Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Longitudinal single-cell transcriptomic and TCR profiling reveals early tumour, stromal and immune remodelling during curative radiotherapy

Benjamin Jenkins

benjamin.jenkins@soton.ac.uk

School of Cancer Sciences, University of Southampton https://orcid.org/0000-0002-7588-0044

Wei Jia

School of Cancer Sciences, University of Southampton

Ian Tracy

School of Cancer Sciences, University of Southampton

Vikash Dodhia

School of Cancer Sciences, University of Southampton

Karwan Moutasim

University of Southampton

Matthew Gillespie

School of Cancer Sciences, University of Southampton

Ian Reddin

University of Southampton https://orcid.org/0000-0001-5478-7855

Cecilia Liberatoscioli

University Hospital Southampton NHS Foundation Trust

Joseph Davies

Dorset Cancer Centre, Poole Hospital NHS Foundation Trust

Maxine Flubacher

Dorset Cancer Centre, Poole Hospital NHS Foundation Trust

Tim Fenton

School of Cancer Sciences, University of Southampton https://orcid.org/0000-0002-4737-8233

Emma King

School of Cancer Sciences, University of Southampton

Christopher Hanley

School of Cancer Sciences, Faculty of Medicine, University of Southampton https://orcid.org/0000-0003-3816-7220

Gareth Thomas

University of Southampton https://orcid.org/0000-0003-3832-7335

Article

Keywords:

Posted Date: November 19th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11468576/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

Longitudinal single-cell transcriptomic and TCR profiling reveals early tumour, stromal and immune remodelling during curative radiotherapy

Benjamin H Jenkins[1,2,*], Wei Jia[1,2], Ian Tracy[1,2], Vikash H Dodhia[1,4], Karwan A Moutasim[1,3], Matthew S Gillespie[1,2], Ian Reddin[1,2], Cecilia Liberatoscioli[3], Joseph Davies[4], Maxine Flubacher[4], Tim R Fenton[1,2], Emma V King[1,2,4], Christopher J Hanley[1,2,5] and Gareth J Thomas[1,2,3,5,**]

[1]School of Cancer Sciences, University of Southampton, Southampton, UK

[2]NIHR Experimental Cancer Medicine Centre, University of Southampton, Southampton, UK

[3]University Hospital Southampton NHS Foundation Trust, Southampton, UK

[4]Dorset Cancer Centre, Poole Hospital NHS Foundation Trust, Poole, UK

[5]These authors jointly supervised this work

[*]Correspondence: Benjamin.Jenkins@soton.ac.uk

[**]Correspondence: G.Thomas@soton.ac.uk

Summary:

Radiotherapy (RT) is a cornerstone of cancer treatment, yet early tumour ecosystem remodelling during treatment remains poorly understood because serial on-treatment biopsies are rarely available. Using single-cell transcriptomic and paired TCR sequencing of matched biopsies collected before and during RT in HPV-positive head and neck cancer, we define early cellular responses to treatment. RT rapidly depleted malignant cells in most patients. Residual tumour cells acquired an HPVoff state characterised by viral silencing, restored p53 activity, NOTCH activation, and a partial EMT/differentiation programme associated with radioresistance. Consistent with this phenotype, pre-treatment expression of the differentiation marker involucrin was associated with inferior survival. Fibroblasts persisted during treatment and adopted wound-healing and stress-adaptive programmes, whereas CAF states were maintained only in the presence of residual tumour. Concurrently, RT depleted tumour-associated CD8+ T-cell clones and increased TCR diversity through influx of naïve CD4+ cells. These findings reveal coordinated tumour, stromal, and immune remodelling associated with RT response and resistance.

Introduction:

Radiotherapy (RT) is estimated to be used in around 50-70% of all cancer patients, either as standalone therapy or in combination with other modalities such as surgery, chemotherapy or

immunotherapy[1]. Beyond its direct cytotoxic effects through the induction of irreparable DNA damage, RT significantly modulates the tumour microenvironment (TME), altering vascular integrity, fibroblast signalling and immune composition[2]. Clinical responses are highly heterogeneous, and the early cellular adaptations that determine sensitivity or resistance in human tumours remain poorly understood.

Preclinical studies using cell lines and animal models have identified multiple mechanisms that contribute to radio-resistance. Tumour-intrinsic factors, including enhanced DNA repair, evasion of apoptosis and maintenance of stem-like cellular states promote survival following radiation[3]. In parallel, extrinsic microenvironment factors, such as hypoxia and signalling from cancer-associated fibroblasts, provide additional protection against RT-induced cell death[4]. However, our understanding of the early tumour responses in patients is limited by the logistical and ethical challenges of obtaining serial tumour biopsies during RT. Consequently, most insights are derived from post-treatment or recurrent disease, leaving a critical gap in knowledge regarding the dynamic, tissue-level responses that emerge during RT and underpin treatment sensitivity or resistance.

Human papillomavirus (HPV)-positive head and neck squamous carcinoma (HNSCC) provides a clinically relevant setting in which to study early RT response. HPV-positive (+) tumours account for approximately 25-30% of HNSCC cases and are associated with significantly improved survival compared with HPV-negative disease. These tumours are typically non-keratinising and immune-inflamed, with high levels of tumour infiltrating T-cells[5], and are highly sensitive to radiotherapy, which is commonly delivered as definitive treatment to the primary tumour and regional cervical lymph nodes. This favourable RT response has prompted multiple studies of dose de-escalation aimed at reducing the substantial burden of late treatment-related toxicity, particularly fibrosis-related morbidity that affects approximately 30-50% of patients[6]. Despite this radiosensitivity, a subset of patients (~8%) develop locoregional recurrence within the radiation field[7], highlighting the need for improved understanding of early adaptive responses and mechanisms of resistance.

Here, we performed integrated single-cell transcriptomic and paired T-cell receptor (TCR) sequencing of matched pre-treatment and early on-treatment tumour biopsies from nine patients with HPV+ head and neck cancer, sampled one hour after fraction 10 of RT. This longitudinal, high-resolution analysis captures immediate tumour-intrinsic and microenvironmental adaptations to RT in human cancers. We identify early tumour cell programmes associated with RT resistance alongside heterogeneous immune and stromal remodelling during treatment. Together, these

findings provide mechanistic insight into the earliest cellular responses to RT and reveal therapeutic opportunities to optimise RT and inform rational combination strategies.

Results:

Radiotherapy (RT)-induced alterations in the tumour-microenvironment of HPV+ HNSCC

To characterise acute RT-induced remodelling of the TME, we performed paired single-cell RNA sequencing (scRNA-seq) and T cell receptor sequencing (scTCR-seq) on longitudinal tumour biopsies from nine patients with HPV+ HNSCC (**Figure 1A**). Samples were obtained at diagnosis (pre-treatment, pre-RT; n=9), and after two-weeks of RT (on-RT n=8) following delivery of 21.7Gy in 10 fractions of a definitive radiotherapy course (65Gy in 30 fractions over 6 weeks). Additional samples included irradiated tumour-adjacent normal mucosa (NAT-RT; n=1) and post treatment biopsies collected four (n=1) and nine (n=1) months after therapy. Overall, 20 samples were analysed, including 16 matched pre- and on-RT biopsies. Patients received radiotherapy alone (n=3), concurrent chemotherapy (CRT; n=4), or neoadjuvant chemotherapy followed by CRT (n=2). In patients undergoing CRT, platinum-based chemotherapy was administered primarily for radiosensitisation (**Figure 1A; Table S1**). At clinical follow up, one patient (HN641) developed locoregional recurrence (seven months post-RT) and subsequently died, whereas the remaining patients remained disease-free (two-years post the date of diagnosis).

After quality control 129,878 tumour-derived cells were integrated with a reference dataset of normal oropharyngeal mucosa[8] generating a harmonised atlas of 162,099 cells. Unsupervised clustering revealed 18 major cell populations spanning five tissue states (normal mucosa, pre-RT, on-RT, NAT-RT and post-RT; **Figure 1B**). Cell identities were assigned using canonical marker genes (**Figure S1A-B; Table S2**). To identify malignant cells, epithelial populations were re-clustered with epithelial cells from normal mucosa, with malignant identity determined by transcriptome-inferred copy number variation (CNV) supported by HPV transcript expression (**Figure S2A**; see methods).

Figure 1.

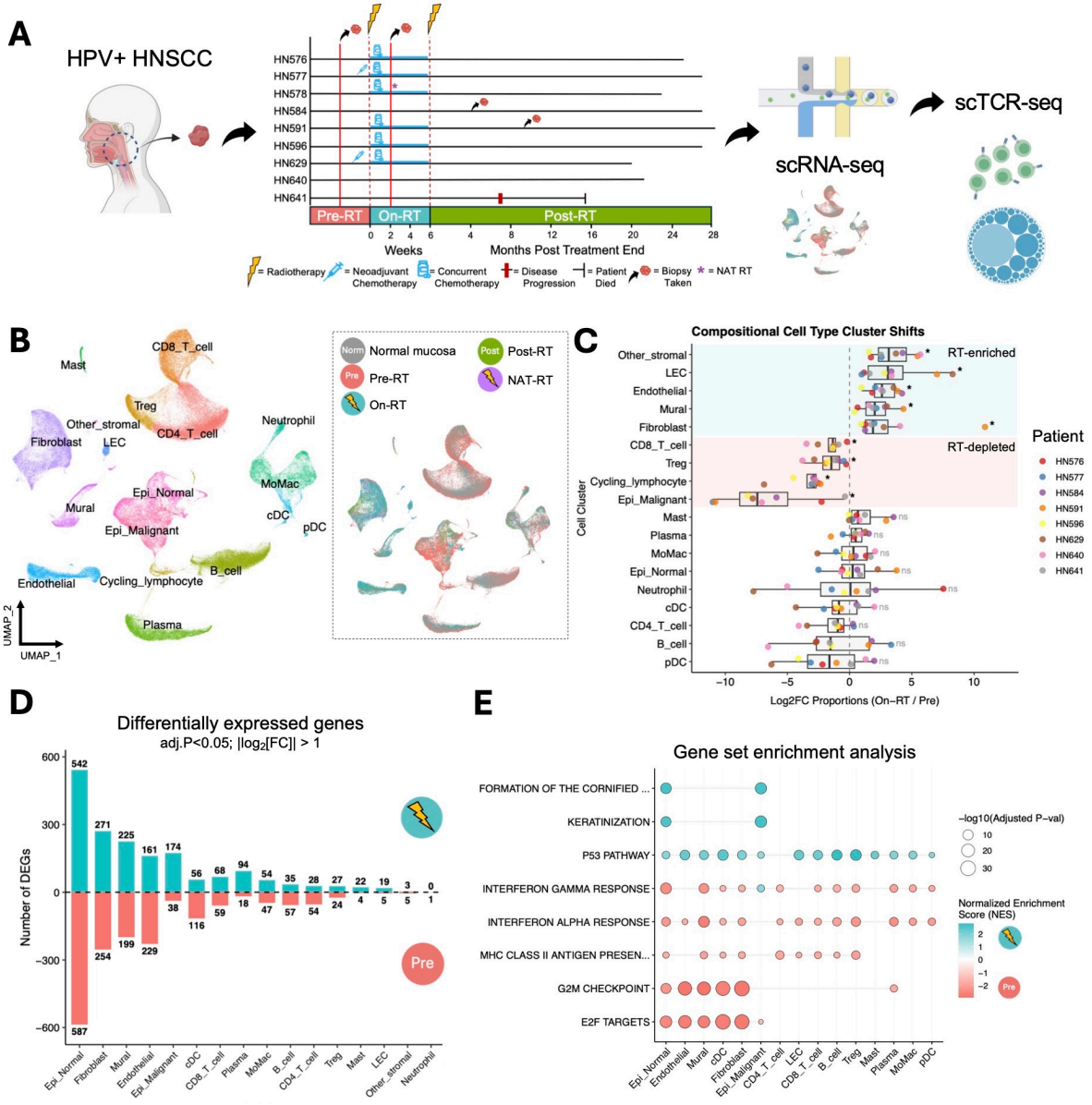


Figure 1. Radiotherapy (RT)-induced alterations in the tumour-microenvironment of HPV+ HNSCC. A. Schematic of the human papillomavirus-positive (HPV+) head and neck squamous cell carcinoma (HNSCC) patient cohort and workflow. Matched tumour biopsies were obtained before radiotherapy (pre-RT; n=9) and during radiotherapy after two weeks of treatment (on-RT; n=8). Additional biopsies included one RT-treated normal adjacent tissue (NAT-RT) sample obtained at two weeks and two post-treatment biopsies (post-RT) collected at four and nine months after RT. Patients received RT alone (n=3), concurrent chemotherapy (CRT; n=4) or neoadjuvant chemotherapy followed by CRT (n=2). Samples were profiled by scRNA-seq and scTCR-seq. B. UMAP plots of the integrated dataset comprising 129,878 tumour-derived cells and 32,221 normal mucosal cells[8], coloured by major cell type (left) and treatment condition (right). C. Boxplots showing log2 fold change in cell type proportions between patient-matched pre-RT and on-RT samples. Statistically significant changes were determined using the Wilcoxon signed-rank test, *adj.P<0.05. D. Barplot showing the

number of differentially expressed genes (DEGs; adj.P<0.05 and |log2[FC]| > 1) in each cell type, determined by DESeq2 analysis of pseudobulk samples (on-RT vs pre-RT). E. Gene set enrichment analysis (GSEA) results showing selected significant gene sets across cell types (adj.P<0.05). P values adjusted using the Benjamini–Hochberg method.

RT-induces depletion of lymphoid populations and enrichment of stromal compartments

To assess RT-induced TME remodelling, we calculated changes in cell-type proportions between matched pre-RT and on-RT samples for each patient. RT was associated with a significant reduction in malignant cell abundance (adj.p<0.05), although the magnitude of this response varied considerably between patients (**Figure 1C**). In contrast, normal epithelial cell abundance remained stable. Among immune populations, CD8+ T cells, regulatory T cells (Treg) and cycling lymphocytes were significantly reduced on-treatment (all adj.P<0.05; **Figure 1C; Figure S1B-C**). Overall, T cells decreased by 61.45 ± 19.84% (mean ± SD; epithelial cells excluded) (**Figure S1B**), with greater reductions in CD8+ T cells (66.57 ± 17.63%) and Tregs (67.27 ± 18.93%) than in CD4+ T cells (57.66 ± 30.73%).

Macrophage/monocytes and plasma cell populations were largely unchanged, whereas B cells and dendritic cells exhibited substantial inter-patient variability (**Figure 1C; Figure S1C**). In contrast, stromal populations, including fibroblasts, mural cells, vascular and lymphatic endothelial cells (LEC) and other connective tissue-associated cell types (Schwann cells, perineural fibroblasts and myoblasts) were significantly enriched during treatment (adj.p<0.05; **Figure 1C; Figure S1C**).

RT triggers coordinated multicellular stress responses and epithelial differentiation

Differential expression analysis across major cell lineages revealed extensive transcriptional remodelling during RT. Normal epithelial cells showed the greatest response, with 542 upregulated and 587 downregulated genes ($|\log_2[FC]|>1$; adj.p<0.05; **Figure 1D, Table S3**). Mesenchymal populations, including fibroblasts, endothelial cells and mural cells, also exhibited substantial transcriptional changes on-RT (**Figure 1D; Table S3**).

Gene set enrichment analysis (GSEA) identified several conserved responses across cell types (**Figure 1E**). Downregulated pathways included cell cycle progression, interferon-γ, interferon-α and MHC class-II antigen presentation (**Figure 1E; Table S3-4**). Conversely, p53 signalling emerged as a dominant RT-induced response, with significant pathway enrichment in 14 of 18 cell

subpopulations, accompanied by similar activation of DNA damage and cellular stress pathways (**Figure 1E; Table S4**). Consistent with this response, the p53 target genes, *AEN*, *BBC3*, *DDB2* and *RPS27L* were upregulated across all non-epithelial lineages (**Figure S1D; Table S3**). Additional stress response genes, including *GDF15*, *CDKN1A and RRM2B,* were induced in stromal and myeloid cells, with increased expression of *IL6, HSPA4L, and TYMS* in stromal populations, supporting a coordinated multicellular response to radiation-induced damage (**Figure S1D**).

Normal epithelial cells upregulated genes associated with keratinisation and cornified envelope formation (e.g. *KRT6A/B, KRT16, KLK5, IVL, TGM1, SPRR1A*) and innate epithelial defence (e.g. *PI3, BPIFB1, LCN2*), consistent with epithelial differentiation and a wound-healing-like response (**Figure 1E**; **Table S3-4**). Notably, malignant cells analysed separately also showed significant enrichment of epithelial differentiation-associated gene sets on-RT (**Table S4**).

Collectively, these data show that RT induces a coordinated multicellular response characterised by widespread activation of p53 and cellular stress responses, suppression of proliferative and interferon pathways, and remodelling of both normal and malignant epithelial cell states.

Malignant cells stratify by HPV expression and RT response

To resolve epithelial cell-state changes during treatment, we subsetted epithelial cells and performed higher resolution analysis. Malignant cells formed distinct patient-specific clusters reflecting substantial inter-patient transcriptional heterogeneity (**Figure 2A-C; Figure S2A).** Consistent with previous reports, malignant cells further segregated into discrete HPVon and HPVoff states, defined by the presence or absence of detectable HPV transcripts, respectively (**Figure 2A-C; Figure S2A**; ref.[9]).

In pre-RT samples, HPVoff cells comprised 21.39 ± 7.52% (mean ± SD) of malignant cells (range 7.27–30.62%). To validate the presence of HPVoff cells, we performed in situ hybridisation for HPV E6/E7 in an independent cohort of ten HPV-positive HNSCC samples. The proportion of tumour cells showing HPV E6/E7 signal varied substantially between cases (42.00 ± 21.50%), ranging from 20-80% of malignant cells (**Figure S2B-C**), supporting the existence of transcriptionally distinct HPVon and HPVoff tumour cell populations.

Figure 2.

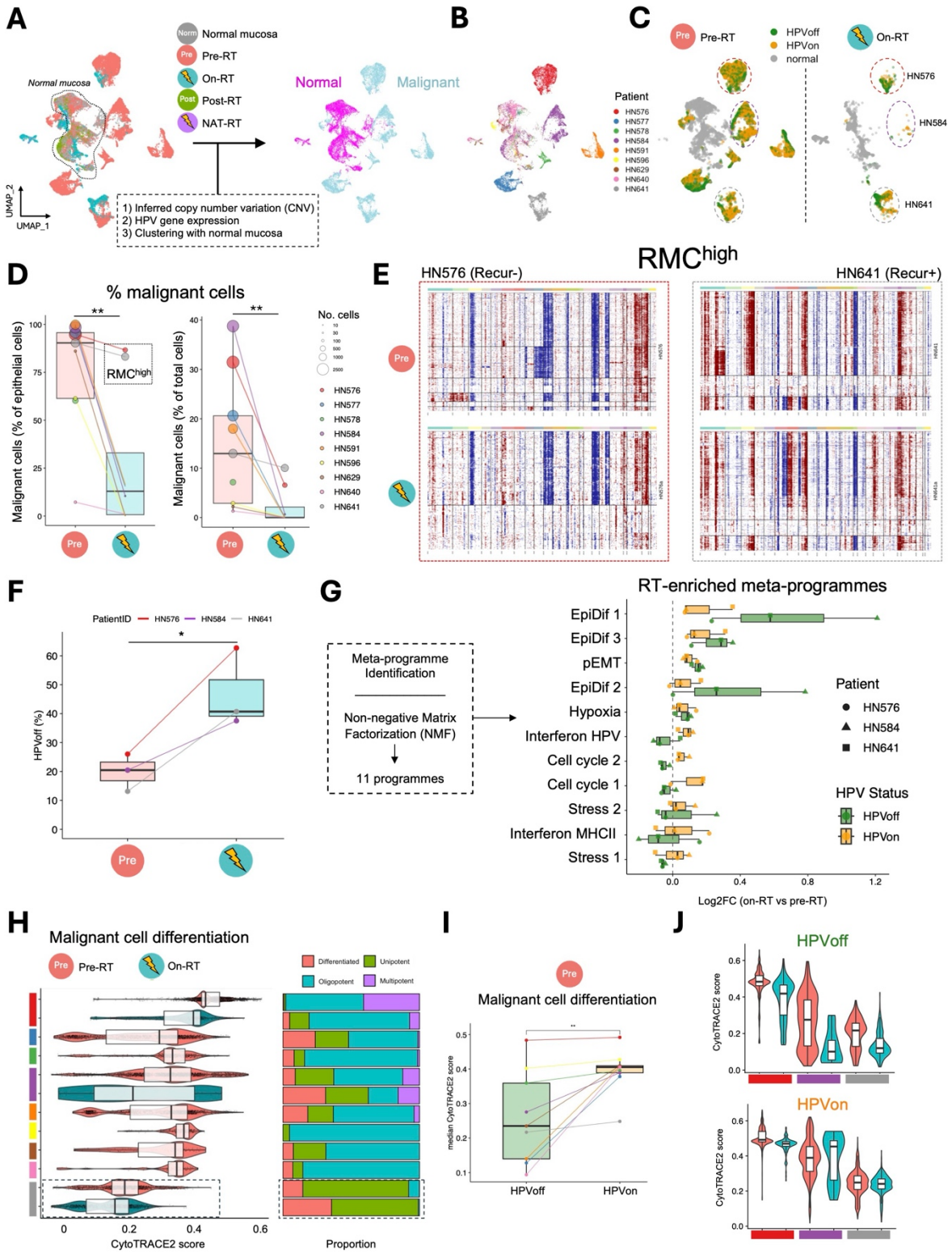


Figure 2. RT enriches differentiated HPVoff tumour states. A. Schematic illustrating the strategy used to distinguish

malignant and normal epithelial cells. B. UMAP plot of epithelial cells annotated by patient. C. UMAP plot of epithelial cells stratified by treatment and annotated by HPV expression status. D. Proportion of malignant cells in pre-RT and on-RT samples, expressed as percentage of epithelial cells (left) and total cells (right). Statistical significance was determined by Wilcoxon signed-rank test, **p<0.01. E. InferCNV heatmaps for two residual malignant cell high (RMC^{high}) patient samples (HN576 and HN641), split by pre-RT and on-RT. F. Proportion of HPVoff cells (within malignant cells) per sample pre-RT and on-RT, in patients with >10 malignant cells detected on-RT. Statistical significance was determined by paired t-test, *p<0.05. G. Boxplot showing enrichment scores for malignant cell meta-programmes, in samples from patients with <10 malignant cells detected on-RT, stratified by HPVon/off status, expressed as log2(fold-change) in mean scores on-RT versus pre-RT. H. Malignant cell CytoTRACE2 differentiation scores per sample, shown as individual cell scores (left) and proportion of discrete potency categories (right), with patient HN641 (RMC^{high}Recur+) highlighted by boxes. I. Median CytoTRACE2 differentiation scores in pre-RT malignant cells across HPVoff and HPVon cells, statistical significance determined by Wilcoxon signed-rank test, **p<0.01. J. Malignant cell CytoTRACE2 differentiation scores according to HPVon/off status for the three patients with >10 malignant cells on-RT.

Consistent with the marked sensitivity of HPV+ HNSCC to RT, malignant cells were almost completely eliminated by the two-week on-RT timepoint in 5 of 8 patients, with fewer than 10 residual malignant cells (RMC) detected (**Figure 2D**). Nevertheless, despite all patients achieving an initial clinical complete response, RMCs remained detectable in 3 of 8 cases. Two tumours exhibited substantial persistence of malignant cells (>5% of all cells; RMC^{high}), whereas a third contained only rare residual malignant cells (32 RMC detected on-RT; **Figure 2D**). CNV analysis confirmed the persistence of genomically aberrant malignant clones in RMC^{high} tumours, and indicated enrichment of specific CNV subclones during treatment (**Figure 2E**).

Notably, the patient who subsequently developed locoregional recurrence (HN641; RMC^{high}Recur+), exhibited the highest on-RT residual tumour burden and one of the smallest reductions relative to baseline (**Figure 2D**). In contrast, the second RMC^{high} patient (HN576; RMC^{high}Recur−), did not recur during follow-up, suggesting that continued RT was sufficient to eliminate RMCs in this case.

HPV gene expression persisted during RT but was significantly reduced within RMCs, reflected by both a smaller proportion of HPVon cells and lower HPV transcript abundance within the remaining HPVon population (**Figure S2D-F**). Consistent with this, RMCs showed a ~2.5-fold enrichment of HPVoff states during treatment (p<0.05; **Figure 2F**). In HN641 (RMC^{high}Recur+), HPV gene expression was restricted to *E6, E7 and E1* transcripts, with no detectable expression of *E2, E4, E5* or late viral genes, a pattern consistent with HPV integration (**Figure S2G**; ref.[10]). Together, these findings

indicate that tumour persistence during RT is associated with enrichment of HPVoff malignant cell states.

RT enriches differentiated HPVoff tumour states

To define malignant cell states across samples, we applied non-negative matrix factorization (NMF) and identified 11 conserved meta-programmes (**Figure S3A; Table S5**). These encompassed proliferation (Cell cycle 1 & 2), epithelial differentiation (EpiDif 1-3), stress response (Stress 1 & 2), interferon signalling (Interferon-MHCII & Interferon-HPV), Hypoxia and partial epithelial to mesenchymal transition (pEMT), and closely corresponded to previously described transcriptional states in HNSCC (**Figure S3B**; ref. [9]).

We next asked whether specific malignant cell states were associated with HPV gene expression and RT response. At baseline (pre-RT), HPVon cells were enriched for proliferative and interferon-HPV meta-programmes, whereas HPVoff cells were enriched for EpiDif and pEMT programmes ($p<0.05$; **Figure S3C-D**), indicating that HPVoff cells occupy a more differentiated and transcriptionally plastic state. Following RT, RMCs showed increased EpiDif and pEMT programme activity, particularly within HPVoff populations, suggesting preferential enrichment of differentiated and stress adapted states during treatment (**Figure 2G; Figure S3D**).

To independently assess changes in differentiation, we performed CytoTRACE2 analysis. RMCs exhibited significantly lower CytoTRACE2 scores than pre-RT malignant cells, consistent with progressive differentiation during RT **(Figure 2H)**. Because EpiDif programmes correlated with a previously described epithelial senescence programme (EpiSen; **Figure S3B**; ref.[9]), we next investigated whether these transcriptional changes reflected treatment-induced senescence. Although, EpiSen scores increased following RT (**Figure S4A**), canonical senescence gene signatures were not increased in RMCs, and the EpiDif1 programme significantly correlated with CytoTRACE2 scores ($R=-0.74$) but not with senescence signatures (**Figure S4B-C**). Together these findings indicate that RT-associated EpiDif programmes primarily reflect epithelial differentiation rather than senescence.

To determine whether differentiation was independently associated with HPVon/off status and RT exposure, we performed linear mixed-effects modelling. CytoTRACE2 scores were significantly predicted by both HPVoff status ($\beta = -0.105$, $p < 0.0001$) and RT exposure ($\beta = -0.029$, $p < 0.0001$;

Table S6). Accordingly, HPVoff cells were intrinsically more differentiated than HPVon cells and displayed a greater degree of differentiation on-RT (**Figure 2I-J**).

Differentiation states nevertheless varied substantially between tumours. Notably, the two RMC^{high} cases occupied opposite ends of the differentiation spectrum. HN641 (RMC^{high}Recur+) was the most differentiated tumour pre-RT, with most malignant cells classified as unipotent (77.5%) and a subset already differentiated (14.5%). In contrast, HN576 (RMC^{high}Recur−) was the least differentiated tumour in the cohort, with malignant cells predominantly classified as multipotent (41.0%) or oligopotent (57.0%; **Figure 2H**). Despite these marked baseline differences, both tumours exhibited reduced CytoTRACE2 scores during RT, consistent with increased differentiation (**Figure 2H**).

Pathway analysis further distinguished the two RMC^{high} tumours (**Figure S4E-F**). HN576 (RMC^{high}Recur−) was enriched for inflammatory and pro-survival signalling pathways, including TNFα, NFκB and IL6-JAK-STAT3 signalling. In contrast, HN641 (RMC^{high}Recur+) displayed reduced inflammatory activity together with increased oxidative phosphorylation and DNA repair programmes, resulting in a transcriptional profile that more closely resembled normal epithelium (**Figure S4E-F**).

RT promotes transition from proliferative HPVon to differentiated HPVoff states

To investigate treatment-induced state transitions at higher resolution we focussed on the case that recurred, HN641 (RMC^{high}Recur+), which was also characterised by the largest RT-induced expansion of HPVoff cells. Subclustering identified four malignant and two normal subpopulations (**Figure 3A; Figure S5A; Table S7)**. RT selectively depleted clusters 0 (*CALB*+, *ASS1*+, *MUC4*+) and 4 (*PIF1*+, *TOP2A*+, *MKI67*+) while enriching clusters 1 (*NTS*+, *UCHL1*+, *KRT13*+, HPVon) and 2 (*SERPINB7*+, *GDF15*+, *IVL*+, HPVoff) (**Figure 3A**-**C**). Cluster 2 (HPVoff) showed the greatest relative expansion on-RT, increasing sixfold (4.64% to 28.75%), while cluster 1 (HPVon) also expanded 3.6-fold (13.42% to 48.54%), becoming the predominant on-RT population (**Figure 3A, 3C**).

Differential expression and pathway enrichment analysis of on-RT compared with pre-RT malignant cells from HN641, revealed induction of keratinisation and cornified envelope genes (e.g. *IVL, SPRR2A, KRT16, TGM1, PI3)*, p53 signalling targets (e.g. *CDKN1A, FDXR, PHLDA3)*, inflammatory response genes (e.g. *IL1A, IL1B, CXCL8, IL4R)* and NOTCH signalling components (e.g. *NOTCH1, NOTCH3* and *JAG1)* during RT (**Figure 3D-E; Table S7**).

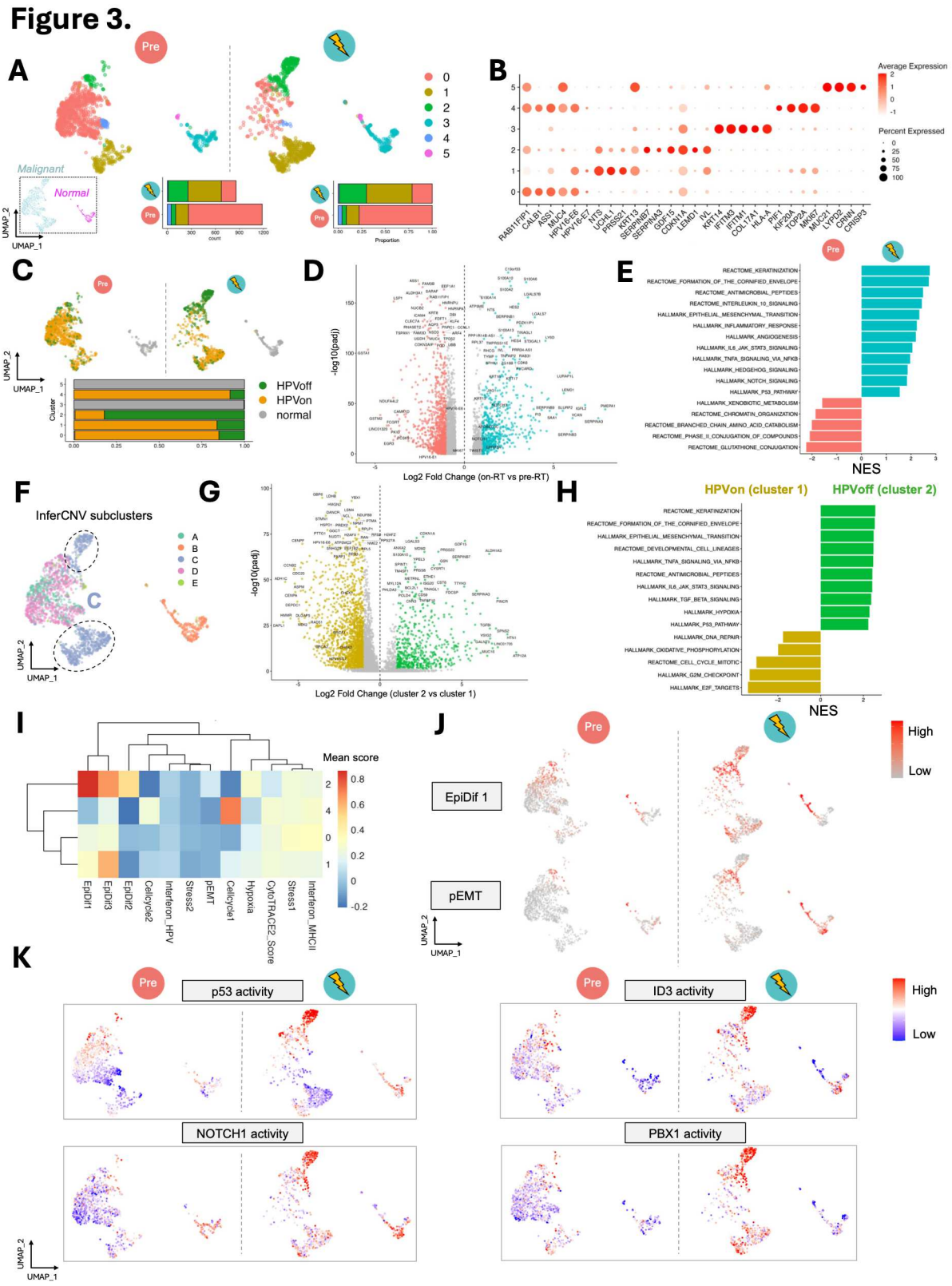


Figure 3. RT promotes transition from proliferative HPVon to differentiated HPVoff states. A. UMAP plots of HN641 epithelial cells annotated by cluster (and split by treatment; top) or annotated by normal/malignant cell identity (bottom),

accompanied by barplots of cluster abundance (absolute and proportion) across pre-RT and on-RT. B. Dot plot of cluster marker genes. C. UMAP plot of HN641 epithelial cells annotated by HPVon/off status, split by treatment (top), with corresponding barplots showing HPVon/off proportions across clusters (bottom). D. Volcano plot showing differentially expressed genes (DEGs) between on-RT and pre-RT HN641 malignant cells. E. Gene set enrichment analysis (GSEA) results showing selected significant gene sets (on-RT versus pre-RT; adj.P<0.05). F. UMAP plot of HN641 epithelial cells annotated by InferCNV subclusters. G. Volcano plot showing DEGs between cluster 2 (HPVoff) and cluster 1 (HPVon) malignant cells. H. GSEA results showing selected significant terms (cluster 2 versus cluster 1; adj.P<0.05). I. Heatmap of mean meta-programme enrichment and CytoTRACE2 scores across malignant cell clusters. J. UMAP plot displaying enrichment scores for EpiDif1 and pEMT meta-programmes across HN641 epithelial cells. K. UMAP plot showing p53 pathway activity and NOTCH1, ID3 and PBX1 transcription factor activities across HN641 epithelial cells.

Transcriptome-inferred CNV analysis demonstrated that the two major on-RT populations (HPVon/cluster 1 and HPVoff/cluster 2) shared a common CNV profile (subclone C), arguing against independent clonal origins and instead supporting RT-induced phenotypic state transitions (**Figure 3F; Figure S5B**). Compared with HPVon RMCs, HPVoff RMCs were enriched for epithelial differentiation, TNFα signalling, hypoxia and EMT (**Figure 3G-H; Table S7**). These differences remained significant when analysis was restricted to pre-RT cells, indicating that HPVon and HPVoff populations represent biologically distinct tumour states rather than treatment-induced artefacts (**Table S7**). The HPVoff cluster additionally expressed several genes previously implicated in radioresistance, including *PINCR*, *ALDH1A3*, and *GDF15*[11–13]. In contrast, the HPVon population retained a proliferative phenotype characterised by enrichment of G2M progression, cell cycle and DNA repair programmes.

Consistent with these observations, EpiDif1 and pEMT meta-programmes were selectively enriched in HPVoff cells (**Figure 3I-J; Figure S5C**). HPVoff cells also exhibited restoration of p53 signalling, with pathway activity exceeding that observed in normal epithelium, whereas HPVon cells remained relatively p53-suppressed, consistent with HPV E6 activity (**Figure 3K; Figure S5D-E**). Transcription factor analysis identified NOTCH1 activation as a defining feature of the differentiated HPVoff cells and normal epithelium, with further increases during RT (**Figure 3K; Figure S5E-F**). In contrast, plasticity-associated regulators ID3 and PBX1 were most active in HPVoff tumour cells relative to HPVon tumour cells (and normal epithelium), consistent with acquisition of an adaptive cellular state (**Figure 3K; Figure S5E-F**).

Together, these data support a model in which RT drives a transition from proliferative HPVon states to differentiated HPVoff states. These HPVoff populations are characterised by restoration of p53

activity, activation of NOTCH driven differentiation programmes and engagement of pEMT, inflammatory and hypoxia-associated pathways. Importantly, shared CNV profiles indicate that HPVon and HPVoff populations arise from a common clonal background, supporting RT-induced phenotypic state transitions rather than selection of genetically distinct subclones. Similar transcriptional programmes were observed in HN576 (RMC^{high}Recur-) suggesting that differentiation-associated HPVoff states may represent a conserved mechanism of tumour persistence during RT in HPV+ HNSCC (**Figure S6; Table S8**).

Involucrin identifies a differentiated tumour state associated with poor RT response

Given the association between RT-induced tumour persistence and differentiated HPVoff states, we next investigated whether tumour differentiation state was associated with clinical outcome in HPV+ HNSCC. Because EpiDif1 scores strongly correlated with CytoTRACE2-derived differentiation scores (**Figure S4C**), we used EpiDif1 as a surrogate measure of epithelial differentiation. In an independent cohort of RT-treated HPV+ HNSCC from the cancer genome atlas (TCGA; n=55), tumours with high EpiDif1 scores were associated with significantly poorer progression-free survival (PFS) than those with low scores (log-rank p=0.034; **Figure 4A; Figure S7A**), indicating that pre-existing differentiation programmes may be associated with reduced RT sensitivity.

To identify clinically useful biomarkers of this differentiated state, we interrogated genes within the EpiDif1 programme. Candidate genes were required to be epithelial-specific, highly expressed in the recurrent RMC^{high}Recur+ tumour (HN641); enriched during RT and demonstrate high RNA-protein concordance (**Figure 4B**). This approach identified a single candidate, involucrin (*IVL*), a well-established marker of squamous differentiation (**Figure S7B**). *IVL* was highly enriched in HN641 relative to all other tumours within the cohort (**Figure 4C**). Within HN641, *IVL* was detected in both HPVoff and HPVon populations but was highest in the HPVoff cells and increased further during RT across both cellular states (**Figure 4D; Figure S7C**), consistent with progressive enrichment of a differentiated tumour programme.

Spatial transcriptomic profiling demonstrated marked intratumoural heterogeneity in *IVL* expression, with discrete *IVL*-high tumour regions and cells variably distributed across HPV+ HNSCC samples (**Figure 4E; Figure S7D**). Given that HN641 (RMC^{high}Recur+) also exhibited a viral gene

Figure 4.

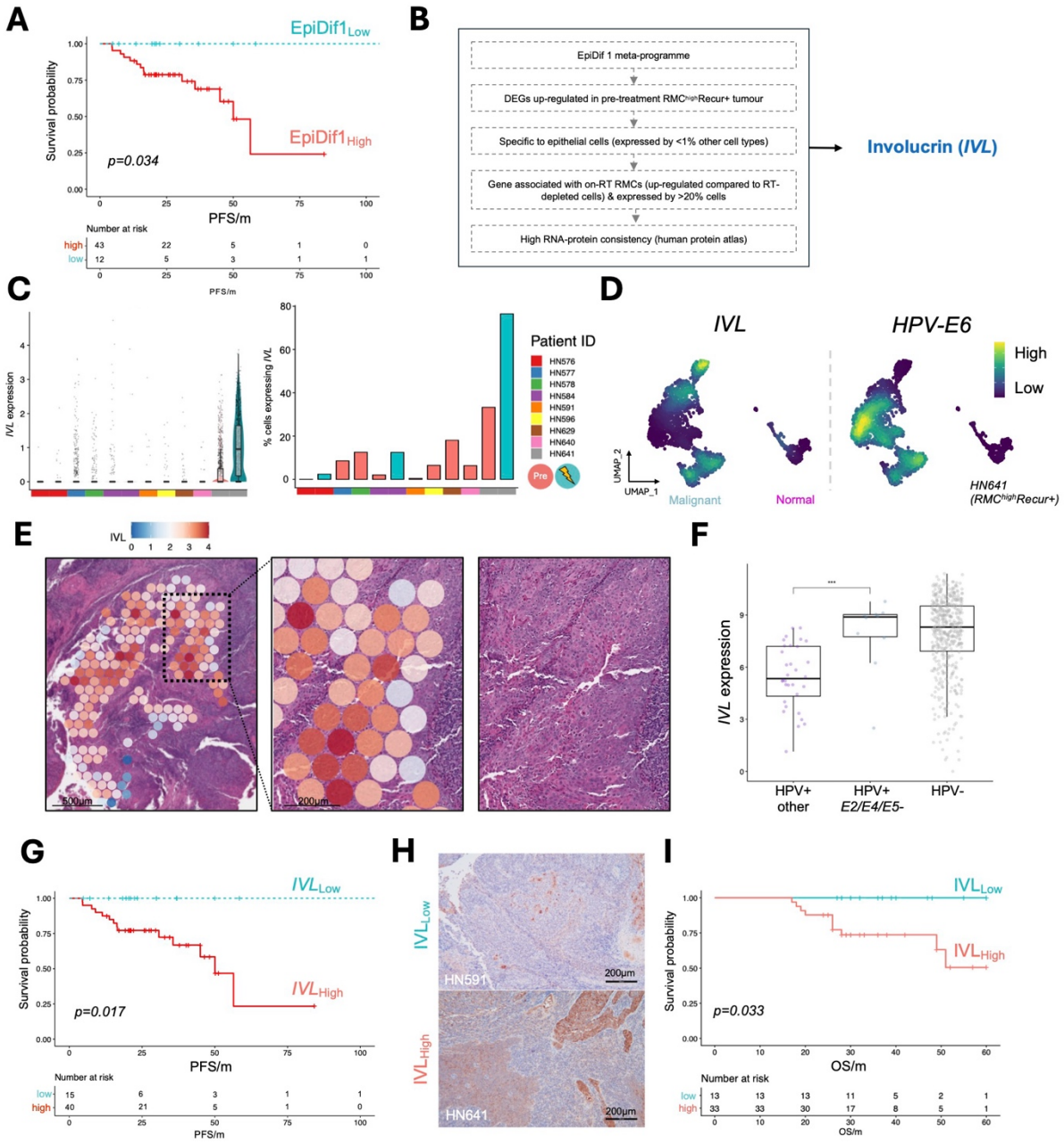


Figure 4. Involucrin identifies a differentiated tumour state associated with poor RT response. A. Kaplan-Meier (progression free survival; PFS) plot showing HPV-positive HNSCC (n=55) treated with RT (TCGA cohort) stratified by EpiDif1 (ssGSEA) scores. B. Schematic depicting the identification of involucrin as a potential biomarker of radioresistance in HPV-positive HNSCC. C. Violin plot of *IVL* expression (left) and barplot of percentage of *IVL* expressing cells (right) across patient malignant cells. D. UMAP plot of HN641 epithelial cells (Figure 3A) showing *IVL* and *HPV-E6* expression. E. Spatial feature plot of *IVL* expression across spatial transcriptomics spots (limited to epithelial cells; 10x Visium) in a HPV-positive HNSCC section. F. Boxplot of *IVL* expression across the TCGA-HNSC cohort subdivided by HPV gene expression. Samples were split based on absence of *E2, E4* and *E5* transcripts across HPV-positive HNSCC cases (as quantified previously; ref.[14]).

Statistically significant changes were determined by Welch's t-test, ***p<0.001. G. Kaplan-Meier PFS plot showing HPV-positive HNSCC (n=55) treated with RT (TCGA cohort) stratified by *IVL* expression. H. Examples of high and low immunohistochemistry (IHC) staining of IVL in an independent cohort of HPV-positive HNSCC (pre-treatment, treated with curative intent RT). I. Kaplan-Meier (overall survival) plot showing HPV-positive HNSCC (n=46; independent cohort) stratified by IVL IHC staining. Significance in Kaplan–Meier plots were assessed using the log-rank test.

expression pattern consistent with HPV integration (*E2/E4/E5* loss; **Figure S2F**), we next examined whether *IVL* was associated with integration status. HPV+ HNSCC tumours from TCGA cohort were classified according to previously reported viral gene expression and integration patterns[14]. Tumours exhibiting solely integrated (*E2/E4/E5*-negative) features displayed significantly higher *IVL* expression than tumours retaining episomal HPV (E2/E4/E5-positive; p<0.001), with expression levels approaching those observed in HPV-negative HNSCC (**Figure 4F**).

We next evaluated the prognostic significance of *IVL* expression in HPV+ HNSCC treated with RT. Within TCGA cohort, elevated *IVL* expression was associated with significantly poorer PFS (log-rank *p*=0.017; **Figure 4G**) and showed a similar trend for overall survival (**Figure S7E**). To independently validate these observations, we performed IVL immunohistochemistry on pre-treatment tumour samples from 46 patients with HPV+ HNSCC treated with curative-intent RT and followed for a minimum of 2 years (**Figure 4H**). Patients with IVL-high tumours exhibited significantly worse overall survival than IVL-low tumours (log-rank *p*=0.033; **Figure 4I**).

Collectively, these findings link the differentiated HPVoff tumour state identified during RT to adverse clinical outcome and implicate epithelial differentiation and HPV integration as features of treatment-resistant disease, establishing IVL as a clinically tractable biomarker of radioresistance in HPV+ HNSCC.

RT-driven fibroblast heterogeneity and repair programmes

Having identified RT-induced tumour cell state transitions, we next investigated how RT reshaped the stromal compartment. Mesenchymal populations were among the most extensively remodelled cellular compartments during treatment (**Figure 1C-D**). As epithelial and lymphoid populations contracted following RT, stromal cells constituted a substantially larger proportion of on-RT samples than pre-RT samples (median 32.9% versus 7.6%; p<0.01; **Figure S8A**). However, given that scRNA-

seq provides relative rather than absolute estimates of cell abundance within, the observed stromal enrichment may simply reflect loss of tumour and immune cells rather than on-RT expansion.

Re-clustering of mesenchymal cells identified eight major populations, including fibroblasts, mural cells, vascular endothelial cells, LEC, Schwann cells, *CLDN1*+ perineural fibroblasts, myoblasts and proliferating stromal cells (**Figure 5A; Figure S8B; Table S9**). Schwann cells, perineural fibroblasts, and myoblasts were relatively rare (pre-RT mean: $0.10 \pm 0.15\%$) and disproportionately enriched on-RT (mean $1.14 \pm 1.95\%$; **Figure 5B**). This enrichment likely reflects increased detection sensitivity, rather than the de novo emergence of these populations.

Fibroblasts exhibited the greatest number of differentially abundant graph neighbourhoods enriched on-RT (**Figure 5B**). Subsequent fibroblast subclustering identified ten transcriptionally distinct states (**Figure 5C-D; Figure S8C; Table S9**), including four *PI16*+ fibroblast subsets, FRC-like fibroblasts, *IL11*+ inflammatory CAFs (iCAFs), and myofibroblastic CAFs (myCAFs) consistent with previously described fibroblast populations in HNSCC (**Figure S8D**; ref.[8]). Notably, myCAF significantly decreased on-RT ($p<0.05$, **Figure S8E**).

CAF phenotypes persist only in the presence of residual tumour

Fibroblast remodelling differed according to treatment response. RT-responsive tumours ($RMC^{neg/low}$) showed marked depletion of both myCAF and *IL11*+ iCAF subtypes, whereas these canonical CAF states persisted in tumours retaining substantial malignant cell populations (RMC^{high}) (**Figure 5E**; **Figure S8E**). Consistent with this observation, 4 of 5 $RMC^{neg/low}$ tumours (<10 RMCs) displayed reduced expression of CAF-associated markers, including *FAP, POSTN and LRRC15*, together with re-emergence of homeostatic (universal) fibroblast markers such as *DPT, CFD* and *APOD,* indicative of partial stromal normalisation (**Figure 5F**). These findings suggest that maintenance of canonical CAF phenotypes depends on continued tumour-stroma interactions and is rapidly lost following effective tumour clearance.

RT enriches wound-repair and stress-adaptive fibroblast states

RT selectively enriched three fibroblast populations not previously identified in treatment naïve HNSCC scRNA-seq datasets: *APCDD1*+, *APOE*+, and *PI16*+*SFRP1*+ fibroblasts ($p<0.05$; **Figure 5G; Figure S8E**). These populations demonstrated activation of p53 and oestrogen signalling pathways

(**Figure 5H**), and were present but exceptionally rare in prior (treatment naïve) pan-cancer atlases (**Figure S8F**)[8]. Conversely, *APCDD1+* and *APOE+* subpopulations showed strong transcriptional

Figure 5.

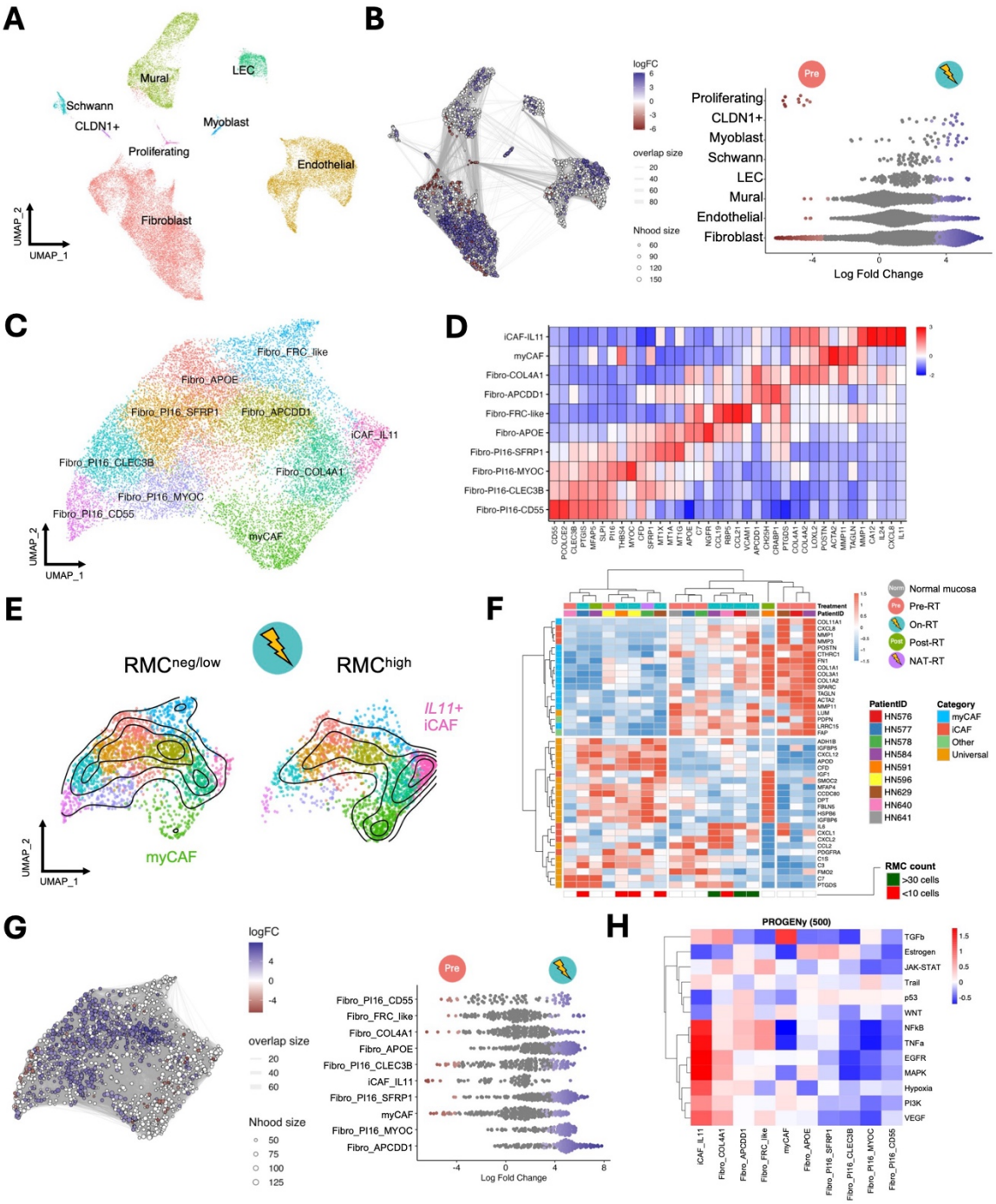


Figure 5. RT-driven fibroblast heterogeneity and repair programmes. A. UMAP plot of stromal cells annotated by broad cell type. B. MiloR differential abundance analysis showing stromal cell significant cell neighbourhoods (SpatialFDR<0.05) between pre-RT and on-RT conditions visualised by UMAP plot (left) and beeswarm plot (right). C. UMAP plot of fibroblasts

annotated by cluster identity. D. Heatmap of marker genes for fibroblast clusters. E. UMAP plot of on-RT fibroblasts split by RMC$^{neg/low}$ and RMC^{high} status, with cell density overlaid. F. Heatmap of average myCAF, iCAF and universal fibroblast gene expression per sample. G. MiloR differential abundance analysis showing fibroblast significant cell neighbourhoods (SpatialFDR<0.05) between pre-RT and on-RT conditions visualised by UMAP plot (left) and beeswarm plot (right). H. Heatmap of PROGENy pathway activities for fibroblast clusters.

similarity to fibroblasts associated with skin homeostasis and tissue repair (**Figure S8F**; ref.[15–17]). The *APCDD1*+ population displayed features of papillary fibroblasts, including high WNT signalling activity and expression of *CRABP1*, a regulator of retinoic acid-mediated tissue repair (**Figure 5H; Figure S8G**; ref.[18,19]). *APOE*+ fibroblasts resembled perivascular fibroblasts described in healthy skin[16]. In contrast, *PI16+SFRP1+* fibroblasts exhibited marked upregulation of metallothionein genes (*MT1X, MT1A, MT1G*; **Figure 5D**), consistent with enhanced resistance to oxidative stress and suggesting a previously unrecognised RT-associated fibroblast state. Together, these findings identify multiple RT-associated fibroblast states linked to tissue repair and cellular stress adaptation.

RT induces cytokine remodelling without robust fibroblast senescence

Despite activation of stress-response pathways, acute RT did not induce a canonical fibroblast senescence programme. RT triggered extensive cytokine remodelling across fibroblast populations, characterised by increased expression of *IL6*, *CXCL12, GDF15, CXCL14* and *BMP5,* alongside reduced expression *CXCL9, CXCL11* and *INHBA* (**Figure S8H; Figure S9A-B**). Although *CDKN1A* (p21) was consistently induced across fibroblast populations, *CDKN2A* (p16) expression remained limited (**Figure S9C-D**), and SenMayo gene set activity was not significantly increased (**Figure S9E-F**; ref.[20]). While selected senescence-associated and SASP-related genes were induced, these changes were insufficient to define a stable senescent phenotype.

Interestingly, one late post-treatment sample (HN591; 9 months post-RT) contained a population of *CDKN1A+/CDKN2A*+ myofibroblasts, whereas an earlier post-treatment sample (HN584; 4 months post-RT) did not (**Figure S9D; Figure S9G**), suggesting that RT-induced fibroblast senescence may emerge only after prolonged latency. Notably, these *CDKN1A+/CDKN2A*+ cells displayed a strongly myofibroblastic phenotype characterised by high expression of extracellular matrix genes, including

POSTN, COL1A1 and *COL3A1*, yet lacked expression of canonical inflammatory SASP genes such as *IL6, CXCL8* and *CXCL1* (**Figure 5F; Figure S9F**).

Collectively, these findings demonstrate that RT drives extensive fibroblast remodelling characterised by loss of established CAF subsets, emergence of wound-repair and stress-adaptive states and broad cytokine reprogramming. Canonical CAF phenotypes were depleted following effective tumour clearance but persisted in RMC^{high} tumours, suggesting on-going tumour-stroma interactions during treatment resistance. In contrast to the immediate transcriptional stress response observed during RT, robust fibroblast senescence was largely absent acutely and may emerge only after prolonged latency.

RT increases T-cell repertoire diversity through contraction of expanded CD8+ clones and influx of naïve CD4+ T cells

To define RT-induced remodelling of the T-cell compartment, which exhibited the largest treatment-associated shifts in abundance (**Figure 1C**), we integrated paired scRNA-seq and scTCR-seq datasets (**Tables S10-S11**). Subclustering identified 19 transcriptionally distinct T and natural killer (NK) cell states, spanning naïve, memory and regulatory CD4+ T cells, multiple CD8+ effector and tissue-resident memory (TRM) populations, NK cells and rare innate-like lymphocyte subsets (**Figure 6A-B; Table S12**). CD8+ T cells were dominated by *GZMK*^{high}, *IFNG*^{high}, and TRM populations, whereas CD4+ T cells comprised naïve/central memory, resident memory, follicular helper and TH17 subsets. TCR sequences also identified invariant T cell populations, including mucosal-associated invariant T (MAIT) cells and invariant NK T (iNKT) cells. MAIT cells (n=440) primarily localised to the CD8+ CD103- TRM cluster consistent with a distinct tissue-resident phenotype (**Figure S10A**).

Differential abundance analysis confirmed widespread depletion of antigen-experienced T-cell populations during RT, including CD8+ T cells, Tregs, helper CD4+ T cells and cycling T cells (**Figure 6C**). In contrast, naïve-like CD4+ T cells ($p<0.01$) and a subset of TRM CD4+ cell ($p<0.05$) populations were relatively enriched (**Figure S10B**). Together, these finding indicate a marked shift from antigen-experienced towards a naïve-dominated T-cell compartment during RT.

In terms of clonal dynamics, expanded clonotypes (multiplets; >1 cell of a given clonotype within a sample) were predominantly observed within CD8+ T cell populations, particularly CD103+ TRM and

Figure 6.

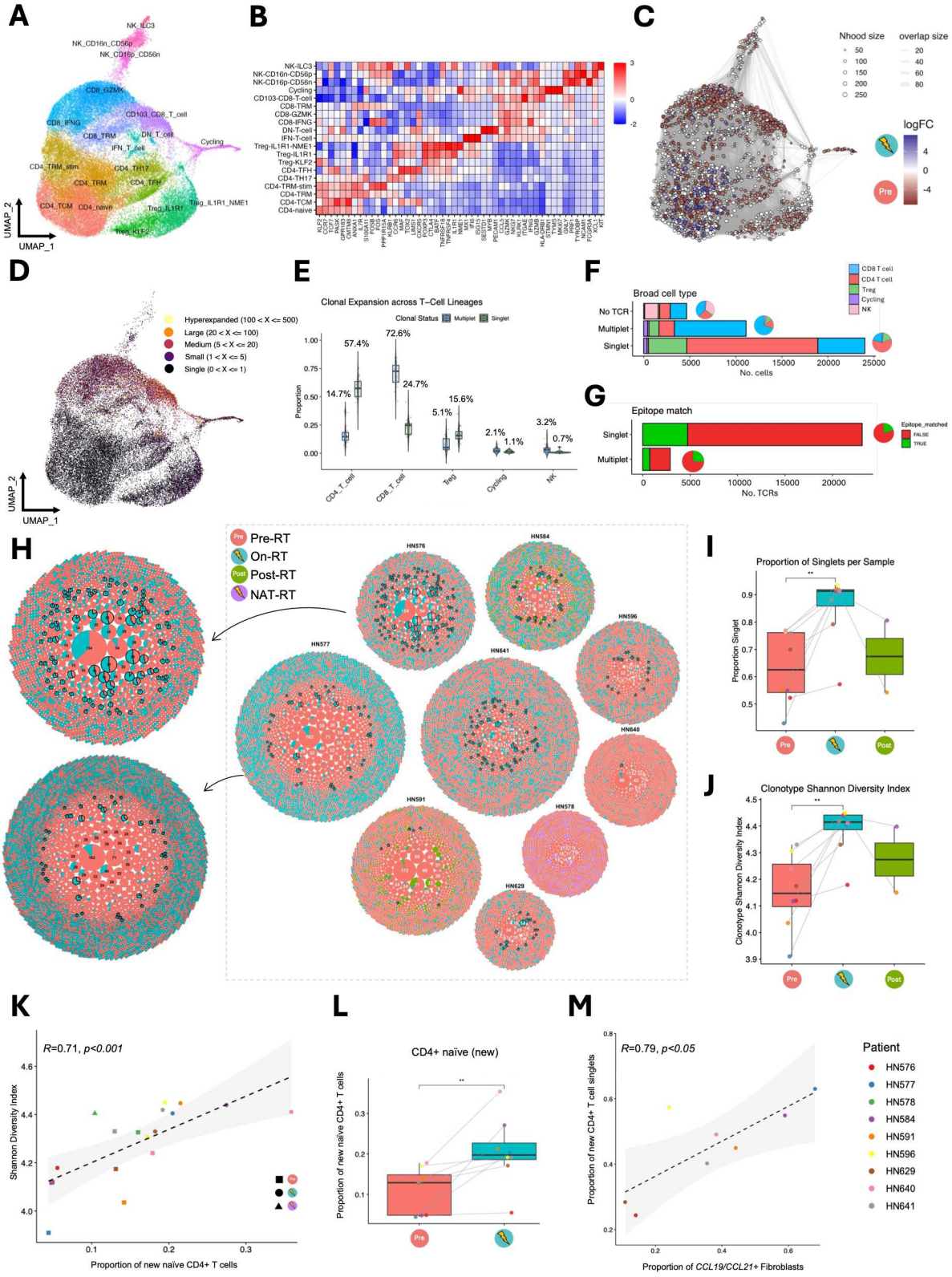

A
NK_ILC3
NK_CD16n_CD56p
NK_CD16p_CD56n
CD8_GZMK
CD103_CD8_T_cell
CD8_IFNG
CD8_TRM
DN_T_cell
Cycling
IFN_T_cell
CD4_TRM_stim
CD4_TH17
CD4_TRM
CD4_TFH
CD4_TCM
CD4_naive
Treg_IL1R1
Treg_IL1R1_NME1
Treg_KLF2
UMAP_2
UMAP_1
B
NK-ILC3
NK-CD16n-CD56p
NK-CD16p-CD56n
Cycling
CD103-CD8-T-cell
CD8-TRM
CD8-GZMK
CD8-IFNG
DN-T-cell
IFN-T-cell
Treg-IL1R1-NME1
Treg-IL1R1
Treg-KLF2
CD4-TFH
CD4-TH17
CD4-TRM-stim
CD4-TRM
CD4-TCM
CD4-naive
C
Nhood size
overlap size
logFC
Pre
D
Hyperexpanded (100 < X <= 500)
Large (20 < X <= 100)
Medium (5 < X <= 20)
Small (1 < X <= 5)
Single (0 < X <= 1)
UMAP_2
UMAP_1
E
Clonal Expansion across T−Cell Lineages
Clonal Status
Multiplet
Singlet
Proportion
57.4%
72.6%
24.7%
14.7%
15.6%
5.1%
2.1%
1.1%
3.2%
0.7%
CD4_T_cell
CD8_T_cell
Treg
Cycling
NK
F
Broad cell type
CD8 T cell
CD4 T cell
Treg
Cycling
NK
No TCR
Multiplet
Singlet
No. cells
G
Epitope match
Epitope_matched
FALSE
TRUE
Singlet
Multiplet
No. TCRs
H
Pre-RT
On-RT
Post-RT
NAT-RT
HN576
HN584
HN596
HN577
HN641
HN640
HN591
HN578
HN629
I
Proportion of Singlets per Sample
Proportion Singlet
Pre
Post
J
Clonotype Shannon Diversity Index
Clonotype Shannon Diversity Index
Pre
Post
K
R=0.71, p<0.001
Shannon Diversity Index
Proportion of new naïve CD4+ T cells
L
CD4+ naïve (new)
Proportion of new naïve CD4+ T cells
Pre
M
R=0.79, p<0.05
Proportion of new CD4+ T cell singlets
Proportion of CCL19/CCL21+ Fibroblasts
Patient
HN576
HN577
HN578
HN584
HN591
HN596
HN629
HN640
HN641

Figure 6. RT depletes clonally expanded CD8+ effector T cell populations and enriches naïve CD4+ T cells. A. UMAP plot of T and NK cells annotated by cell (sub)types. B. Heatmap of key marker genes relating to T/NK cell clusters. C. MiloR differential abundance analysis showing T/NK significant cell neighbourhoods (SpatialFDR<0.05) between pre-RT and on-RT conditions visualised by UMAP plot. D. UMAP plot displaying clonotype expander subtypes distributed across clonotype bins. E. Violin-boxplots showing the proportions of singlets (clonotype=1 cell per sample) and multiplets (clonotype >1 cell per sample) across broad cell types with median percentage labelled. F. Barplots showing the number of cells of each broad cell type split by multiplet status. G. Barplots showing the number of unique clonotypes with database-matched epitope specificity, split by multiplet status. H. Summary clonotype plots for each patient, displaying all unique clonotypes and respective sizes. Clonotype size (number of cells) is represented by circle size, coloured by treatment, with larger clones centrally located and singlets at the periphery. The top 10 largest clonotypes are labelled with cell number. Clonotypes more frequent on-RT are indicated by bold circle lines. Two patient clonotype examples are shown larger. I. Boxplots showing the proportion of singlets per sample. J. Boxplots showing sample clonal diversity scores represented by Shannon diversity index. K. Pearson correlation between Shannon diversity index and the proportion of new naïve CD4+ T cells per sample. L. Proportion of new CD4+ naïve T cells per sample. M. Pearson correlation between the proportion of new singlet CD4+ T cells and the proportion of *CCL19* or *CCL21*-expressing fibroblasts per sample. Statistically significant changes were determined using the Wilcoxon signed-rank test, **p<0.01.

GZMK^{high} subsets (**Figure 6D; Figure S10C-D**). Across samples, CD8+ T-cells accounted for the majority of expanded clonotypes (median 72.6%), whereas CD4+ T cells represented the largest population among unexpanded singlets (57.4%; **Figure 6E-F**), largely driven by naïve CD4+ T cells (**Figure S10C**).

Annotation against public TCR sequences and epitope prediction tools (VDJdb, McPAS-TCR, TCRex) (**Table S11**) identified a limited number of pathogen-associated clonotypes (CMV, Mycobacterium tuberculosis, influenza, SARS-CoV-2; **Figure S11A-B**), whereas most expanded clones (74.4%) lacked known antigen specificities (**Figure 6G**) reflecting the incomplete coverage of current reference datasets. Although exact matches to HPV-specific TCRs were not identified, sequence similarity analysis suggested candidate clones with potential HPV reactivity (**Figure S11C**). GLIPH2 analysis of on-RT multiplets identified convergent TCR clusters within and across patients, suggesting responses to shared antigens despite sequence diversity (**Figure S11D**).

RT induced marked but heterogeneous clonal remodelling across patients (**Figure 6H**; **Figure S11E**). In most cases, treatment was associated with contraction of expanded clonotypes and emergence of new, predominantly singlet clonotypes. Consequently, the proportion of singlet clonotypes increased significantly during RT (p<0.01; **Figure 6I**) accompanied by a corresponding increase in TCR diversity measured by Shannon index (p<0.01; **Figure 6J**). Notably, increased clonotype diversity

strongly correlated with the influx of naïve CD4+ T-cells (R=0.71; p<0.001; **Figure 6K-L; Figure S11F, Figure S12A**) suggesting that diversification of the T-cell repertoire during RT is driven primarily by recruitment of previously unprimed T-cell populations. Consistent with this interpretation, newly detected clonotypes were predominantly derived from CD4+ subsets, whereas conserved clonotypes (present in both matched pre-RT and on-RT samples) were enriched in CD8+ populations, particularly *GZMK*^{high} and CD103+ TRM cells (**Figure S12B**).

Residual tumours exhibit attenuated immune remodelling during RT

Although RT induced substantial immune remodelling in most patients, this response was attenuated in tumours retaining substantial malignant cell populations (RMC^{high}). These tumours exhibited smaller shifts in T-cell composition (L1 distance), limited increases in repertoire diversity and greater retention of pre-existing clonotypes (**Figure S12C-D**).

This pattern was particularly evident in HN576 (RMC^{high}Recur-), in which expanded CD103+ CD8+ T cell clones persisted during RT, including several large clonotypes that remained detectable throughout treatment together with newly emerging expanded clones (**Figure 6H; Figure S12A**). In contrast, RT-responsive tumours (RMC$^{neg/low}$) were characterised by extensive loss of pre-existing clonotypes and replacement by more diverse repertoires. These findings suggest that successful tumour clearance is accompanied by extensive remodelling of the local T-cell repertoire, whereas persistent tumours retain pre-existing immune architectures during treatment.

RT-induced fibroblast remodelling may support recruitment of naïve CD4+ T cells

To investigate mechanisms underlying naïve CD4+ T-cell recruitment, we examined expression of chemokine receptors and their cognate ligands. *CCR7* expression was highest within CD4+ T-cell populations (**Figure 6B**), which encodes the receptor for chemokines CCL19 and CCL21. Fibroblasts represented the dominant TME source of *CCL19* and *CCL21* (especially on-RT; **Figure S12E**), with multiple fibroblast subsets expressing *CCL19*, whereas *CCL21* expression was largely restricted to FRC-like fibroblasts (**Figure 5D; Table S9**). Indeed, in on-RT samples, the influx of newly detected singlet CD4+ T cells correlated strongly with the proportion of *CCL19* or *CCL21* expressing fibroblasts (R=0.79, p<0.05; **Figure 6M**). Although the overall proportion of *CCL19/CCL21*-positive fibroblasts remained unchanged following RT (**Figure S12F**), the primary source of CCR7 ligands

shifted from FRC-like fibroblasts (pre-RT) to RT-enriched *APCDD1*+ fibroblasts whose contribution to *CCL19*-expressing cells increased ~1.9-fold during RT ($p<0.05$; **Figure S12G**), becoming the new dominant CCR7-ligand source. These findings suggest that recruitment of naïve CD4+ T cells is coupled to RT-induced stromal remodelling. However, whether CCR7 ligand expression by fibroblasts acts as the primary driver of this influx or conversely, is amplified by arriving CD4+ T cells remains unclear.

Together, these findings demonstrate that acute RT induces profound remodelling of the T-cell compartment characterised by depletion of expanded tumour-associated CD8+ clones and increased TCR diversity driven by influx of naïve CD4+ T cells. This immune restructuring was attenuated in tumours with persistent malignant cells and was associated with RT-induced changes in fibroblast-derived chemokine signalling, suggesting coordinated stromal-immune remodelling during treatment.

Discussion:

RT is central to the curative treatment of many cancers and exerts both cytotoxic and immune-modulating effects. However, the early cellular programs that determine durable tumour control versus treatment resistance remain poorly understood, in part because serial on-treatment biopsies are rarely available. Using longitudinal scRNA-seq and scTCR-seq of serial tumour biopsies obtained before and during RT, we delineate coordinated tumour, stromal and immune adaption to treatment, revealing multicellular dynamics associated with early response to RT in HPV+ HNSCC.

HPV+ HNSCC provides a particularly informative model in which to study RT response because of its intrinsic radiosensitivity, immune-inflamed microenvironment and constitutive viral antigen expression. Clinically, this radiosensitivity has resulted in widespread interest in RT de-escalation strategies aimed at reducing treatment-related toxicity, particularly late fibrosis, which affects ~30-50% of patients[6]. Ongoing trials are evaluating reduced-dose RT alone or in combination with surgery, chemotherapy or immune checkpoint blockade[21]. However, despite the generally favourable prognosis of HPV+ HNSCC, a subset of patients experience treatment failure and locoregional recurrence. The biological basis of this heterogenous response remains poorly understood, creating a major challenge for the safe implementation of RT de-escalation strategies.

A central finding of this study was the rapid depletion of malignant cells during treatment. Most tumours demonstrated near-complete loss of malignant cells after two weeks of RT, whereas substantial residual disease was restricted to two patients, one of whom subsequently recurred. Tumour persistence was associated with expansion of an HPVoff state characterised by loss of viral transcription, restoration of p53 signalling, activation of NOTCH, and engagement of differentiation and partial EMT programmes. Although HPV transcript loss has been reported in pre-treatment HNSCC tumour cells and associated with drug resistance and invasion in HNSCC cell lines[9], its emergence during RT has not previously been recognised as a potential resistance mechanism. Several lines of evidence support the conclusion that this phenotype reflects adaptive cellular reprogramming rather than selection of pre-existing resistant clones. HPVoff and HPVon populations shared highly similar CNV profiles, indicating a common clonal origin, whereas RT induced progressive enrichment of differentiation-associated programmes and increased cellular differentiation scores. NOTCH activation emerged as a defining feature of HPVoff cells and may contribute directly to suppression of HPV oncogene expression.

In normal keratinocytes, NOTCH promotes cell-cycle exit and epithelial differentiation[22], while experimental studies have demonstrated reciprocal regulation between NOTCH signalling and HPV E6/E7 expression[23]. Consistent with this model, loss of viral oncogene expression was accompanied by restoration of p53 activity. Importantly, p53 activation did not primarily induce apoptosis but instead coincided with differentiation, stress-adaptive and repair programmes, highlighting the context-dependent outcomes of p53 signalling in epithelial tissues[13,24,25].

The resulting phenotype resembles a partially differentiated, slow-cycling state that retains considerable plasticity. Concurrent enrichment of pEMT and stress-response pathways suggests that RT drives tumour cells into a hybrid state that combines features of differentiation with preserved adaptive potential[26]. Such a state may enable survival during therapy while maintaining the capacity for subsequent reactivation. Notably, the early differentiation marker involucrin (IVL) was induced during RT and associated with poorer survival across independent HPV+ cohorts treated with RT. Given the increasing interest in RT de-escalation for HPV+ disease, involucrin-negative tumours may represent a subgroup with particularly favourable radiosensitivity and warrant evaluation as candidates for treatment de-escalation strategies.

Our findings further suggest that HPV genomic status may influence this adaptive trajectory. Previous studies have linked viral integration to increased epithelial differentiation and worse

clinical outcome[27,28]. In this context, integrated HPV genomes may enable viral silencing and transition into a differentiated, therapy-tolerant state, whereas episomal HPV may favour continued E6/E7 expression, proliferation and RT sensitivity. While these hypotheses require direct functional validation, they provide a framework linking viral biology to therapeutic response.

Beyond tumour cells, RT induced extensive remodelling of the stromal compartment. Fibroblasts persisted throughout treatment, and adopted transcriptional programmes associated with wound repair, stress adaption and tissue regeneration. Several RT-enriched fibroblast states closely resembled populations previously identified in normal skin and wound healing, rather than tumour associated fibroblast states observed across cancers[15–17], suggesting reactivation of tissue repair programmes. These findings suggest that RT actively reprograms the tumour stroma towards tissue repair programmes rather than simply preserving tumour-associated fibroblast subtypes.

Notably, maintenance of canonical CAF states appeared dependent upon continued tumour-cell presence. myCAF and iCAF populations were depleted in tumours that responded effectively to RT but persisted in tumours with residual malignant cells. Conversely, successful tumour clearance was associated with loss of canonical CAF programmes and emergence of homeostatic fibroblast states. Together, these findings support a dynamic and reversible model of CAF identity shaped by ongoing tumour-stroma interactions[29–31].

At a mechanistic level, RT-induced fibroblast states exhibited features consistent with stress adaptation, including activation of p53 signalling and upregulation of metallothionein genes associated with oxidative stress tolerance[32,33]. In particular, the metallothionein 1 expressing *PI16+SFRP1+* fibroblasts may represent a previously unrecognised RT-associated state with enhanced capacity to tolerate radiation-induced injury[34]. RT also induced extensive cytokine remodelling, characterised by increased expression of *IL6, CXCL12*, and *GDF15* together with reduced expression of *CXCL9* and *CXCL10*, indicating a substantial reorganisation of local immune signalling networks[35–37]. Notably, fibroblasts remained the dominant source of CCR7 ligands within the TME, and the abundance of *CCL19/CCL21*-expressing fibroblasts correlated with recruitment of naïve CD4+ T cells during RT. Although causality cannot be established from the current study, these observations suggest that stromal remodelling may actively contribute to immune reconstitution during treatment.

Despite stress activation, fibroblasts did not exhibit a uniform senescence programme during early treatment. While *CDKN1A* (p21) expression was broadly induced, consistent with transient cell cycle

arrest, induction of *CDKN2A* (p16) and canonical SASP features were limited, indicating dominance of acute stress response/repair programmes over stable senescence[38]. However, one post-treatment sample showed enrichment of p16+/p21+ myofibroblasts, suggesting delayed emergence of senescence-associated fibroblast states during tissue repair[39]. This temporal separation implies that radiation-induced fibrosis may reflect progressive stromal evolution rather than immediate radiation injury, consistent with the known role of senescent fibroblasts and SASP signalling in fibrotic disease[39]. Furthermore, the absence of canonical SenMayo and SASP gene expression in p16+/p21+ fibroblasts aligns with recent cross-tissue profiling of cellular senescence, which reveals extensive transcriptional heterogeneity and challenges the concept of a universal senescence phenotype[40].

In the immune compartment, a central observation was the marked depletion of effector T-cell populations during RT, including CD8+ T cells, helper CD4+ subsets, T regulatory cells, and proliferating T cells. This was accompanied by a relative enrichment of naïve-like CD4+ T cells, indicating a shift away from antigen-experienced states. Consistent with this depletion, we observed widespread contraction of expanded clonotypes during treatment, particularly within CD8+ TRM and effector-like populations that are typically enriched for tumour reactivity[41]. This raises the possibility that concurrent checkpoint immunotherapy may be suboptimal in this setting, potentially contributing to the lack of efficacy observed in the phase 3 KEYNOTE-412 and JAVELIN Head and Neck 100 trials[42,43].

This loss of expanded clones was accompanied by an increase in repertoire diversity and a higher proportion of singlet clonotypes. Notably, this diversification was strongly correlated with the influx of naïve CD4+ T cells, suggesting that increased TCR diversity reflects recruitment of previously unprimed T cells rather than diversification of pre-existing tumour-specific responses. These findings challenge the assumption that increased TCR diversity necessarily reflects enhanced anti-tumour immunity, and instead indicate that diversity metrics must be interpreted in the context of underlying cellular composition[44].

Our data further highlight distinct roles for CD8+ and CD4+ T-cell compartments during RT. Expanded clonotypes were predominantly derived from CD8+ T cells, particularly tissue-resident and effector-like subsets, supporting their role as the principal tumour-reactive population prior to treatment[41]. In contrast, CD4+ T cells were largely unexpanded and naïve-like, and accounted for the majority of newly detected clonotypes during RT. This divergence suggests that RT preferentially

depletes tumour-experienced CD8+ clones while simultaneously permitting repopulation by CD4+ T cells of uncertain antigen specificity. Whether these newly recruited CD4+ populations contribute to productive anti-tumour immunity or represent bystander infiltration remains a central open question.

Importantly, the extent and nature of immune remodelling varied according to treatment response. RT-responsive tumours were characterised by pronounced clonal turnover, including loss of pre-existing expanded clones and replacement with more diverse repertoires. In contrast, tumours with high residual malignant cell burden showed attenuated remodelling, with relative preservation of pre-existing clonotypes and reduced recruitment of naïve T cells. These observations suggest that effective RT is associated with extensive immune remodelling, whereas resistant tumours maintain a more conserved immune composition throughout treatment.

Collectively, our data support a model in which RT induces a rapid cycle of immune disruption and reconstitution, characterised by depletion of tumour-resident T cells, recruitment of naïve populations, and reshaping of the clonal repertoire. Crucially, the outcome of this process appears to depend on whether anti-tumour immunity is effectively re-primed or whether dysfunctional T-cell states persist. These findings have important translational implications. First, early increases in TCR diversity should not be interpreted in isolation as a biomarker of response. Second, they highlight a potential window for therapeutic intervention during RT-induced immune remodelling in which interventions that enhance antigen presentation and T-cell priming, including vaccination or innate immune agonists, could influence the trajectory of newly recruited T-cells and shape the developing anti-tumour immune response. Finally, the depletion of tumour-reactive clones highlights the importance of treatment sequencing and suggests that the efficacy of checkpoint blockade may depend on its timing relative to RT.

In summary, our findings redefine the early effects of RT in HPV+ HNSCC as a process of multicellular tumour ecosystem remodelling rather than a purely tumour-intrinsic or immune-mediated phenomenon. RT simultaneously drives tumour cell plasticity, stromal reprogramming toward wound-repair states, and marked restructuring of the T-cell compartment. In contrast, treatment resistance is characterised by phenotypic adaptation within tumour cells, persistence of tumour-supportive stromal programmes and attenuated immune remodelling, indicating that RT response is determined by interconnected changes across multiple cellular compartments. These observations challenge conventional interpretations of immune biomarkers during RT and suggest

that strategies aimed at enhancing tumour antigen presentation and T-cell priming may be most effective during periods of immune reconstitution. More broadly, our results raise the possibility that immune checkpoint blockade may be more effective when delivered prior to RT, to expand tumour-reactive clones, or after RT, once de novo priming and clonal reconstitution have occurred.

Resource availability

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Prof. Gareth Thomas (G.Thomas@soton.ac.uk).

Materials availability

This study did not generate any new materials.

Data and code availability

scRNA-seq and TCR-seq data generated in this study will be deposited on Gene Expression Omnibus (GEO) and be publicly available as of the date of publication. The scRNA-seq data will also be available at CZ CELLxGENE for convenient data exploration. The normal mucosa dataset was previously generated by our group[8] and is available through CZ CELLxGENE (https://cellxgene.cziscience.com/collections/3c34e6f1-6827-47dd-8e19-9edcd461893f). Additional publicly available scRNA-Seq datasets were obtained from the following sources: HNSCC and pan-cancer fibroblast datasets (https://doi.org/10.5281/zenodo.14284357); oral mucosa (GSE164241); Skin fibroblast atlas dataset (https://cellatlas.io/studies/skin-fibroblast/dataset/60/cherita). Previously generated 10x Visium spatial transcriptomics data from 10 HNSCC patients (n=7 HPV+; n=3 HPV-) is available from https://doi.org/10.5281/zenodo.14284038. TCGA data for the Head and Neck Cancer (HNSC) cohort were accessed via https://xenabrowser.net/. All code used for this study will be available on GitHub. Any additional information required to reanalyse the data reported in this paper is available from the lead contact upon request.

Acknowledgements

This work was funded through a Cancer Research UK Programme grant (DRVRPG-Jun\100004; GJT/CJH) and the CRUK & NIHR Experimental Cancer Medicine Centre (ECMC) Southampton (A15581 & ECMCQQR-2022/100018). The authors gratefully acknowledge Professor Sean Lim (funded Cancer Research UK Advanced Clinician Scientist Fellowship; A27179) and Mrs Anna H Turaj, for their support with single-cell sequencing, the University of Southampton Tissue Bank and the IRIDIS High-Performance Computing Facility at the University of Southampton. Finally, the authors gratefully acknowledge the patients who generously consented to provide tissue samples during radiotherapy, without whom this research would not have been possible.

Supplemental information

Data S1. Tables S1-S12.

Document S1. Figures S1-S12.

Author contributions

Conceptualization: EVK, CJH, GJT; Data curation: BHJ, IT; Formal analysis: BHJ, VHD, KAM, IR; Funding acquisition: EVK, CJH, GJT; Investigation: BHJ, IT, VHD, KAM, MSG, IR; Methodology: BHJ, WJ, IT, VHD, KAM, CL, JD, MF, EVK, CJH, GJT; Project administration: WJ, IT, EVK, CJH, GJT; Resources: BHJ, WJ, IT, VHD, KAM, CL, JD, MF, EVK; Software: BHJ, IT; Supervision: TRF, EVK, CJH, GJT; Visualization: BHJ, VHD, KAM; Writing – original draft: BHJ, CJH, GJT; Writing – review & editing: BHJ, WJ, IT, VHD, KAM, MSG, IR, CL, JD, MF, TRF, EVK, CJH, GJT.

Declaration of interests

The authors declare no competing interests.

Materials and methods:

Experimental model and study participant details

Sample collection and patient cohort

Ethical approval for the study was granted by UK National Research Ethics Service (REC No. 09/H0501/90) and all participants provided written informed consent. Patients with HPV-positive head and neck squamous cell carcinoma (HNSCC) treated at Poole Hospital NHS Foundation Trust (Poole, Dorset, UK) comprised the scRNA-seq cohort (**Table S1**). All tumours were p16-positive, with HPV status confirmed by alignment to a hybrid HPV-human reference genome (eight HPV16, one HPV33; see methods below). Biopsies were collected at baseline (pre-RT; n=9) and at two weeks during treatment (on-RT; n=8), with additional samples obtained from irradiated normal mucosa (n=1) and post-treatment tissue at four and nine months (n=2). Samples were transported on ice to the Southampton laboratory (Southampton, UK) in serum-free Dulbecco's Modified Eagle Medium (DMEM; Sigma-Aldrich). In total, 20 samples were analysed, including 16 matched pre- and on-RT biopsies, with patients receiving radiotherapy alone (n=3), concurrent chemotherapy (CRT; n=4), or neoadjuvant chemotherapy followed by CRT (n=2).

All patients achieved an initial complete response, confirmed by clinical examination, cross-sectional imaging, and repeat biopsies when clinically indicated. One patient (HN641) developed a locoregional recurrence seven months after completing radiotherapy. The remaining eight patients remained disease-free at 2-year follow up.

Method details

scRNA-seq

scRNA-seq sample processing

The methodology for preparing HNSCC samples for the 10x Chromium pipeline was a two-stage digestion protocol designed to improve capture rates of stromal cells relative to immune cell subsets, previously published[8].

scRNA-seq

For each sample, 10,000 single cells were captured on an Illumina 10x Chromium Controller™ system using the Illumina single cell 5' gene-expression kit (v2 Dual Index #1000263), Human TCR

amplification kit (#1000252), Next-GEM Chip-K (#1000286), Dual-index kit-TT set-A (#1000215), and library construction kit (#1000190). Sample capture, sample indexing, and library preparation were carried out according to manufacturer's instructions. Size distribution, quality control, and quantification of the libraries was assessed using High Sensitivity DNA chips (Agilent Technologies #5067 – 4626) and KAPA library quantification qPCR kit (Roche #07960140001). Prepared libraries were pooled and sequenced by Novogene, UK. Illumina guidelines on sequencing depth /reads-per-cell were followed and 150-base pair, paired-end sequencing on a Novaseq6000™ was performed.

Sequence alignment and annotation

FASTQ files were aligned to the Human reference genome (GRCh38–2020-A) which had the HPV genome concatenated to both the FASTA and GTF reference files (using cellranger multi v7.1.0, 10x Genomics), as previously outlined[8]. Human-HPV references were made using the cellranger mkref command (cellranger v6.1.1). scRNA/TCR-Seq data was processed with cellranger multi (cellranger v7.1.0) generating feature-barcode matrices.

Ambient RNA removal and empty droplet identification

CellBender (v0.3.0) was used to remove ambient RNAs and identify cell-containing versus empty droplets[45]. Inspection of cellranger count output files revealed poor cell calling for two on-RT samples (HN591a and HN640a). The raw_feature_bc_matrix.h5 files generated from cellranger were used as input for CellBender remove-background with the following parameters (--cuda --expected-cells 10000 --total-droplets-included 15000 --fpr 0.01 --epochs 150), run using IRIDIS High-Performance Computing Facility (University of Southampton). To convert the CellBender filtered output files to Seurat compatible .h5 files, ptrepack (PyTables) was used.

Quality control, normalisation, doublet detection and integration

CellBender-filtered count matrices were loaded into R (v4.5.2) and analysed using Seurat (v5.3.0). Individual Seurat objects were created per sample (min.cells = 3, min.features = 200) and cells filtered to possess percentage mitochondrial genes <10%. Samples were normalised using SCTransform, with subsequent dimensionality reduction and clustering steps: RunPCA, RunUMAP

(dims=1:20), FindNeighbours and FindClusters. Poor quality clusters marked by very low numbers of expressed genes (nFeature) and counts (nCount) values were removed.

Doublets were identified using DoubletFinder (v2.0) with recommended parameters, and results added to metadata. Final doublet annotation combined DoubletFinder predictions with manual curation based on clustering at increasing resolutions to examine gene/UMI counts and differentially expressed genes.

Normal mucosa epithelial cells were also reprocessed using CellBender and subjected to the same quality-control procedures as described previously[8].

Integration of all cells

To avoid excluding genes that were detected across most, but not all, samples during SCT-based variable-feature selection, standard log-normalisation (NormalizeData) was applied prior to RPCA integration. Both datasets were merged and normalised using NormalizeData followed by CellCycleScoring. The dataset was split per sample, in which FindVariableFeatures (nfeatures = 3000) and ScaleData (vars.to.regress = S.Score, G2M.Score, percent_mt) was run. The dataset was integrated using Seurat v5 RPCAIntegration (IntegrateLayers) with default parameters and visualised using RunPCA, RunUMAP (dims=1:30), FindNeighbours and FindClusters[46].

Integration of major cell lineages

The dataset was split into major cell lineages including T/NK cells, stromal cells, myeloid cells and B/plasma cells. To increase the resolution before subclustering, each major cell lineage was then reintegrated separately as previously outlined (except for removing cell cycle regression). Further subclustering within each major cell lineage was performed by specifying the new dimensional reduction during further UMAP dimensionality reduction and clustering (dims = 1:20).

Integration of epithelial cells

Due to expected large transcriptional heterogeneity across patient malignant cells, epithelial cells were broadly integrated (as above) across datasets (newly generated and previously published normal mucosa datasets) to promote the clustering of transcriptionally similar normal epithelial cell

states while retaining patient-specific clustering of malignant cells. Further analysis of specific patient epithelial cells was then carried out by reintegration across samples.

Clusters were characterised by identification of marker genes (using differential expression analysis) as previously described[8].

Sample level analysis

When splitting cells into samples and finer groupings, subsets with <10 cells were excluded from sample-level analyses (e.g. samples with <10 malignant cells in on-RT samples omitted from further abundance/pathway analysis).

Differential expression analysis (DEA) and gene set enrichment analysis (GSEA)

For scRNA-seq DEA across single cells in intra-patient samples, FindMarkers was used (Wilcoxon rank sum test), with genes requiring expression in ≥ 25% of cells and log2FC ≥ 0.5 (adjusted p value < 0.05).

To determine DEGs across pre-RT and on-RT conditions across patients (for each cell type), pseudobulk DEA was carried out. Firstly, cell subsets (across patient samples) with fewer than 10 cells and genes expressed in fewer than 50 individual cells and less than 4 samples were removed. AggregateExpression was used to generate aggregated raw counts matrices. DESeq2 (v1.48.2) was then used to determine DEGs (padj<0.05) between on-RT and pre-RT samples while accounting for patient-level variation (design = ~ PatientID + Treatment).

GSEA on pseudobulk DEA outputs was performed on unfiltered DEGs using the fgsea (v1.34.2) package (minSize = 15, maxSize = 500, padj < 0.05). scRNA-seq DEA outputs (min.pct = 0.1) were evaluated for GSEA using clusterProfiler (v4.16.0) GSEA with default parameters. MSigDB Hallmarks and Reactome gene sets were acquired using the msigdbr (v26.1.0) package.

Gene set enrichment scores, pathway and transcription factor analysis

In scRNA-seq data, enrichment scores were calculated using Seurat's AddModuleScore function (nbin = 24, ctrl = 100). Pseudobulk scRNA-seq and bulk RNA-seq samples were scored using single sample GSEA (ssGSEA) from the GSVA package (v2.2.1), with normalisation.

PROGENy pathway analysis[47] was carried out using the decoupleR package (v2.14.0). The PROGENy model gene weights were retrieved using get_progeny (organism=human, top=500) in which normalised log-transformed counts were extracted and used as input for determining regulatory activities using the Multivariate Linear Model (mlm) method. Mean (scaled) activity scores per group across pathways were then calculated and plotted via heatmap.

Transcription factor activity analysis was also carried out through decoupleR[48]. The transcription factor target network was retrieved using get_collectri (organism='human', split_complexes=FALSE) and normalised log-transformed counts used as input for determining enrichment scores through the Univariate Linear Model (ulm) method. Scaled activity scores were then used for downstream visualisation.

Differential abundance testing

The fold change in cell type proportions (relative to total cells) between on-RT and pre-RT samples was determined per patient. Patient HN578 was excluded due to missing an on-RT sample (NAT RT). Fold change values were log2 transformed, with a small pseudocount (10^{-4}) added. Conserved cell type changes across patients were assessed using Wilcoxon signed-rank tests against zero.

To confirm cell type abundance changes, miloR (v2.4.1; ref.[49]) was used to determine differentially abundant cell neighbourhoods on KNN graphs. For the whole dataset, buildGraph was used with k=30 and d=30, followed by makeNhoods with: prop=0.05, k=30, d=30, refined = TRUE and refinement_scheme = graph. testNhoods was then used to perform differential neighbourhood abundance testing (design = ~ PatientID + Treatment, fdr.weighting=graph-overlap) contrasting on-RT and pre-RT conditions (spatial FDR < 0.15). For fibroblasts and T/NK cells, k=20 and d=20 was used and prop=0.1 or 0.05 (respectively). testNhoods was then used to perform differential neighbourhood abundance testing (design = ~ 0 + Treatment, fdr.weighting=graph-overlap) contrasting on-RT and pre-RT conditions (spatial FDR < 0.05).

Transcriptome (inferred) copy number variation (CNV) analysis

To determine CNVs we used inferCNV (v1.22.0; https://github.com/broadinstitute/inferCNV). Stromal cells (fibroblasts, endothelial cells and mural cells) were used as reference cell types. CreateInfercnvObject was run using epithelial cell raw counts, grouped by patient (using annotations_file), with gene_order_file= hg38_gencode_v27.txt. Subsequently, infercnv::run was performed with the following parameters (cutoff=0.1, cluster_by_groups=TRUE, denoise=TRUE, num_threads = 20, HMM=TRUE, analysis_mode = subclusters, tumor_subcluster_partition_method = leiden, leiden_resolution = 0.0005, HMM_type=i6).

Quantifying the magnitude of inferred transcriptional CNV deviations

A continuous CNV burden score was derived from inferCNV output to quantify the magnitude of inferred copy-number alterations in individual epithelial cells. InferCNV values were centred on the diploid reference state, and genes were grouped by chromosome using annotations from the inferCNV HMM output. The mean CNV deviation was calculated for each chromosome and cell, and an overall per-cell CNV burden score was calculated as the Euclidean norm (square root of the sum of squared chromosome-level deviations) across all chromosomes. The scores reflect the overall magnitude of copy-number alteration irrespective of the direction (gain or loss) of individual changes; higher scores therefore indicate greater overall CNV burden.

An alternative approach, CopyKAT (v1.1.0), was used as a complementary method to estimate expression-based CNV profiles and corroborate inferCNV findings. Reference cell subsets were selected as fibroblasts, endothelial cells and mural cells. CNV profiles of epithelial cells were estimated per patient. Raw counts were extracted from both epithelial and reference stromal cells and merged. These were then used when running the copykat function, with the following parameters: id.type=S, ngene.chr=5, win.size=25, KS.cut=0.1, distance=euclidean, n.cores=20.

Malignant-normal cell discrimination

To distinguish normal and malignant epithelial cells, epithelial cells from the newly generated dataset and normal mucosal reference samples[1] were integrated and analysed together rather than across patients alone. Cells subsequently formed both patient-specific clusters and clusters comprising cells from multiple patients and normal mucosal samples. Each cluster was evaluated according to three criteria: 1) representation of cells from multiple patients and normal mucosa; 2)

inferred genomic instability, based on CNV scores and the proportion of cells predicted to be aneuploid; and 3) expression of HPV transcripts.

Malignant cells were defined by the presence of predominantly patient-specific cells together with high CNV scores, aneuploid prediction, and/or detection of HPV transcripts. HPV transcript expression was not an essential criterion for malignant cell identification. This annotation was subsequently used to assess residual cell burden during radiotherapy. Samples containing few or no residual malignant cell presence were labelled as $RMC^{neg/low}$ ($\leq$ 32 cells).

Non-negative matrix factorization (NMF)

To identify robust gene programs across multiple samples, NMF was performed using the GeneNMF (v0.9.2) package[50]. The Seurat object was first split per sample, then multiNMF was run (k=4:9, min.exp = 0.05, nfeatures = 3000). Gene programs that are consistently observed across samples were determined using getMetaPrograms (metric = cosine, specificity.weight = 5, weight.explained = 0.5, nMP=12, min.confidence = 0.7). One metaprogramme with a low silhouette score was removed using dropMetaPrograms.

CytoTRACE2

To estimate developmental potential of single cells, CytoTRACE2 (v1.1.0; ref.[51]) was implemented by first extracting raw counts, then passing these through the cytotrace2 function (species = human). The output was then added back to the Seurat object.

Linear mixed-effects model

To investigate the associations between HPVon/off status and RT exposure with CytoTRACE2 scores, a linear mixed-effects model was fitted using the lmer function (lme4 v2.0-6 package), with restricted maximum likelihood (REML) estimation. The model included HPVon/off status and treatment as fixed effects, with PatientID used as a random intercept. Fixed-effect significance was assessed using t-tests with Satterthwaite's approximation for degrees of freedom, implemented using lmerTest (v3.2-1). Model performance was assessed using marginal and conditional R^2 (performance v0.17.1 package).

Fibroblast marker genes and label transfer

To examine transcriptional changes in fibroblasts, we acquired gene sets attributed to different fibroblast phenotypes across the literature, including: myCAF, universal fibroblasts[52] and inflammatory CAF[53]. Genes from signatures with normalised mean expression values <0.1 were removed.

To investigate the degree of similarity of identified fibroblasts to fibroblasts of previous scRNA-seq datasets (across tissue/cancer types), Seurat's label transfer approach was used. The integrated fibroblasts object was used as the reference, with other scRNA-seq dataset objects iteratively used as query within FindTransferAnchors (dims = 1:20, reference.reduction = pca). TransferData was then used to transfer predictions across single-cell datasets (dims = 1:20). Previous scRNA-seq HNSCC and pan-cancer fibroblast datasets (https://doi.org/10.5281/zenodo.14284357) and oral mucosa fibroblast datasets were obtained (GSE164241; ref.[54]). The skin fibroblast atlas dataset[17] was acquired (https://cellatlas.io/studies/skin-fibroblast/dataset/60/cherita) and down-sampled to a maximum of 5000 cells per disease subtypes (Patient_status). To project query dataset cells onto the integrated fibroblast computed UMAP visualisation, the UMAP model of the reference dataset was returned, and the MapQuery function was utilised specifying the previously determined transfer anchors.

T cell receptor analysis (scTCR-seq)

Data processing

The filtered_contig_annotations.csv files generated through the cellranger multi pipeline were retrieved for each sample and analysed using the scRepertoire (v2.4.0) package[55]. Contig data for all samples were loaded into R and joined into a list. Contigs were then combined into clones using combineTCR. TCR data was then added back into the T/NK cell Seurat object using combineExpression based on overlapping barcodes. Default clonotype bins were used to label clonotypes based on cell frequencies: Hyperexpanded (100 < X <= 500), Large (20 < X <= 100), Medium (5 < X <= 20), Small (1 < X <= 5), Single (0 < X <= 1). Clonotypes were also broadly split into singlets and multiplets based on whether in each sample, there was only 1 cell or >1 cell of a given clonotype, respectively. The TCR clonal diversity was determined using clonalDiversity and selecting the Shannon metric, in which down-sampling of diversity metrics using 100 bootstraps (n.boots = 100) is performed and mean values determined. Clonotypes were also labelled based on whether

they were detected across both pre-RT and on-RT samples (per patient), or whether they were unique to either condition (conserved or new, respectively).

TCR annotation & epitope prediction

Annotation against public TCR sequences and was carried out by matching CDR3 clonotype sequences with those found in the VDJdb[56] and McPAS-TCR[57] databases. Furthermore, TCRex (https://tcrex.biodatamining.be), an epitope prediction tool[58], was also used to annotate clonotypes using default parameters (version 2023-06-26; BPR threshold = 0.0001).

Invariant subset identification

Invariant subsets including MAIT and iNKT cells were identified based on their restricted diversity of the TRA chains (https://www.10xgenomics.com/support/software/cell-ranger/latest/hidden/cr-5p-vdj-algorithm-inktmait). MAIT cells were identified by *TRAV1-2* + *TRAJ33/12/20*. iNKT cells were identified by *TRAV10* + *TRAJ18*.

HLA-subtype identification

arcasHLA (v0.6.0) was used to infer Human Leukocyte Antigen (HLA) genotypes from scRNA-seq data[59]. Firstly, sample_alignments.bam files were retrieved from cellranger count outputs for each patient. Next, the latest IMGT/HLA database (v3.63) was acquired. arcasHLA extract (--unmapped) was used to extract chromosome 6 reads and associated HLA sequences and output these as paired FASTQ files. The most likely genotype was then predicted using arcasHLA genotype. arcasHLA merge and convert (--resolution 1) were subsequently used to combine jsons and organise the resolution of allele calls, respectively.

TCR sequence similarity analysis

Pairwise similarity between known HPV-associated TRB CDR3 sequences (from databases, n=33) and target TRB CDR3 sequences were assessed using the Levenshtein (edit) distance. For each query–target pair, the Levenshtein distance was normalised by the length of the longer CDR3 sequence, and a similarity score was calculated as 1− normalised distance. Amino acids within CDR3 sequences were coloured according to chemical properties, including acidic (D, E), basic (K, R, H), hydrophobic (A, V, L, P, F), neutral (Q, N), and polar (G, S, T, Y, C).

GLIPH2

Clustering of TCRs to highlight potential convergent epitope binding was carried out using the gliph2 algorithm[60] [implemented in turboGliph (v0.99.2)], which was run on the CDR3 beta sequences using default parameters and the human_v2.0/ref_CD48_v2.0.txt reference database.

Quantification of T cell remodelling

To quantify the extent of microenvironmental remodelling of T cells across samples, cells were first annotated by both conserved/new status and scRNA-seq cluster identity, and proportions were calculated per sample. Cell-type proportions were then compared between matched patient samples (pre-RT and on-RT). Overall compositional change between paired samples was quantified using the L1 distance, calculated as the sum of the absolute differences in cell-type proportions across all cell types (higher L1 values indicating greater overall changes in cellular composition between pre-RT and on-RT samples).

Bulk RNA-seq analysis

Normalised (TPM) RNA-seq data from the cancer genome atlas (TCGA) was accessed for the GDC TCGA Head and Neck Cancer (HNSC) cohort via https://xenabrowser.net/. Clinical data, including HPV status and survival information was retrieved from cBioPortal[61] and samples annotated. To assess the relationship between involucrin expression and HPV gene expression in an independent cohort, we leveraged previously determined HPV viral gene expression within the HNSC-TCGA cohort[14] and added this to sample metadata. HPV+ samples with available metadata were then split depending on the detected expression of *E2*, *E4* or *E5*.

Spatial transcriptomics analysis

Previously generated spatial transcriptomics data (10x Visium) for 10 HNSCC patients (n=7 HPV+; n=3 HPV-) was accessed through https://doi.org/10.5281/zenodo.14284038. HPV+ sections (n=7) were then selected and spots pathologist annotated (GJT) to exclude adjacent normal regions of tissue. Furthermore, RCTD[62] deconvoluted normalised epithelial abundance (as carried out previously; ref.[8]) was used to only include spots with ≥0.70 values (high spot-level epithelial

abundance). The count data was normalised using SCTransform and expression levels and distribution of involucrin visualised.

HPV RNA in situ hybridisation

Formalin-fixed, paraffin-embedded (FFPE) tumour sections from ten HPV-positive head and neck squamous cell carcinomas (HNSCCs) were analysed using the automated RNAScope HPV-HR assay (Advanced Cell Diagnostics), performed on the Leica BOND platform to detect HPV E6/E7 transcripts. Briefly, 4-µm sections were deparaffinised, pre-treated and hybridised with an HPV E6/E7 probe. Probe detection was visualised using the RNAscope Multiplex Detection Reagent Kit (Advanced Cell Diagnostics). Appropriate positive and negative control probes were included in each assay. HPV E6/E7 signal was assessed in morphologically defined tumour cells by two head & neck pathologists (KAM, GJT) and quantified as the percentage of tumour cells showing positive expression.

RNA-protein consistency

To evaluate concordance between mRNA and protein expression, candidate genes were cross-referenced against the Human Protein Atlas (www.proteinatlas.org; ref.[63]). Protein-level validation was determined using the reported immunohistochemistry data reliability scores for each gene.

Immunohistochemistry (IHC)

Involucrin immunohistochemistry was performed on pre-treatment tumour samples from 46 patients with HPV+ HNSCC treated with curative intent RT and followed for a minimum of two years. Whole tissue sections or 1mm tumour cores incorporated into a tissue microarray (TMA) were stained with Involucrin Antibody SY5 (MA5-11803, Thermo Fisher, US). Involucrin staining was scored using a histochemical-score from 0 to 300 and samples were split into negative (0 – 20) and positive groups (20+).

Survival analysis

For survival analysis the survival (v3.8-9) and survminer (v0.5.2) R packages were used. Patients were split into high and low expression groups using the optimal cut points calculated using surv_cutpoint (for PFS). By identifying the biologically relevant threshold most associated with tumour progression (PFS), we applied the same cut point in overall survival analysis. Kaplan-Meier survival analysis was then performed for PFS and overall survival, with log-rank test statistics calculated.

Quantification and statistical analysis

Statistical analysis was performed in the R environment (v4.5.2). Statistical comparisons were performed using two-sided Wilcoxon rank-sum tests for unpaired data or Wilcoxon signed-rank tests for paired data. For parametric comparisons, two-sided Welch's t-tests were used for unpaired data and paired t-tests for paired data. Multiple comparisons were investigated by adjusting the p-value using the Benjamini–Hochberg method. Correlation analysis was carried out using the Pearson correlation coefficient (two-sided). Survival analysis was carried out using Kaplan-Meier curves with log-rank test. $P < 0.05$ was considered to indicate a statistically significant difference.

References

1. Jaffray, D. A. *et al.* Global Task Force on Radiotherapy for Cancer Control. *Lancet Oncol.* **16**, 1144–1146 (2015).
2. Liu, S. *et al.* Radiotherapy remodels the tumor microenvironment for enhancing immunotherapeutic sensitivity. *Cell Death Dis.* **14**, 679 (2023).
3. Wu, Y., Song, Y., Wang, R. & Wang, T. Molecular mechanisms of tumor resistance to radiotherapy. *Mol. Cancer* **22**, 96 (2023).
4. Barker, H. E., Paget, J. T. E., Khan, A. A. & Harrington, K. J. The Tumour Microenvironment after Radiotherapy: Mechanisms of Resistance and Recurrence. *Nat. Rev. Cancer* **15**, 409–425 (2015).
5. Ward, M. J. *et al.* Tumour-infiltrating lymphocytes predict for outcome in HPV-positive oropharyngeal cancer. *Br. J. Cancer* **110**, 489–500 (2014).
6. Muehlebach, M. E. *et al.* Radiation-Induced Fibrosis (RIF) in Head and Neck Squamous Cell Carcinoma (HNSCC): A Review. *Cells* **14**, 1969 (2025).

7. Contrera, K. J. *et al.* Locoregional and distant recurrence for HPV-associated oropharyngeal cancer using AJCC 8 staging. *Oral Oncol.* **111**, 105030 (2020).
8. Jenkins, B. H. *et al.* Single cell and spatial analysis of immune-hot and immune-cold tumours identifies fibroblast subtypes associated with distinct immunological niches and positive immunotherapy response. *Mol. Cancer* **24**, 3 (2025).
9. Puram, S. V. *et al.* Cellular states are coupled to genomic and viral heterogeneity in HPV-related oropharyngeal carcinoma. *Nat. Genet.* **55**, 640–650 (2023).
10. Molina, M. A., Steenbergen, R. D. M., Pumpe, A., Kenyon, A. N. & Melchers, W. J. G. HPV integration and cervical cancer: a failed evolutionary viral trait. *Trends Mol. Med.* **30**, 890–902 (2024).
11. Kurth, I. *et al.* Cancer stem cell related markers of radioresistance in head and neck squamous cell carcinoma. *Oncotarget* **6**, 34494–34509 (2015).
12. Li, Y.-L. *et al.* GDF15 contributes to radioresistance and cancer stemness of head and neck cancer by regulating cellular reactive oxygen species via a SMAD-associated signaling pathway. *Oncotarget* **8**, 1508–1528 (2016).
13. Chaudhary, R. *et al.* Prosurvival long noncoding RNA PINCR regulates a subset of p53 targets in human colorectal cancer cells by binding to Matrin 3. *eLife* **6**, e23244 (2017).
14. Tang, K.-W., Alaei-Mahabadi, B., Samuelsson, T., Lindh, M. & Larsson, E. The landscape of viral expression and host gene fusion and adaptation in human cancer. *Nat. Commun.* **4**, 2513 (2013).
15. Tabib, T., Morse, C., Wang, T., Chen, W. & Lafyatis, R. SFRP2/DPP4 and FMO1/LSP1 Define Major Fibroblast Populations in Human Skin. *J. Invest. Dermatol.* **138**, 802–810 (2018).
16. Tabib, T. *et al.* Myofibroblast transcriptome indicates SFRP2hi fibroblast progenitors in systemic sclerosis skin. *Nat. Commun.* **12**, 4384 (2021).
17. Steele, L. *et al.* A single-cell and spatial genomics atlas of human skin fibroblasts reveals shared disease-related fibroblast subtypes across tissues. *Nat. Immunol.* **26**, 1807–1820 (2025).
18. Philippeos, C. *et al.* Spatial and Single-Cell Transcriptional Profiling Identifies Functionally Distinct Human Dermal Fibroblast Subpopulations. *J. Invest. Dermatol.* **138**, 811–825 (2018).
19. Abbasi, S. *et al.* Distinct Regulatory Programs Control the Latent Regenerative Potential of Dermal Fibroblasts during Wound Healing. *Cell Stem Cell* **27**, 396-412.e6 (2020).

20. Saul, D. *et al.* A new gene set identifies senescent cells and predicts senescence-associated pathways across tissues. *Nat. Commun.* **13**, 4827 (2022).
21. Chen, A. M. De-escalation for Human Papillomavirus-Positive Oropharyngeal Cancer: A Look at the Prospective Evidence. *Curr. Oncol. Rep.* **27**, 355–361 (2025).
22. Lowell, S., Jones, P., Le Roux, I., Dunne, J. & Watt, F. M. Stimulation of human epidermal differentiation by Delta–Notch signalling at the boundaries of stem-cell clusters. *Curr. Biol.* **10**, 491–500 (2000).
23. Talora, C., Sgroi, D. C., Crum, C. P. & Dotto, G. P. Specific down-modulation of Notch1 signaling in cervical cancer cells is required for sustained HPV-E6/E7 expression and late steps of malignant transformation. *Genes Dev.* **16**, 2252–2263 (2002).
24. Hafner, A., Bulyk, M. L., Jambhekar, A. & Lahav, G. The multiple mechanisms that regulate p53 activity and cell fate. *Nat. Rev. Mol. Cell Biol.* **20**, 199–210 (2019).
25. Liu, Y., Leslie, P. L. & Zhang, Y. Life-and-Death Decision-Making by p53 and Implications in Cancer Immunotherapy. *Trends Cancer* **7**, 226–239 (2021).
26. Labrie, M., Brugge, J. S., Mills, G. B. & Zervantonakis, I. K. Therapy resistance: opportunities created by adaptive responses to targeted therapies in cancer. *Nat. Rev. Cancer* **22**, 323–339 (2022).
27. Zhang, Y. *et al.* Subtypes of HPV-positive head and neck cancers are associated with HPV characteristics, copy number alterations, PIK3CA mutation, and pathway signatures. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **22**, 4735–4745 (2016).
28. Nulton, T. J., Kim, N.-K., DiNardo, L. J., Morgan, I. M. & Windle, B. Patients with integrated HPV16 in head and neck cancer show poor survival. *Oral Oncol.* **80**, 52–55 (2018).
29. Öhlund, D. *et al.* Distinct populations of inflammatory fibroblasts and myofibroblasts in pancreatic cancer. *J. Exp. Med.* **214**, 579–596 (2017).
30. Biffi, G. *et al.* IL1-Induced JAK/STAT Signaling Is Antagonized by TGFβ to Shape CAF Heterogeneity in Pancreatic Ductal Adenocarcinoma. *Cancer Discov.* **9**, 282–301 (2019).
31. Ford, K. *et al.* NOX4 Inhibition Potentiates Immunotherapy by Overcoming Cancer-Associated Fibroblast-Mediated CD8 T-cell Exclusion from Tumors. *Cancer Res.* **80**, 1846–1860 (2020).
32. Cherian, M. G., Jayasurya, A. & Bay, B.-H. Metallothioneins in human tumors and potential roles in carcinogenesis. *Mutat. Res. - Fundam. Mol. Mech. Mutagen.* **533**, 201–209 (2003).
33. Si, M. & Lang, J. The roles of metallothioneins in carcinogenesis. *J. Hematol. Oncol.J Hematol Oncol* **11**, 107 (2018).

34. Lei, G. *et al.* Radiotherapy promotes cuproptosis and synergizes with cuproptosis inducers to overcome tumor radioresistance. *Cancer Cell* **43**, 1076-1092.e5 (2025).
35. Karin, N. CXCR3 Ligands in Cancer and Autoimmunity, Chemoattraction of Effector T Cells, and Beyond. *Front. Immunol.* **11**, 976 (2020).
36. Galbo, P. M., Zang, X. & Zheng, D. Molecular Features of Cancer-associated Fibroblast Subtypes and their Implication on Cancer Pathogenesis, Prognosis, and Immunotherapy Resistance. *Clin. Cancer Res.* **27**, 2636–2647 (2021).
37. Melero, I. *et al.* Neutralizing GDF-15 can overcome anti-PD-1 and anti-PD-L1 resistance in solid tumours. *Nature* **637**, 1218–1227 (2025).
38. Kumari, R. & Jat, P. Mechanisms of Cellular Senescence: Cell Cycle Arrest and Senescence Associated Secretory Phenotype. *Front. Cell Dev. Biol.* **9**, 645593 (2021).
39. Nguyen, H. Q. *et al.* Ionizing radiation-induced cellular senescence promotes tissue fibrosis after radiotherapy. A review. *Crit. Rev. Oncol. Hematol.* **129**, 13–26 (2018).
40. Jones, R. C. *et al.* Tabula Sapiens 2.0: A comprehensive transcriptomic atlas of human cell types. *Cell* **0**, (2026).
41. Duhen, T. *et al.* Co-expression of CD39 and CD103 identifies tumor-reactive CD8 T cells in human solid tumors. *Nat. Commun.* **9**, 2724 (2018).
42. Machiels, J.-P. *et al.* Pembrolizumab plus concurrent chemoradiotherapy versus placebo plus concurrent chemoradiotherapy in patients with locally advanced squamous cell carcinoma of the head and neck (KEYNOTE-412): a randomised, double-blind, phase 3 trial. *Lancet Oncol.* **25**, 572–587 (2024).
43. Aboaid, H. *et al.* Advances and challenges in immunotherapy in head and neck cancer. *Front. Immunol.* **16**, 1596583 (2025).
44. Porciello, N., Franzese, O., D'Ambrosio, L., Palermo, B. & Nisticò, P. T-cell repertoire diversity: friend or foe for protective antitumor response? *J. Exp. Clin. Cancer Res. CR* **41**, 356 (2022).
45. 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).
46. Hao, Y. *et al.* Dictionary learning for integrative, multimodal, and massively scalable single-cell analysis. *Nat. Biotechnol.* **42**, 293–304 (2024).
47. Schubert, M. *et al.* Perturbation-response genes reveal signaling footprints in cancer gene expression. *Nat. Commun.* **9**, 20 (2018).

48. Badia-i-Mompel, P. *et al.* decoupleR: ensemble of computational methods to infer biological activities from omics data. *Bioinforma. Adv.* **2**, vbac016 (2022).
49. Dann, E., Henderson, N. C., Teichmann, S. A., Morgan, M. D. & Marioni, J. C. Differential abundance testing on single-cell data using k-nearest neighbor graphs. *Nat. Biotechnol.* **40**, 245–253 (2022).
50. Yerly, L. *et al.* Wounding triggers invasive progression in human basal cell carcinoma. 2024.05.31.596823 Preprint at https://doi.org/10.1101/2024.05.31.596823 (2025).
51. Kang, M. *et al.* Improved reconstruction of single-cell developmental potential with CytoTRACE 2. *Nat. Methods* **22**, 2258–2263 (2025).
52. Buechler, M. B. *et al.* Cross-tissue organization of the fibroblast lineage. *Nature* **593**, 575–579 (2021).
53. Elyada, E. *et al.* Cross-Species Single-Cell Analysis of Pancreatic Ductal Adenocarcinoma Reveals Antigen-Presenting Cancer-Associated Fibroblasts. *Cancer Discov.* **9**, 1102–1123 (2019).
54. Williams, D. W. *et al.* Human oral mucosa cell atlas reveals a stromal-neutrophil axis regulating tissue immunity. *Cell* **184**, 4090-4104.e15 (2021).
55. Borcherding, N., Bormann, N. L. & Kraus, G. scRepertoire: An R-based toolkit for single-cell immune receptor analysis. *F1000Research* **9**, 47 (2020).
56. Goncharov, M. *et al.* VDJdb in the pandemic era: a compendium of T cell receptors specific for SARS-CoV-2. *Nat. Methods* **19**, 1017–1019 (2022).
57. Tickotsky, N., Sagiv, T., Prilusky, J., Shifrut, E. & Friedman, N. McPAS-TCR: a manually curated catalogue of pathology-associated T cell receptor sequences. *Bioinformatics* **33**, 2924–2929 (2017).
58. Gielis, S. *et al.* Detection of Enriched T Cell Epitope Specificity in Full T Cell Receptor Sequence Repertoires. *Front. Immunol.* **10**, (2019).
59. Orenbuch, R. *et al.* arcasHLA: high-resolution HLA typing from RNAseq. *Bioinformatics* **36**, 33–40 (2020).
60. Huang, H., Wang, C., Rubelt, F., Scriba, T. J. & Davis, M. M. Analyzing the Mycobacterium tuberculosis immune response by T-cell receptor clustering with GLIPH2 and genome-wide antigen screening. *Nat. Biotechnol.* **38**, 1194–1202 (2020).
61. Cerami, E. *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2**, 401–404 (2012).

62. Cable, D. M. *et al*. Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat. Biotechnol*. **40**, 517–526 (2022).
63. Uhlén, M. *et al*. Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- JenkinsetalSupplementalInformation.pdf
- JenkinsetalSupplementalInformation.pdf
- JenkinsetalSupplementaryTablesS1.xlsx