

Distinct immune cell dynamics associated with immune-related adverse events during combined chemoradiation and immune checkpoint inhibitor therapy

Received: 3 April 2025

Accepted: 4 December 2025

Published online: 22 December 2025

 Check for updates

Lei Zhang^{1,7}, Xiaokai Fan^{2,3}, Jun Ma^{4,7}, Jun Zhang^{2,3,7}, Ying Wei^{2,3}, Bin Hu¹, Di Zhou¹, Junjun Zhou¹, Yongrui Bai¹, Xiumei Ma¹, Jian Song⁵  , Jianming Tang⁶  , Haiyan Chen¹   & Ying Jing^{2,3}  

Combining chemoradiotherapy with immunotherapy increases the risk of immune-related adverse events (irAEs), but the underlying mechanisms remain poorly understood. To address this, we conducted a longitudinal single-cell multi-omics analysis of patients with locally advanced cervical cancer. Here we show that the proportions of CD4⁺ and CD8⁺ terminally differentiated effector memory or effector T cells are elevated in patients with irAEs. Chemoradiotherapy reduces B cell clonality while increasing the abundance and somatic hypermutation frequencies of IgA⁺ and IgG⁺ B cells in irAE patients. In the myeloid compartment, combined treatment expands specific monocyte subclusters associated with irAEs. Spatial transcriptomics and immunofluorescence analyses further reveal that these irAE-associated immune cells aggregate within the tumor microenvironment. Finally, we develop the predictive models for irAEs and integrate them, along with all datasets, into a user-friendly data portal. Our findings suggest that chemoradiotherapy and immunotherapy exert distinct effects on different immune cells, contributing to irAE development.

Cervical cancer is the third most common cause of death among women globally¹, with its incidence continuing to rise annually in China². Although early-stage cervical cancer is generally curable, approximately 40–50% of patients are diagnosed with locally advanced cervical cancer (LACC, International Federation of Gynecology and Obstetrics stage IB3-IVA), and this contributes significantly to the high mortality associated with this disease^{3,4}. The current

standard treatment for LACC is concurrent chemoradiotherapy. However, 30–50% of these patients experience disease recurrence or progression within five years^{5,6}. Combining chemoradiotherapy with immunotherapy is being explored as a potential strategy to enhance treatment efficacy and improve the prognosis for patients with LACC⁷. However, this combination also brings an increased risk of immune-related adverse events (irAEs), as the immune responses triggered by

¹Department of Radiation Oncology, Ren Ji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China. ²Center for Intelligent Medicine Research, Greater Bay Area Institute of Precision Medicine (Guangzhou), School of Life Sciences, Fudan University, Guangzhou, China. ³State Key Laboratory of Genetic Engineering, Center for Evolutionary Biology, School of Life Sciences, Fudan University, Shanghai, China. ⁴Eye Institute, Eye & ENT Hospital, Shanghai Medical College, Fudan University, Shanghai, China. ⁵Institute of Cardiovascular Sciences, Guangxi Academy of Medical Sciences, & The People's Hospital of Guangxi Zhuang Autonomous Region, Nanning, China. ⁶Department of Radiation Oncology, The First Hospital of Lanzhou University, Lanzhou University, Lanzhou, China. ⁷These authors contributed equally: Lei Zhang, Jun Ma, Jun Zhang. ✉ e-mail: jsong@gxams.org.cn; ldyy_tangjm@lzu.edu.cn; chenhaiyan1220@163.com; jingying@ipm-gba.org.cn

radiotherapy may synergize with those from immunotherapy^{8,9}. In addition to divergent treatment efficacies observed in recent clinical trials, the incidence of grade 3 or higher treatment-related toxicities remains consistently high^{5,6,10–12}.

irAEs are caused by excessive activation of the immune system against host tissues, potentially leading to serious consequences, such as discontinuation of treatment, permanent tissue damage and even fatal outcomes^{13,14}. These reactions, which vary from mild to life-threatening, can significantly diminish the clinical benefits to the patients¹⁵. Nevertheless, a few studies have identified molecular characteristics associated with treatment efficacy rather than irAEs. For example, enhanced stromal immune infiltration¹⁰, higher tumor aneuploidy scores¹⁶ and the presence of circulating antigen-experienced CD103⁺ T cells¹⁷ have all been linked to improved patient outcomes. Notably, immune responses specifically related to irAEs in this context remain largely unexplored. Understanding the mechanisms driving irAEs and identifying relevant biomarkers could improve the prediction and management of these adverse events, ultimately enhancing treatment safety and clinical outcomes for patients undergoing combination immunotherapy.

To address existing knowledge gaps, we conducted a clinical trial (NCT06378840, an observational standard of care treatment study) aimed to evaluate the role of immune cells in treatment-naïve patients with LACC receiving sequential chemoradiotherapy and ICI therapy. We performed single-cell transcriptome and V(D)J sequencing on matched tumor and peripheral blood samples, which were collected at four distinct treatment stages. This approach allowed us to investigate how different treatments affected both tumor behavior and systemic immune responses, with a specific focus on identifying irAE-specific changes triggered by various treatment approaches. By utilizing these longitudinal data, our deep learning models were able to identify representative irAE characteristics of immune cells at various treatment stages, and this was further validated in ten independent irAE datasets. This study provides a comprehensive single-cell immune profile related to irAEs and establishes reliable predictive models for irAEs. These models may help identify patients with higher risks of developing irAEs and could potentially assist in optimizing patient selection for those undergoing combined chemoradiation and ICI therapies.

Results

Longitudinal single-cell multi-omics atlas of LACC tumor and blood samples

We prospectively enrolled thirty patients with treatment-naïve LACC and collected 96 tumor biopsies and peripheral blood samples. Due to the prospective nature of our study and the need to maintain high cell integrity and RNA quality, we exclusively used fresh samples and did not include any frozen samples. Among the patients, twenty received sequential chemoradiotherapy combined with anti-programmed cell death-1 (PD-1) antibodies (radio+ICI group), while the remaining ten underwent chemoradiotherapy followed by chemotherapy alone (radio group) (Fig. 1a, b, Supplementary Table 1a). All patients were closely monitored for irAE development throughout their treatment, with a median follow-up time of 524 days and a median time to irAE of 139 days (Supplementary Table 1b). The majority of patients experienced at least one irAEs, ranging from mild (grade 1) to severe (grade 3), and these affected various organ systems (Supplementary Table 1b, Methods). Blood and tumor samples were collected at key time points: pre-chemoradiotherapy (pre-radio), during chemoradiotherapy (on-radio), pre- and during ICI treatment (pre-imm and on-imm) for radio+ICI group and pre- and during chemotherapy (pre-chemo and on-chemo) for radio group (Fig. 1a). Specifically, pre-radio samples were collected before initiation of therapy. On-radio samples were collected after the 10th fraction of external beam irradiation (1.8 Gy per fraction). Pre-imm and pre-chemo samples were obtained after completing

chemoradiotherapy but before administering ICI or chemotherapy. On-imm samples were collected before the 4th cycle of ICI, and on-chemo samples were taken before the 3rd cycle of chemotherapy (Fig. 1a). This sample collection strategy allowed us to track local and systemic immunological dynamics throughout the treatment process.

We employed 10× Genomics Chromium single-cell RNA sequencing (scRNA-seq) in conjunction with single-cell V(D)J sequencing (scV(D)J-seq) of T cell receptor (TCR) and B cell receptor (BCR) clonotypes to analyze the molecular characteristics and alterations of the tumor microenvironment (TME) and systemic immune system in patients with LACC during treatment (Fig. 1b, Supplementary Table 1a). After quality control and removal of doublets, the scRNA-seq assay yielded 708,924 cells, representing 13 major cell types and 48 fine-grained cell subclusters based on the canonical marker gene expression (Fig. 1c, Supplementary Fig. 1a). These cell types were shared across patients, tissue types, sample collection time points and treatment modalities (Supplementary Fig. 1b–e). We identified the tumor cells from epithelial cells by using the inferred copy number variation (CNV), which aligned with the data obtained from scRNA-seq and whole exome sequencing data from another study¹⁸, with the cells generally being clustered by patient (Supplementary Fig. 1b, f).

Cellular dynamics in response to chemoradiotherapy and ICI treatment in relation to irAEs

Next, we analyzed the cellular abundance of major cell types in relation to treatments and clinical endpoints, defined by the occurrence of irAEs and disease relapse. Chemoradiotherapy significantly reduced tumor cell proportions in tumor samples and increased macrophages in both tumor and peripheral blood samples, with monocytes showing a notable increase exclusively in peripheral blood samples (Supplementary Fig. 2a, b), consistent with well-documented effects of radiotherapy^{19,20}. Following ICI treatment, T cells, macrophages and monocytes showed slight decreases in blood samples (Supplementary Fig. 2b). Comparisons between irAE and non-irAE patients revealed elevated proportions of B cells, T cells and macrophages in blood samples collected from irAE patients prior to ICI administration. After ICI treatment, T cells increased, while macrophages decreased in irAE patients (Fig. 1d, Supplementary Fig. 3d, e). No significant differences in immune cell proportions were observed in tumor samples at pre-radio and on-radio time points between the irAE and non-irAE patients (Supplementary Fig. 3a–b). Collectively, these findings indicate that chemoradiotherapy reduces tumor cells and increases macrophages across tumor and peripheral blood samples. Notably, irAE-related immune cell changes were observed in the peripheral blood samples collected before and after ICI treatment, suggesting that systemic immune responses appear to be more closely linked to irAEs. However, pre-radio tumor samples from patients who remained relapse-free two years after receiving first therapy (Methods) showed increased proportions of several immune cells, including B cells, T cells and macrophages. Meanwhile, blood samples from non-relapsed patients collected prior to ICI treatment displayed elevated levels of B cells and T cells, with macrophages decreased before and after ICI treatment (Fig. 1e, Supplementary Fig. 3f–i). These results indicate that treatment response-related alterations were predominantly identified in baseline tumor samples and immune responses within the TME may be more strongly associated with treatment outcomes.

Increased CD8.Temra and CD4.Temra cells during combined treatment are associated with irAEs

Our aforementioned analyses revealed an increased proportion of T cells in the blood and tumor associated with irAEs and disease relapse, respectively. Next, we compared the abundances and functions of T cell subclusters between patients in different irAE and disease relapse subgroups. We used canonical immune markers²¹ and curated gene signatures²² to identify T cell subclusters and these

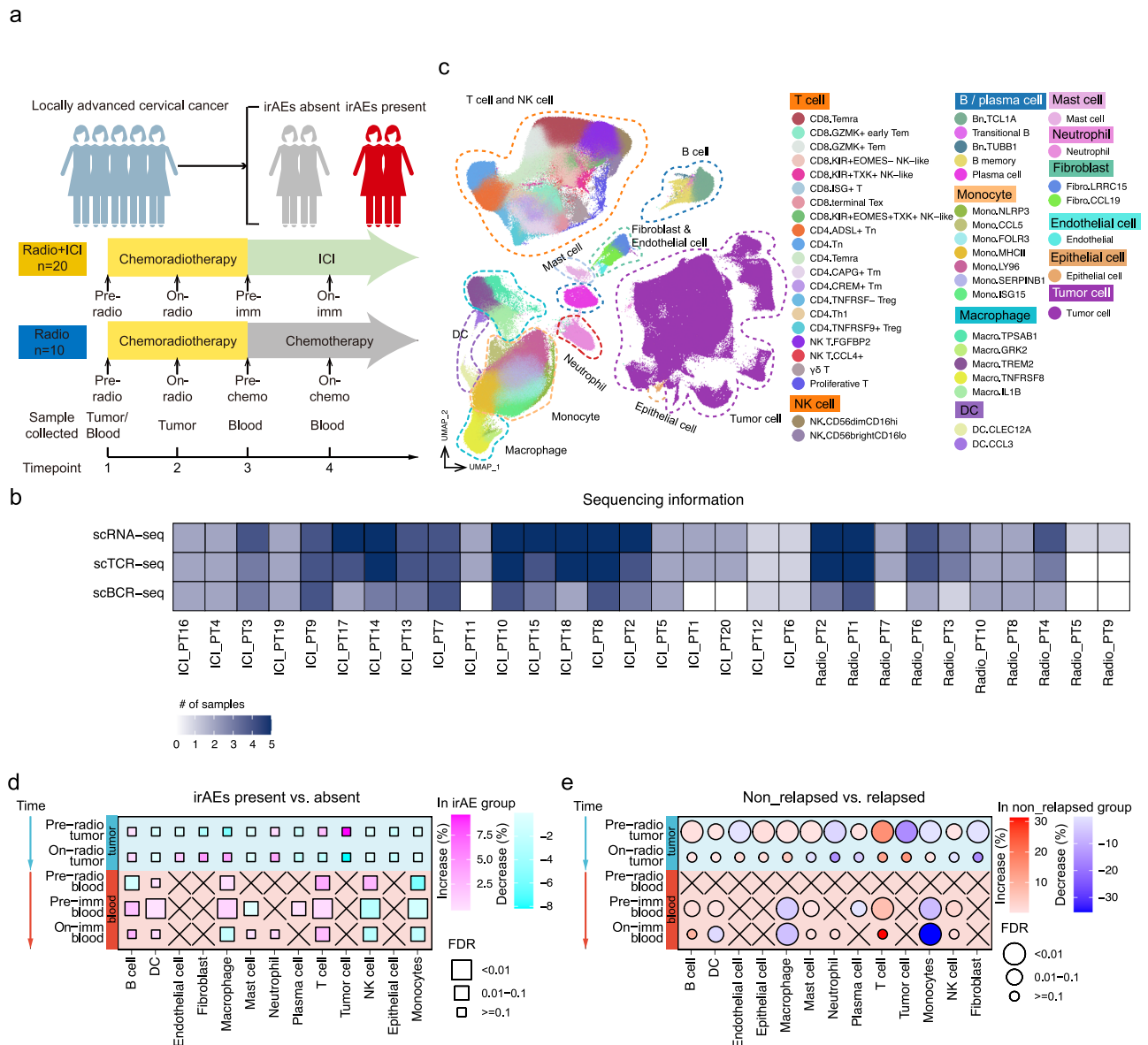
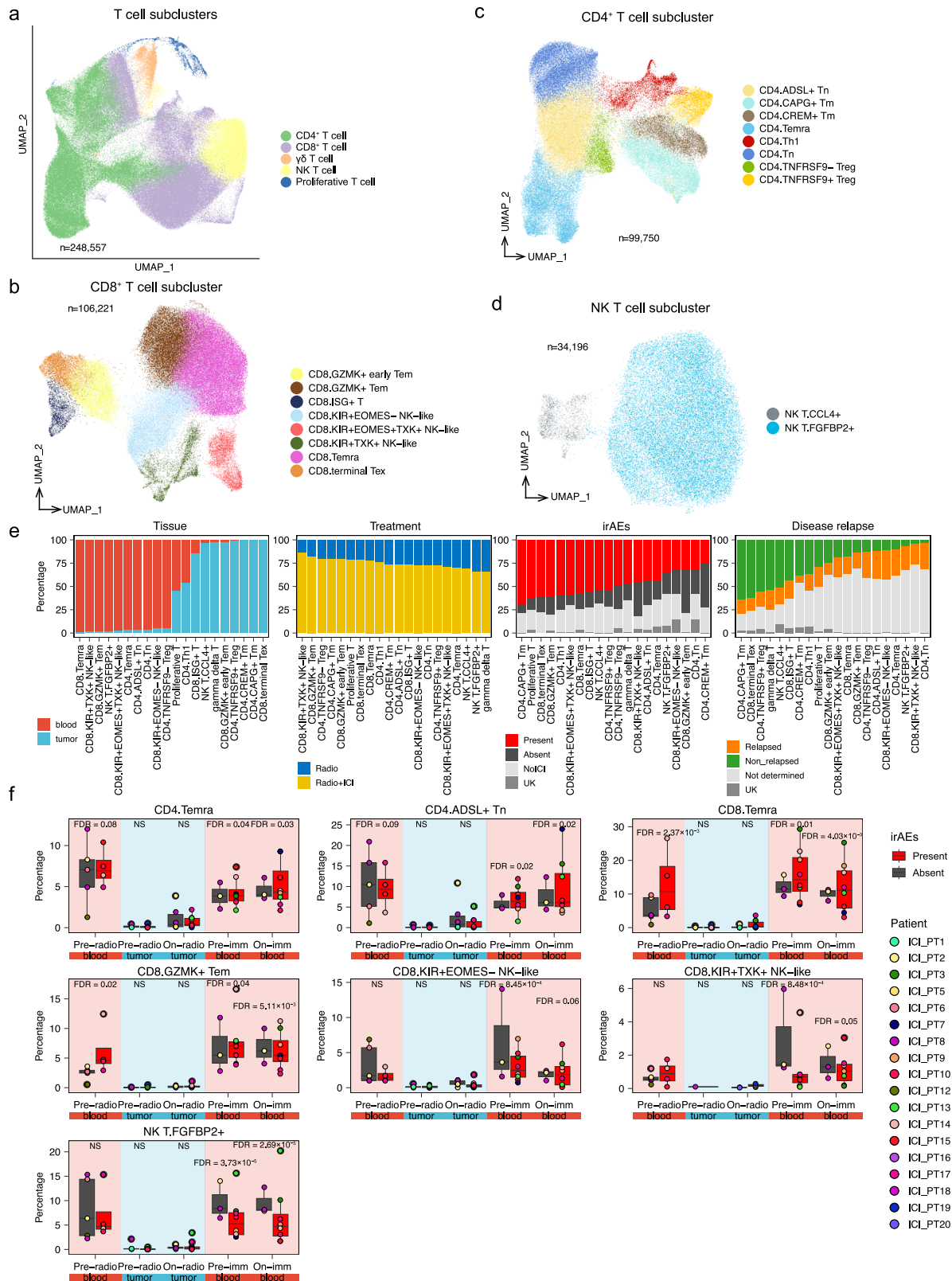


Fig. 1 | Dynamic single-cell multi-omics landscape of LACC patients during chemoradiotherapy and ICI. a A schematic diagram of the treatment design and sample collection. Patients with LACC were divided into two groups: those receiving chemoradiotherapy alone (Radio) and those receiving chemoradiotherapy combined with immune checkpoint inhibitors (Radio+ICI). Fresh tumor and blood samples were collected at four time points per patient for single-cell multi-omics analysis. **b** Characteristics of the single-cell multi-omics cohort, including the irAE onset, anatomic sites of irAEs, irAE grade, disease relapse status two years after entire treatment and number of samples for each sequencing (Supplementary Table 1). **c** Uniform manifold approximation and projection showing the major and fine-grained cell clusters. **d, e** Heatmaps illustrating the association between the percentage of major cell clusters and irAEs (**d**) or disease relapse (**e**). The magenta and cyan squares indicate the higher and lower percentages in irAE patients, respectively. The red and blue dots indicate the higher and lower percentages in non-relapsed patients, respectively. The size of the squares and dots represents statistical significance, with large, medium and small dots

indicating *FDR* of <0.01, 0.01-0.1 and ≥ 0.1 , respectively. × indicates the data was not applicable. *P* values were calculated using the two-sided Dirichlet-multinomial regression model and *FDR* values were calculated by the Benjamini–Hochberg method. Sample sizes: pre-radio and on-radio tumor (irAEs absent: $n = 6$; irAEs present: $n = 12$; non_relapsed: $n = 8$; relapsed $n = 2$); pre-radio PBMC (irAEs absent: $n = 5$; irAEs present: $n = 4$); pre-imm and on-imm PBMC (irAE absent: $n = 3$; irAE present: $n = 8$; non_relapsed: $n = 3$; relapsed: $n = 2$). All sample sizes reported are biological replicates, with the patient serving as the independent unit of study. LACC locally advanced cervical cancer, ICI immune checkpoint inhibitor, irAEs immune-related adverse events, Radio chemoradiotherapy, pre-radio before chemoradiotherapy and ICI, on-radio during chemoradiotherapy, pre-imm after chemoradiotherapy but before ICI, treatment in the Radio+ICI group, on-imm during ICI treatment in the Radio+ICI group, pre-chemo after chemoradiotherapy but before chemotherapy in the Radio group, on-chemo during chemotherapy in the Radio group, NK natural killer, DC dendritic cell. Source data are provided as a Source Data file.

included eight CD4⁺, eight CD8⁺, two natural killer T (NK T) as well as one each of $\gamma\delta$ T and proliferative T subclusters (Fig. 2a–d, Supplementary Figs. 4a, 5a–e). While most T cell subclusters were shared across samples from different patients as well as during treatments and at different time points, some of these, such as CD4⁺ terminally

differentiated effector memory or effector T cells (CD4.Temra), CD8⁺ terminally differentiated effector memory or effector T cells (CD8.Temra) and CD4⁺ adenylosuccinate lyase-positive naive T cells (CD4.ADSL⁺ Tn), showed a strong distribution preference in peripheral blood (Fig. 2a, Supplementary Fig. 4b). For CD4⁺ T cells in patients with



irAEs, the proportions of CD4.Temra, with cytotoxicity phenotype, and CD4.ADSL⁺ Tn, with a highly expressed oxidative phosphorylation signature, were significantly reduced in baseline pre-radio blood samples but increased in pre-imm blood samples after chemotherapy. Consistent with the reported role of memory CD4⁺ T cells in irAEs²², CD4.Temra increased following ICI treatment in both irAE

groups, with higher abundances observed in patients with irAEs (Fig. 2e, f, Supplementary Fig. 5e).

For CD8⁺ T cell subclusters, we observed significantly increased proportions of CD8.Temra, with highly expressed cytotoxicity signature, and CD8⁺ granzyme K-positive effector memory T cells (CD8.GZMK⁺ Tem), with highly expressed senescence and stress

Fig. 2 | Associations of T cells with irAEs. a–d The uniform manifold approximation and projection (UMAP) plots of 248,557 high-quality T cells (a) and CD8⁺ T cell subcluster (b), CD4⁺ T cell subcluster (c) and NK T cell subcluster (d). Each dot represents individual cell and colored by their cell subclusters. **e** Stacked bar plots showing the distribution of the T cell subclusters in different tissue type, treatment, irAE status and disease relapse. **f** Proportional analysis of T cell subclusters comparing patients with and without irAEs at different time points. In all boxplots, the center line represents the median, the bounds of the box indicate the 25th and 75th percentiles (interquartile range, IQR), and the whiskers extend to 1.5 × IQR. Each dot represents an individual patient. *P* values were calculated using the two-sided

Dirichlet-multinomial regression model and *FDR* values were calculated by the Benjamini–Hochberg method. Sample sizes: pre-radio and on-radio tumor (irAEs absent: *n* = 6; irAEs present: *n* = 12); pre-radio PBMC (irAEs absent: *n* = 5; irAEs present: *n* = 4); pre-imm and on-imm PBMC (irAE absent: *n* = 3; irAE present: *n* = 8). All sample sizes reported are biological replicates, with the patient serving as the independent unit of study. Tn naive T cell, Tm memory T cell, NK natural killer, Tem effector memory T cells, Th T helper cell, Tex exhausted T cell, Temra terminally differentiated effector memory or effector T cells, ICI immune checkpoint inhibitor, irAEs immune-related adverse events, Radio chemoradiotherapy, UK, unknown, NoICI not treated by ICI. Source data are provided as a Source Data file.

response signatures, in both the pre-radio (baseline) and pre-imm (after chemoradiotherapy but before ICI treatment) blood samples of irAE patients. However, these two subclusters showed opposite alterations in on-imm (after chemoradiotherapy and during ICI treatment) blood samples from irAE patients (Fig. 2e, f, Supplementary Fig. 5d). Interestingly, two NK-like CD8⁺ T cell subclusters, with relatively higher expression of cytotoxicity signature (Supplementary Fig. 5d), and one NK T subcluster, CD8⁺ killer cell immunoglobulin-like receptor-positive tyrosine kinase-positive NK-like T cell (CD8.KIR⁺TXK⁺ NK-like), CD8⁺ killer cell immunoglobulin-like receptor-positive eomesodermin-negative NK-like T cell (CD8.KIR⁺EOMES⁻ NK-like) and fibroblast growth factor binding protein 2-positive NK T (NK T.FGF2⁺), decreased in both pre-imm and on-imm samples (Fig. 2e). Our findings indicated that during chemoradiotherapy and ICI treatment, the proportions of CD4⁺ Temra and CD8⁺ Temra, with cytotoxicity and memory phenotypes, were elevated in patients experiencing irAEs, indicating the involvement of both CD4⁺ and CD8⁺ T cells in irAE development.

Although the overall T cell proportions were significantly higher in the pre-radio tumor sample of non-relapse patients, we only observed proportion increments in three CD4⁺ T cell subpopulations. These included CD4⁺ capping actin protein, gelsolin like-positive memory T cells (CD4.CAPG⁺ Tm), CD4⁺ tumor necrosis factor receptor superfamily member 9-positive regulatory T cells (CD4.TNFRSF9⁺ Treg) and CD4⁺ Temra (Supplementary Fig. 5f), which exhibited highly expressed cytotoxicity or exhaustion signatures (Supplementary Fig. 5e).

Differential TCR diversity and clonal expansion in CD4⁺ and CD8⁺ T Cells of irAE patients following ICI treatment

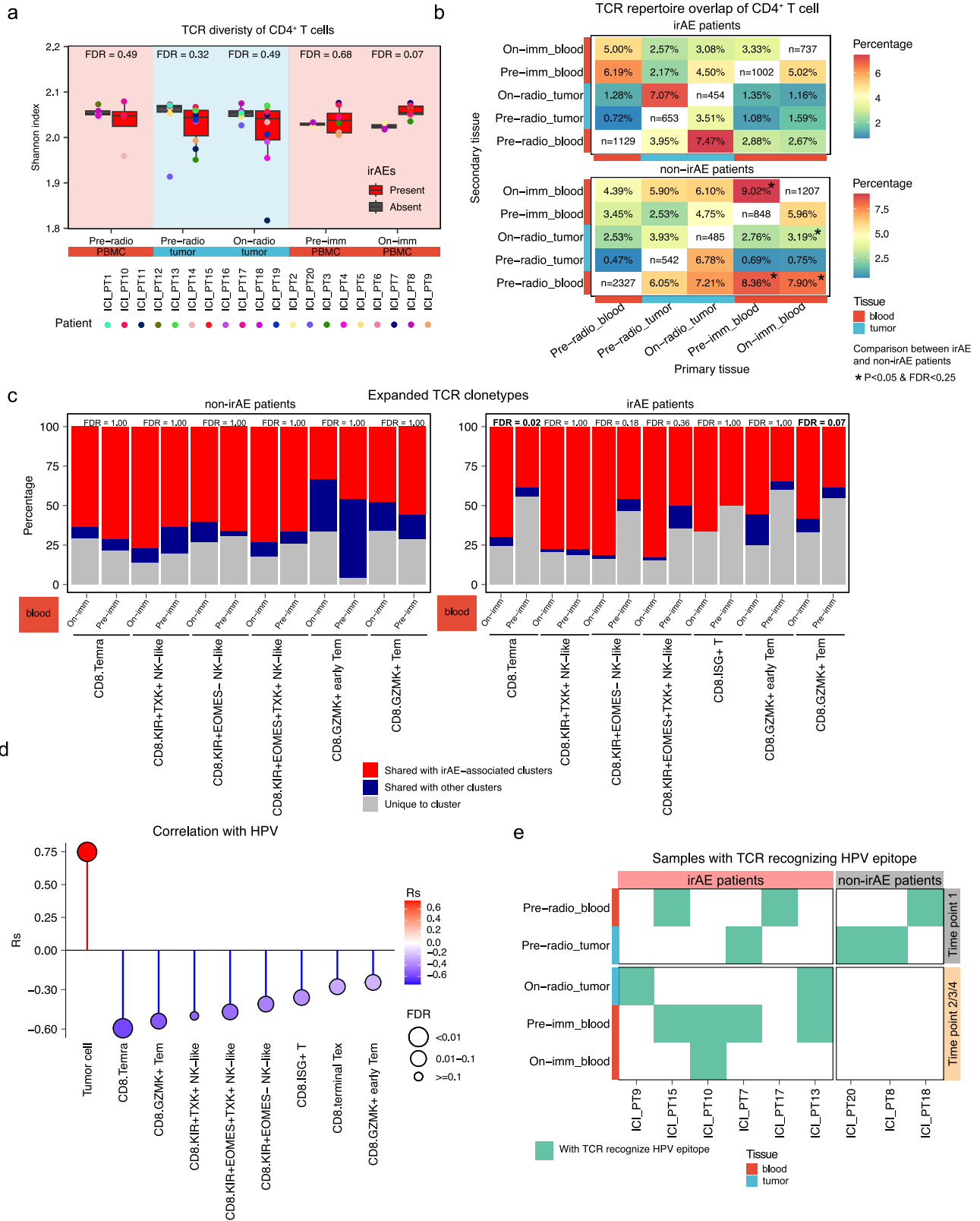
Using TCR sequences (clonotypes) as molecular barcodes enables mapping of T cell clonal expansions and diversity resulting from chemoradiotherapy and ICI administration. We obtained 161,711 cells with both α and β TCR chains from the CD4⁺ and CD8⁺ subsets, including 79,811 with unique and 81,900 with repeated TCRs (Supplementary Fig. 6a–d). The repeated TCRs are indicative of clonal expansion. TCR diversity, assessed by the Shannon indexes, did not show a significant correlation with disease relapse for either CD8⁺ or CD4⁺ T cells (Supplementary Fig. 6e), suggesting limited disease relapse-related differences in TCR within current dataset. Therefore, subsequent analysis focused on irAEs. Notably, CD4⁺ T cells from irAE patients displayed significantly higher TCR diversity in on-imm blood samples compared to non-irAE patients, with no significant differences observed at the other time points (Fig. 3a). Conversely, the TCR diversity of CD8⁺ T cells shows no significant association with irAEs (Supplementary Fig. 6f). These findings suggest that irAE patients have a more polyclonal repertoire of CD4⁺ T cells but not CD8⁺ T cells. Clonal expansions were observed in both CD4⁺ and CD8⁺ T cells across blood and tumor samples, with divergent percentages of overlapping clones (0.47 to 9.02% and 4.42 to 41.48% for CD4⁺ and CD8⁺ T cells, respectively; Fig. 3b, Supplementary Fig. 6g). In alignment with the decreased TCR diversity, overlapping CD4⁺ T cell clonotypes between pre-imm and on-imm blood samples, as well as between pre-imm and pre-radio

blood samples, were significantly lower in irAE patients compared to non-irAE patients. In contrast, no significant differences were observed in overlapping CD8⁺ T cells clonotypes (Fig. 3b, Supplementary Fig. 6g).

Given the high fractions of expanded CD8⁺ T cells, we next investigated whether TCRs from clonally expanded CD8⁺ T cells were shared between subclusters that were increased in irAE patients. In irAE patients, CD8⁺ Temra and CD8⁺ GZMK⁺ Tem exhibited significantly higher percentages of clonally expanded TCRs shared with CD8⁺ Temra after ICI treatment (Fig. 3c). However, no significant changes were observed in CD4⁺ T cells (Supplementary Fig. 6h). We next asked whether the expanded TCRs recognized specific epitopes. As HPV abundance was negatively associated with the proportions of most CD8⁺ T subclusters (Fig. 3d, Supplementary Fig. 6i), we explored whether CD8⁺ T cells harbored TCRs recognizing HPV epitopes. Interestingly, six irAE patients had HPV-specific TCRs in samples collected at all time points, whereas only three non-irAE patients had HPV-specific TCRs in samples collected at baseline (Fig. 3e). These results suggest that expanded HPV-recognizing T cells may play a role in irAE development. In summary, our findings indicate that CD4⁺ T cells displayed greater TCR diversity and fewer overlapping clonotypes across time points and tissues obtained from irAE patients. Meanwhile, CD8⁺ T cells exhibited a higher proportion of TCRs shared with irAE-associated subclusters after ICI treatment, suggesting distinct roles of CD4⁺ and CD8⁺ T cells in irAEs.

Lower BCR clonality after chemoradiotherapy in irAE patients

Previous studies have reported associations between B cells and irAEs, such as clonality²⁴ and abundance of specific B cells²⁵. Based on this, we investigated the abundance and immunoglobulin (Ig) repertoires of B-lineage cells in patients with and without irAEs. Unsupervised clustering identified five distinct B cell subclusters: two naive B cell subclusters (tubulin beta 1class VI-positive naive B cell, Bn-TUBB1, and T-cell leukemia/lymphoma protein 1A-positive naive B cell, Bn-TCL1A), one memory B cell subcluster, one transitional B subcluster and one plasma cell (PC) subcluster (Fig. 4a, Supplementary Fig. 7a). Unlike T cells, B cell subclusters exhibited fewer significant frequency changes across patient groups. Proportions of Bn-TCL1A significantly increased following ICI treatment, especially in irAE patients (Fig. 4b, Supplementary Fig. 7b, c). The percentages of the top BCR clonotypes at on-radio, pre-imm and on-imm time points increased in non-irAE patient but not in irAE patients, indicating a more even distribution of clonal size in irAE patients following chemoradiotherapy and ICI treatment (Fig. 4c). Next, we analyzed Ig heavy-chain isotypes based on the Ig constant region sequences. As expected, naive B and transitional B cell subclusters were dominated by IgM isotypes, memory B cell exhibited a variety of isotypes and the PCs were highly enriched in IgG and IgA isotypes (Supplementary Fig. 8a, b). Comparisons between irAE and non-irAE patients revealed significant increases in the proportion of IgA⁺ and IgG⁺ B cells in both blood and tumor samples following chemoradiotherapy, accompanied by a decrease in IgM⁺ B cells (Fig. 4d). To further investigate the functional status of B cells, we analyzed somatic hypermutation (SHM), where high SHM levels



indicate strong antigen-specific affinity²⁶. As expected, most naive B cells exhibited no SHM, memory and transitional B cells displayed moderate SHM frequencies, and PCs had higher SHM frequencies (Fig. 4e). Consistent with the observed isotype dynamics, SHM frequencies of IgA⁺ and IgG⁺ B cells increased in both blood and tumor samples from irAE patients following chemoradiotherapy (Fig. 4f). Overall, these findings indicate that chemoradiotherapy reduced B cell

clonality while increasing the abundance and SHM frequencies of IgA⁺ and IgG⁺ B cells in blood and tumor samples of patients with irAEs. This dynamic may play a role in the formation of autoantibodies, ultimately leading to irAEs.

Analyses of disease relapse status showed that memory B cells were more abundant in pre-radio tumor samples from patient without disease relapse (Supplementary Fig. 7d). However, the distribution of

Fig. 3 | irAE-related changes in T cell clonotypes. **a** Boxplots showing TCR diversity in CD4⁺ T cells, measured by the Shannon indexes. In all boxplots, the center line represents the median, the bounds of the box indicate the 25th and 75th percentiles, and the whiskers extend to 1.5 × IQR. Each dot represents an individual patient. *P* values were determined by using the two-sided Wilcoxon rank-sum test and *FDR* values were calculated by the Benjamini–Hochberg method. **b** Overlap of the CD4⁺ T cell TCR repertoire across samples collected at different time points for patients with and without irAEs. A TCR clonotype was defined as shared when it was detected and shared in T cells from at least two samples from any patient. The heat map shows the percentage of TCR clonotypes from samples at one time point that overlapped with those from another time point. The primary samples are plotted on the x-axis, and the secondary samples are plotted on the y-axis. Statistical comparisons were performed between irAE and non-irAE patients for the proportion of overlapping TCR clonotypes identified between specific sample pairs. For clonotype overlaps between pre-imm PBMC and on-imm PBMC, *P* = 0.014 and *FDR* = 0.23; for overlaps between on-imm PBMC and on-radio tumor, *P* = 0.046 and *FDR* = 0.23; for overlaps between pre-imm PBMC and pre-radio tumor, *P* = 0.034 and *FDR* = 0.23; and for overlaps between on-imm PBMC and pre-radio tumor, *P* = 0.034 and *FDR* = 0.23. *P* values were determined by using the two-sided Wilcoxon rank-sum test, and *FDR* values were calculated by the Benjamini–Hochberg method. Sample sizes for a–b: pre-radio tumor (irAEs absent: *n* = 6; irAEs present: *n* = 10); on-radio tumor (irAEs absent: *n* = 5; irAEs present: *n* = 10); pre-radio PBMC (irAEs absent: *n* = 5; irAEs present: *n* = 4); pre-imm and on-imm PBMC (irAE absent:

n = 3; irAE present: *n* = 8). **c** Analysis of the expanded TCR clonotypes shared across clusters in patients with and without irAEs. For each sample and subcluster, the number of expanded TCR clonotypes from the CD8⁺ T cells was calculated and summed up for the indicated irAE groups. A comparison was made between blood samples collected obtained at the pre-imm and on-imm time points. *P* values were determined by using the two-sided Wilcoxon rank-sum test and *FDR* values were calculated by the Benjamini–Hochberg method. The expanded clonotypes shared with an irAE-associated cluster (CD8.Temra) are shown in red, while those shared with other CD8⁺ T cell subclusters are shown in blue. Sample sizes for **c**: pre-radio tumor (irAEs absent: *n* = 6; irAEs present: *n* = 10); on-radio tumor (irAEs absent: *n* = 5; irAEs present: *n* = 10); pre-radio PBMC (irAEs absent: *n* = 5; irAEs present: *n* = 3); pre-imm and on-imm PBMC (irAE absent: *n* = 3; irAE present: *n* = 8). **d** Spearman correlation between HPV abundance (determined by count per million reads) and the percentage of CD8⁺ T cell subcluster. The red and blue lollipops indicate positive and negative correlations, respectively. The shade of the dot represents the correlation coefficient (*R*s). The size of each dot corresponds to the significance level, with the large, medium, and small dots representing *FDR* < 0.01, 0.01–0.1, and ≥ 0.1, respectively. The sample size for the correlation analysis is 51. **e**, Samples with TCR clonotypes recognizing the HPV epitope. An assessment of the publicly available databases revealed 13 samples with HPV-specific TCRs. All sample sizes reported are biological replicates, with the patient serving as the independent unit of study. TCR T cell receptor, HPV human papillomavirus. Source data are provided as a Source Data file.

BCR clonotypes were similar between patients with and without relapse (Supplementary Fig. 8c). IgA⁺ and IgM⁺ B cells exhibited inverse changes in tumor samples from non-relapsed patients after chemoradiotherapy (Supplementary Fig. 8d). SHM frequencies remained largely unchanged in relapsed and non-relapsed patients (Supplementary Fig. 8e). Collectively, although memory B cell increased in non-relapsed patients, clonotypes and SHM did not change significantly.

Divergent myeloid cell reprogramming induced by a combination of chemoradiation and ICI therapies

Building on our previous study suggesting the distinct roles of macrophages in the TME related to irAEs and treatment efficacies in lung squamous cell carcinoma¹⁵, we further investigated myeloid cells. We identified five major myeloid cell populations, further categorized into 14 unique subclusters based on published signatures and cell markers²⁷ (Fig. 5a–d, Supplementary Figs. 9, 10). The frequencies of myeloid cell subclusters differed significantly between irAE and non-irAE groups across various time points. Notably, two CX3CR1-positive macrophage subclusters displayed contrasting trends at the pre-imm time point. Tumor necrosis factor receptor superfamily member 8-positive macrophage (Macro.TNFRSF8), characterized by a high inflammatory response signature, significantly increased in irAE patients, whereas G protein-coupled receptor kinase 2-positive macrophages (Macro.GRK2) decreased (Fig. 5e, Supplementary Fig. 9d). Among monocyte subclusters, folate receptor 3-positive (Mono.FOLR3) monocyte, which expressed the highest inflammatory response signature, were enriched in pre-imm and on-imm blood samples from irAE patients (Fig. 5e). Conversely, four monocyte subclusters with lower expression of inflammatory response, C-C motif chemokine ligand 5-positive (Mono.CCL5), lymphocyte antigen 96-positive (Mono.LY96), serpin family B member 1-positive (Mono.SERPIN1) and the major histocompatibility complex class II-positive (Mono.MHCII) monocyte, were reduced in irAE patients both before and after ICI treatment (Fig. 5e).

When analyzing gene signature expression related to myeloid cell functions, most macrophage and monocyte subclusters in irAE patients' pre-imm blood samples exhibited significantly higher levels of inflammation- and antigen presentation-related pathways (Fig. 5f). In contrast, at the on-imm time point, only subclusters with increased frequencies in irAE patients maintained elevated inflammation- and angiogenesis-related signature expression. However, phagocytosis-related signatures were consistently higher in most myeloid

subclusters in irAE patients (Fig. 5f). Other gene signatures expressed varied across samples collected at pre-radio and on-radio time points (Supplementary Fig. 10h). In summary, chemoradiotherapy combined with ICI treatment increased the frequencies of monocyte subclusters with heightened inflammatory responses and enhanced the phagocytic capabilities of myeloid cells in patients with irAEs. These findings suggest that the inflammatory status of monocytes may contribute to irAE development.

Two macrophage subclusters with higher phagocytosis-related signature expression, tryptase alpha/beta 1-positive (Macro.TPSAB1) and triggering receptor expressed on myeloid cells 2-positive (Macro.TREM2) macrophage, were enriched in baseline tumor samples from non-relapsed patients, but interleukin 1 beta-positive macrophage (Macro.IL1B), with higher expression of inflammatory response, was reduced (Supplementary Figs. 9d, 10g). Mono.CCL15 and Mono.LY96 decreased in pre-imm blood samples of non-relapsed patients (Supplementary Fig. 10g). Myeloid cell function-related analysis found that in non-relapsed patients, angiogenesis-related signatures were elevated in most macrophage and monocyte subclusters both before and after chemoradiotherapy and ICI treatment (Supplementary Fig. 10i). Altogether, these results suggested that angiogenic potential of myeloid cells may positively influence treatment efficacy.

Colocalization of cells associated with irAEs within the TME in response to chemoradiotherapy

We performed Spearman correlation analysis on the frequencies of cell subclusters in pre-radio and on-radio tumors in order to explore the relationships between cellular dynamics related to irAEs. In the on-radio tumor samples, we found strong correlations among cell subclusters that showed significant changes in samples collected after the on-radio time point, such as in the CD4.Temra, CD8.Temra and CD4.ADSL⁺ Tn (Supplementary Fig. 11a). However, the correlation coefficients were similar in the pre-radio tumor samples (Supplementary Fig. 11b). In order to further investigate the spatial organization of irAE-associated cells within the TME of irAE patients, we performed spatial transcriptomic analysis using the 10× Genomics Xenium platform, which profiled 380 genes. Subcellular resolution of the Xenium data²⁸ allowed us to capture and visualize the spatial localization of different cell subclusters within the TME (Supplementary Fig. 12). We observed a significant decrease in malignant cells and an increase in fibroblasts in on-radio tumors when compared to the pre-radio tumors, and this aligned with the results from the typical

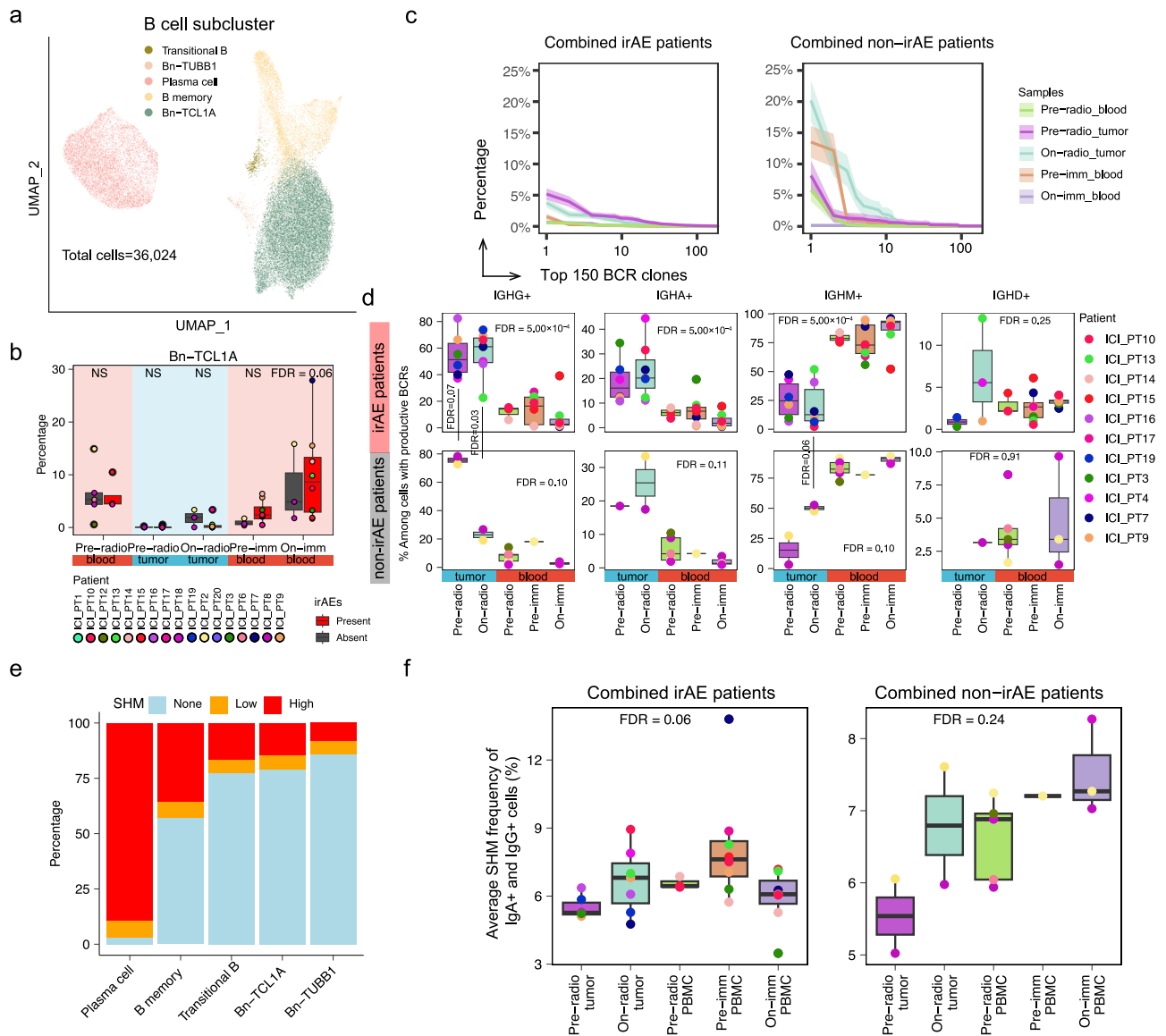


Fig. 4 | Characterization of B cell isotypes and clonotypes in response to chemoradiotherapy in irAE versus non-irAE patients. **a** A uniform manifold approximation and projection plot of B cells captured across all samples, colored by B cell subclusters. **b** Boxplots comparing the cellular fractions of B cell subclusters within different samples from irAE and non-irAE patients. Each dot represents an individual patient. The two-sided Dirichlet-multinomial regression model was used to determine the *P* values. Sample sizes: pre-radio tumor (irAEs absent: *n* = 5; irAEs present: *n* = 9); on-radio tumor (irAEs absent: *n* = 3; irAEs present: *n* = 7); pre-radio PBMC (irAEs absent: *n* = 5; irAEs present: *n* = 4); pre-imm and on-imm PBMC (irAE absent: *n* = 3; irAE present: *n* = 8). **c** Clone size distribution of the BCR repertoires across samples collected at different time points in irAE (left) or non-irAE patients (right). The top 150 BCR clonotypes are displayed. The solid line represents the mean abundance curve generated using the estimateAbundance function from the alakazam package. Shaded areas indicate the 95% confidence interval across biological replicates. **d** Box plots comparing the cellular fractions of IgA⁺, IgM⁺, IgG⁺, and IgD⁺ cells within different samples grouped by irAE statuses. Each dot represents an individual patient. *P* values were determined by the two-

sided Kruskal-Wallis test. The lines across two boxplots in panel **d** indicates the comparisons between irAE and non-irAE patients. **e** The distribution of SHM frequencies across the B cell subclusters. **f** Box plot showing average SHM frequencies in samples collected at different time points. Each dot represents an individual patient. Colors correspond to patient IDs as shown in panel **d**. *P* values were calculated by the two-sided Kruskal-Wallis test. Sample sizes for **c**, **d** and **f**: pre-radio tumor (irAEs absent: *n* = 2; irAEs present: *n* = 6); on-radio tumor (irAEs absent: *n* = 2; irAEs present: *n* = 7); pre-radio PBMC (irAEs absent: *n* = 5; irAEs present: *n* = 3); pre-imm PBMC (irAE absent: *n* = 1; irAE present: *n* = 8); on-imm PBMC (irAE absent: *n* = 3; irAE present: *n* = 7). In all boxplots in **b**, **d** and **f**, the center line represents the median, the bounds of the box indicate the 25th and 75th percentiles, and the whiskers extend to 1.5 × IQR. All *FDR* values were calculated by the Benjamini-Hochberg method. All sample sizes reported are biological replicates, with the patient serving as the independent unit of study. Bn naive B cell, BCR B cell receptors, SHM somatic hypermutation. Source data are provided as a Source Data file.

radiation-induced changes observed in the tumors²⁹ (Fig. 6). Notably, the immune cells associated with irAEs, including B cell, CD4⁺ T cells and CD8⁺ T cells, were aggregated in the on-radio tumors of irAE patients, whereas this aggregation was absent in the on-radio tumors obtained from non-irAE patients (Fig. 6a, c). To corroborate this observation, we performed the multiplex immunofluorescence (mIF) staining using

markers for total T cells (CD3), CD4⁺ T cells (CD4), CD8⁺ T cells (CD8), B cells (CD20), macrophages (CD68) and malignant cells (pan-CK) (Fig. 6e). We then analyzed the spatial distribution of immune cells and measured the intercellular distances between B cells, CD4⁺ T cells and CD8⁺ T cells using HALO software (Fig. 6f). The results indicated an increased proximity between immune cells in irAE patients compared to

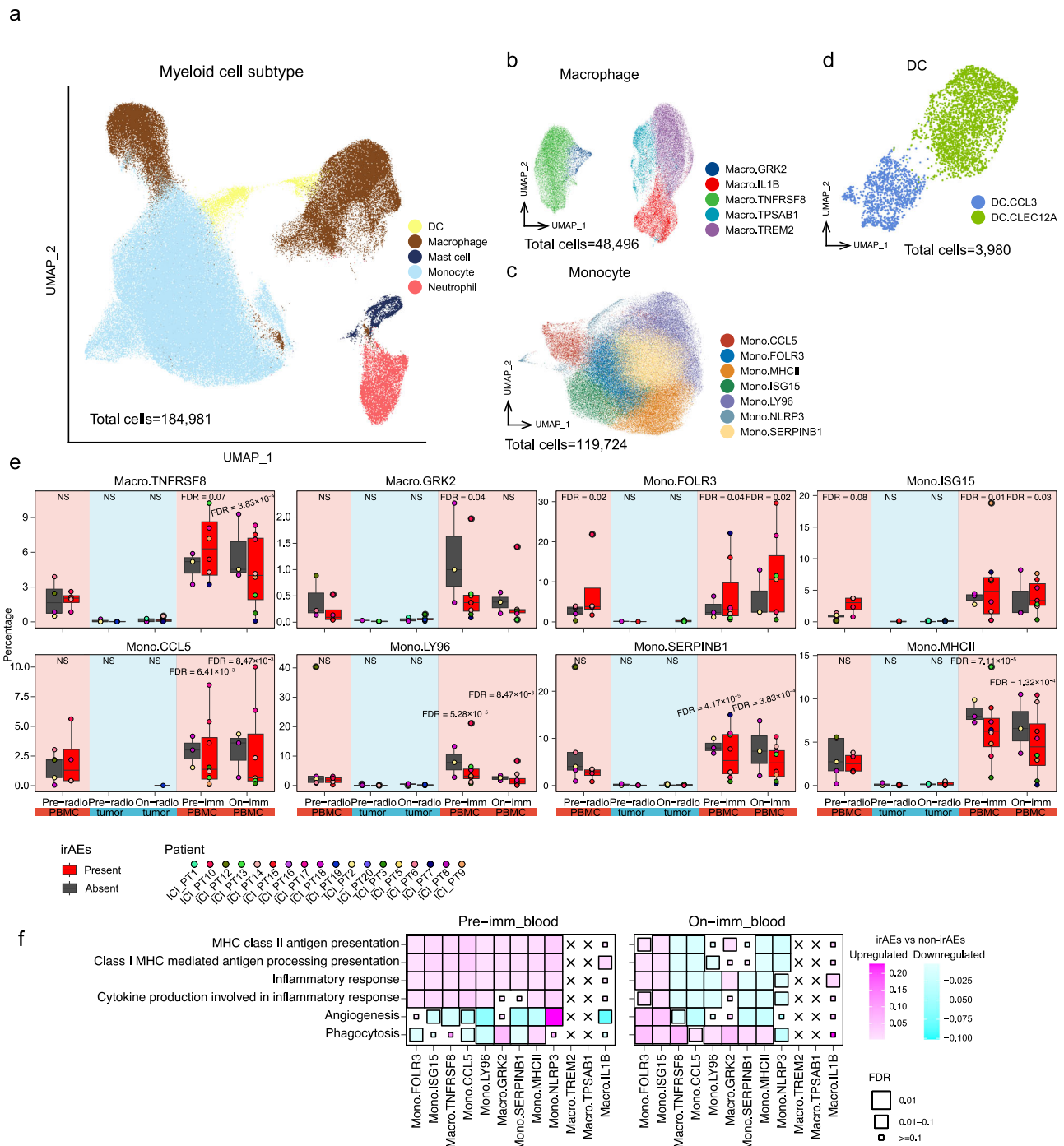
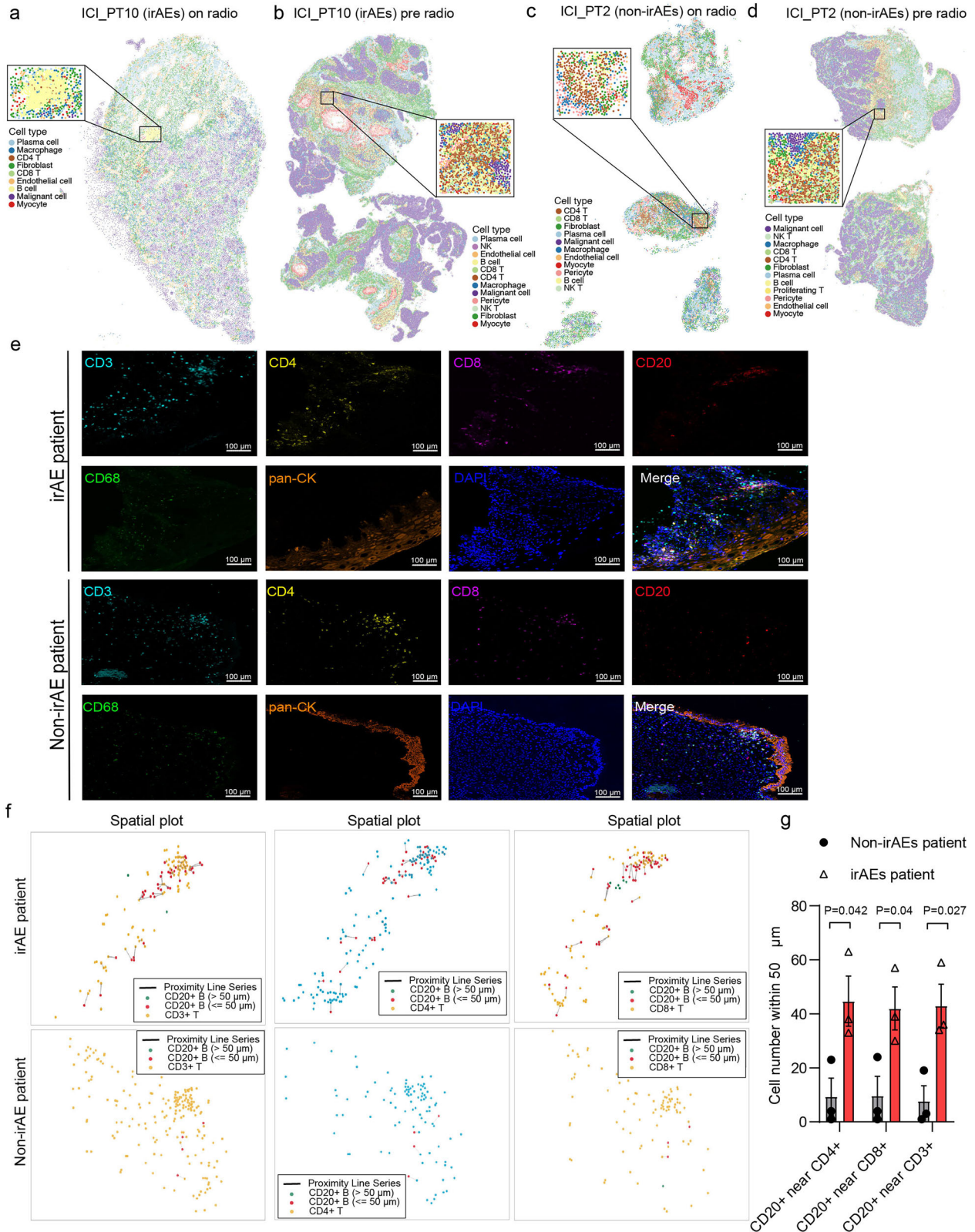


Fig. 5 | Inflammatory and phagocytosis gene expression in myeloid cells are associated with irAEs. **a–d** A uniform manifold approximation and projection plot of all myeloid cells (**a**), including macrophages (**b**), monocytes (**c**) and DCs (**d**), captured from all the samples and colored by their myeloid cell subclusters. **e** Boxplots comparing the cellular fractions of myeloid cell subclusters within different samples from irAE and non-irAE patients. *P* values were calculated using the two-sided Dirichlet-multinomial regression model, and *FDR* values were calculated using the Benjamini–Hochberg method. In all boxplots, the center line represents the median, the bounds of the box indicate the 25th and 75th percentiles, and the whiskers extend to $1.5 \times IQR$. **f**, A comparison of the myeloid cell function-related signaling pathway scores in irAE versus non-irAE patients. The left and right panels show analyses in the pre-imm blood and on-imm blood samples, respectively. *P*

values were determined by the two-sided Wilcoxon rank-sum test and compute *FDR* values using the Benjamini–Hochberg method. Magenta and cyan indicate upregulation in the irAE and non-irAE patients, respectively. The shade of each square indicates the difference in expression levels between groups, and the size of each square corresponds to the significance level, with the large, medium, and small squares representing $FDR < 0.01$, $0.01-0.1$, and ≥ 0.1 , respectively. \times indicates the data was not applicable. Sample sizes: pre-radio and on-radio tumor (irAEs absent: $n = 6$; irAEs present: $n = 12$); pre-radio PBMC (irAEs absent: $n = 5$; irAEs present: $n = 4$); pre-imm and on-imm PBMC (irAE absent: $n = 3$; irAE present: $n = 8$). All sample sizes reported are biological replicates, with the patient serving as the independent unit of study. DC dendritic cells, Macro macrophage, Mono monocyte. Source data are provided as a Source Data file.



non-irAE patients. Furthermore, we quantified the number of CD20⁺ in close proximity to CD4⁺, CD8⁺ and CD3⁺ cells (distance ≤ 50 μm) and observed significantly more such cell proximities in tumor samples from irAE patients (Fig. 6g), suggesting increased immune cell aggregations and interactions in the TME. This spatial organization of key immune cells may trigger subsequent immune responses, potentially contributing to the development of irAEs.

The predictive model based on the molecular characteristics of immune cells during ICI treatment showed superior performance in predicting irAEs

Given that the majority of irAE-related alterations were observed in peripheral blood samples and that different cell types exhibited divergent molecular features, we aimed to develop predictive models for irAEs using the scRNA-seq data from pre-radio, pre-imm

Fig. 6 | Xenium data reveal irAE-associated immune cell localization in the TME after chemoradiotherapy. **a** Spatial distribution of cell types in the on-radio tumor samples of ICI_PT10. A magnified view highlights a representative area containing the immune cells aggregates. **b** A spatial distribution of the cell types in the pre-radio tumor samples of ICI_PT10. A magnified view highlights a representative area containing immune cells adjacent to each other. **c, d** A spatial distribution of the cell types in the on-radio tumor (**c**) and pre-radio tumor (**d**) samples of ICI_PT2. Zoom-in views of representative areas are shown for areas containing the immune cells. The color of each dot denotes the cell type in each panel. **e** Representative images of multiplex immunofluorescence (mIF) staining of tumor samples after chemoradiotherapy from patients with or without irAEs. The markers visualized include CD3 (cyan, T cell marker), CD4 (yellow, CD4⁺ T cell marker), CD8 (purple,

CD8⁺ T cell marker), CD20 (red, B cell marker), CD68 (blue, macrophage marker), pan-CK (orange, epithelial marker) and DAPI (blue, nuclei). Scale bar = 100 μ m. The experiment was conducted with three independent biological replicates ($n = 3$ patients for each group). **f** Spatial plots illustrating intercellular distances of designated cells within tumor tissues, analyzed using HALO software (V3.6.4134.263). Each dot represents an individual cell, with lines connecting cells that are proximal to each other (distance ≤ 50 μ m). **g** Bar plot showing the number of designated cells with distance ≤ 50 μ m for each patient with ($n = 3$) or without irAEs ($n = 3$). Each dot represents an individual patient. Data are presented as mean values \pm SEM and statistical comparisons between the two groups were performed using an unpaired two-sided Student's *t*-test. Source data are provided as a Source Data file.

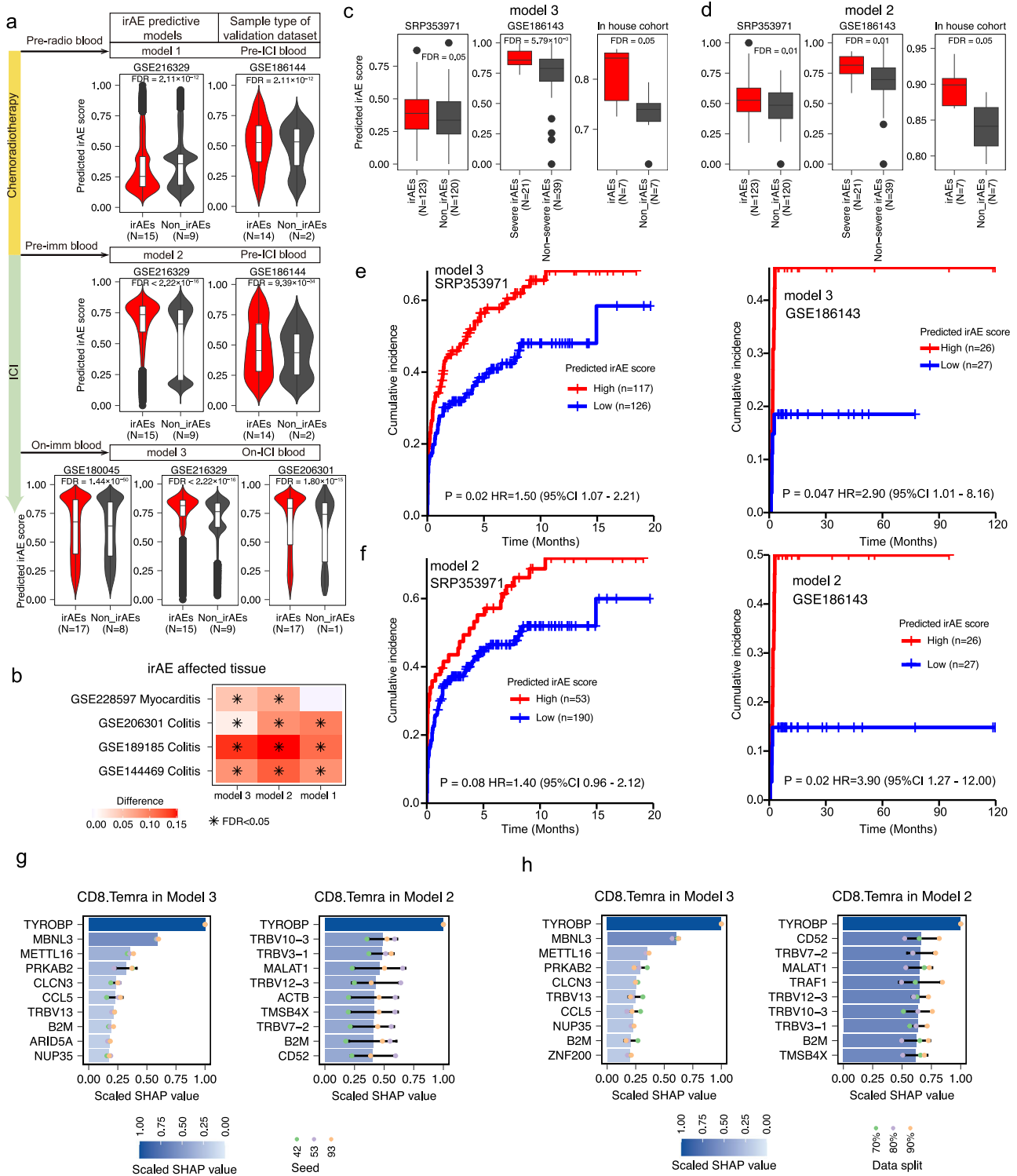
and on-imm blood samples. Initially, we employed a method that based on differentially expressed genes (DEG) in cell subclusters enriched in irAE patients (DEG-based irAE signatures, Methods). However, we found that none of the signatures consistently showed increased levels across the validation cohort, regardless of whether scRNA-seq or bulk RNA sequencing (RNA-seq) datasets were used (Supplementary Fig. 13). Cumulative incidence analysis suggested that low irAE scores indicated higher irAE risk or failed to achieve statistical significance (Supplementary Fig. 13f, g). These results suggest that DEG-based irAE signatures may have limitations in capturing the full molecular complexity of irAEs. Therefore, we explored deep learning strategies to enhance irAE prediction. We developed models based on the scRNA-seq data obtained from pre-radio, pre-imm, and on-imm blood samples, referred to as models 1, 2, and 3, respectively. Immune cell subclusters at different time points were classified into irAE-related, significantly enriched in irAE patients, and non-irAE-related groups, significantly enriched in non-irAE patients. Genes were ranked for both groups based on their expressions in individual cells, and 'cell sentences' were formulated using gene names and their ranks. These 'cell sentences' were then input into the bidirectional encoder representation from transformer (BERT) language model, forming the foundation of our irAE prediction models (Methods). We performed subsampling analyses to assess the robustness of these models. We randomly partitioned these labeled cells into training sets consisting of 70%, 80% and 90% of the data, with corresponding testing sets of 30%, 20% and 10%, respectively. Model performance was evaluated using area under the receiver operating characteristic curve (AUC), Precision-Recall-AUC (PRAUC), accuracy (ACC) and F1 score, based on the testing data. Across all four metrics, the models demonstrated stable performance (Supplementary Fig. 14a), indicating robustness.

To ensure these models accurately captured representative features of irAEs at various treatment stages, we validated them using samples (Supplementary Table 2) collected at time points comparable to those used for corresponding model development. Specifically, we used blood samples collected prior to ICI (pre-ICI blood) to validate models 1 and 2, blood samples collected during ICI treatment (on-ICI blood) to validate model 3 and the irAE-affected tissue, where almost all irAE-related immune responses occurred, to validate all models. In pre-ICI blood from two independent datasets^{30,31}, the model 2, not model 1, consistently predicted higher irAE scores in patients who developed irAEs (Fig. 7a, Supplementary Fig. 14b). In three additional datasets^{31–33} with on-ICI blood samples, the predicted irAE scores from the model 3 were also significantly higher in irAE patients (Fig. 7a, Supplementary Fig. 14b). Interestingly, in four datasets collected tissues affected by irAEs^{33–36}, such as colitis and myocarditis, both the model 2 and 3 consistently predicted higher irAE scores in irAE-affected tissues compared to control samples (Fig. 7b, Supplementary Fig. 14c). These findings demonstrate that the deep learning algorithm effectively captured irAE-related molecular features across different treatment time point.

To broaden the applicability of irAE models and identify the most effective model, we applied them to bulk RNA-seq data obtained from blood samples collected before ICI administration. The time points before treatment are critical for irAE prediction, as these stages provide valuable opportunities to implement interventions or alternative treatment designs to benefit patients. Our analysis included an in-house cohort ($n = 14$, Supplementary Table 3) with pre-imm bloods samples from patients treated with combined radiotherapy and ICI, as well as two publicly available datasets^{30,37} with larger sample sizes ($n = 53$ and $n = 243$, respectively). In these three datasets, the model 2 and 3 predicted significantly higher irAE scores in pre-ICI blood samples from irAE patients, with the model 3 showing the most pronounced differences (Fig. 7c, d). Cumulative incidence analyses of the time to any irAE onset³⁷ or severe irAE onset³⁰, based on the follow-up duration after ICI administration, revealed that patients with higher predicted irAE scores from the model 3 had a significantly shorter time to irAE onset, with hazard ratios of 1.50 and 2.90, respectively, whereas the model 2 did not show the same predictive potential (Fig. 7e, f). These results indicate the predictive potential of the model 3 in forecasting patient susceptibility to irAEs.

Interpreting model predictions via cell subcluster analysis

To better comprehend these models, we performed ablation analyses to quantify the contributions of distinct cell subclusters to model performance. We ablated a single cell subcluster by excluding its data from the input and training the model on the remaining data. The ablated model was then applied to the full dataset to generate predictions, from which we computed AUC, PRAUC, ACC, and F1 score. Consistent with previous observations, in Model 3, ablations of CD8.Temra, Mono.FOLR3 and NK T.FGFBP2 resulted in the top declines across four metrics (Supplementary Fig. 14d–g). In Model 2, removal of Macro.TNFRSF8, CD8.Temra and NK T.FGFBP2 caused the top reductions across these metrics (Supplementary Fig. 14d–g). Next, we exported the top 100 genes (Supplementary Table 4, determined by key word extraction approach for BERT model³⁸) from model 2 and model 3 and found that these genes were enriched in pathways associated with activated immune responses, such as T cell activation and cytokine production (Supplementary Fig. 14h). Third, we computed SHAP values, an established interpretability practice for models employing language embeddings³⁹ and deep neural network learning models³⁷. To ensure robustness, we used three different random seeds (Fig. 7g & Supplementary Fig. 15a) and three data split proportions (Fig. 7h & Supplementary Fig. 15b). In model 3, genes such as tyrosine kinase binding protein (TYROBP), C-C motif chemokine ligand 5 (CCL5), and TCR variable genes (e.g. TRBV13) consistently in the top 10 within CD8.Temra, CD8.GZMK⁺ Tem and CD4.Temra (Fig. 7g, h & Supplementary Fig. 15). TYROBP is expressed by lymphocytes and mediates signaling for multiple cell-surface receptors⁴⁰. In addition, TYROBP is highly expressed in certain CD8 T cell subclusters, which are reported to be activated in autoimmune disease^{41,42}. TYROBP has also been linked to M2 tumor-associated macrophage polarization in



pancreatic cancer⁴³. CCL5, which induces leukocyte activation and chemotaxis⁴⁴, has been reported to be upregulated in myocarditis-associated CD8 T cells⁴⁵, colitis-associated CD4 T cells⁴⁶ and severe irAE-associated CD4 T cells³⁰. In model 2, TYROBP also rank top in some cell clusters, such as CD8.Temra and CD8.GZMK+ Tem (Fig. 7g, h and Supplementary Fig. 15). Taken together, these results suggest that CD8.Temra likely play a meaningful role in underlying irAE mechanisms.

Therefore, we employed FACS to quantify CD8.Temra within the CD8⁺ T cell population of PBMCs from two independent ICI patient cohorts (Supplementary Table 5). The first cohort consisted of patients

from our scRNA-seq dataset, from whom we thawed freshly frozen PBMC samples. The second cohort included patients with various cancer types undergoing ICI therapy, from whom fresh blood samples were collected. CD8.Temra cells were defined as CD45RA⁺ CCR7⁻ within the CD45⁺ CD3⁺ CD8⁺ T cell gate (Fig. 8a). Our analysis revealed significantly higher proportions of CD8.Temra in the irAE patients across both cohorts (Fig. 8b-e). These results confirm the association between irAEs and increased CD8.Temra cells are consistent with our cell-type and gene-level importance analysis. To facilitate use of irAE models, we have developed a user-friendly data portal, irAE-atlas (<https://iraeatlas.ipm-gba.org.cn>). This website not only allows users to

Fig. 7 | Validation of the irAE predictive models using independent scRNA-seq and bulk RNA-seq datasets. **a** Violin plots displaying predicted irAE scores in blood samples collected before (pre-ICI blood) or during ICI treatment (on-ICI blood) from irAE or non-irAE patients. Scores were calculated using the irAE models developed from scRNA-seq data of blood samples collected at three time points: pre-radio (before chemoradiotherapy and ICI, referred to as model 1), pre-imm (after chemoradiotherapy but before ICI, referred to as model 2), and on-imm (after chemoradiotherapy and during ICI treatment, referred to as model 3). *P* values were determined using two-sided Wilcoxon rank-sum test, and *FDR* values were calculated using the Benjamini–Hochberg method. **b** Heatmap showing changes in predicted irAE scores (calculated by models 1–3) between irAE-affected tissues and normal controls. Red and blue dots indicate increased and decreased scores in irAE-affected tissue, respectively. *P* values were determined using the two-sided Wilcoxon rank-sum test, and *FDR* values were calculated using the Benjamini–Hochberg method. * represents *FDR* < 0.05. Sample sizes: GSE206301 (colitis: *n* = 13; control: *n* = 5); GSE144469 (Colitis: *n* = 6; control: *n* = 5); GSE189185 (Colitis: *n* = 4; control: *n* = 3); GSE228597 (myocarditis: *n* = 15; control: *n* = 8). **c, d** Boxplots comparing predicted irAE scores from the model 3(c) and model 2 (d) using bulk RNA-seq data of pre-ICI blood samples from irAE and non-irAE patients

across publicly available datasets and an in-house cohort. For GSE186143 and in-house cohort, *P* values were determined using the two-sided Wilcoxon rank-sum test, while for SRP353971, the one-sided Wilcoxon rank-sum test was applied. *FDR* values were computed using the Benjamini–Hochberg method. In all boxplots from **a, c** and **d**, the center line represents the median, the bounds of the box indicate the 25th and 75th percentiles, and the whiskers extend to $1.5 \times \text{IQR}$. **e, f** Cumulative incidence analysis of severe irAE or any irAE onset. Patients were stratified based on predicted irAE scores from model 3 (**e**) and model 2 (**f**). *P* values were determined by log-rank test for cumulative incidence analysis. All sample sizes in **a–f** reported are biological replicates, with the patient serving as the independent unit of study. **g, h** SHAP values quantifying the contribution of each input gene to irAE predictions of models across three different random seeds (**g**) and three data split proportions (**h**). Raw SHAP values were averaged within each cell subcluster and scaled to 0–1 range. Data are presented as mean \pm SD. Bar plots show mean scaled SHAP values; error bars indicate standard deviation, and each dot represents one independent model run using a different random seed or split proportion. HR hazard ratio, SHAP shapley additive explanations. Source data are provided as a Source Data file.

calculate the predicted irAE scores for their own samples but also serves as one of the most comprehensive single-cell transcriptome data resources for irAEs.

Discussion

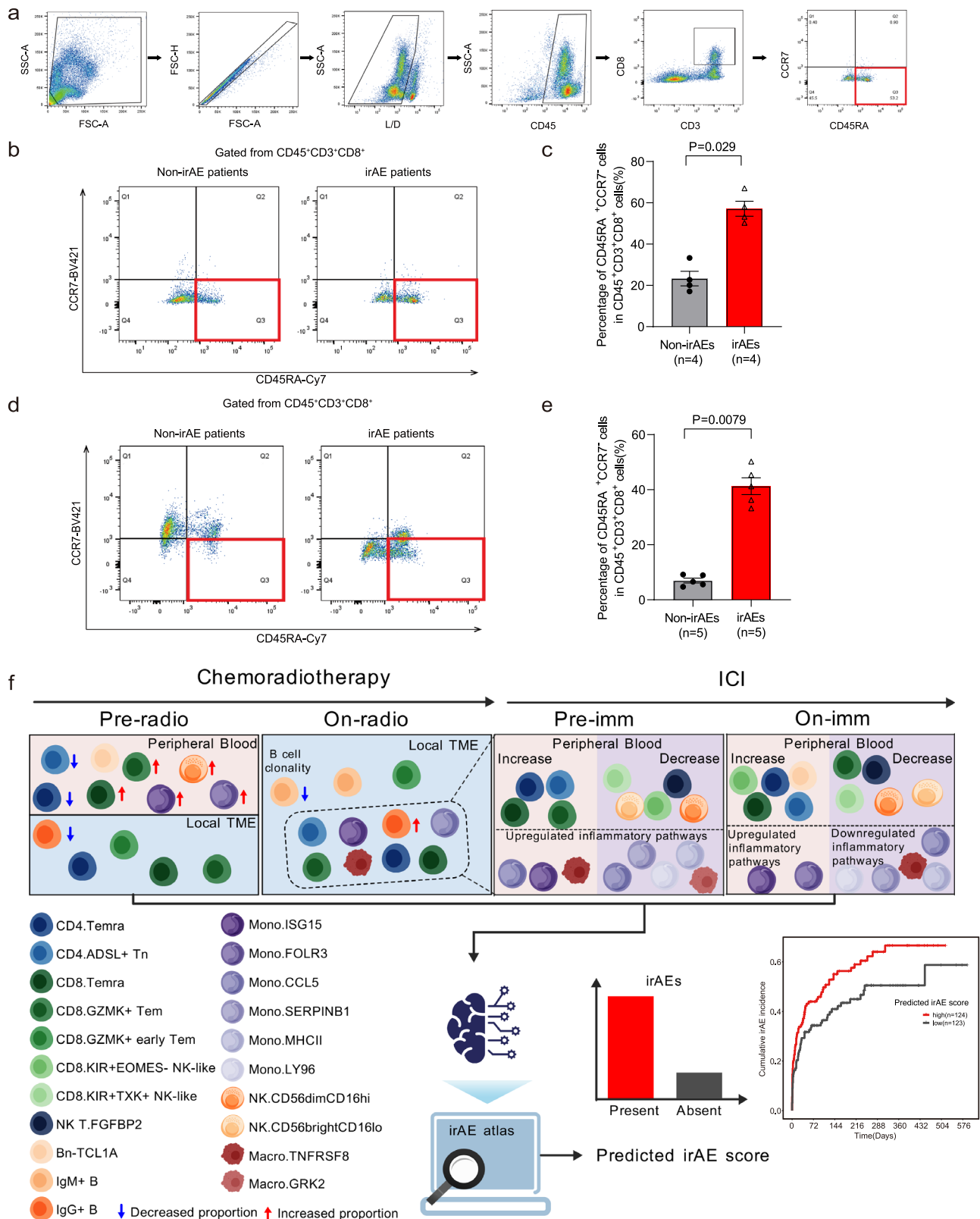
Radiotherapy has attracted significant interest as a potential combination therapy with immunotherapies for LACC. However, the high incidence of toxicities and inconsistent response rates observed in recent clinical trials underscore the need for a deeper understanding of immune cell states and compositions throughout the treatment process. Such insights are essential for identifying critical irAE-related molecular characteristics in response to chemoradiotherapy and ICI administration. In this study, we systematically charted immune cells in peripheral blood and tumor tissues throughout the treatment course using a combination of scRNA-seq, scVDJ-seq and the 10 \times Genomics Xenium platform. These advanced techniques enabled comprehensive analyses of single-cell transcriptomes, TCR and BCR repertoires, as well as cellular spatial localizations. Consistent with previous studies^{31,47,48}, we observed a marked increase of CD4⁺ Temra and CD8⁺ Temra cells in irAE patients following chemoradiation and ICI therapies, emphasizing the central role of T cells in the development of irAEs. Interestingly, the effects of chemoradiotherapy and ICIs on B cell and myeloid cells differed significantly in irAE patients. Chemoradiotherapy led to an increased percentage of B cells with IgA/IgG isotypes, reduced BCR clonality and elevated expression of inflammation-related pathways of myeloid cells in irAE patients. In contrast, addition of ICI treatment led to a higher proportion of Bn-TCL1A cells and an increase in phagocytosis-related gene signatures in myeloid cells (Fig. 8f). However, in non-relapse patient, baseline tumor samples revealed an enrichment of CD4⁺ T cell with cytotoxicity and exhaustion phenotypes, B cell with memory phenotype and macrophages with phagocytic phenotype. These findings suggest that chemoradiotherapy and ICIs reshape key immune cells in the systemic immune system, potentially leading to cell killing, autoantibody production and inflammatory response that contribute to the development of irAEs. For treatment efficacy, cell-killing-related functions of key immune cells in baseline TME could serve as indicators of improved patient prognosis. Overall, this study provides valuable insights into the complex immune responses elicited by combination treatment and offers possible directions for optimizing clinical benefit for cancer patients.

The high spatial resolution obtained from the Xenium data allowed us to visualize the localization of various cell subclusters within the TME, revealing the spatial aggregation of irAE-associated cells following chemoradiotherapy. This observation suggests a

connection between the local and systemic immune dynamics in response to chemoradiotherapy and ICIs, emphasizing the potential role of the TME in the development of irAEs. Our results also suggest that prior anti-cancer therapies, which alter the organization of the TME, may influence the occurrence of irAEs. We hypothesize that chemoradiotherapy promotes the accumulation of irAE-associated cells within the TME and reshapes their expression profiles. Subsequent ICI treatment may either exacerbate or mitigate these changes, likely depending on their cytotoxicity and the inflammatory response status of these cells, ultimately leading to irAEs. Further validation of this hypothesis through experiments in appropriate animal models is needed.

In addition, we examined the relationship between our irAE predictive models and the efficacy of ICI treatment. Interestingly, the predicted irAEs scores from three models showed no significant differences between patients who experienced durable clinical benefits and those who did not (Supplementary Fig. 16), suggesting unique predictive potential of these irAE predictive models. This finding indicates that the immune mechanisms driving irAEs may be mechanistically distinct from those responsible for anti-tumor efficacy, which is consistent with our observations in lung squamous cell carcinoma patients receiving ICIs¹⁵. We also investigated molecular features of endocrine irAEs (eirAEs) and the musculoskeletal and connective tissue systems (mcirAEs), which account for larger number of patients among organ-specific irAEs (Supplementary Table 6), and found that pre-radio PBMC samples from eirAE patients had higher proportions of CD8⁺GZMK⁺ Tem, CD8⁺ Temra and NK.CD56^{dim}CD16^{hi}, consistent with general irAE patterns (Supplementary Fig. 17a). However, pre-imm PBMC samples from patients with mcirAEs exhibited reduced percentages of specific monocyte subsets, including Mono.LY96, Mono.MHCII and Mono.SERPINB1 (Supplementary Fig. 17b), which showed relatively lower expression levels of the inflammatory response pathway. These findings indicate distinct immune profiles associated with different irAE types. CD8⁺ T cells, which cell cytotoxicity potentials, were correlate with eirAEs, while monocytes with lower inflammatory response levels demonstrated a negative association with mcirAEs (Supplementary Fig. 17c). Further validation is need to confirm model specificity and accuracy.

To improve the accessibility of our prediction methodologies, we have integrated our irAE predictive models into an easy-use online tool, available at <https://iraeatlas.ipm-gba.org.cn>. This tool may allow users to input ranked gene names from bulk RNA-seq data of blood samples in order to evaluate potential irAE risks in their patients. They can utilize the online tool to calculate raw irAE scores from bulk RNA-seq data. To interpret these scores in terms of irAE risks, a reference



dataset is needed. If they have hundreds of patients with known irAE status, they can scale the raw irAE scores to a 0 to 1 range, establish their own reference, and use the average score of irAE patients as a cutoff. If only individual patients are available, they can use SRP353971 as the reference dataset, where 0.39 is the average score of irAE patients. For a new patient, the raw irAE score can be calculated, integrated into the reference dataset, and scaled. Comparing a

patient's irAE score against the cutoff enables risk assessment, where a higher score may indicate increased likelihood of irAEs and the need for close monitoring. We will continue to update our data portal to provide an expanded reference resource for the research community.

Our study does have several limitations. First, although we validated the performance of the predictive models across cohort with different cancer types, the predictive models were developed using

Fig. 8 | Validation of the association between CD8.Temra and irAEs. **a** Gating strategy for CD8.Temra identification. All PBMCs were gated to exclude doublets and dead cells. Immunocyte subsets were selected based on CD45⁺ expression, followed by CD3⁺CD8⁺ T cell gating. CD8.Temra cells were identified as CD45RA⁺CCR7⁺ within the CD45⁺CD3⁺CD8⁺ T cell population. **b** Representative flow cytometry plots showing CD45RA and CCR7 expression in CD45⁺CD3⁺CD8⁺ T cells from irAE and non-irAE patients after combined chemoradiotherapy and ICI treatment. The red box indicates the CD8.Temra population (CD45RA⁺ CCR7⁺). **c** Bar plots comparing proportions of CD8.Temra cells between irAE and non-irAE groups ($n = 4$). Data are presented as mean values \pm SEM and P values were calculated using two-sided Mann-Whitney tests. Each dot represents an individual patient. **d** Representative flow cytometry plots showed CD45RA and CCR7 expression in CD45⁺CD3⁺CD8⁺ T cells from irAE and non-irAE patients after combined chemotherapy and ICI treatment, with the CD8.Temra quadrant highlighted in the red

rectangle. **e** Bar plot comparing CD8.Temra proportions between irAE and non-irAE groups ($n = 5$). Each dot represents an individual patient. Data are presented as mean values \pm SEM and P values were calculated using two-sided Mann-Whitney tests. All sample sizes reported are biological replicates, with the patient serving as the independent unit of study. **f** Illustration of irAE-related alterations in response to chemoradiotherapy and ICIs. Marked increases of CD4.Temra and CD8.Temra were observed in irAE patients following chemoradiation and ICI therapies, emphasizing the central role of T cells in the development of irAEs. The effects of chemoradiotherapy and ICIs on B cell and myeloid cells differed significantly in irAE patients. Chemoradiotherapy led to an increased percentage of B cells with IgA/IgG isotypes, reduced BCR clonality and elevated expression of inflammation-related pathways of myeloid cells in irAE patients. Created with <https://gdp.renlab.cn>⁵⁹. Source data are provided as a Source Data file.

scRNA-seq data from patients with LACC, necessitating future prospective studies across other cancer types to optimize the irAE predictive models. Secondly, although we analyzed 96 samples at the single-cell level, the sample size of patients remains relatively limited. However, it may require several years of multicenter efforts to prospectively collect longitudinal samples from a large number of patients with irAEs. But the number of irAEs is increasing exponentially in recent years, and potential severity and fatality of irAEs has been realized by the research community recently^{14,33}. Therefore, we employed deep learning methods to efficiently extract representative irAE-related features for irAE predictive models and validated them using 10 datasets. Further studies with a larger sample size are necessary, particular for predicting different types of irAEs. Due to limited follow-up time or unsuitable patient conditions, some patients lacked samples of one or two time points. As a result, certain comparisons, such as those between the on-imm blood sample of relapsed and non-relapsed patients, were based on a small sample size (e.g., only two patients in the relapsed group). Therefore, the conclusions drawn from such comparisons should be validated in future studies. Finally, due to insufficient follow-up time at this stage, we have only a subset of patients with disease relapse information. As a result, treatment efficacy-related molecular characteristics require further validation, which will be addressed in the near future.

Our longitudinal immune cell landscape throughout combined chemoradiotherapy and ICI treatment provides a comprehensive resource for the irAE research community, defining the irAE-related molecular hallmarks and transcriptional signatures of different immune cells in response to each treatment. This information enabled us to develop a robust deep learning model for predicting irAEs in patients. Our work provides knowledge regarding the interactions between chemoradiotherapy and ICI treatment. Additionally, it demonstrates how these treatment regimens can impact the immune responses associated with irAEs. We have also developed an online tool that may potentially help clinicians in assessing irAE risks for ICI patients.

Methods

Patient populations

In the radio group, patients received weekly intravenous cisplatin for 5 weeks during EBRT, with or without subsequent chemotherapy, depending on physician recommendations. According to the clinical trial protocol (NCT06378840, an observational standard of care treatment study), 17 patients in the radio+ICI group received intravenous nab-paclitaxel and cisplatin weekly for 5 weeks during EBRT, followed by sinitilimab 200 mg every 3 weeks for a total of 8 cycles. Participants received anti-PD-1 therapy at no cost as part of their clinical treatment. Three additional patients who were not included in the clinical trial but also received ICI (zimerelimab or penpulimab) after radical chemoradiotherapy were included in the radio+ICI group

for analysis (Supplementary Table 1b). Disease relapse was evaluated using imaging techniques such as CT, MRI and PET-CT, along with histologically confirmation. It was defined as the recurrence of tumors in the pelvic region, which includes the uterus, cervix, vagina, paracervical tissue, pelvic wall, and pelvic or paraaortic lymph nodes, as well as in other areas, such as distant organs or lymph node located above the renal veins.

We monitored the irAE development of all patients throughout their treatment. We determined the immune relatedness of adverse events by clinical symptoms, laboratory tests and image examinations as well as alleviation of manifestations after discontinuation of ICI. Gastrointestinal irAEs developed following ICI therapy and more than one month after chemoradiotherapy. Laboratory tests were employed to rule out potential infections. Discontinuation of ICI and administration of corticosteroid effectively alleviated the symptoms. For the skin, endocrine, musculoskeletal and connective tissue and hematologic irAEs, they did not occur during chemoradiotherapy, and most were relieved following corticosteroid treatment (Supplementary Table 6). The grading of irAEs was determined for each patient based on the National Cancer Institute Common Terminology Criteria for Adverse Events (CTCAE), version 5. All the anatomical sites where symptoms of irAEs were evident were recorded (Supplementary Table 1).

Single-cell RNA sequencing

For cervical cancer tissues, fresh tissue samples were dissociated at 37 °C with gently shaking (50 rpm) using 0.25% trypsin and 10 μ g/mL DNase I in PBS containing 5% fetal bovine serum (FBS; Thermo Fisher, Cat. no. SV30087.02) for approximately 40 min. The resulting cell suspension was filtered through a 40 μ m nylon strainer to remove debris and clumps. After removing red blood cells by 1 \times Red Blood Cell Lysis Solution (Thermo Fisher, Cat. no. 00-4333-57), the dissociated cells were then washed and resuspended in 1 \times DPBS with 2% FBS. For PBMCs, we diluted 2 mL of blood with an equal volume of DPBS and carefully layered this onto 4 mL of lymphocyte separation medium in a 15 mL tube. After centrifugation at 800 g for 20 min, we transferred the mononuclear cell layer to a new tube and added 5 mL of DPBS with 2% FBS. The cells were then centrifuged, treated with red blood cell lysis solution, washed and resuspended in 1 \times DPBS with 2% FBS. We assessed the cell viability by using 0.4% trypan blue (Thermo Fisher, Cat. no. 14190144) staining and a Countess[®] II Automated Cell Counter (Thermo Fisher). The cell suspension was then processed using the 10 \times Genomics Chromium controller to generate gel beads in emulsion (GEMs) and this produced barcoded cDNA by following the manufacturer's instruction (CG000206 RevD). Sequencing libraries were quantified using a High Sensitivity DNA Chip (Agilent) on a Bioanalyzer 2100 and the Qubit High Sensitivity DNA Assay (Thermo Fisher Scientific). Libraries were sequenced on a NovaSeq6000 (Illumina) platform using 2 \times 150 chemistry.

Single-cell 5' and V(D)J enriched library preparation

We prepared single-cell V(D)J libraries using the 10× single cell immune profiling solution kit. We utilized the Chromium Single-Cell V(D)J Enrichment Kit to prepare full-length TCR VDJ segments and full-length BCR VDJ segments, adhering to the manufacturer's guidelines. The process began with washing and resuspension of the cells in PBS containing 0.04% BSA to achieve a concentration of 100–800 cells/μL. We then captured 500–7000 cells in droplets, with estimated multiplet rates of 0.4–5.4%. After encapsulating cells for reverse transcription and barcoding, we broke the emulsions formed and purified the cDNA using Dynabeads MyOne Silane. The cDNA underwent PCR amplification under specific conditions. For the 5' gene expression library, we used 2.4–50 ng of amplified cDNA, which was fragmented, end-repaired and size-selected with SPRI select beads before PCR amplification with sample indexing primers. Similarly, for V(D)J-enriched libraries, we amplified V(D)J transcripts from 2 μL of cDNA, followed by fragmentation, end-repair and size selection. We sequenced single-cell RNA libraries on Illumina's HiSeq 4000 platforms with a minimum goal of 25,000 reads per cell. For V(D)J libraries, we aimed for at least 5000 reads per cell, and these were sequenced on the Illumina HiSeq 4000 system.

scRNA-seq, scTCR-seq and scBCR-seq data processing

We processed the raw scRNA-seq data using Cell Ranger software (version 7.1.0). The sequencing reads were aligned to the human reference genome GRCh38 and quantified using default parameters. We then removed ambient RNAs and any technical artifacts by using the CellBender⁴⁹ 'remove-background' function with default parameters and identified and removed the doublets via DoubletFinder⁵⁰ v2.0.3. To ensure data quality, we applied the following filtering criteria: cells with fewer than 200 expressed genes and those with more than 10% of reads from mitochondrial genes were excluded and removed. After filtering and normalization, our final dataset had an average of 1844 genes and 5586 unique transcripts per cell.

Dimensionality reduction, clustering and annotation

We merged all the samples and applied R package harmony v1.2.0 to remove batch effects. The Seurat (v5.0.1) R toolkit was used to complete dimensionality reduction and clustering. The clusters on a 2D map were further visualized by using uniform manifold approximation and projection (UMAP). We assigned cell type labels to clusters based on canonical marker genes: immune cells (PTPRC), T cells (CD3D), NK cells (NKG7), B cells (CD79A), plasma cells (MZB1), monocytes (FCN1), macrophages (CD68, CD163), neutrophils (CSF3R), dendritic cells (CD1C), mast cells (TPSAB1), fibroblasts (COL1A1), endothelial cells (VWF) and epithelial cells (KRT6A).

Copy number analysis and malignant cell determination

We utilized the R package inferCNV (version 1.2.1), which detects CNVs and provides insights into genomic alterations at the individual cell level, to conduct CNV analysis at the single-cell level. We also used normal epithelial cells as reference cells¹⁸ and excluded the sex chromosomes from the analysis. We further distinguished cancer cells from normal epithelial cells within the epithelial cell population by analyzing their cluster distribution, marker gene expression and inferred copy number variations.

scTCR-seq analysis

We utilized the Cell Ranger VDJ pipeline (version 7.1.0) to analyze scRNA-seq data. This process involved assembling T-cell receptor sequences, identifying CDR3 regions and TCR genes, and annotating T-cell clusters. Cells were selected based on their identification as T cells and the presence of both productive TCR α and β chains. Clonal expansion was defined as identical chains from at least three cells. Clonal diversity, abundance and overlap of the TCR repertoire were

quantitatively measured using the clonalDiversity, quantContig and clonalOverlap functions in the scRepertoire package⁵¹ (version 1.12.0), respectively.

scBCR-seq analysis

To assemble BCR sequences and identify CDR3 sequences and BCR genes, we used the Cell Ranger VDJ pipeline (version 7.1.0). We selected cells based on their classification as B cells in the scRNA-seq data and the presence of functional BCR heavy and light chains. Clonal expansion was defined as identical BCR heavy and light chains present in at least three cells. In order to analyze SHM, we reconstructed germline sequences of BCRs using CreateGermlines.py function and calculated number of mutations in each sequence using observedMutations function from Shazam v1.1.0⁵². We assessed BCR clonal abundances using the estimateAbundance functions from the alaka-zam (version 1.3.0) package⁵².

Xenium tissue preparation and RNA assay

We performed the Xenium in situ assay using the 10× Genomics Xenium platform according to the manufacturer's guidelines. We prepared 5 μm sections of formalin-fixed, paraffin-embedded tissue microarrays on Xenium slides. The tissues were de-paraffinized and permeabilized to expose mRNA. We then targeted mRNAs using 380 specific probes and 2 controls: one for non-specific binding and another for genomic DNA to ensure the signals originated from mRNA. We incubated tissue slides overnight at 50 °C using a 10 nM probe solution. After washing, we ligated the probes for two hours at 37 °C and annealed a rolling circle amplification (RCA) primer. We then amplified the circularized probes to create multiple copies of gene-specific barcodes for each RNA-binding event. This process involved incubation for one hour at 4 °C, followed by two hours at 37 °C. Following amplification, we washed the slides and chemically quenched background fluorescence to reduce autofluorescence from tissue components. The prepared slides were then placed into imaging cassettes for analysis using a Xenium analyzer.

Xenium data processing

The Xenium analyzer captured multi-layered (Z-stack) images for each cycle across all channels. We processed and combined these images to construct a comprehensive spatial map of transcripts in the tissue section. Using the DAPI (a DNA stain) image as a reference, we integrated image stacks from various fields of view and color channels to produce a single, representative image of the entire tissue section. We then used Squidpy (v1.6.0) and Seurat (v5.0.1) software packages to analyze and visualize the Xenium results, respectively. Cells with at least one transcript expressed were kept for subsequent analysis. We assigned cell type labels to clusters based on the canonical marker genes used in scRNA-seq analysis.

Generation of the irAE signatures based on DEGs

We adopted methods from Yan et al., who generated an activate TLS signature from scRNA-seq data⁵³ and Xue et al., who developed TIMELASER tumor subtypes from scRNA-seq data⁵⁴. Consistent with the process used for developing deep learning models, we first identified irAE-associated cell subclusters in blood samples collected at pre-radio, pre-imm and on-imm time points using two criteria: 1) a significantly increased proportion in irAE patients, with *FDR* < 0.1 as the cutoff, and 2) a total percentage exceeding 5% in the tissue at the specified time points. We then performed DEG analysis between irAE and non-irAE groups within all irAE-associated cell subclusters, retaining genes with a fold change > 2 and adjusted *P* value < 0.01. To minimize unexpected noise and expression artefacts from single-cell dissociation, we excluded genes associated with mitochondria, heat-shock protein, ribosome and dissociation⁵⁴. This process yields three lists of upregulated genes in irAE-enriched immune cells from pre-

radio, pre-imm, and on-imm blood samples, respectively, referred to as signatures 1, 2, and 3.

Generation of the irAE predictive models

Given the majority of irAE-related alterations were observed in peripheral blood samples and the diverse molecular features across different cell types, we employed deep learning strategies to develop predictive models for irAEs based on the scRNA-seq data from pre-radio, pre-imm and on-imm blood samples. The model development process involved four key steps: 1) Cell classification. Immune cell subclusters were categorized as irAE-associated (P_1) or non-irAE associated (P_0) cells based on their distributions. Subclusters significantly enriched in the irAE group were classified P_1 , while those with lower proportions were classified as P_0 . Statistical significance was determined using the two-sided Dirichlet-multinomial regression model, with *FDR* values calculated by the Benjamini–Hochberg method (*FDR* < 0.1 was used as a significant cutoff). Subclusters comprising less than 5% of the tissue were excluded. 2) Gene ranking and sentence formation. For each cell in these two subgroups, non-zero genes were ranked by raw expression counts, and the top 260 genes were selected to form a ‘sentence’ based on their gene names and ranks. The choice of 260 genes was made through multiple rounds of testing, as excessively long lists added noise, while shorter lists risked omitting critical information. Each sentence s_i for cell c_i was represented as $\{g_1, g_2, g_3, \dots, g_n\}$, where n is gene rank (max=260 in this case) and g_n denotes the name of n -th genes. The input data for BERT model was $I = \{(s_1, y_1), (s_2, y_0) \dots, (s_i, y_1)\}$, where y_1 indicated a cell from P_1 and y_0 indicated a cell from P_0 . Using these ‘sentences’ for cells or samples allowed better generalization across datasets generated from different platforms. 3) Prototype learning. The embedding function f_φ was used to train BERT model, which constructs a network of gene-to-gene relationships. A basic form of prototypical networks⁵⁵, a meta-learning approach introduced by Jake Snell, was implemented to fairly learn from various cell subclusters in P_1 and P_0 . For each class ($k = 0$ for P_0 and $k = 1$ for P_1), we compute an M -dimensional representation $x_k \in \mathbb{R}^M$ where $k \in \{0, 1\}$ through the embedding function $f_\varphi: \mathbb{R}^D \rightarrow \mathbb{R}^M$ with learnable parameters φ .

$$x_k = \frac{1}{|S_k|} \sum_{(s_i, y_i) \in S_k} f_\varphi(s_i)$$

where S_k is the set of sentences in class k .

The Euclidean distance function d was used to compute the distance of a query cell c from the prototypes x_k :

$$p_\varphi(y = k, |c) = \frac{\exp(-d(f_\varphi(c), x_k))}{\sum_k \exp(-d(f_\varphi(c), x_k))}$$

4) IrAE score calculation. Predicted irAE scores were calculated by generating gene sentences s_n for cell n and applying the embedding function f_φ to obtain embedding vectors $e_n = f_\varphi(s_n)$. Using the Euclidean distance function d , we computed $r1 = d(e_n, x_1)$ and $r0 = d(e_n, x_0)$, representing the distances to irAE (P_1) and non-irAE (P_0). The relative distance $r1 - r0$ was used to assess irAE risk, with smaller relative distances indicating higher irAE risks.

Predicted irAE scores (e_i) for cell i were normalized to a 0–1 scaling using:

$$e_i = 1 - \frac{d_i - \min(d)}{\max(d) - \min(d)}$$

where d_i is the relative distance of cell i and $\min(d)$ and $\max(d)$ are the minimum and maximum relative distances in the entire dataset, respectively.

The BERT model architecture was implemented using the Transformers (v4.47.1) package by Hugging Face. Prototypical networks was built using PyTorch (v2.5.1) in python (v3.11.11). PyTorch was employed for efficient gene ordering and sentence generation.

For the public scRNA-seq datasets (Supplementary Table 2), we used the cell type annotations when the authors provide these. When the cell type annotation was not provided, we used the Python package CellTypist (v1.6.1) to annotate the cells. We calculated the relative distance for each individual cell and normalized them within each dataset.

Gene and subcluster importance estimation

To evaluate gene contributions to predictions of irAEs, we calculated the SHAP value using the Explainer function of the SHAP package (version 0.48.0)⁵⁶. After initializing our prototype network with pre-trained weights, we implemented a gene-centric tokenization approach, treating each gene symbol as an indivisible token. SHAP values were then rescaled from 0 to 1. For each gene, we averaged the SHAP values across cells within each subcluster and ranked the genes according to their averaged SHAP values. An ablation analysis was performed by deleting cell subclusters one at a time from each irAE predictive model and comparing four metrics (F1 score, ACC, AUC, and PRAUC) to evaluate the impact of each cell type’s removal. AUC, ACC, PRAUC and F1 score were calculated by R package MLmetrics (version 1.1.3).

Cumulative incidence analysis for irAEs

We acquired the bulk RNA-seq data and the corresponding clinical information for patients with cancer who received ICI therapy from publicly available datasets in the Gene Expression Omnibus (GEO) and Sequence Read Archive (SRA) repositories (accession numbers GSE186143 and SRP353971, respectively). Predicted irAE scores were calculated as described in previous steps for each sample by the specified models. We then used the Kaplan-Meier analysis to assess the time to onset of irAEs in the specified patient groups. *P* values were determined by log-rank test for cumulative incidence analysis. Patients in SRP353971 were divided into high and low groups based on the optimal cutoff derived from the Youden index of the ROC curve. Due to the limited number of patients in GSE186143, they were divided into high and low groups based on the median values.

Bulk RNA-seq

Blood samples were collected from 14 LACC patients after chemoradiation but before the administration of ICIs for bulk RNA-seq. Total RNA was extracted using TRIzol reagent (Thermo Fisher Scientific, 15596018CN). Following library construction and amplification, sequencing was performed on the Illumina NovaSeq 6000 platform, with 150 bp paired-end reads, by Shanghai Biotechnology Corporation in accordance with the manufacturer’s guidelines. The clean data generated from the FastQ raw data were subsequently aligned to the human GRCh38 reference genome using Hisat2 software (version 2.0.4). Gene expression levels were quantified using featureCounts and transcript per million (TPM) values were calculated with the calculateTPM function from the R package scuttle.

Multiplex immunofluorescence (mIF) staining

LACC tissue samples were fixed in 4% paraformaldehyde, embedded in paraffin, and sectioned to a thickness of 4 μ m. After deparaffinization, the sections were rehydrated using graded alcohols and transferred to the antigen retrieval solution. Following this, the sections were blocked with 5% BSA and then incubated at room temperature for 40 min with the following primary antibodies: CD3 (1:200, Abcam, ab16669), CD4(1:200, Abcam, ab133616), CD8 (1:200, Abcam, ab237709), CD20 (1:200, Abcam, ab64088), CD68(1:400, Cell

Signaling Technology, 97778), pan-CK (1:200, Abcam, ab234297). After washing three times with PBS-T over 15 min, the sections were exposed to the corresponding secondary antibody for 20 min at room temperature. Subsequent labeling was performed using the Opal 6-Plex Detection Kits (AKOYA Biosciences, NEL821001KT) according to the manufacturer's protocols. The same procedure was repeated for the additional antibodies. Nuclear staining was achieved using DAPI Solution (Thermo Fisher Scientific, 62248) with a 10-minute dark incubation at room temperature. Images were captured using a 3DHISTECH scanner (Hungary), and intercellular distances were analyzed using HALO software (V3.6.4134.263).

Flow cytometry

Peripheral blood mononuclear cells (PBMCs) were resuspended in FACS buffer (PBS with 2% fetal bovine serum (FBS)) and centrifuged at $300 \times g$ for 5 min at 4 °C. After supernatant removal, 100 μ L of L/D stain (0.1%, BioLegend [423101]) was added to each tube and incubated for 15 min at room temperature. Cells were then washed with 0.5 mL PBS, centrifuged at $350 \times g$ (5 min, 4 °C), and treated with 50 μ L of human FcR blocking reagent (0.4%) for 15 min at room temperature to reduce nonspecific binding. Surface antibodies (BioLegend: hCD45RA [304127], hCCR7 [353207], hCD8 [344739], hCD3 [300308], hCD45 [368504]) were diluted 1:100 in 50 μ L staining solution, added to cells, and incubated for 30 min at 4 °C in the dark. After incubation, cells were washed with 0.5 mL FACS buffer, centrifuged ($350 \times g$, 5 min, 4 °C), resuspended in 0.5 mL PBS, and filtered. Cell populations were acquired using a BD FACSAria™ III flow cytometer.

Statistical Analysis

We performed all statistical analyses utilizing R (version 4.3.0). We employed a Dirichlet-multinomial regression model⁵⁷ to compare cell percentages within each sample. Because the total percentages of all cell subcluster equals 100, an increase in the percentage of one cell subcluster will inevitably cause a decrease in the percentages of other cell subclusters. The Dirichlet-multinomial regression model evaluates differences in cellular distribution between irAE groups (e.g., irAE patients versus non-irAE patients) while considering the proportions of all of the other cell types. The other comparisons between two groups were made with the two-sided Wilcoxon rank-sum test unless otherwise stated, and program enrichment significance among subclusters was determined using the two-sided Kruskal-Wallis test. For cumulative incidence analysis, we employed the Kaplan-Meier method and log-rank to identify the statistical difference. We considered *FDR* values < 0.1 and/or *P* values < 0.05 to be statistically significant.

Ethical statement

We collected cervical cancer tissue samples and peripheral blood samples from 30 patients with LACC between Jan 17, 2022, and Jan 22, 2024. All patients provided written consent before joining the study. The Ethics Committee of Renji Hospital, Shanghai Jiao Tong University School of Medicine granted approval for this research. This study included two patient groups: ten patients who received radical chemoradiotherapy alone or followed by chemotherapy were included in the radio group. Twenty patients who underwent chemoradiotherapy sequential with anti-programmed cell death-1 (PD-1) antibodies were in the radio+ICI group. Radical radiotherapy consisted of 45 Gy external beam radiotherapy (EBRT) delivered in 5 fractions per week, followed by image-guided brachytherapy which consisted of 30–35 Gy in 5–6 fractions.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data supporting the findings have been deposited in the Genome Sequence Archive database under accession code [HRA008063](https://doi.org/10.1038/s41467-025-67689-2) and [HRA010222](https://doi.org/10.1038/s41467-025-67689-2). These data are available under restricted access in compliance with the Regulation on the Management of Human Genetic Resources of the People's Republic of China. To obtain access, researchers must agree to adhere to the responsibilities outlined for data users (<https://ngdc.cncb.ac.cn/gsa-human/policy>). Access requests should be submitted through the GSA-Human platform and will be reviewed by the Data Access Committee and database administrator. For inquiries, researchers may contact the corresponding author, Haiyan Chen (chenhaiyan1220@163.com) and Ying Jing (jingying@ipm-gba.org.cn). Requests for raw data are typically processed within 1–6 months under the guidelines of Genome Sequence Archive for noncommercial and the data access will be granted for 3 months. The scRNA-seq dataset used to validate irAE models are available in the GEO database under accession code [GSE186144](https://doi.org/10.1038/s41467-025-67689-2), [GSE216329](https://doi.org/10.1038/s41467-025-67689-2), [GSE213486](https://doi.org/10.1038/s41467-025-67689-2), [GSE206301](https://doi.org/10.1038/s41467-025-67689-2) and [GSE144469](https://doi.org/10.1038/s41467-025-67689-2). The bulk RNA-seq dataset used to validate irAE models are available in the GEO database under accession code [GSE186143](https://doi.org/10.1038/s41467-025-67689-2) and the SRA under accession code [SRP353971](https://doi.org/10.1038/s41467-025-67689-2). Source data are provided with this paper.

Code availability

All software used in the study is publicly available as described in the “Methods” section. The source codes for the development of irAE models are available <https://github.com/Jinglab/irAEmodel> and Zenodo (<https://doi.org/10.5281/zenodo.17446943>)⁵⁸.

References

- Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer statistics, 2022. *CA Cancer J Clin* **72**, 7–33 (2022).
- Chen, W. et al. Cancer statistics in China, 2015. *CA Cancer J Clin* **66**, 115–132 (2016).
- Monk, B. J., Tewari, K. S. & Koh, W. J. Multimodality therapy for locally advanced cervical carcinoma: state of the art and future directions. *J Clin Oncol* **25**, 2952–2965 (2007).
- Gennigens, C., De Cuypere, M., Hermesse, J., Kridelka, F. & Jerusalem, G. Optimal treatment in locally advanced cervical cancer. *Expert Rev Anticancer Ther* **21**, 657–671 (2021).
- Lorusso, D. et al. Pembrolizumab or placebo with chemoradiotherapy followed by pembrolizumab or placebo for newly diagnosed, high-risk, locally advanced cervical cancer (ENGOT-cx11/GOG-3047/KEYNOTE-A18): a randomised, double-blind, phase 3 clinical trial. *Lancet* **403**, 1341–1350 (2024).
- Monk, B. J. et al. Durvalumab versus placebo with chemoradiotherapy for locally advanced cervical cancer (CALLA): a randomised, double-blind, phase 3 trial. *Lancet Oncol* **24**, 1334–1348 (2023).
- Galluzzi, L., Aryankalayil, M. J., Coleman, C. N. & Formenti, S. C. Emerging evidence for adapting radiotherapy to immunotherapy. *Nat Rev Clin Oncol* **20**, 543–557 (2023).
- Yan, B. et al. Autoantibodies drive heart damage caused by concomitant radiation and PD-1 blockade. *Cancer Immunol Res* **11**, 546–555 (2023).
- Cheema, A. K. et al. Radiotherapy induces innate immune responses in patients treated for prostate cancers. *Clin Cancer Res* **29**, 921–929 (2023).
- Rodrigues, M. et al. Nivolumab plus chemoradiotherapy in locally advanced cervical cancer: the NICOL phase 1 trial. *Nat Commun* **14**, 3698 (2023).
- Monk, B. et al. O001/#504 Durvalumab, in combination with and following chemoradiotherapy, in locally advanced cervical cancer: results from the phase 3 international, randomized, double-blind, placebo-controlled calla trial. *International Journal of Gynecologic Cancer* **32**, A2–A3 (2022).

12. Li, K. et al. Neoadjuvant chemotherapy plus camrelizumab for locally advanced cervical cancer (NACI study): a multicentre, single-arm, phase 2 trial. *Lancet Oncol* **25**, 76–85 (2024).
13. Postow, M. A., Sidlow, R. & Hellmann, M. D. Immune-related adverse events associated with immune checkpoint blockade. *N Engl J Med* **378**, 158–168 (2018).
14. Jing, Y., Yang, J., Johnson, D. B., Moslehi, J. J. & Han, L. Harnessing big data to characterize immune-related adverse events. *Nat Rev Clin Oncol* **19**, 269–280 (2022).
15. Chen, M. et al. Divergent tumor and immune cell reprogramming underlying immunotherapy response and immune-related adverse events in lung squamous cell carcinoma. *J. Immunother. Cancer* **11**, <https://doi.org/10.1136/jitc-2023-007305> (2023).
16. Spurr, L. F. et al. Highly aneuploid non-small cell lung cancer shows enhanced responsiveness to concurrent radiation and immune checkpoint blockade. *Nat Cancer* **3**, 1498–1512 (2022).
17. Altorki, N. K. et al. Neoadjuvant durvalumab plus radiation versus durvalumab alone in stages I–III non-small cell lung cancer: survival outcomes and molecular correlates of a randomized phase II trial. *Nat Commun* **14**, 8435 (2023).
18. Fan, J. et al. Multiomic analysis of cervical squamous cell carcinoma identifies cellular ecosystems with biological and clinical relevance. *Nat Genet* **55**, 2175–2188 (2023).
19. Tadepalli, S. et al. Rapid recruitment and IFN-I-mediated activation of monocytes dictate focal radiotherapy efficacy. *Sci Immunol* **8**, eadd7446 (2023).
20. Lynch, C., Pitroda, S. P. & Weichselbaum, R. R. Radiotherapy, immunity, and immune checkpoint inhibitors. *Lancet Oncol* **25**, e352–e362 (2024).
21. Zheng, L. et al. Pan-cancer single-cell landscape of tumor-infiltrating T cells. *Science* **374**, abe6474 (2021).
22. Chu, Y. et al. Pan-cancer T cell atlas links a cellular stress response state to immunotherapy resistance. *Nat Med* **29**, 1550–1562 (2023).
23. Johnson, D. B. et al. A case report of clonal EBV-like memory CD4(+) T cell activation in fatal checkpoint inhibitor-induced encephalitis. *Nat Med* **25**, 1243–1250 (2019).
24. Das, R. et al. Early B cell changes predict autoimmunity following combination immune checkpoint blockade. *J Clin Invest* **128**, 715–720 (2018).
25. Patel, A. J. et al. Regulatory B cell repertoire defects predispose lung cancer patients to immune-related toxicity following checkpoint blockade. *Nat Commun* **13**, 3148 (2022).
26. Shinnakasu, R. et al. Regulated selection of germinal-center cells into the memory B cell compartment. *Nat Immunol* **17**, 861–869 (2016).
27. Cheng, S. et al. A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells. *Cell* **184**, 792–809.e723 (2021).
28. Janesick, A. et al. High resolution mapping of the tumor micro-environment using integrated single-cell, spatial and in situ analysis. *Nat Commun* **14**, 8353 (2023).
29. Cytlak, U. M. et al. Immunomodulation by radiotherapy in tumour control and normal tissue toxicity. *Nat. Rev. Immunol.* **22**, 124–138 (2022).
30. Lozano, A. X. et al. T cell characteristics associated with toxicity to immune checkpoint blockade in patients with melanoma. *Nat Med* **28**, 353–362 (2022).
31. Bukhari, S. et al. Single-cell RNA sequencing reveals distinct T cell populations in immune-related adverse events of checkpoint inhibitors. *Cell Rep Med.* **4**, 100868 (2023).
32. Zhu, H. et al. Identification of pathogenic immune cell subsets associated with checkpoint inhibitor-induced myocarditis. *Circulation* **146**, 316–335 (2022).
33. Thomas, M. F. et al. Single-cell transcriptomic analyses reveal distinct immune cell contributions to epithelial barrier dysfunction in checkpoint inhibitor colitis. *Nat Med* **30**, 1349–1362 (2024).
34. Luoma, A. M. et al. Molecular pathways of colon inflammation induced by cancer immunotherapy. *Cell* **182**, 655–671.e622 (2020).
35. Gupta, T. et al. Tracking in situ checkpoint inhibitor-bound target T cells in patients with checkpoint-induced colitis. *Cancer Cell* **42**, 797–814.e715 (2024).
36. Blum, S. M. et al. Immune responses in checkpoint myocarditis across heart, blood and tumour. *Nature* **636**, 215–223 (2024).
37. Sung, C. et al. Integrative analysis of risk factors for immune-related adverse events of checkpoint blockade therapy in cancer. *Nat Cancer* **4**, 844–859 (2023).
38. Nimisha, M., Shamitha, T. & Mishra, G. Comparative analysis of embedding models for keyphrase extraction: A KeyBERT-based approach. *2023 4th IEEE global conference for advancement in technology (GCAT)*, 1–6 (2023).
39. Zaslavsky, M. E. et al. Disease diagnostics using machine learning of B cell and T cell receptor sequences. *Science* **387**, eadp2407 (2025).
40. Turnbull, I. R. & Colonna, M. Activating and inhibitory functions of DAP12. *Nat Rev Immunol* **7**, 155–161 (2007).
41. Li, J. et al. KIR(+)CD8(+) T cells suppress pathogenic T cells and are active in autoimmune diseases and COVID-19. *Science* **376**, eabi9591 (2022).
42. Choi, S. J. et al. KIR(+)CD8(+) and NKG2A(+)CD8(+) T cells are distinct innate-like populations in humans. *Cell Rep* **42**, 112236 (2023).
43. Zhong, D. et al. TYROBP promotes the spread of pancreatic cancer by causing M2 TAM polarization. *J Gastroenterol Hepatol* **39**, 2926–2939 (2024).
44. Proudfoot, A. E. Chemokine receptors: multifaceted therapeutic targets. *Nat Rev Immunol* **2**, 106–115 (2002).
45. Huang, Y. V. et al. Novel therapeutic approach targeting CXCR3 to treat immunotherapy myocarditis. *Circ Res* **136**, 473–490 (2025).
46. Lo, J. W. et al. Immune checkpoint inhibitor-induced colitis is mediated by polyfunctional lymphocytes and is dependent on an IL23/IFN γ axis. *Nat Commun* **14**, 6719 (2023).
47. Berner, F. et al. Autoreactive napsin A-specific T cells are enriched in lung tumors and inflammatory lung lesions during immune checkpoint blockade. *Sci Immunol* **7**, eabn9644 (2022).
48. Jing, Y. et al. Multi-omics prediction of immune-related adverse events during checkpoint immunotherapy. *Nat Commun* **11**, 4946 (2020).
49. Fleming, S. J. et al. Unsupervised removal of systematic background noise from droplet-based single-cell experiments using CellBender. *Nat Methods* **20**, 1323–1335 (2023).
50. McGinnis, C. S., Murrow, L. M. & Gartner, Z. J. Doubletfinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. *Cell Syst* **8**, 329–337.e324 (2019).
51. Borchering, N., Bormann, N. L. & Kraus, G. scRepertoire: An R-based toolkit for single-cell immune receptor analysis. *F1000Res* **9**, 47 (2020).
52. Gupta, N. T. et al. Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. *Bioinformatics* **31**, 3356–3358 (2015).
53. Yan, Y. et al. Multi-omic profiling highlights factors associated with resistance to immuno-chemotherapy in non-small-cell lung cancer. *Nat Genet.* <https://doi.org/10.1038/s41588-024-01998-y> (2024).
54. Xue, R. et al. Liver tumour immune microenvironment subtypes and neutrophil heterogeneity. *Nature* **612**, 141–147 (2022).
55. Snell, J., Swersky, K. & Zemel, R. S. Prototypical networks for few-shot learning. *Adv. Neural Inf. Process. Syst.* **30**, 4077–4087 (2017).
56. Lundberg, S. M. L., Su-In. *A unified approach to interpreting model predictions.* (2017).
57. Smillie, C. S. et al. Intra- and inter-cellular rewiring of the human colon during ulcerative colitis. *Cell* **178**, 714–730.e722 (2019).
58. Lei, et al. Distinct immune cell dynamics associated with immune-related adverse events during combined chemoradiation and

immune checkpoint inhibitor therapy. *irAE model* <https://zenodo.org/records/17446943> (2025).

59. Jiang, S. et al. Generic diagramming platform (GDP): a comprehensive database of high-quality biomedical graphics. *Nucleic Acids Res* **53**, D1670–D1676 (2025).

Acknowledgements

This work was supported by several funding sources: the Shanghai Pujiang Program (No. 22PJ1409000 to Y.J.), the National Natural Science Foundation of China (Nos. 82272747 to L.Z., and 81903129 to H.C.), the Incubating Program for Clinical Innovation of Renji Hospital (No. PYDY-DZX-009 to Y.B.), the Science and Technology Commission of Shanghai Municipality (Nos. 21ZR1438500 to Y.B. and 23ZR1438900 to X.M.), and the Beijing Xisike Clinical Oncology Research Foundation (No. Y-XD202001/zb-0011 to Y.B.). We are grateful to Dr. Dev Sooranna, Imperial College London, for English language edits of the manuscript.

Author contributions

Y.J., L.Z., and H.Y.C. conceptualized the project and designed experiments. L.Z., J.M., and J.M.T. conducted experiments. Y.J., J.Z., X.K.F., and Y.W. performed the bioinformatic analyses. Y.J., L.Z., H.Y.C., and J.S. wrote the manuscript with inputs from all authors. B.H., D.Z., J.J.Z., Y.R.B., and X.M.M. contributed to the interpretation of the data and provided samples and reagents.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-67689-2>.

Correspondence and requests for materials should be addressed to Jian Song, Jianming Tang, Haiyan Chen or Ying Jing.

Peer review information *Nature Communications* thanks Justin Balko, Ann Klopp and the other anonymous reviewer for their contribution to the peer review of this work. [A peer review file is available].

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025