Supplemental information

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

Supplementary Figures

Figure S1.

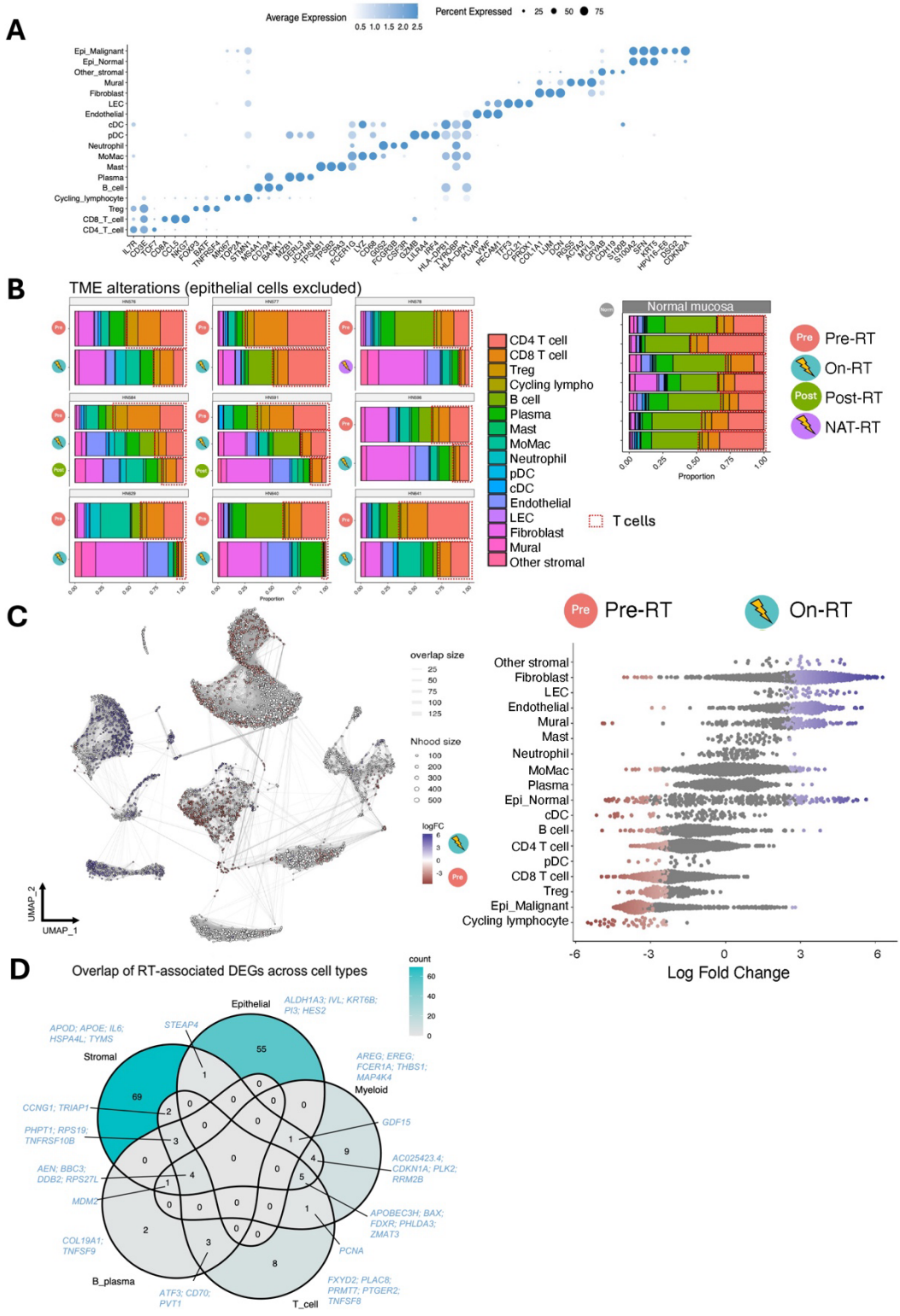


Figure S1 (related to Figure 1). A. Dot plot showing marker genes used to annotate the broad cell clusters shown in Figure 1B. **B**. Bar plots showing cell type proportions across samples, with epithelial cells excluded. T cell proportions (CD4 T cells, CD8 T cells and Treg; red boxes) show a substantial decrease during RT. **C**. MiloR differential abundance analysis showing significant cell neighbourhoods across cell types. SpatialFDR<0.15. **D**. Venn diagram showing overlapping differentially expressed genes (log_2[FC]>1, adj.P<0.05) upregulated in response to RT across broad cell lineages. Genes included if upregulated in ≥50% of cell types within each broad lineage.

Figure S2.

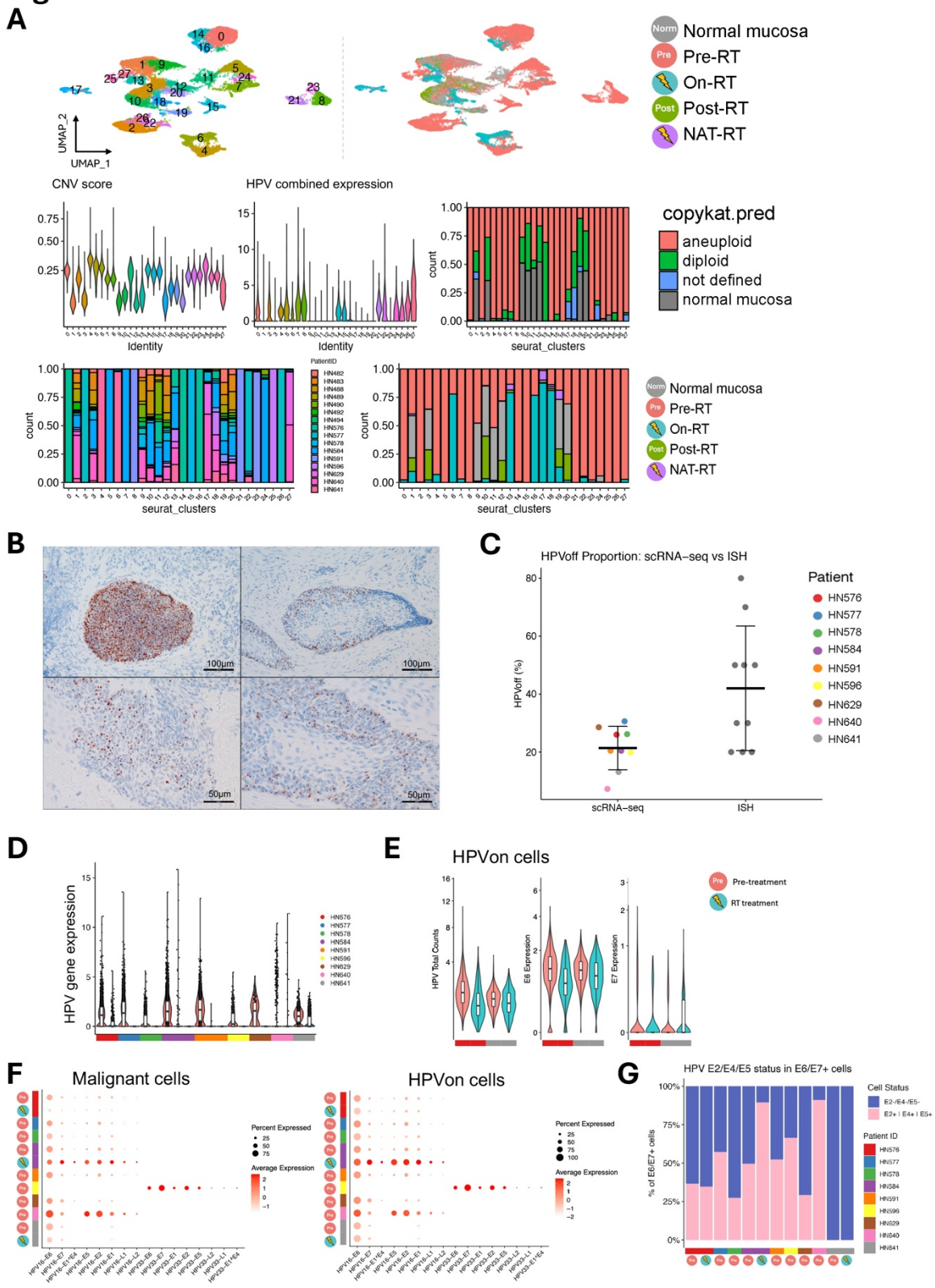

A
Norm Normal mucosa
Pre Pre-RT
On On-RT
Post Post-RT
NAT NAT-RT
UMAP_1
UMAP_2
CNV score
HPV combined expression
Identity
seurat_clusters
count
copykat.pred
aneuploid
diploid
not defined
normal mucosa
PatientID
B
100μm
50μm
C
HPVoff Proportion: scRNA–seq vs ISH
HPVoff (%)
scRNA–seq
ISH
Patient
HN576
HN577
HN578
HN584
HN591
HN596
HN629
HN640
HN641
D
HPV gene expression
E
HPVon cells
HPV Total Counts
E6 Expression
E7 Expression
Pre-treatment
RT treatment
F
Malignant cells
HPVon cells
Percent Expressed
Average Expression
G
HPV E2/E4/E5 status in E6/E7+ cells
% of E6/E7+ cells
Cell Status
E2-/E4-/E5-
E2+ | E4+ | E5+
Patient ID

Figure S2 (related to Figure 2). A. Series of analyses informing malignant-normal cell discrimination. Top, UMAP plots annotated by Seurat cluster and treatment conditions. Middle (from left to right), copy number variation scores calculated using inferCNV, combined expression of all HPV transcripts per cluster, bar plot comprising of copykat predictions per cluster. Bottom, patient and source of sample breakdown per cluster. **B.** Representative high-power RNAScope™ images from HPV-positive HNSCCs demonstrating heterogeneous expression of HPV E6/E7 transcripts. **C.** Percent of HPVoff malignant cells in pre-RT scRNA-seq (n=9) and HPV ISH (n=10) HNSCC cohorts (mean ± SD). **D**. Normalised HPV gene expression (all transcripts) per sample across epithelial cells. **E**. Normalised expression of all HPV transcripts, *E6* and *E7* across HPVon cells in RMC^{high} patients across pre-RT and on-RT samples. **F**. Expression of individual HPV genes across patient samples in all malignant cells (left) and in HPVon malignant cells only (right). **G**. Proportion of *E2*, *E4* or *E5* expressing *E6/E7*+ malignant cells across samples.

Figure S3.

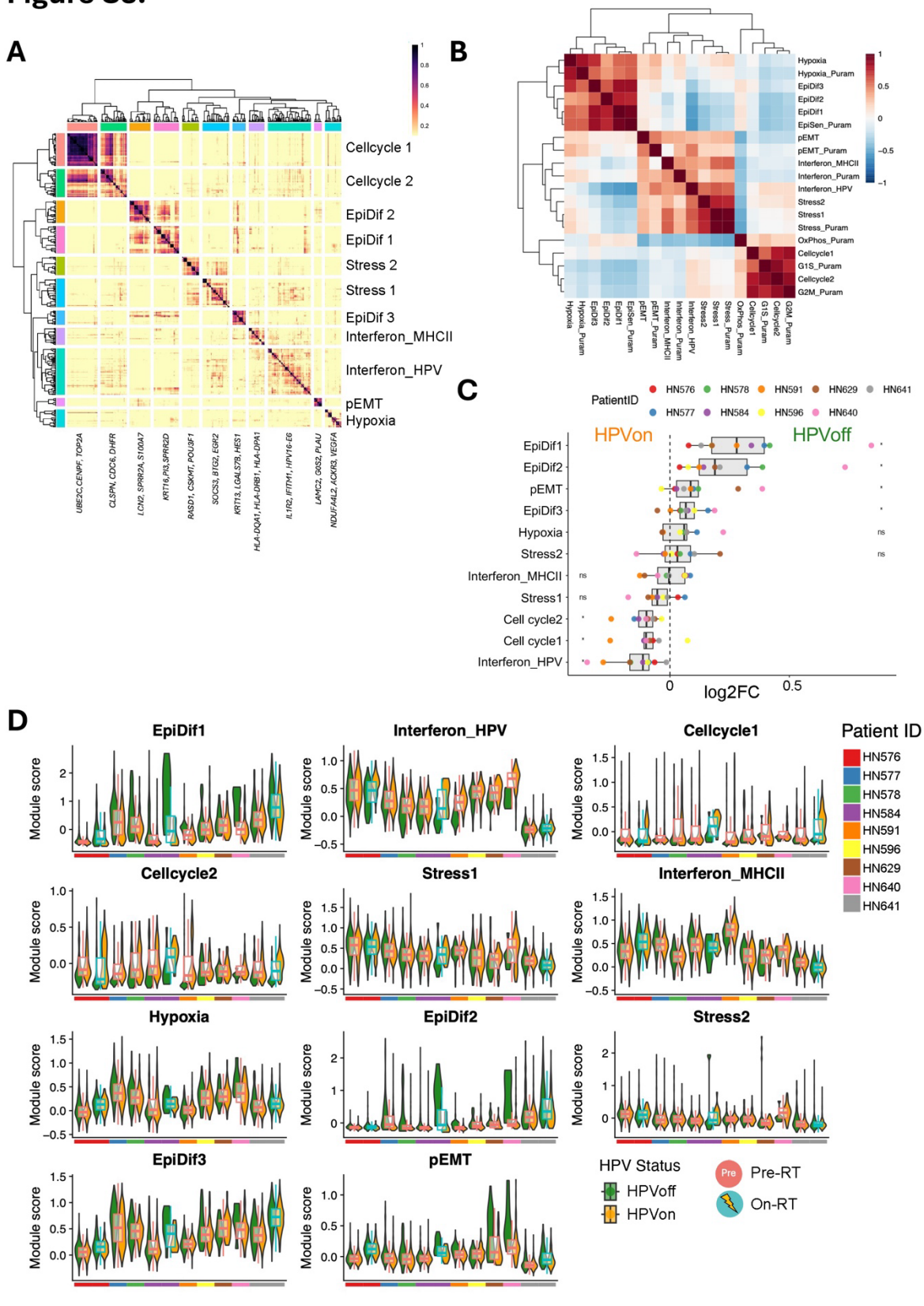

A
Cellcycle 1
Cellcycle 2
EpiDif 2
EpiDif 1
Stress 2
Stress 1
EpiDif 3
Interferon_MHCII
Interferon_HPV
pEMT
Hypoxia
UBE2C,CENPF,TOP2A
CLSPN,CDC6,DHFR
LCN2,SPRR2A,S100A7
KRT16,PI3,SPRR2D
RASD1,CSKMT,POU3F1
SOCS3,BTG2,EGR2
KRT13,LGALS7B,HES1
HLA-DQA1,HLA-DRB1,HLA-DPA1
IL1R2,IFITM1,HPV16-E6
LAMC2,G0S2,PLAU
NDUFA4L2,ACKR3,VEGFA
B
Hypoxia
Hypoxia_Puram
EpiDif3
EpiDif2
EpiDif1
EpiSen_Puram
pEMT
pEMT_Puram
Interferon_MHCII
Interferon_Puram
Interferon_HPV
Stress2
Stress1
Stress_Puram
OxPhos_Puram
Cellcycle1
G1S_Puram
Cellcycle2
G2M_Puram
C
PatientID
HN576
HN578
HN591
HN629
HN641
HN577
HN584
HN596
HN640
HPVon
HPVoff
EpiDif1
EpiDif2
pEMT
EpiDif3
Hypoxia
Stress2
Interferon_MHCII
Stress1
Cell cycle2
Cell cycle1
Interferon_HPV
ns
log2FC
0
0.5
D
EpiDif1
Interferon_HPV
Cellcycle1
Cellcycle2
Stress1
Interferon_MHCII
Hypoxia
EpiDif2
Stress2
EpiDif3
pEMT
Module score
Patient ID
HN576
HN577
HN578
HN584
HN591
HN596
HN629
HN640
HN641
HPV Status
HPVoff
HPVon
Pre
Pre-RT
On-RT

Figure S3 (related to Figure 2). A. Non-negative matrix factorisation (NMF) clustered heatmap reflecting pairwise cosine similarity between individual gene programs. Example genes for each metaprogramme shown. **B**. Clustered heatmap reflecting Pearson correlations between identified metaprogrammes and previously described HNSCC metaprogrammes[1]. **C**. Fold change in mean enrichment scores of metaprogrammes in HPVon and HPVoff cells per patient (pre-treatment malignant cells only). Wilcoxon rank-sum test. P values adjusted using the Benjamini–Hochberg method. *p<0.05. **D.** Metaprogramme enrichment across all malignant cells split by HPVon/off status.

Figure S4.

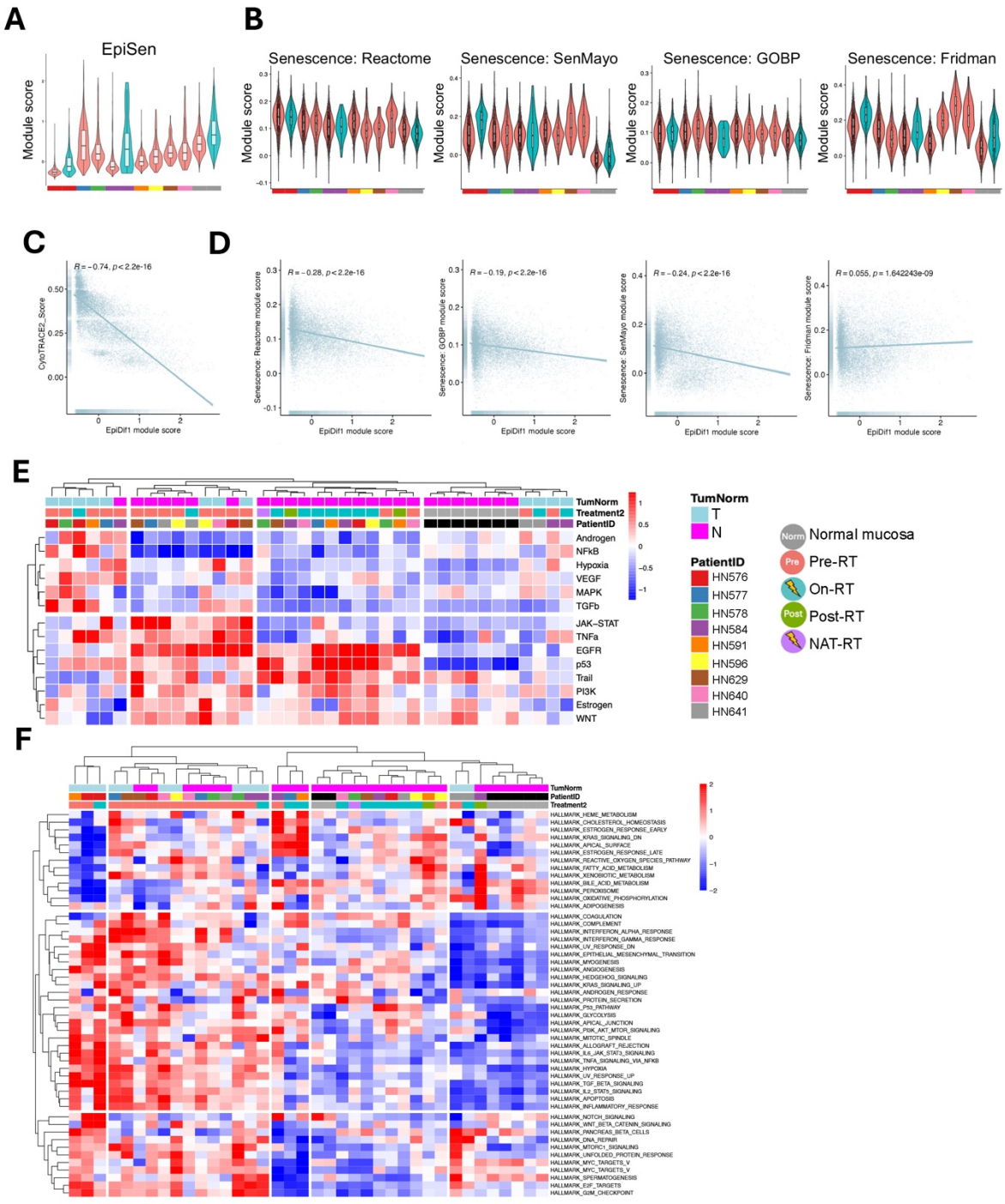


Figure S4 (related to Figure 2). A. Module score enrichment of epithelial senescence (EpiSen) metaprogramme[1] across malignant cells. **B**. Module score enrichment for various senescence gene sets (MSigDB Reactome, GOBP, Fridman and SenMayo) across malignant cells. **C**. Pearson correlation between EpiDif1 module scores and CytoTRACE2 scores in malignant cells. **D**. Pearson correlation between EpiDif1 module scores and various senescence gene set module scores in

malignant cells. **E**. PROGENy pathway analysis across all samples of normal and malignant cells. **F**. Enrichment of MSigDB hallmark pathways across all epithelial cells.

Figure S5.

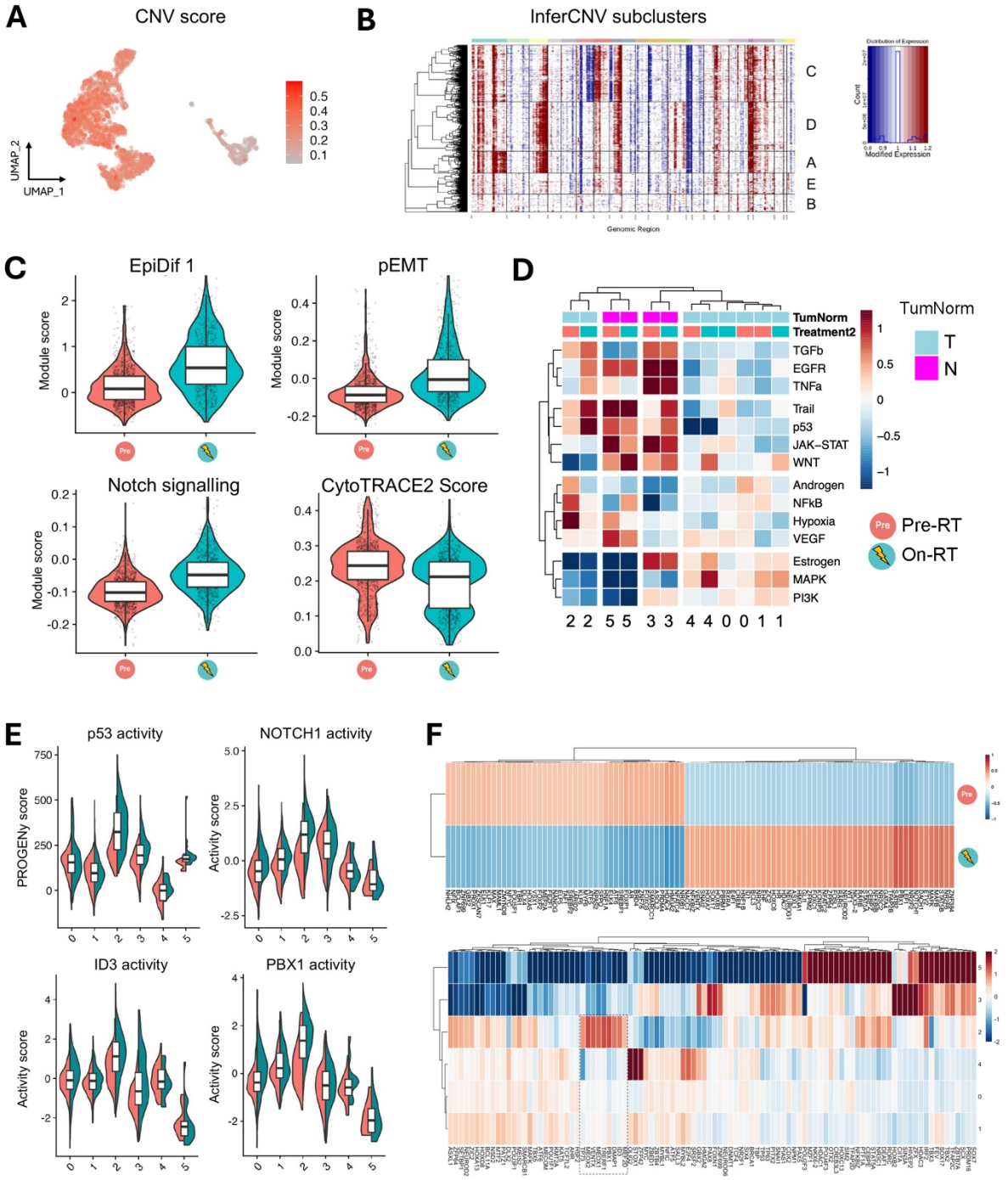


Figure S5 (related to Figure 3). A. UMAP plot of HN641 epithelial cells coloured by InferCNV scores. **B**. InferCNV heatmap for HN641 epithelial cells displaying CNV clusters. **C**. Enrichment scores for

EpiDif1, pEMT, notch signalling (MSigDB hallmark) and CytoTRACE2 scores between pre-RT and on-RT malignant cells in HN641. **D**. Heatmap of PROGENy pathway activities for each cluster split by pre-RT and on-RT. **E**. p53 pathway and NOTCH1, ID3 and PBX1 transcription factor activity scores for each HN641 cluster split by pre-RT and on-RT. **F**. Full output of top 100 transcripton factors with greatest activity in HN641 between pre-RT and on-RT (top) or across clusters (bottom).

Figure S6.

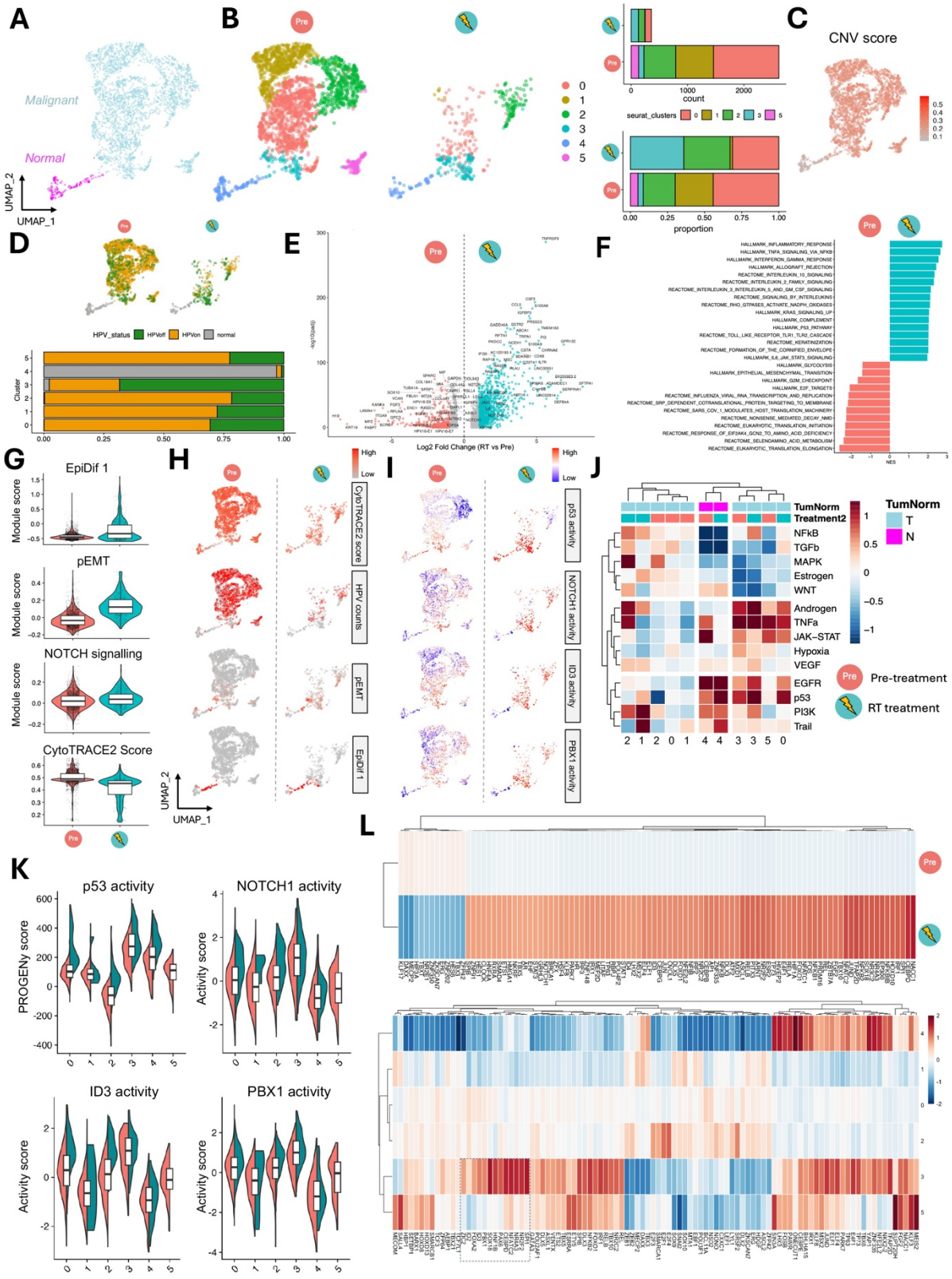

A
Malignant
Normal
UMAP_2
UMAP_1
B
Pre
0
1
2
3
4
5
count
seurat_clusters
proportion
C
CNV score
D
HPV_status
HPVoff
HPVon
normal
Cluster
E
Log2 Fold Change (RT vs Pre)
F
NES
G
EpiDif 1
pEMT
NOTCH signalling
CytoTRACE2 Score
Module score
H
High
Low
CytoTRACE2 score
HPV counts
pEMT
EpiDif 1
I
p53 activity
NOTCH1 activity
ID3 activity
PBX1 activity
J
TumNorm
Treatment2
NFkB
TGFb
MAPK
Estrogen
WNT
Androgen
TNFa
JAK–STAT
Hypoxia
VEGF
EGFR
p53
PI3K
Trail
T
N
Pre-treatment
RT treatment
K
p53 activity
NOTCH1 activity
ID3 activity
PBX1 activity
PROGENy score
Activity score
L

Figure S6 (related to Figure 3). A. UMAP plot of HN576 epithelial cells annotated as normal or malignant. **B**. Left, UMAP of HN576 epithelial cells split by pre-RT and on-RT. Right, bar plots of cluster abundance (absolute and proportion) across pre-RT and on-RT. **C**. UMAP of HN576 epithelial cells coloured by CNV score. **D**. Top, UMAP of HN576 epithelial cells split by pre-RT and on-RT and annotated by HPVon/off status. Bottom, bar plot of HPVon/off proportions per cluster. **E**. Volcano plot showing differentially expressed genes (DEGs) between on-RT and pre-RT malignant cells. **F**. Gene set enrichment analysis (GSEA) of selected significant terms (on-RT versus pre-RT). **G**. Enrichment scores for EpiDif1, pEMT, notch signalling (MSigDB hallmark) and CytoTRACE2 scores between pre-RT and on-RT malignant cells in HN576. **H.** UMAP displaying CytoTRACE2 scores, HPV counts and enrichment scores for EpiDif1 and pEMT metaprogrammes across HN576 epithelial cells. **I**. UMAP showing p53 pathway activity and NOTCH1, ID3 and PBX1 transcription factor activities across HN576 epithelial cells. **J**. Heatmap of PROGENy pathway activities for each cluster split by pre-RT and on-RT. **K**. p53 pathway and NOTCH1, ID3 and PBX1 transcription factor activity scores for each HN576 cluster split by pre-RT and on-RT. **L**. Full output of top 100 transcripton factors with greatest activity in HN576 between pre-RT and on-RT (top) or across clusters (bottom).

Figure S7.

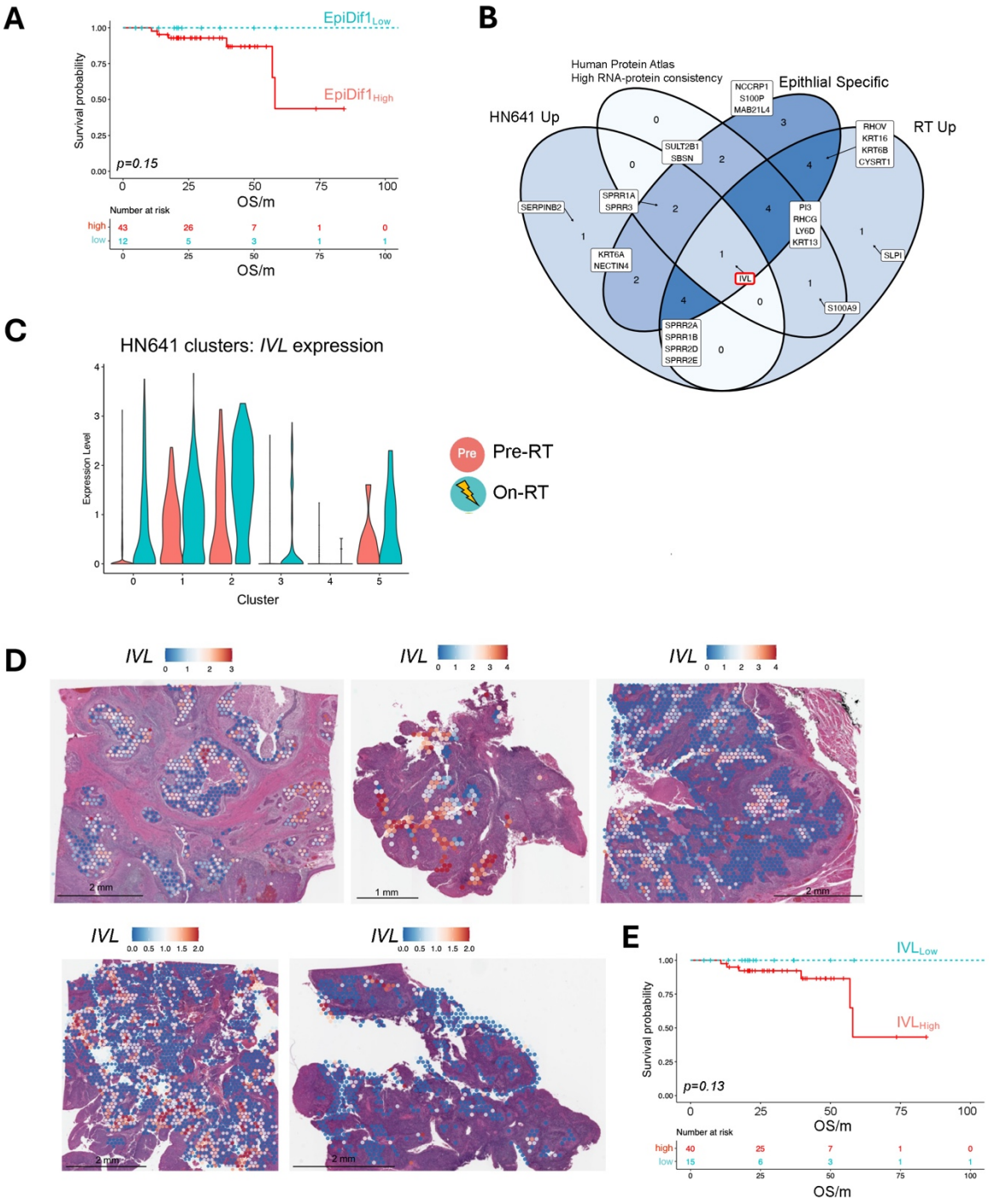


Figure S7 (related to Figure 4). A. Kaplan-Meier (overall survival; OS) plot showing HPV-positive HNSCC (n=45) treated with RT (TCGA cohort) stratified by EpiDif1 (ssGSEA) scores. **B**. Venn diagram of EpiDif1 signature genes relating to Figure 4B. **C**. *IVL* expression across HN641 epithelial cell clusters, split by pre-RT and on-RT cells (Figure 3A). **D**. *IVL* expression across spatial transcriptomics

spots (10x Visium) in five HPV-positive HNSCC sections. Expression limited to epithelial cells. **E.** Kaplan-Meier OS plot showing HPV-positive HNSCC (n=45) treated with RT (TCGA cohort) stratified by *IVL* expression.

Figure S8.

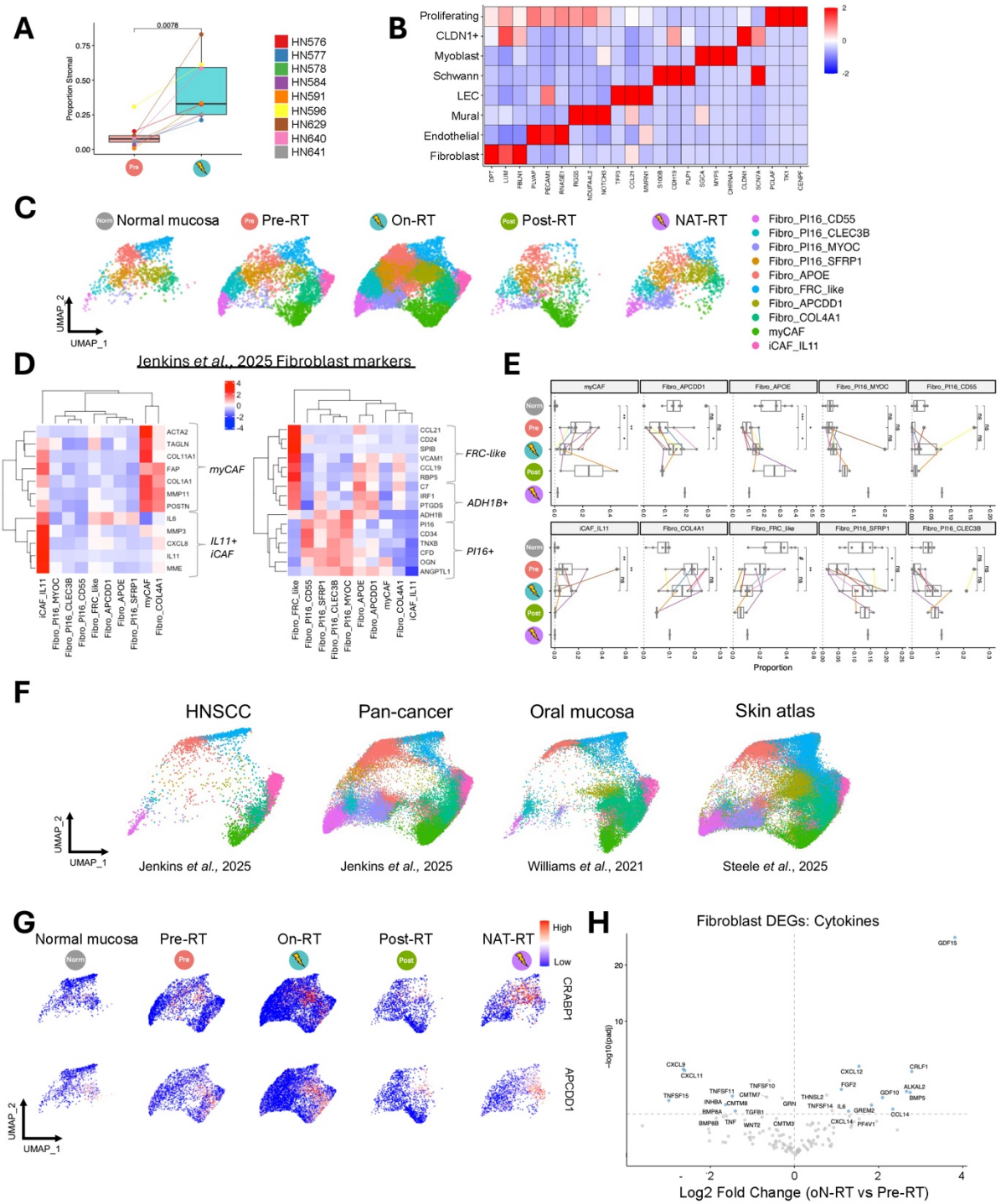


Figure S8 (related to Figure 5). A. Proportion of stromal cells (relative to total cells) in pre-RT and on-

RT samples. Wilcoxon signed-rank test. **B**. Heatmap of marker genes for broad stromal cell types. **C**. UMAP plot of fibroblast clusters split by condition. **D**. Heatmap of marker genes distinguishing previously described HNSCC fibroblast clusters[2]. **E**. Proportion of each fibroblast cluster across samples. Wilcoxon rank-sum tests were used for unpaired comparisons and Wilcoxon signed-rank tests for paired comparisons. #p<0.10, *p<0.05, **p<0.01, ***p<0.001. **F**. Reference mapping (Seurat) of various datasets across HNSCC, pan-cancer, oral mucosa and skin. **G**. UMAP plots of fibroblasts split by treatment condition showing the expression of *CRABP1* and *APCDD1*. **H**. Volcano plot showing pseudobulk differentially expressed genes (Table S3) in fibroblasts on-RT vs pre-RT filtered for GOMF_CYTOKINE ACTIVITY MSigDB gene set. |log2(FC)| > 1; padj<0.1.

Figure S9.

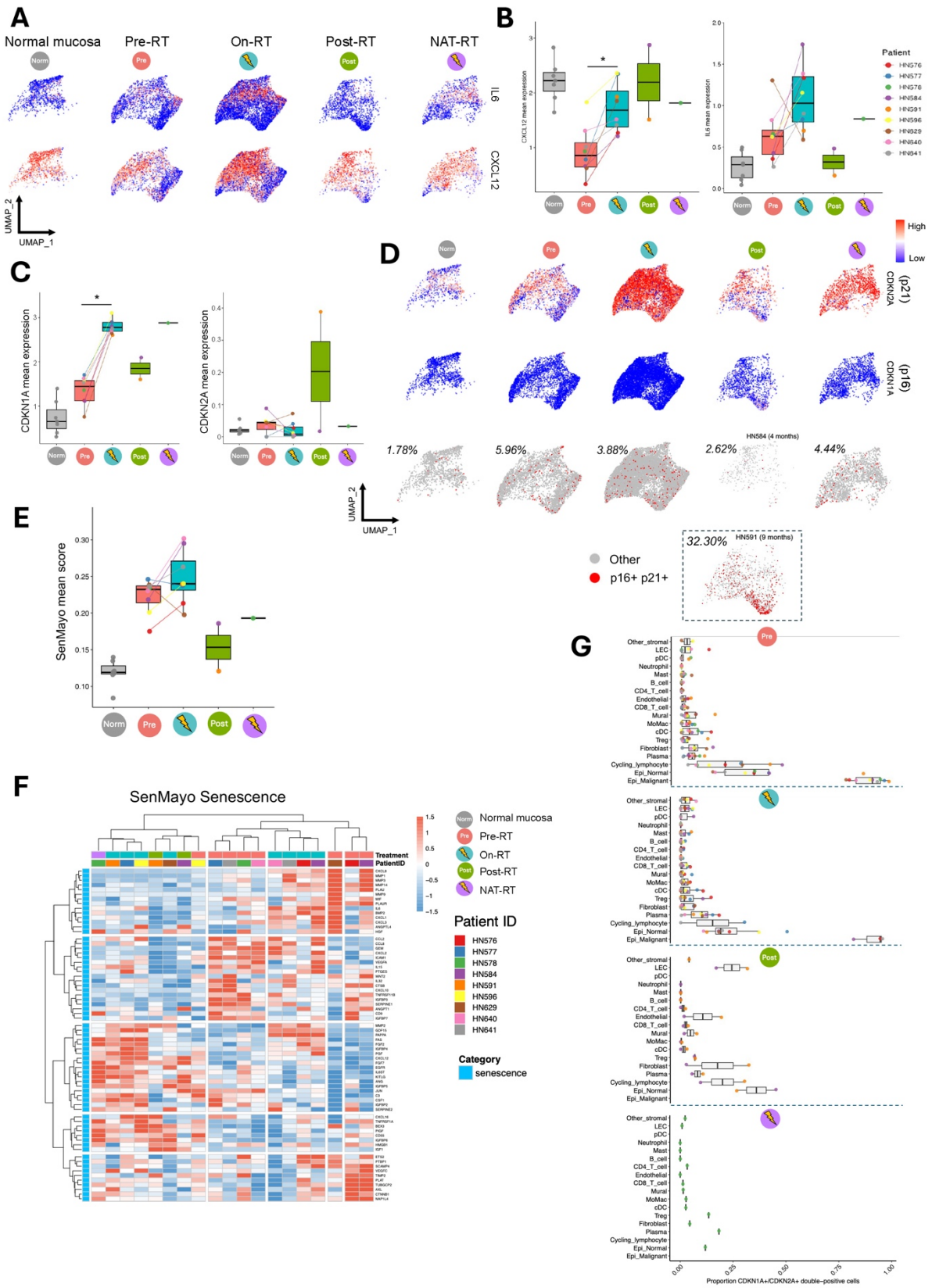

A
Normal mucosa
Pre-RT
On-RT
Post-RT
NAT-RT
IL6
CXCL12
UMAP_2
UMAP_1
B
CXCL12 mean expression
IL6 mean expression
Patient
HN576
HN577
HN578
HN584
HN591
HN596
HN629
HN640
HN641
C
CDKN1A mean expression
CDKN2A mean expression
D
High
Low
CDKN2A (p21)
CDKN1A (p16)
1.78%
5.96%
3.88%
2.62%
4.44%
HN584 (4 months)
32.30%
HN591 (9 months)
Other
p16+ p21+
E
SenMayo mean score
F
SenMayo Senescence
Treatment
PatientID
Normal mucosa
Pre-RT
On-RT
Post-RT
NAT-RT
Patient ID
Category
senescence
G
Other_stromal
LEC
pDC
Neutrophil
Mast
B_cell
CD4_T_cell
Endothelial
CD8_T_cell
Mural
MoMac
cDC
Treg
Fibroblast
Plasma
Cycling_lymphocyte
Epi_Normal
Epi_Malignant
Proportion CDKN1A+/CDKN2A+ double-positive cells

Figure S9 (related to Figure 5). A. UMAP plots of fibroblasts split by treatment condition showing the expression of *IL6* and *CXCL12*. **B**. Average sample-level expression of *IL6* and *CXCL12* across conditions. Wilcoxon signed-rank test. *p<0.05. **C**. Average sample-level expression of *CDKN1A* (p21) and *CDKN2A* (p16) across conditions. Wilcoxon signed-rank test. *p<0.05. **D**. Top two rows, UMAP plots of fibroblasts split by treatment condition showing the expression of *CDKN1A* (p21) and *CDKN2A* (p16). Bottom row, UMAP plot with cells expressing both *CDKN1A* (p21) and *CDKN2A* (p16) marked red, with percentage of double-positive cells labelled. Post-treatment condition split showing differing percentages in each sample. **E**. Average module scores for SenMayo senescence gene set across treatment conditions. **F**. Heatmap showing average expression of SenMayo senescence genes across samples. **G**. Proportion of cells expressing both *CDKN1A* (p21) and *CDKN2A* (p16) across broad cell types, per sample.

Figure S10.

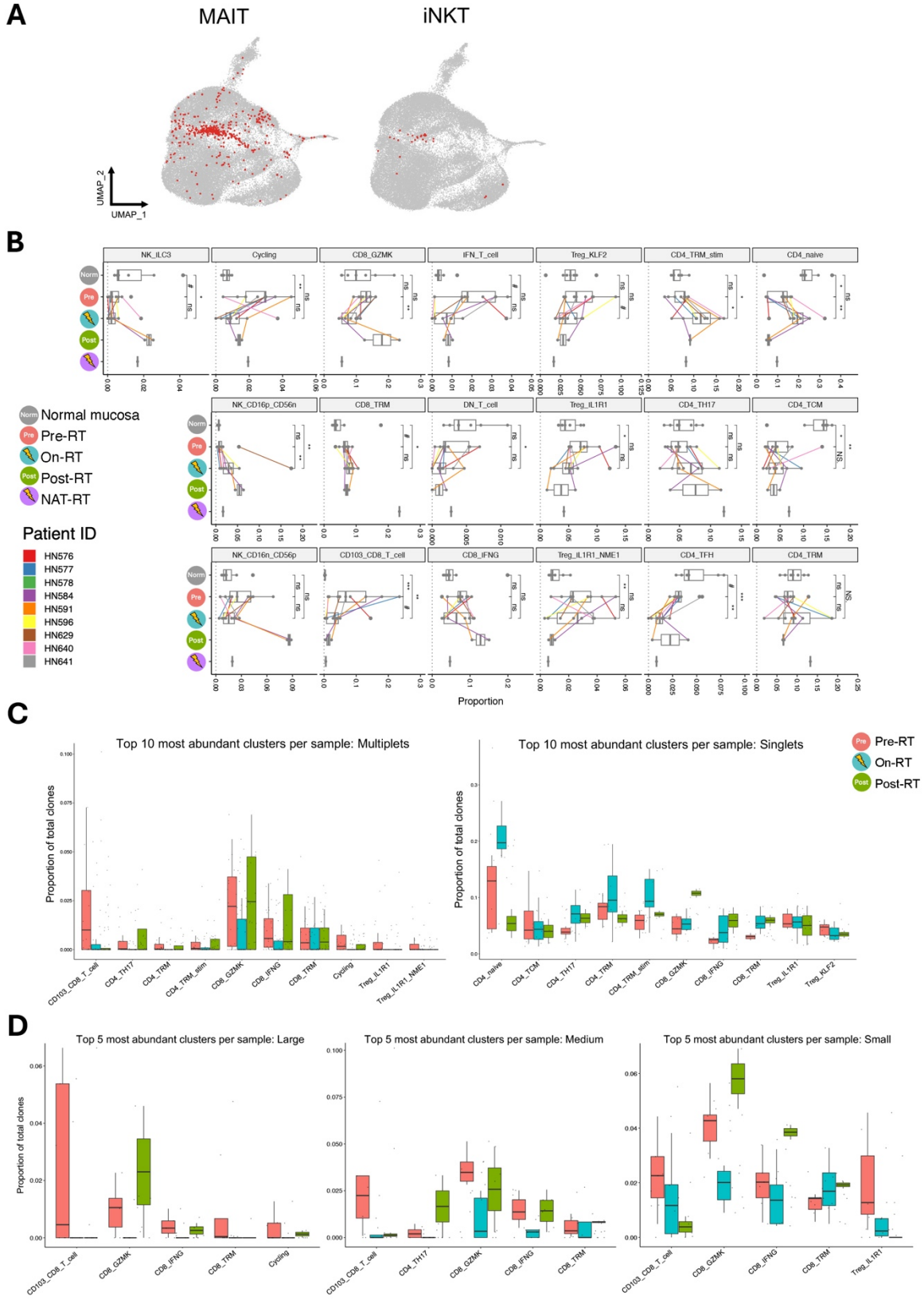

A
MAIT
iNKT
UMAP_2
UMAP_1
B
NK_ILC3
Cycling
CD8_GZMK
IFN_T_cell
Treg_KLF2
CD4_TRM_stim
CD4_naive
NK_CD16p_CD56n
CD8_TRM
DN_T_cell
Treg_IL1R1
CD4_TH17
CD4_TCM
NK_CD16n_CD56p
CD103_CD8_T_cell
CD8_IFNG
Treg_IL1R1_NME1
CD4_TFH
CD4_TRM
Normal mucosa
Pre-RT
On-RT
Post-RT
NAT-RT
Patient ID
HN576
HN577
HN578
HN584
HN591
HN596
HN629
HN640
HN641
Proportion
C
Top 10 most abundant clusters per sample: Multiplets
Top 10 most abundant clusters per sample: Singlets
Proportion of total clones
Pre-RT
On-RT
Post-RT
D
Top 5 most abundant clusters per sample: Large
Top 5 most abundant clusters per sample: Medium
Top 5 most abundant clusters per sample: Small
Proportion of total clones

Figure S10 (related to Figure 6). A. UMAP plot of T/NK cells with mucosal-associated invariant T (MAIT) cells and invariant natural killer T (iNKT) cells marked in red. Identified using TCR sequences (see methods). **B**. Proportion of each T/NK cell cluster across samples. Wilcoxon rank-sum tests were used for unpaired comparisons and Wilcoxon signed-rank tests for paired comparisons. #p<0.10, *p<0.05, **p<0.01, ***p<0.001. **C**. Top 10 most abundant clusters per sample split into multiplets (left) and singlets (right). **D**. Top 5 most abundant clusters per sample for Large (20 < X <= 100), Medium (5 < X <= 20) and Small (1 < X <= 5), clonotypes.

Figure S11.

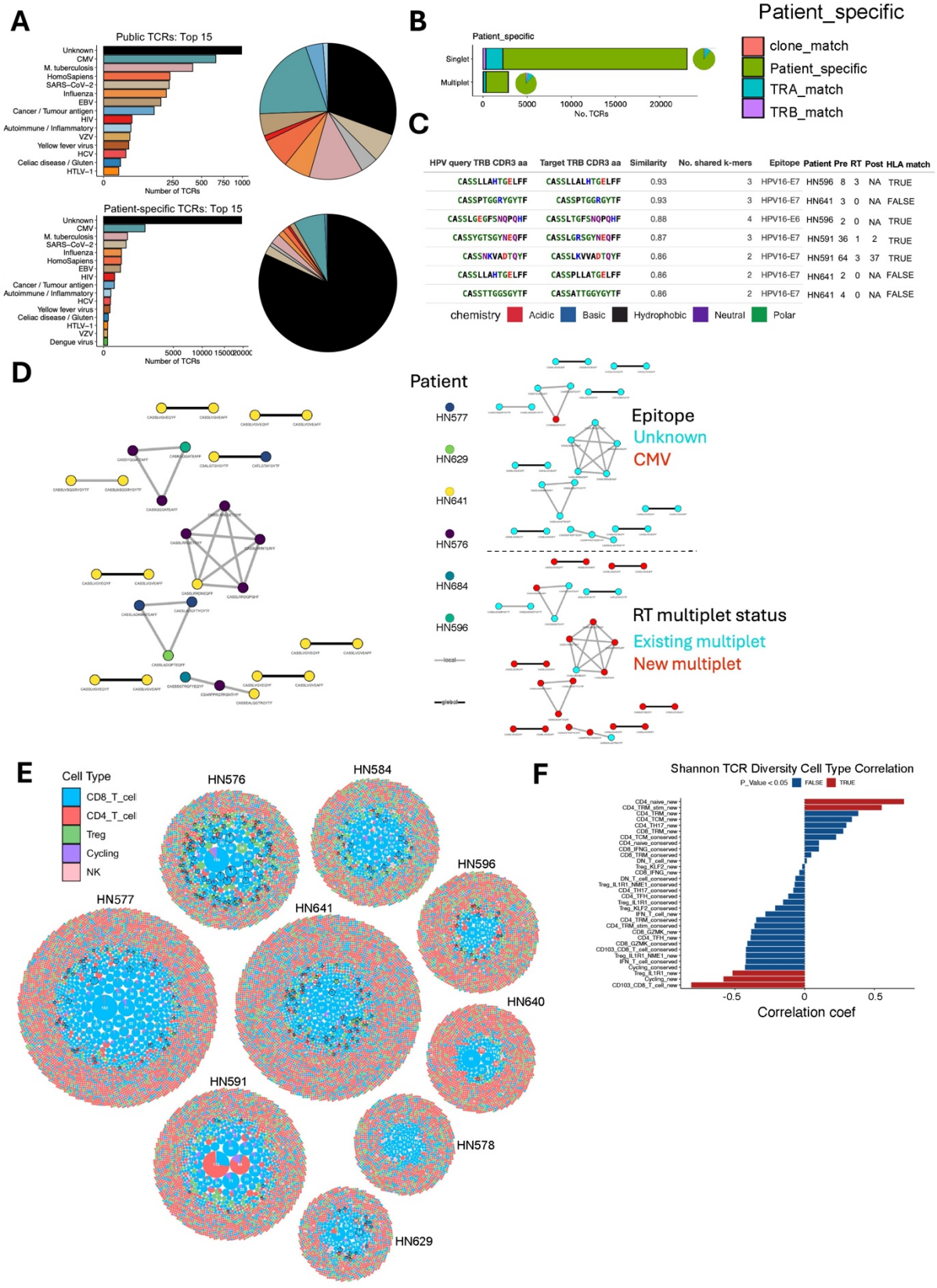

A
Public TCRs: Top 15
Unknown
CMV
M. tuberculosis
HomoSapiens
SARS–CoV–2
Influenza
EBV
Cancer / Tumour antigen
HIV
Autoimmune / Inflammatory
VZV
Yellow fever virus
HCV
Celiac disease / Gluten
HTLV–1
0
250
500
750
1000
Number of TCRs
Patient-specific TCRs: Top 15
Unknown
CMV
M. tuberculosis
SARS–CoV–2
Influenza
HomoSapiens
EBV
HIV
Cancer / Tumour antigen
Autoimmune / Inflammatory
HCV
Yellow fever virus
Celiac disease / Gluten
HTLV–1
VZV
Dengue virus
0
5000
10000
15000
20000
Number of TCRs
B
Patient_specific
Singlet
Multiplet
0
5000
10000
15000
20000
No. TCRs
Patient_specific
clone_match
Patient_specific
TRA_match
TRB_match
C
HPV query TRB CDR3 aa
Target TRB CDR3 aa
Similarity
No. shared k-mers
Epitope
Patient
Pre
RT
Post
HLA match
CASSLLAHTGELFF
CASSLLALHTGELFF
0.93
3
HPV16-E7
HN596
8
3
NA
TRUE
CASSPTGGRYGYTF
CASSPTGGRGYTF
0.93
3
HPV16-E7
HN641
3
0
NA
FALSE
CASSLGEGFSNQPQHF
CASSLTGFSNQPQHF
0.88
4
HPV16-E6
HN596
2
0
NA
TRUE
CASSYGTSGYNEQFF
CASSLGRSGYNEQFF
0.87
3
HPV16-E7
HN591
36
1
2
TRUE
CASSNKVADTQYF
CASSLKVVADTQYF
0.86
2
HPV16-E7
HN591
64
3
37
TRUE
CASSLLAHTGELFF
CASSPLLATGELFF
0.86
2
HPV16-E7
HN641
2
0
NA
FALSE
CASSTTGGSGYTF
CASSATTGGYGYTF
0.86
2
HPV16-E7
HN641
4
0
NA
FALSE
chemistry
Acidic
Basic
Hydrophobic
Neutral
Polar
D
Patient
HN577
HN629
HN641
HN576
HN684
HN596
local
global
Epitope
Unknown
CMV
RT multiplet status
Existing multiplet
New multiplet
E
Cell Type
CD8_T_cell
CD4_T_cell
Treg
Cycling
NK
HN576
HN584
HN596
HN577
HN641
HN640
HN591
HN578
HN629
F
Shannon TCR Diversity Cell Type Correlation
P_Value < 0.05
FALSE
TRUE
CD4_naive_new
CD4_TRM_stim_new
CD4_TRM_new
CD4_TCM_new
CD4_TH17_new
CD8_TRM_new
CD4_TCM_conserved
CD4_naive_conserved
CD8_IFNG_conserved
CD8_TRM_conserved
DN_T_cell_new
Treg_KLF2_new
CD8_IFNG_new
DN_T_cell_conserved
Treg_IL1R1_NME1_conserved
CD4_TH17_conserved
CD4_TFH_conserved
Treg_IL1R1_conserved
Treg_KLF2_conserved
IFN_T_cell_new
CD4_TRM_conserved
CD4_TRM_stim_conserved
CD8_GZMK_new
CD4_TFH_new
CD8_GZMK_conserved
CD103_CD8_T_cell_conserved
Treg_IL1R1_NME1_new
IFN_T_cell_conserved
Cycling_conserved
Treg_IL1R1_new
Cycling_new
CD103_CD8_T_cell_new
-0.5
0
0.5
Correlation coef

Figure S11 (related to Figure 6). A. Top 15 most frequent TCR-epitope database matches across public (>1 patient) and patient-specific clonotypes. **B**. Number of unique clonotypes with a shared TCR-sequence across the cohort, split by multiplet status. **C**. Dataset TRB sequences with closest match to known HPV-epitope binding (in exisiting databases). TRB CDR3 sequence similarity was calculated using length-normalised Levenshtein distance. Pre, RT and Post reflecting number of cells with TRB sequence across conditions. HLA status inferred using arcasHLA. 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). **D**. GLIPH2 analysis of on-RT multiplets across samples, annotated by patient, epitope-match and new/existing clonotype multiplet status. **E**. Summary clonotype plots for each patient, displaying all unique clonotypes and respective sizes. Size of clonotype (number of cells) reflected by size of circles, coloured by broad cell type, with the largest clones in the centre 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. **F**. Pearson correlation output for each cell type proportion (split into new and conserved based on clonotype presence across pre-RT and on-RT) against shannon diversity index. Sample-level correlation.

Figure S12.

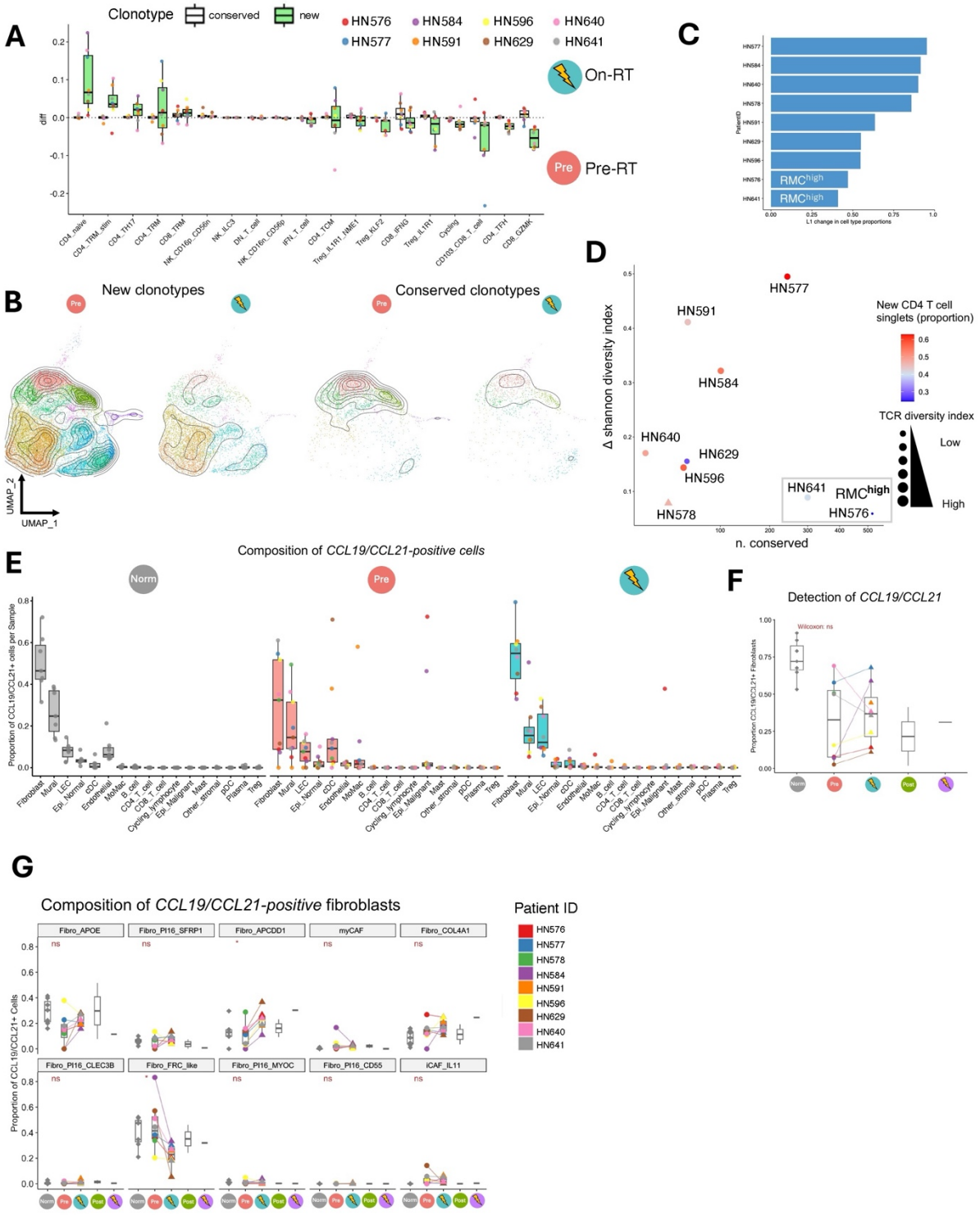


Figure S12 (related to Figure 6). A. Change in per sample cluster proportions (on-RT – pre-RT) for each patient, split into new and conserved clonotypes. **B**. UMAP plot with cell densities overlaid across conditions, split by new and conserved clonotypes. **C**. L1 scores (sum of the absolute differences in proportions between on-RT and pre-RT samples) across patients. **D**. Scatter plot of on-

RT samples [and NAT RT (HN578)] showing the change in TCR shannon diversity index scores (per sample, pre-RT vs on-RT) against the number of conserved T cell clones. Points coloured by proportion of new CD4+ T cell singlets per sample. Size of points reflective of TCR shannon index score of sample. **E**. Proportion of *CCL19* or *CCL21*-expressing cells across all cell types per sample. **F**. Proportion of *CCL19* or *CCL21*-expressing fibroblasts across samples. **G**. Proportion of *CCL19* or *CCL21*-expressing fibroblasts deriving from each fibroblast cluster per sample. Wilcoxon signed-rank test between pre-RT and on-RT. *p<0.05.

References

1. 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).
2. 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).