou, Z., Xu, B., Ye, D., Nathanson, K.L., June, C.H., et al. (2019). Opposing Functions of Interferon Coordinate Adaptive and Innate Immune Responses to Cancer Immune Checkpoint Blockade. Cell *178*, 933-948.e914. https://doi.org/10.1016/j.cell.2019.07.019.
37. Mathew, D., Marmarelis, M.E., Foley, C., Bauml, J.M., Ye, D., Ghinnagow, R., Ngiow, S.F., Klapholz, M., Jun, S., Zhang, Z., et al. (2024). Combined JAK inhibition and PD-1 immunotherapy for non–small cell lung cancer patients. Science *384*. 10.1126/science.adf1329.
38. Teillaud, J.L., Houel, A., Panouillot, M., Riffard, C., and Dieu-Nosjean, M.C. (2024). Tertiary lymphoid structures in anticancer immunity. Nat Rev Cancer. 10.1038/s41568-024-00728-0.
39. Germain, P.-L., Lun, A., Garcia Meixide, C., Macnair, W., and Robinson, M.D. (2022). Doublet identification in single-cell sequencing data using scDblFinder. F1000Research *10*, 979. 10.12688/f1000research.73600.2.
40. McGinnis, C.S., Murrow, L.M., and Gartner, Z.J. (2019). DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. Cell Systems *8*, 329-337.e324. 10.1016/j.cels.2019.03.003.
41. Andreatta, M., Hérault, L., Gueguen, P., Gfeller, D., Berenstein, A.J., and Carmona, S.J. (2024). Semi-supervised integration of single-cell transcriptomics data. Nature Communications *15*, 872. 10.1038/s41467-024-45240-z.
42. Stunnenberg, H.G., Abrignani, S., Adams, D., de Almeida, M., Altucci, L., Amin, V., Amit, I., Antonarakis, S.E., Aparicio, S., Arima, T., et al. (2016). The International Human Epigenome Consortium: A Blueprint for Scientific Collaboration and Discovery. Cell *167*, 1145-1149. https://doi.org/10.1016/j.cell.2016.11.007.
43. Dunham, I., Kundaje, A., Aldred, S.F., Collins, P.J., Davis, C.A., Doyle, F., Epstein, C.B., Frietze, S., Harrow, J., Kaul, R., et al. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57-74. 10.1038/nature11247.
44. Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.R., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods *16*, 1289-1296. 10.1038/s41592-019-0619-0.
45. Hao, Y., Stuart, T., Kowalski, M.H., Choudhary, S., Hoffman, P., Hartman, A., Srivastava, A., Molla, G., Madad, S., Fernandez-Granda, C., and Satija, R. (2023). Dictionary learning for integrative, multimodal and scalable single-cell analysis. Nat Biotechnol. 10.1038/s41587-023-01767-y.

46. Tang, F., Li, J., Qi, L., Liu, D., Bo, Y., Qin, S., Miao, Y., Yu, K., Hou, W., Li, J., et al. (2023). A pan-cancer single-cell panorama of human natural killer cells. Cell. 10.1016/j.cell.2023.07.034.
47. Cheng, S., Li, Z., Gao, R., Xing, B., Gao, Y., Yang, Y., Qin, S., Zhang, L., Ouyang, H., Du, P., et al. (2021). A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells. Cell *184*, 792-809 e723. 10.1016/j.cell.2021.01.010.
48. Ma, J., Wu, Y., Ma, L., Yang, X., Zhang, T., Song, G., Li, T., Gao, K., Shen, X., Lin, J., et al. (2024). A blueprint for tumor-infiltrating B cells across human cancers. Science *384*, eadj4857. 10.1126/science.adj4857.
49. Luo, H., Xia, X., Huang, L.-B., An, H., Cao, M., Kim, G.D., Chen, H.-N., Zhang, W.-H., Shu, Y., Kong, X., et al. (2022). Pan-cancer single-cell analysis reveals the heterogeneity and plasticity of cancer-associated fibroblasts in the tumor microenvironment. Nature Communications *13*. 10.1038/s41467-022-34395-2.
50. Kuznetsova, A., Brockhoff, P.B., and Christensen, R.H.B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. Journal of Statistical Software *82*, 1 - 26. 10.18637/jss.v082.i13.
51. Andreatta, M., and Carmona, S.J. (2021). UCell: Robust and scalable single-cell gene signature scoring. Computational and Structural Biotechnology Journal *19*, 3796-3798. https://doi.org/10.1016/j.csbj.2021.06.043.
52. Dolgalev, I. (2022). msigdbr: MSigDB Gene Sets for Multiple Organisms in a Tidy Data Format.
53. Fang, Z., Liu, X., and Peltz, G. (2022). GSEApy: a comprehensive package for performing gene set enrichment analysis in Python. Bioinformatics *39*. 10.1093/bioinformatics/btac757.
54. Castanza, A.S., Recla, J.M., Eby, D., Thorvaldsdóttir, H., Bult, C.J., and Mesirov, J.P. (2023). Extending support for mouse data in the Molecular Signatures Database (MSigDB). Nature Methods *20*, 1619-1620. 10.1038/s41592-023-02014-7.
55. Gaujoux, R. (2024). An introduction to NMF package. https://cran.r-project.org/web/packages/NMF/vignettes/NMF-vignette.pdf.
56. Simko, T.W.a.V. (2021). R package 'corrplot': Visualization of a Correlation Matrix.
57. Gu, Z. (2022). Complex heatmap visualization. iMeta *1*, e43. https://doi.org/10.1002/imt2.43.

Figure Legends

Figure 1. Large-scale pan-cancer atlas of longitudinally paired ICI-treated tumors. A. Summary of study design and analysis workflow. **B-D.** UMAP plots showing main components, source, treatment of integrated dataset. **E.** Dotplot showing marker expression for main components. F. Clinical information and composition of main components by patient. **G-J.** Forest plots showing the association between main components, pathways (CD4+T, CD8+T, and macrophages) and time points during ICI treatment, the effect sizes (error bars indicate 95% confidence intervals) were based on the coefficients from linear mixed-effects models.

Figure 2. Deep-phenotyping annotation and characterization. A. Deep-phenotyping annotation workflow. B. Heatmap showing the presence of cell states/subtypes across datasets. Existing (blue) represents cell types were detected in > 3 samples in a specific cohort, or in at least 1 sample when the cohort has no more than 4 samples. **C.** Dot plot showing gene expression of myeloid cells. Colors represent relative mean expression (column-wise scaled) of marker genes, and the size represents the

percentage of cells expressing these genes. **D-F.** Heatmaps show curated signatures across CD4+T cells, CD8+T cells, and macrophages. **G-H.** Heatmaps show relative mean expression (column-wise scaled) selected marker gene expression across conventional dendritic cells and B cells.

Figure 3. Patient-level compositional analysis. A. Forest plot for the associations between cellular abundance (relative frequency, RF) and treatment time points. **B-E.** Forest plots for the associations between cellular abundance (RF) and time points by cohort in selected cell types. **F.** Paired box and violin plots comparing RF between time points for the top-6 cell states exhibiting the strongest changes. Boxes show 25th, median, and 75th percentiles, the RF changes were classified as "Up" ($\Delta > 10\%$), "Down" ($\Delta < -10\%$) during the treatment. **G.** Effect sizes from subgroups analysis of monotherapy (Mono) treated patients and combination therapy (Dual) treated patients. **H.** Heatmap showing the effect sizes (Cohen's d) of compositional change for each cell type (column) in modality-cohort subgroups (row), *$p<0.05$, **$p<0.01$, ***$p<0.001$, based on paired Student's t-test. Gray color indicates that the cell type wasn't detected in < 3 patients in the subgroup.

Figure 4. Tumor ecosystem and Co-regulating patterns upon ICI treatment. A. NMF-based TME immune programs. The color represents the loading (column-wise scaled) of every cell type in each program. **B.** Box plot and violin plot showing the dynamic changes of three immune programs in tumors. Boxes show 25th, median, and 75th percentiles, each dot represents a patient. P values were based on two-side Student's t-test. **C.** Heatmap showing the correlation between immune programs and non-immune compartments. **D.** Heatmap of correlation matrix showing the pairwise Pearson correlations between dynamic changes (ΔRF) of immune cell types per patient, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, based on the correlation. **E.** Undirected graph showing significant positive (FDR<0.05, Pearson's rho>0.3) correlations (edges) between cell types (nodes). **F.** Box and violin plot comparing overall dynamics of *ISG15+* cell subtypes between expanders and non-expanders. Boxes show 25th, median, and 75th percentiles. **G.** Heatmap showing a continuum from early activation/innate immune response to immune suppression. Patients (column) were sorted according to sum of corresponding cell type (row) dynamic changes (column-wise scaled).

Figure 5. Association of Immune Cell Dynamics with Treatment Response. A. Heatmap showing the pairwise Pearson correlation matrix across the patients according to the dynamics of immune cells. **B.** Mosaic plot showing numbers of responders and non-responders in dynamic-derived clusters. **C.** Volcano plot showing differential dynamics of immune cell dynamics between cluster 1 (responder-enriched) and cluster 3 (non-responder enriched), effect sizes were based on Cohen's d and p values were based on

paired t-test. **D.** Heatmap showing the median dynamics (Δ relative frequency, row-wise scaled) of immune cell types (row) in the 3 clusters. **E.** Association between cell types and time points in responders and non-responders, the effect sizes (error bars indicate 95% confidence intervals) were based on the coefficients from linear mixed-effects models. **F.** Heatmap showing the effect sizes (Cohen's d) of compositional change for each cell types (column) in response-cohort subgroups (row), *p<0.05, **p<0.01, ***p<0.001, based on paired Student's t-test. Gray color indicates that the cell type detected in <3 patients from the subgroup.

Supplementary Tables and Figures

Table S1. Patients information.

Figure S1. Marker gene expression and cell type abundance. A-D. Dot plots showing gene expression of CD4+T, CD8+T, NK, B, and non-immune cells. Colors represent relative mean expression (column-wise scaled) of marker genes, and the size represents the percentage of cells expressing these genes. **E-I.** Box plots showing the abundance of CD4+T, CD8+T, NK, B, and non-immune cells

Figure S2. A. Heatmap showing the median (row-wise scaled) proportion (relative frequency, ΔRF) for cell types across cohorts between time points. The gray color indicates that the cell type was absent from the cohort. **B.** Paired box and violin plots comparing RF between time points for cell states. Boxes showed 25th, median, and 75th percentiles, p values were based on paired Student's t-test, the RF changes were classified as "Up" (Δ > 10%), "Down" (Δ < -10%) during the treatment. **C.** Dot plots showing top differentially expressed genes from comparison between CREM-CD4+ memory T cells and other CD4+ memory T cells. Colors represent relative mean expression (column-wise scaled) of marker genes, and the size represents the percentage of cells expressing these genes. **D.** Bar plot showing up (red) and down (down) regulated pathways (overlapped genes > 50%) from comparison between CREM-CD4+ memory T cells and other CD4+ memory T cells, p values were corrected by Benjamini–Hochberg method. **E.** Association between cell types and time points in patients under monotherapy (Mono) and combination therapy (Dual), the effect sizes (error bars indicate 95% confidence intervals) were based on the coefficients from linear mixed-effects models.

Figure S3. Co-regulating patterns upon ICI treatment. A. Undirected graph showing significant negative [false discovery rate (FDR)<0.05, Pearson's rho>0.3] correlations (edges) between immune cell types (nodes). **B.** Heatmap showing the pairwise correlation matrix between dynamic changes (ΔRF) of

TME cell types per patient, *p < 0.05, **p < 0.01, ***p < 0.001, based on the Pearson correlation. **C.** Undirected graph showing significant positive [false discovery rate (FDR)<0.05, Pearson's rho>0.3] correlations (edges) between TME cell types (nodes). **D.** Undirected graph showing significant negative [false discovery rate (FDR)<0.05, Pearson's rho>0.3] correlations (edges) between TME cell types (nodes). **E.** Scatter plot showing correlation between *LYVE1+* macrophages and adipogenic cancer-associated fibroblasts base on the dynamic changes. **F.** Heatmap showing the continuum from early activation/innate immune response to immune suppression for immune cells. Patients (column) were sorted accroding to sum of corresponding cell type (row) dynamic changes.

Figure S4. Association of Immune Cell Conposition with Treatment Response. A. Heatmap showing the pairwise Pearson correlation matrix across the patients according to the immune cell abundance in pre-treatment (pre-Tx) samples. **B.** Mosaic plot showing numbers of responders and non-responders in pre-Tx abundance derived clusters. **C.** Heatmap showing the pairwise Pearson correlation matrix across the patients according to the immune cell abundance in pre-treatment (post-Tx) samples. **D.** Mosaic plot showing numbers of responders and non-responders in post-Tx abundance derived clusters. **E.** Alluvium plot comparing pre-Tx, dynamic, and post-Tx clusters. **F.** Scater plot comparing coefficients from responders (RE) and non-responders (NR) analysis for cell types.