Figure S1

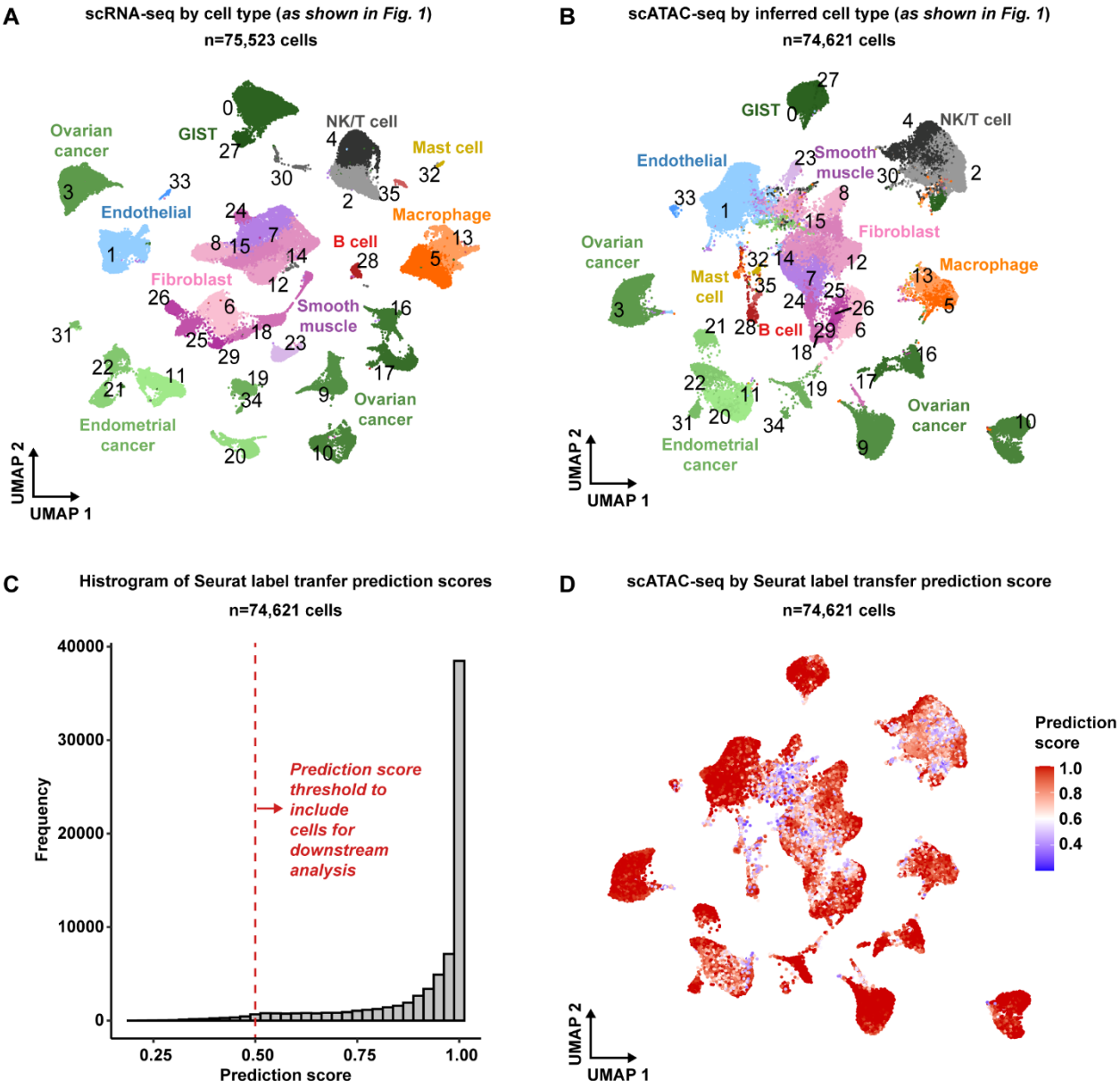
A
scRNA-seq by cell type (as shown in Fig. 1)
n=75,523 cells
Ovarian cancer
GIST
NK/T cell
Mast cell
Endothelial
Macrophage
B cell
Fibroblast
Smooth muscle
Endometrial cancer
Ovarian cancer
UMAP 2
UMAP 1
B
scATAC-seq by inferred cell type (as shown in Fig. 1)
n=74,621 cells
GIST
NK/T cell
Endothelial
Smooth muscle
Fibroblast
Ovarian cancer
Mast cell
Macrophage
B cell
Endometrial cancer
Ovarian cancer
UMAP 2
UMAP 1
C
Histrogram of Seurat label tranfer prediction scores
n=74,621 cells
Frequency
Prediction score threshold to include cells for downstream analysis
Prediction score
D
scATAC-seq by Seurat label transfer prediction score
n=74,621 cells
Prediction score
UMAP 2
UMAP 1

Figure S1. Expanded view of matched scRNA-seq and scATAC-seq for 11 patient tumors, Related to Figure 1.

A) UMAP plot of 75,523 scRNA-seq cells color-coded by cell type across 11 patient tumors. Cell type subclusters as determined by graph-based Louvain clustering are labeled on the UMAP.

B) UMAP plot of 74,621 scATAC-seq cells color-coded by inferred cell type across 11 patient tumors. Inferred cell type subclusters are labeled on the UMAP.

C) Histogram of scATAC-seq inferred cell type subcluster prediction scores. The dashed vertical line in red at 0.5 represents a threshold cutoff for including cells in downstream analysis.

D) UMAP plot of 74,621 scATAC-seq cells, as in **B**, but colored by prediction score.

Figure S2

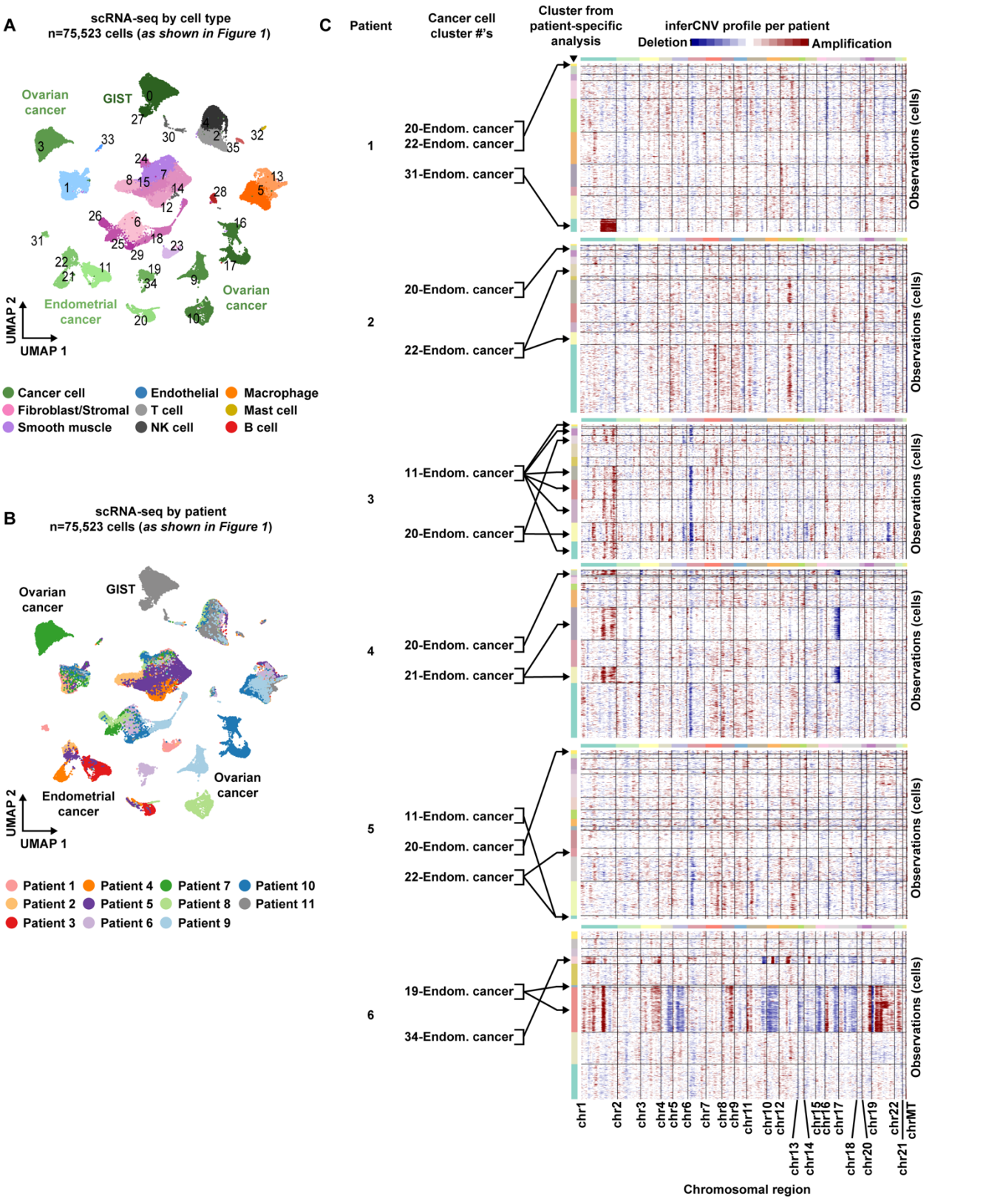
A
scRNA-seq by cell type
n=75,523 cells (as shown in Figure 1)
Ovarian cancer
GIST
Endometrial cancer
Ovarian cancer
UMAP 2
UMAP 1
Cancer cell
Endothelial
Macrophage
Fibroblast/Stromal
T cell
Mast cell
Smooth muscle
NK cell
B cell
B
scRNA-seq by patient
n=75,523 cells (as shown in Figure 1)
Ovarian cancer
GIST
Endometrial cancer
Ovarian cancer
UMAP 2
UMAP 1
Patient 1
Patient 2
Patient 3
Patient 4
Patient 5
Patient 6
Patient 7
Patient 8
Patient 9
Patient 10
Patient 11
C
Patient
Cancer cell cluster #'s
Cluster from patient-specific analysis
inferCNV profile per patient
Deletion
Amplification
1
20-Endom. cancer
22-Endom. cancer
31-Endom. cancer
2
20-Endom. cancer
22-Endom. cancer
3
11-Endom. cancer
20-Endom. cancer
4
20-Endom. cancer
21-Endom. cancer
5
11-Endom. cancer
20-Endom. cancer
22-Endom. cancer
6
19-Endom. cancer
34-Endom. cancer
Observations (cells)
chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr11
chr10
chr12
chr13
chr14
chr15
chr16
chr17
chr18
chr20
chr19
chr22
chr21
chrMT
Chromosomal region

Figure S2. InferCNV results for Patients 1-6 (endometrial cancer), Related to Figure 1.
A) UMAP plot of 75,523 scRNA-seq cells color-coded by cell type and cell type subcluster number.
B) UMAP plot of 75,523 scRNA-seq cells, as in **A**, but color-coded by patient of origin.
C) InferCNV heatmaps of predicted CNV profiles in each patient tumor (*right*). Note that the inferCNV analysis was conducted for each patient tumor individually. Black arrows show the mapping between the full cohort cell type subclusters and the clusters from each patient-specific analysis (*middle*).

Figure S3

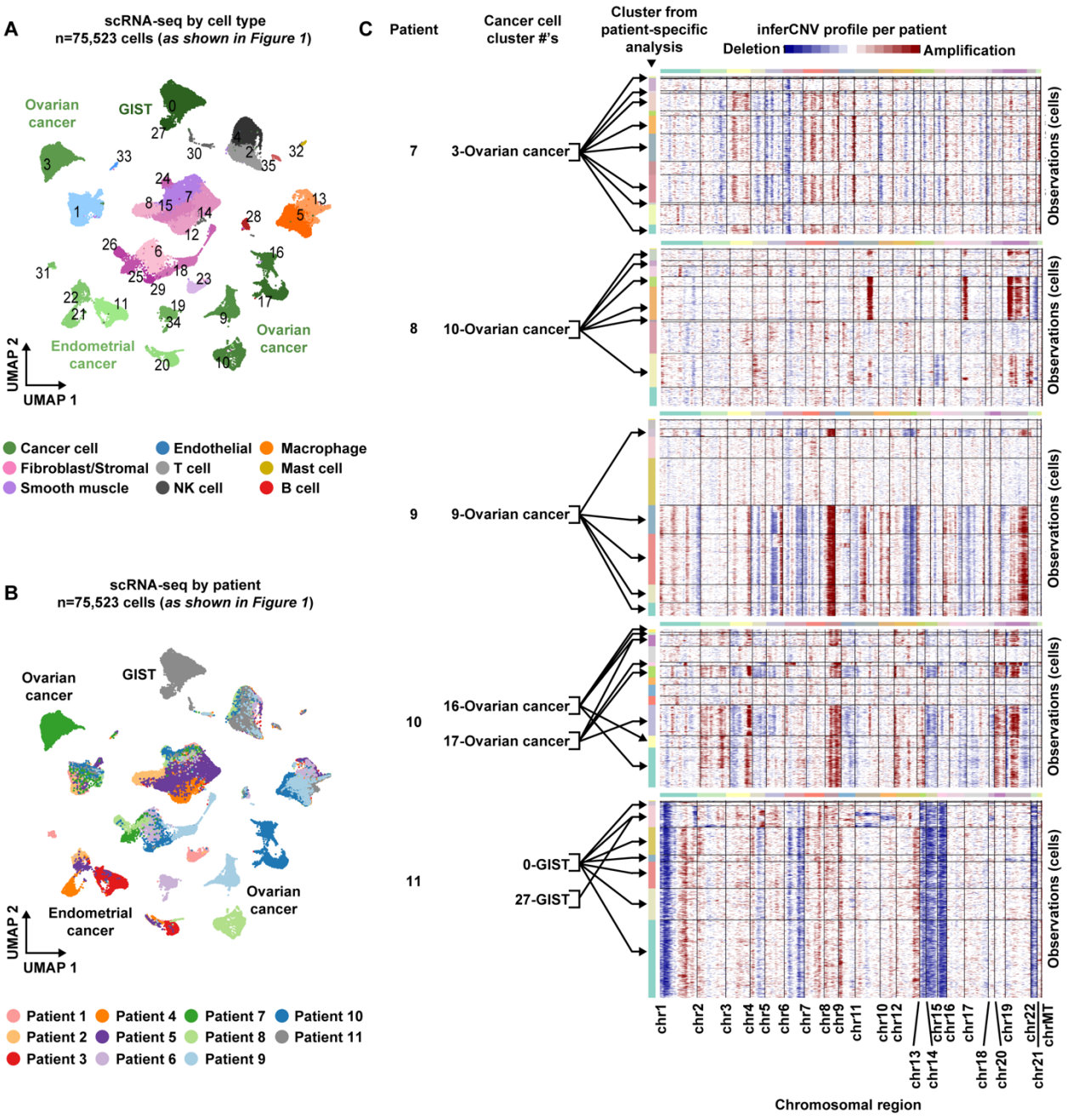
A
scRNA-seq by cell type
n=75,523 cells (as shown in Figure 1)
Ovarian cancer
GIST
Endometrial cancer
Ovarian cancer
UMAP 2
UMAP 1
Cancer cell
Endothelial
Macrophage
Fibroblast/Stromal
T cell
Mast cell
Smooth muscle
NK cell
B cell
B
scRNA-seq by patient
n=75,523 cells (as shown in Figure 1)
Ovarian cancer
GIST
Endometrial cancer
Ovarian cancer
UMAP 2
UMAP 1
Patient 1
Patient 2
Patient 3
Patient 4
Patient 5
Patient 6
Patient 7
Patient 8
Patient 9
Patient 10
Patient 11
C
Patient
Cancer cell cluster #'s
Cluster from patient-specific analysis
inferCNV profile per patient
Deletion
Amplification
7
3-Ovarian cancer
8
10-Ovarian cancer
9
9-Ovarian cancer
10
16-Ovarian cancer
17-Ovarian cancer
11
0-GIST
27-GIST
Observations (cells)
chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr11
chr10
chr12
chr13
chr14
chr15
chr16
chr17
chr18
chr20
chr19
chr22
chr21
chrMT
Chromosomal region

Figure S3. InferCNV results for Patients 7-10 and 11 (ovarian cancer and GIST, respectively), Related to Figure 1.
A) UMAP plot of 75,523 scRNA-seq cells color-coded by cell type and cell type subcluster number.
B) UMAP plot of 75,523 scRNA-seq cells, as in **A**, but color-coded by patient of origin.
C) InferCNV heatmaps of predicted CNV profiles in each patient tumor (*right*). Note that the inferCNV analysis was conducted for each patient tumor individually. Black arrows show the mapping between the full cohort cell type subclusters and the clusters from each patient-specific analysis (*middle*).

Figure S4

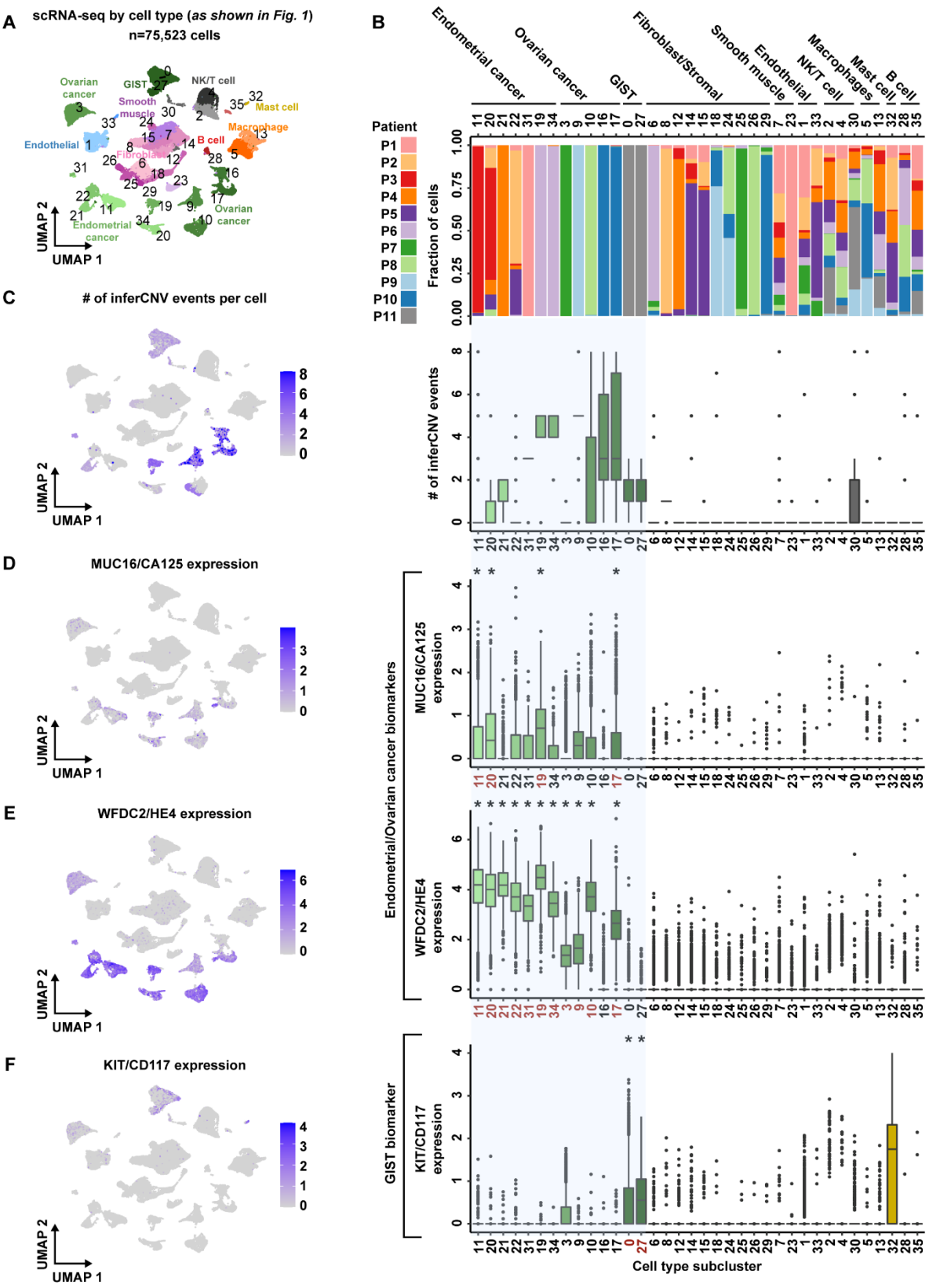
A
scRNA-seq by cell type (as shown in Fig. 1)
n=75,523 cells
Ovarian cancer
GIST
NK/T cell
Smooth muscle
Mast cell
Macrophage
Endothelial
B cell
Fibroblast
Endometrial cancer
UMAP 1
UMAP 2
B
Endometrial cancer
Ovarian cancer
GIST
Fibroblast/Stromal
Smooth muscle
Endothelial
NK/T cell
Macrophages
Mast cell
B cell
Patient
P1
P2
P3
P4
P5
P6
P7
P8
P9
P10
P11
Fraction of cells
# of inferCNV events
C
# of inferCNV events per cell
D
MUC16/CA125 expression
E
WFDC2/HE4 expression
F
KIT/CD117 expression
Endometrial/Ovarian cancer biomarkers
MUC16/CA125 expression
WFDC2/HE4 expression
GIST biomarker
KIT/CD117 expression
Cell type subcluster

Figure S4. Identification of malignant cell type subclusters in scRNA-seq across the full cohort of 11 patient tumors, Related to Figure 1.
A) UMAP plot of 75,523 scRNA-seq cells color-coded by cell type and cell type subcluster number.
B) Stacked bar chart, as in **Fig. 1D**, showing the contribution of each patient to each cell type subcluster in scRNA-seq (*top*). Boxplot showing the number of inferCNV events in each cell type subcluster (*second from top*). The blue shadow highlights cell type subclusters that are predicted to be malignant. Boxplot showing the expression level of ovarian/endometrial cancer biomarker *MUC16*/CA125 across all cell type subclusters. The blue shadow highlights cell type subclusters that are predicted to be malignant. Asterisks denote a statistically significant difference in gene expression (Wilcoxon Rank Sum test, Bonferroni-corrected *p*-value <0.01) between the marked cell type subcluster and the remaining cells outside of the blue shadow (*middle*). Boxplot showing the expression level of ovarian/endometrial cancer biomarker *WFDC2*/HE4 across all cell type subclusters. The blue shadow highlights cell type subclusters that are predicted to be malignant. Asterisks denote a statistically significant difference in gene expression (Wilcoxon Rank Sum test, Bonferroni-corrected *p*-value <0.01) between the marked cell type subcluster and the remaining cells outside of the blue shadow (*second from bottom*). Boxplot showing the expression level of GIST biomarker *KIT*/CD117 across all cell type subclusters. The blue shadow highlights cell type subclusters that are predicted to be malignant. Asterisks denote a statistically significant difference in gene expression (Wilcoxon Rank Sum test, Bonferroni-corrected *p*-value <0.01) between the marked cell type subcluster and the remaining cells outside of the blue shadow (*bottom*).
C) UMAP plot of 75,523 scRNA-seq cells, as in **A**, but colored by total number of inferCNV events per cell.
D) UMAP plot of 75,523 scRNA-seq cells, as in **A**, but colored by the normalized expression level of ovarian/endometrial cancer biomarker *MUC16*/CA125.
E) UMAP plot of 75,523 scRNA-seq cells, as in **A**, but colored by the normalized expression level of ovarian/endometrial cancer biomarker *WFDC2*/HE4.
F) UMAP plot of 75,523 scRNA-seq cells, as in **A**, but colored by the normalized expression level of GIST biomarker *KIT*/CD117.

Figure S5

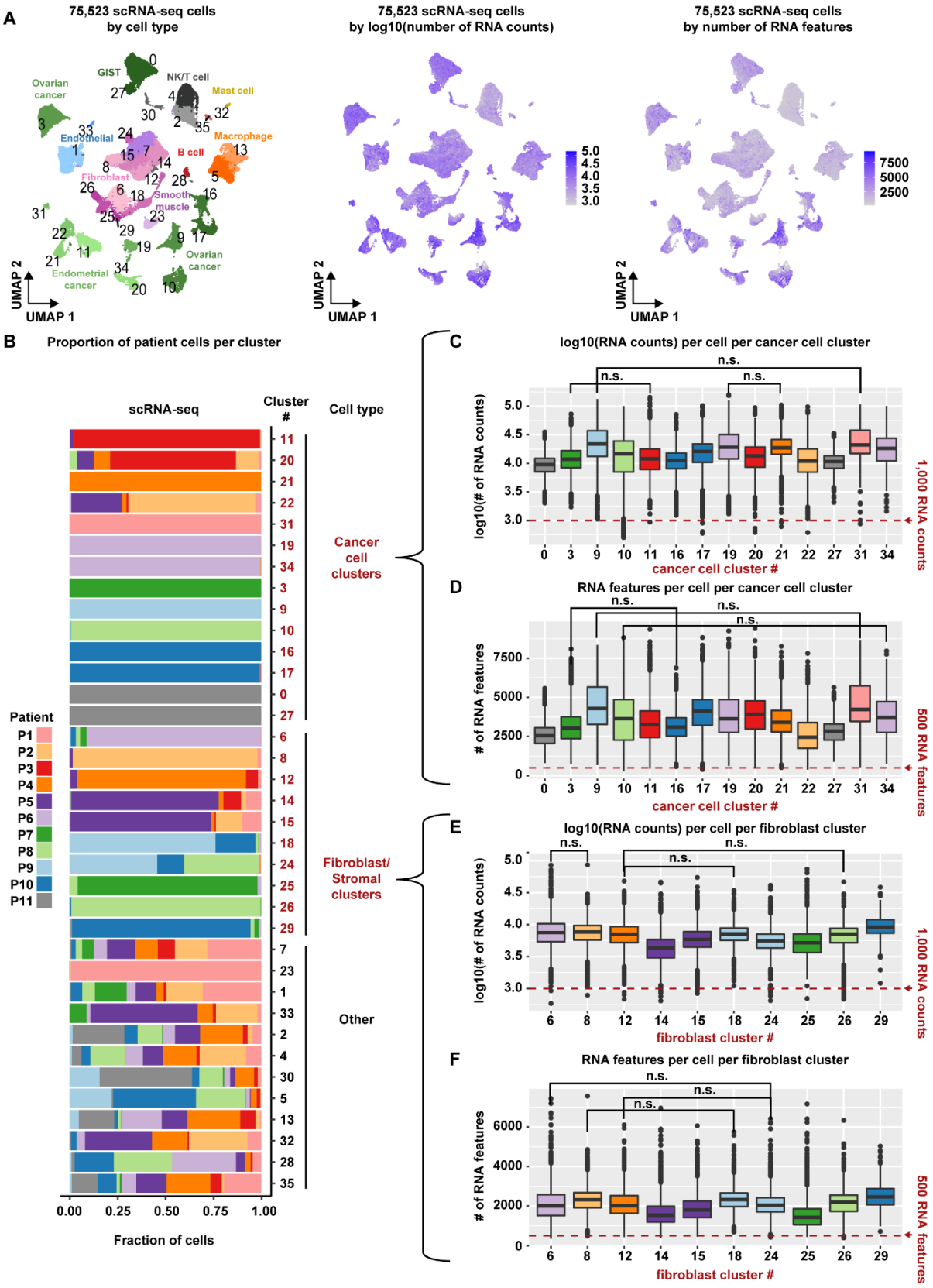
A
75,523 scRNA-seq cells by cell type
75,523 scRNA-seq cells by log10(number of RNA counts)
75,523 scRNA-seq cells by number of RNA features
Ovarian cancer
GIST
NK/T cell
Mast cell
Endothelial
Macrophage
B cell
Fibroblast
Smooth muscle
Ovarian cancer
Endometrial cancer
UMAP 1
UMAP 2
B
Proportion of patient cells per cluster
scRNA-seq
Cluster #
Cell type
Patient
P1
P2
P3
P4
P5
P6
P7
P8
P9
P10
P11
Cancer cell clusters
Fibroblast/ Stromal clusters
Other
Fraction of cells
C
log10(RNA counts) per cell per cancer cell cluster
log10(# of RNA counts)
cancer cell cluster #
1,000 RNA counts
D
RNA features per cell per cancer cell cluster
# of RNA features
cancer cell cluster #
500 RNA features
E
log10(RNA counts) per cell per fibroblast cluster
log10(# of RNA counts)
fibroblast cluster #
1,000 RNA counts
F
RNA features per cell per fibroblast cluster
# of RNA features
fibroblast cluster #
500 RNA features
n.s.

Figure S5. Malignant and fibroblast cell type subclusters in scRNA-seq show high patient-specificity relative to endothelial and immune subclusters, Related to Figure 1.
A) UMAP plot of 75,523 scRNA-seq cells color-coded by cell type and cell type subcluster number (*left*). UMAP plot of 75,523 scRNA-seq cells, as in ***left***, but color-coded by log10(number of RNA counts) per cell (*middle*). UMAP plot of 75,523 scRNA-seq cells, as in ***left***, but color-coded by number of RNA features per cell (*right*).
B) Stacked bar chart, as in **Fig. 1D**, showing the contribution of each patient to each cell type subcluster in scRNA-seq. Malignant or cancer cell subclusters and fibroblast subclusters are marked in red text.
C) Boxplot showing the distribution of log10(RNA counts) per cell for each malignant cell type subcluster. N.S. stands for a statistically insignificant difference between the marked subclusters.
D) Boxplot showing the distribution of RNA features per cell for each malignant cell type subcluster. N.S. stands for a statistically insignificant difference between the marked subclusters.
E) Boxplot showing the distribution of log10(RNA counts) per cell for each fibroblast cell type subcluster. N.S. stands for a statistically insignificant difference between the marked subclusters.
F) Boxplot showing the distribution of RNA features per cell for each fibroblast cell type subcluster. N.S. stands for a statistically insignificant difference between the marked subclusters.

Figure S6

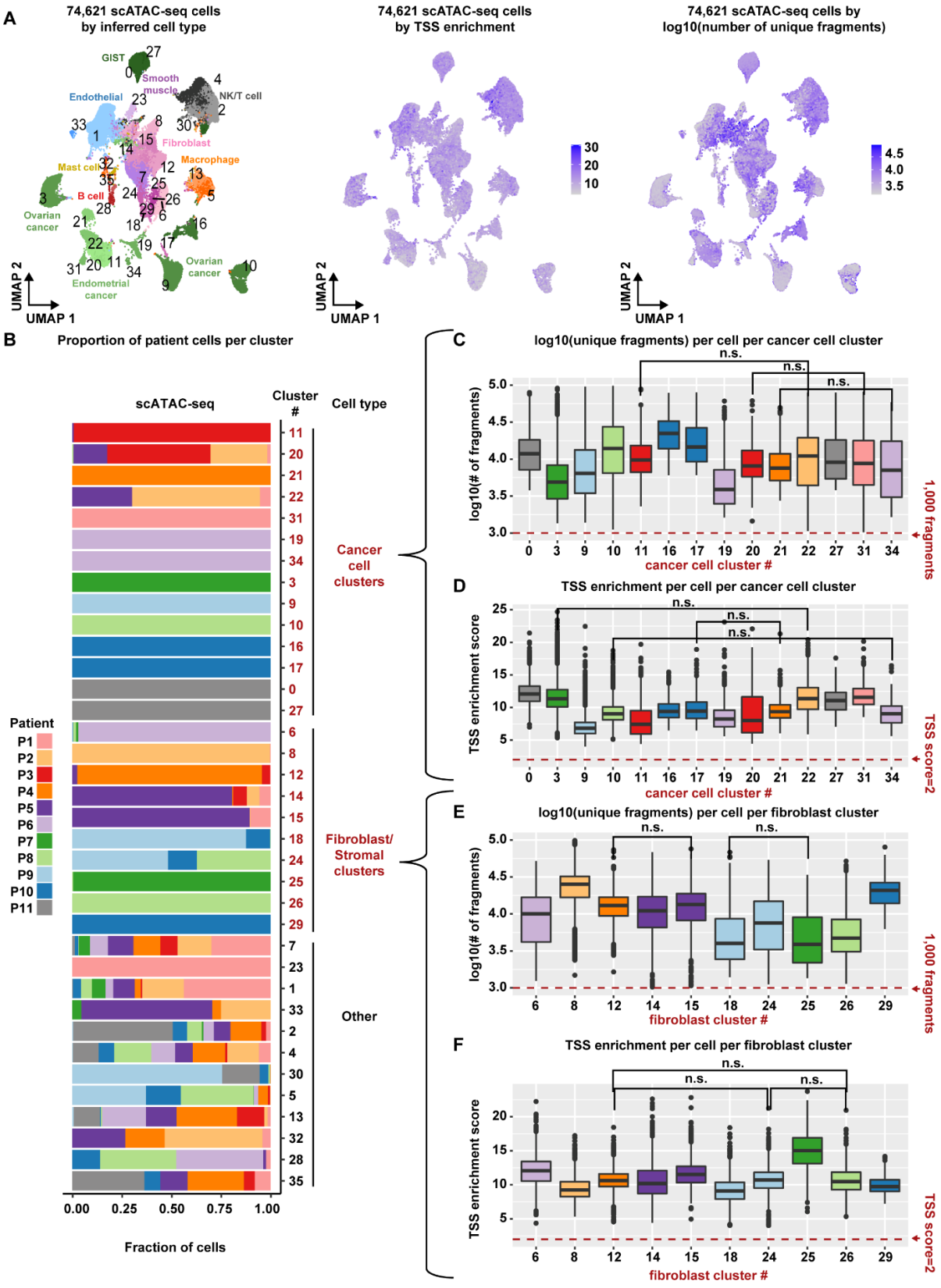
A
74,621 scATAC-seq cells by inferred cell type
74,621 scATAC-seq cells by TSS enrichment
74,621 scATAC-seq cells by log10(number of unique fragments)
GIST
Smooth muscle
Endothelial
NK/T cell
Fibroblast
Mast cell
Macrophage
B cell
Ovarian cancer
Endometrial cancer
UMAP 1
UMAP 2
B
Proportion of patient cells per cluster
scATAC-seq
Cluster #
Cell type
Cancer cell clusters
Fibroblast/ Stromal clusters
Other
Patient
P1
P2
P3
P4
P5
P6
P7
P8
P9
P10
P11
Fraction of cells
C
log10(unique fragments) per cell per cancer cell cluster
log10(# of fragments)
cancer cell cluster #
1,000 fragments
D
TSS enrichment per cell per cancer cell cluster
TSS enrichment score
TSS score=2
E
log10(unique fragments) per cell per fibroblast cluster
fibroblast cluster #
F
TSS enrichment per cell per fibroblast cluster
n.s.

Figure S6. Malignant and fibroblast cell type subclusters in scATAC-seq show high patient-specificity relative to endothelial and immune subclusters, Related to Figure 1.
A) UMAP plot of 74,621 scATAC-seq cells color-coded by inferred cell type and inferred cell type subcluster number (*left*). UMAP plot of 74,621 scATAC-seq cells, as in ***left***, but color-coded by TSS enrichment score (*middle*). UMAP plot of 74,621 scATAC-seq cells, as in ***left***, but color-coded by log10(number of unique fragments) (*right*).
B) Stacked bar chart, as in **Fig. 1D**, showing the contribution of each patient to each inferred cell type subcluster in scATAC-seq. Malignant or cancer cell subclusters and fibroblast subclusters are marked in red text.
C) Boxplot showing the distribution of log10(number of unique fragments) for each malignant cell type subcluster. N.S. stands for a statistically insignificant difference between the marked subclusters.
D) Boxplot showing the distribution of TSS enrichment score for each malignant cell type subcluster. N.S. stands for a statistically insignificant difference between the marked subclusters.
E) Boxplot showing the distribution of log10(number of unique fragments) per cell for each fibroblast cell type subcluster. N.S. stands for a statistically insignificant difference between the marked subclusters.
F) Boxplot showing the distribution of TSS enrichment score for each fibroblast cell type subcluster. N.S. stands for a statistically insignificant difference between the marked subclusters.

Figure S7

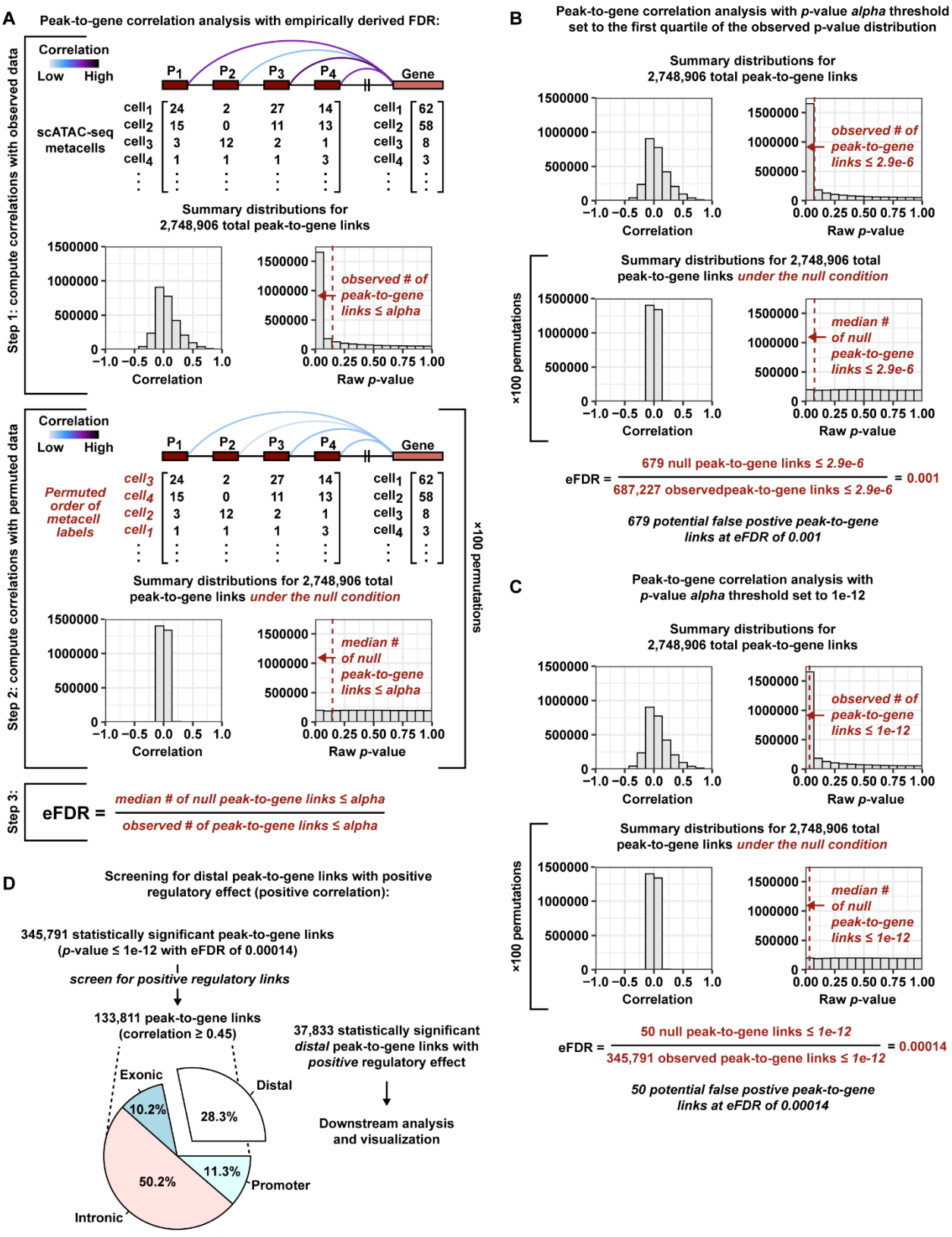
A
Peak-to-gene correlation analysis with empirically derived FDR:
Step 1: compute correlations with observed data
Correlation
Low High
P1 P2 P3 P4 Gene
scATAC-seq metacells
Summary distributions for 2,748,906 total peak-to-gene links
observed # of peak-to-gene links ≤ alpha
Correlation
Raw p-value
Step 2: compute correlations with permuted data
Permuted order of metacell labels
Summary distributions for 2,748,906 total peak-to-gene links under the null condition
median # of null peak-to-gene links ≤ alpha
×100 permutations
Step 3:
eFDR = median # of null peak-to-gene links ≤ alpha / observed # of peak-to-gene links ≤ alpha
B
Peak-to-gene correlation analysis with p-value alpha threshold set to the first quartile of the observed p-value distribution
Summary distributions for 2,748,906 total peak-to-gene links
observed # of peak-to-gene links ≤ 2.9e-6
Summary distributions for 2,748,906 total peak-to-gene links under the null condition
median # of null peak-to-gene links ≤ 2.9e-6
×100 permutations
eFDR = 679 null peak-to-gene links ≤ 2.9e-6 / 687,227 observedpeak-to-gene links ≤ 2.9e-6 = 0.001
679 potential false postive peak-to-gene links at eFDR of 0.001
C
Peak-to-gene correlation analysis with p-value alpha threshold set to 1e-12
Summary distributions for 2,748,906 total peak-to-gene links
observed # of peak-to-gene links ≤ 1e-12
Summary distributions for 2,748,906 total peak-to-gene links under the null condition
median # of null peak-to-gene links ≤ 1e-12
×100 permutations
eFDR = 50 null peak-to-gene links ≤ 1e-12 / 345,791 observed peak-to-gene links ≤ 1e-12 = 0.00014
50 potential false postive peak-to-gene links at eFDR of 0.00014
D
Screening for distal peak-to-gene links with positive regulatory effect (positive correlation):
345,791 statistically significant peak-to-gene links (p-value ≤ 1e-12 with eFDR of 0.00014)
screen for positive regulatory links
133,811 peak-to-gene links (correlation ≥ 0.45)
37,833 statistically significant distal peak-to-gene links with positive regulatory effect
Downstream analysis and visualization
Exonic 10.2%
Distal 28.3%
Promoter 11.3%
Intronic 50.2%

Figure S7. Peak-to-gene correlation analysis with an empirically derived FDR for the full cohort of 11 patient tumors, Related to Figure 2 and STAR Methods.
A) Schematic illustrating the peak-to-gene correlation analysis. **Step 1**: the peak count information from scATAC-seq cells is aggregated into metacells via a KNN algorithm before computing the correlation between every peak and every gene on the same chromosome. The distribution of correlation values and raw p-values are visualized in histograms and the number of observed peak-to-gene link tests <= alpha is recorded. **Step 2**: the peak-to-gene correlations are re-computed for a permuted null version of the dataset where the scATAC-seq metacell labels are shuffled, breaking the link or any potential correlation between peaks and genes. The process is repeated for a total of 100 permutations. For each permutation, the number of null peak-to-gene link tests <= alpha is recorded. As an example, the distribution of null correlation values and null raw p-values are visualized in histograms for one permutation run. **Step 3**: The median number of null peak-to-gene link tests <= alpha across 100 permutations is divided by the number of observed peak-to-gene link tests <= alpha to arrive at an empirically derived FDR (eFDR).
B) The peak-to-gene correlation analysis, as described in **A,** but using an alpha threshold equal to the first quartile in the distribution of raw p-values.
C) The peak-to-gene correlation analysis, as described in **A,** but using an alpha threshold equal to 1e-12.
D) Flow chart demonstrating the screening procedure used to identify distal peak-to-gene links with positive regulatory effects. The pie chart depicts the proportion of positive regulatory peak-to-gene links that are distal peak-to-gene relationships.

Figure S8

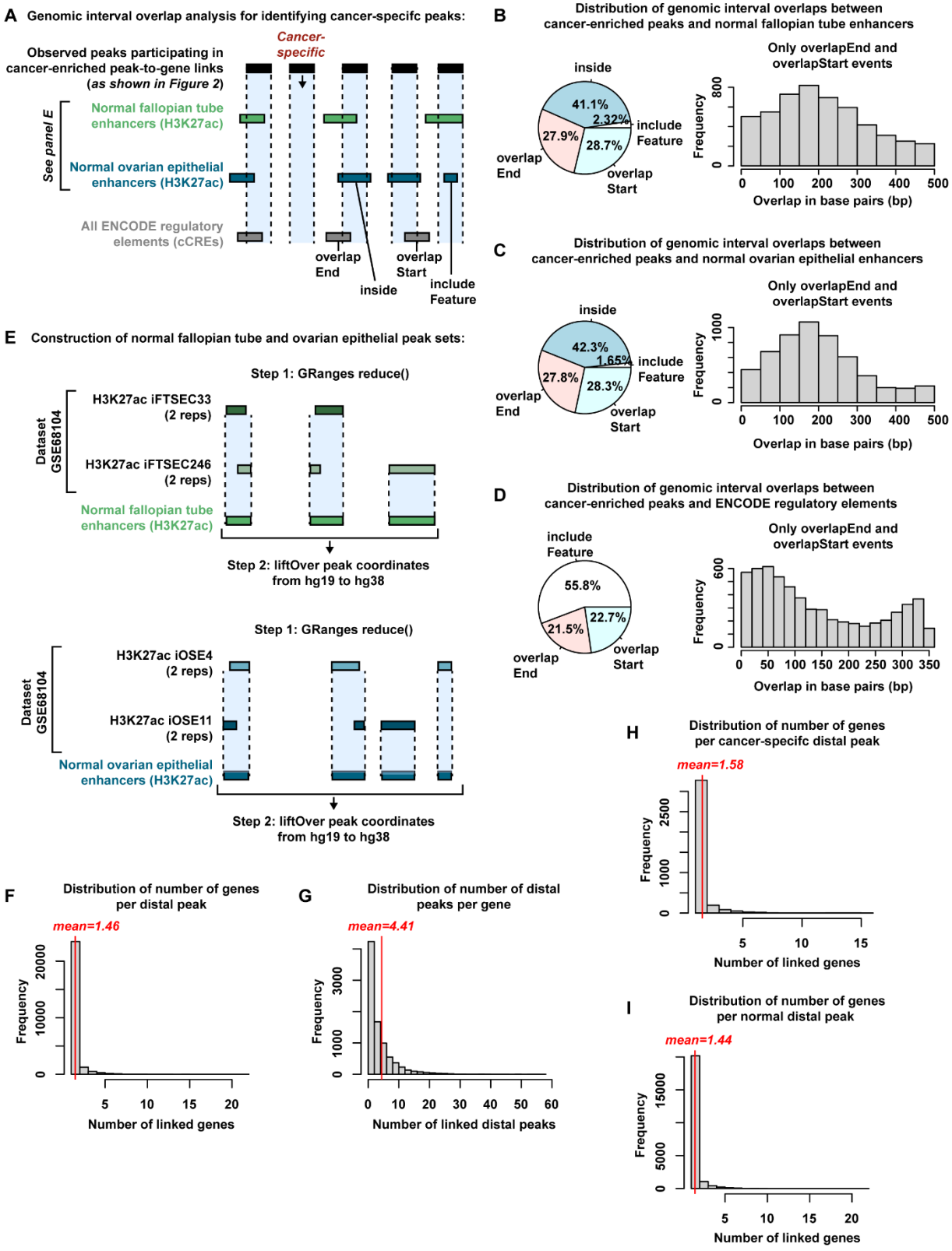
A
Genomic interval overlap analysis for identifying cancer-specifc peaks:
Observed peaks participating in cancer-enriched peak-to-gene links (as shown in Figure 2)
Cancer-specific
See panel E
Normal fallopian tube enhancers (H3K27ac)
Normal ovarian epithelial enhancers (H3K27ac)
All ENCODE regulatory elements (cCREs)
overlap End
inside
overlap Start
include Feature
B
Distribution of genomic interval overlaps between cancer-enriched peaks and normal fallopian tube enhancers
inside
41.1%
2.32%
include Feature
27.9%
28.7%
overlap End
overlap Start
Only overlapEnd and overlapStart events
Frequency
0
800
0 100 200 300 400 500
Overlap in base pairs (bp)
C
Distribution of genomic interval overlaps between cancer-enriched peaks and normal ovarian epithelial enhancers
inside
42.3%
1.65%
include Feature
27.8%
28.3%
overlap End
overlap Start
Only overlapEnd and overlapStart events
Frequency
0
1000
0 100 200 300 400 500
Overlap in base pairs (bp)
D
Distribution of genomic interval overlaps between cancer-enriched peaks and ENCODE regulatory elements
include Feature
55.8%
22.7%
21.5%
overlap End
overlap Start
Only overlapEnd and overlapStart events
Frequency
0
600
0 50 100 150 200 250 300 350
Overlap in base pairs (bp)
E
Construction of normal fallopian tube and ovarian epithelial peak sets:
Step 1: GRanges reduce()
Dataset GSE68104
H3K27ac iFTSEC33 (2 reps)
H3K27ac iFTSEC246 (2 reps)
Normal fallopian tube enhancers (H3K27ac)
Step 2: liftOver peak coordinates from hg19 to hg38
Step 1: GRanges reduce()
Dataset GSE68104
H3K27ac iOSE4 (2 reps)
H3K27ac iOSE11 (2 reps)
Normal ovarian epithelial enhancers (H3K27ac)
Step 2: liftOver peak coordinates from hg19 to hg38
F
Distribution of number of genes per distal peak
mean=1.46
Frequency
0
10000
20000
5 10 15 20
Number of linked genes
G
Distribution of number of distal peaks per gene
mean=4.41
Frequency
0
1000
3000
0 10 20 30 40 50 60
Number of linked distal peaks
H
Distribution of number of genes per cancer-specifc distal peak
mean=1.58
Frequency
0
1000
2500
5 10 15
Number of linked genes
I
Distribution of number of genes per normal distal peak
mean=1.44
Frequency
0
5000
15000
5 10 15 20
Number of linked genes

Figure S8. Genomic coordinate overlap analysis for identifying cancer-specific distal peak-to-gene links across the full cohort of 11 patient tumors, Related to Figure 2 and STAR Methods.

A) Cartoon depicting the genomic coordinate overlap analysis between the observed peaks participating in cancer-enriched distal peak-to-gene links and three other reference peak sets (normal fallopian tube H3K27ac peaks [*green*], normal ovarian epithelial H3K27ac peaks [*blue*], and all ENCODE regulatory elements [*gray*]).

B) Pie chart and histogram summarizing the distribution of genomic overlap events between the cancer-enriched peaks and the normal fallopian tube H3K27ac peaks.

C) Pie chart and histogram summarizing the distribution of genomic overlap events between the cancer-enriched peaks and the normal ovarian epithelial H3K27ac peaks.

D) Pie chart and histogram summarizing the distribution of genomic overlap events between the cancer-enriched peaks and the ENCODE regulatory elements.

E) Cartoon illustrating the procedure for creating the normal fallopian tube enhancer peak set (green, *top*) and normal ovarian epithelial enhancer peak set (blue, *bottom*).

F) Histogram showing the distribution of number of genes per distal peak. The solid red line represents the mean number of genes per distal peak.

G) Histogram showing the distribution of number of distal peaks per gene. The solid red line represents the mean number of distal peaks per gene.

H) Histogram showing the distribution of number of genes per cancer-specific distal peak. The solid red line represents the mean number of genes per cancer-specific distal peak.

I) Histogram showing the distribution of number of genes per normal distal peak. The solid red line represents the mean number of genes per normal distal peak.

Figure S9

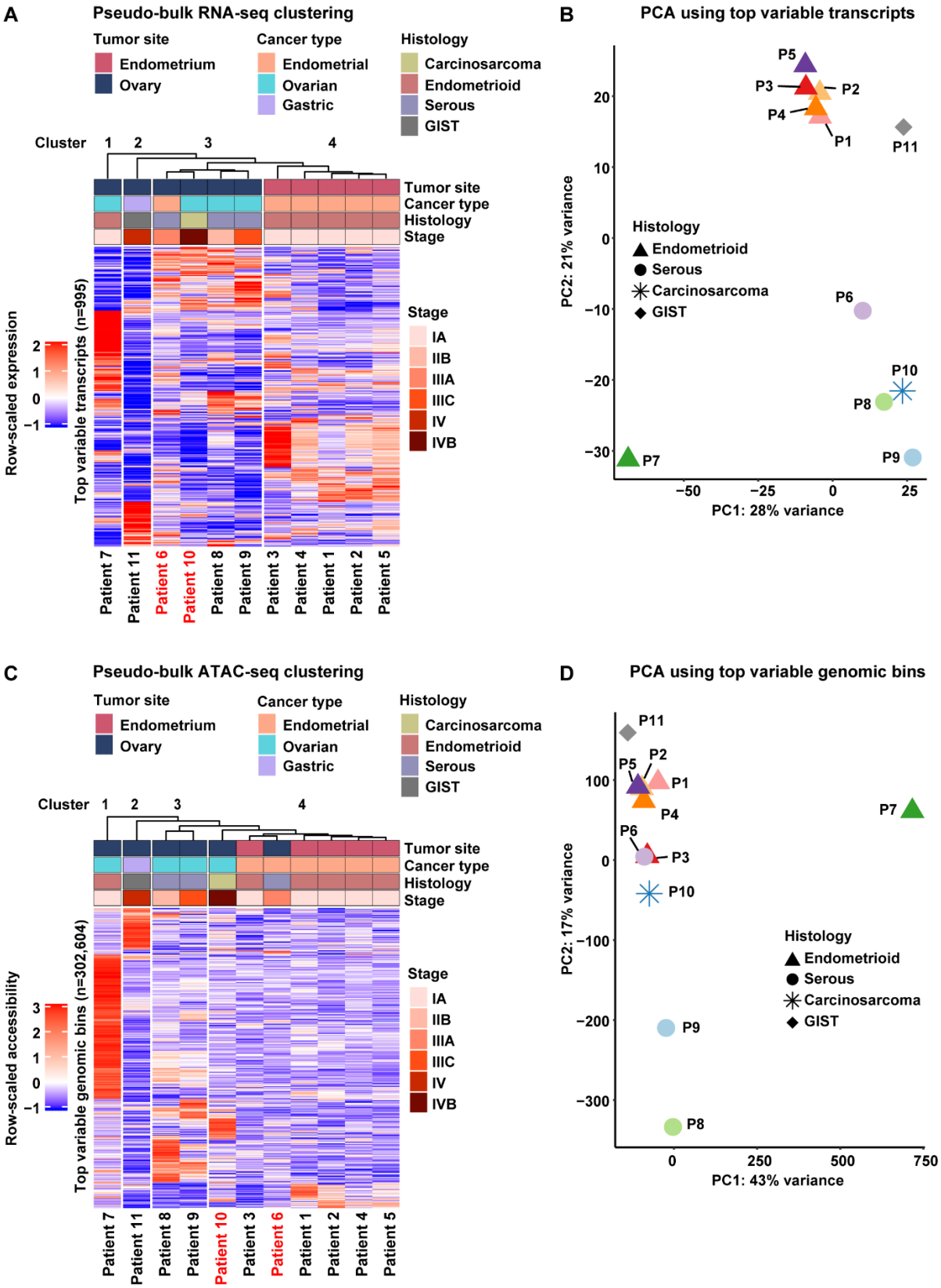
A
Pseudo-bulk RNA-seq clustering
Tumor site
Endometrium
Ovary
Cancer type
Endometrial
Ovarian
Gastric
Histology
Carcinosarcoma
Endometrioid
Serous
GIST
Cluster 1 2 3 4
Tumor site
Cancer type
Histology
Stage
Stage
IA
IIB
IIIA
IIIC
IV
IVB
Row-scaled expression
Top variable transcripts (n=995)
Patient 7
Patient 11
Patient 6
Patient 10
Patient 8
Patient 9
Patient 3
Patient 4
Patient 1
Patient 2
Patient 5
B
PCA using top variable transcripts
PC2: 21% variance
PC1: 28% variance
Histology
Endometrioid
Serous
Carcinosarcoma
GIST
P5
P3
P2
P4
P1
P11
P6
P10
P8
P7
P9
C
Pseudo-bulk ATAC-seq clustering
Tumor site
Endometrium
Ovary
Cancer type
Endometrial
Ovarian
Gastric
Histology
Carcinosarcoma
Endometrioid
Serous
GIST
Cluster 1 2 3 4
Tumor site
Cancer type
Histology
Stage
Stage
IA
IIB
IIIA
IIIC
IV
IVB
Row-scaled accessibility
Top variable genomic bins (n=302,604)
Patient 7
Patient 11
Patient 8
Patient 9
Patient 10
Patient 3
Patient 6
Patient 1
Patient 2
Patient 4
Patient 5
D
PCA using top variable genomic bins
PC2: 17% variance
PC1: 43% variance
Histology
Endometrioid
Serous
Carcinosarcoma
GIST
P11
P5
P2
P1
P4
P7
P6
P3
P10
P9
P8

Figure S9. Cluster analysis of pseudo-bulk transcriptome and chromatin accessibility profiles for 11 patient tumors, Related to Figures 3 & 4.
A) Transcriptome-based hierarchical clustering and row-scaled heatmap using the top 5% variable features (n=985 transcripts). The indicated clusters in the dendrogram are statistically significant as determined by SigClust2.
B) Transcriptome-based PCA using the same features as in panel **A**.
C) Chromatin accessibility-based hierarchical clustering and row-scaled heatmap using the top 5% variable features. Note that only 3,000 randomly sampled features are shown in the heatmap out of 302,604 variable genomic bins. The indicated clusters in the dendrogram are statistically significant as determined by SigClust2.
D) Chromatin accessibility-based PCA using the top 5% variable features (n=302,604 genomic bins).

Figure S10

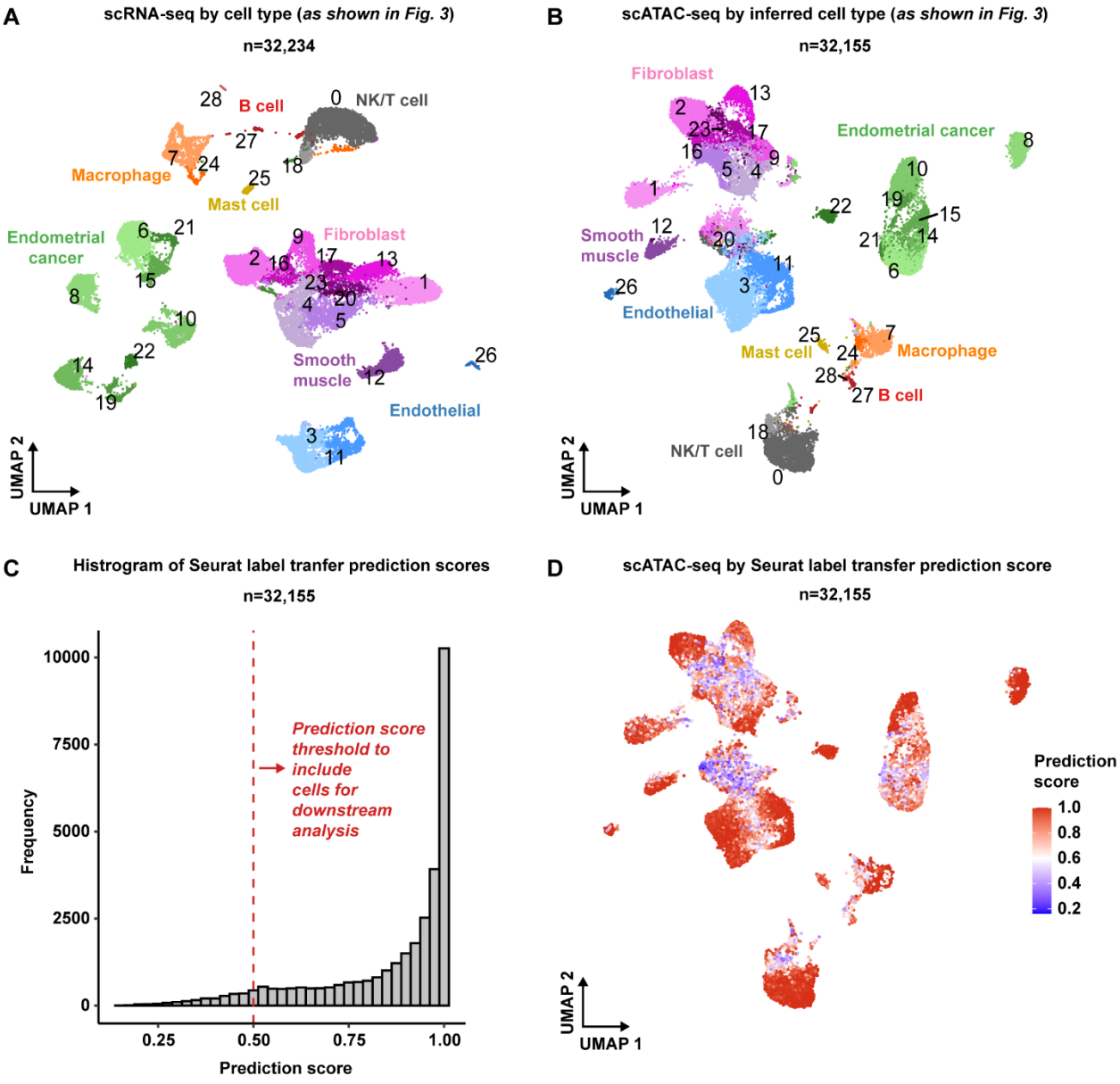
A
scRNA-seq by cell type (as shown in Fig. 3)
n=32,234
B cell
NK/T cell
Macrophage
Mast cell
Endometrial cancer
Fibroblast
Smooth muscle
Endothelial
UMAP 2
UMAP 1
B
scATAC-seq by inferred cell type (as shown in Fig. 3)
n=32,155
Fibroblast
Endometrial cancer
Smooth muscle
Endothelial
Mast cell
Macrophage
B cell
NK/T cell
UMAP 2
UMAP 1
C
Histrogram of Seurat label tranfer prediction scores
n=32,155
Frequency
Prediction score threshold to include cells for downstream analysis
Prediction score
D
scATAC-seq by Seurat label transfer prediction score
n=32,155
Prediction score
UMAP 2
UMAP 1

Figure S10. Expanded view of matched scRNA-seq and scATAC-seq for the Endometrioid Endometrial Cancer patient cohort, Related to Figure 3.
A) UMAP plot of 32,234 scRNA-seq cells color-coded by cell type across the **EEC patient cohort**. Cell type subclusters as determined by graph-based Louvain clustering are labeled on the UMAP.
B) UMAP plot of 32,155 scATAC-seq cells color-coded by inferred cell type across the **EEC patient cohort**. Inferred cell type subclusters are labeled on the UMAP.
C) Histogram of scATAC-seq inferred cell type subcluster prediction scores. The dashed vertical line in red at 0.5 represents a threshold cutoff for including cells in downstream analysis.
D) UMAP plot of 32,155 scATAC-seq cells, as in **B**, but colored by prediction score.

Figure S11

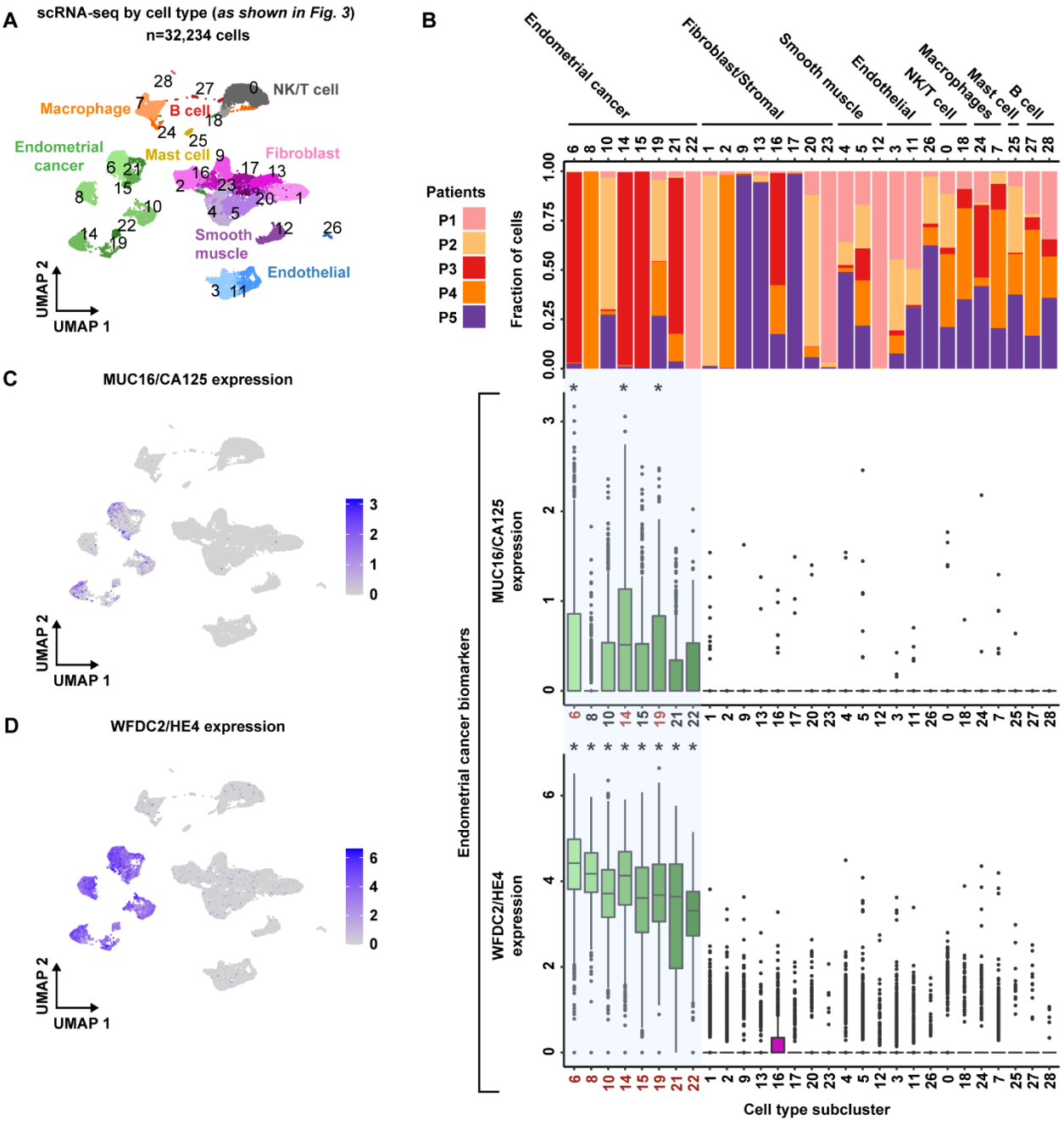
A
scRNA-seq by cell type (as shown in Fig. 3)
n=32,234 cells
Macrophage
B cell
NK/T cell
Endometrial cancer
Mast cell
Fibroblast
Smooth muscle
Endothelial
UMAP 1
UMAP 2
B
Endometrial cancer
Fibroblast/Stromal
Smooth muscle
Endothelial
NK/T cell
Macrophages
Mast cell
B cell
Patients
P1
P2
P3
P4
P5
Fraction of cells
C
MUC16/CA125 expression
D
WFDC2/HE4 expression
Endometrial cancer biomarkers
MUC16/CA125 expression
WFDC2/HE4 expression
Cell type subcluster

Figure S11. Identification of malignant cell type subclusters in scRNA-seq across the Endometrioid Endometrial Cancer patient cohort (Patients 1-5), Related to Figure 3.
A) UMAP plot of 32,234 scRNA-seq cells color-coded by cell type and cell type subcluster number.
B) Stacked bar chart, as in **Fig. 3C**, showing the contribution of each patient to each cell type subcluster in scRNA-seq (*top*). Boxplot showing the expression level of ovarian/endometrial cancer biomarker *MUC16*/CA125 across all cell type subclusters. The blue shadow highlights cell type subclusters that are predicted to be malignant. Asterisks denote a statistically significant difference in gene expression (Wilcoxon Rank Sum test, Bonferroni-corrected *p*-value <0.01) between the marked cell type subcluster and the remaining cells outside of the blue shadow (*middle*). Boxplot showing the expression level of ovarian/endometrial cancer biomarker *WFDC2*/HE4 across all cell type subclusters. The blue shadow highlights cell type subclusters that are predicted to be malignant. Asterisks denote a statistically significant difference in gene expression (Wilcoxon Rank Sum test, Bonferroni-corrected *p*-value <0.01) between the marked cell type subcluster and the remaining cells outside of the blue shadow (*bottom*).
C) UMAP plot of 32,234 scRNA-seq cells, as in **A**, but colored by the normalized expression level of endometrial cancer biomarker *MUC16*/CA125.
D) UMAP plot of 32,234 scRNA-seq cells, as in **A**, but colored by the normalized expression level of endometrial cancer biomarker *WFDC2*/HE4.

Figure S12

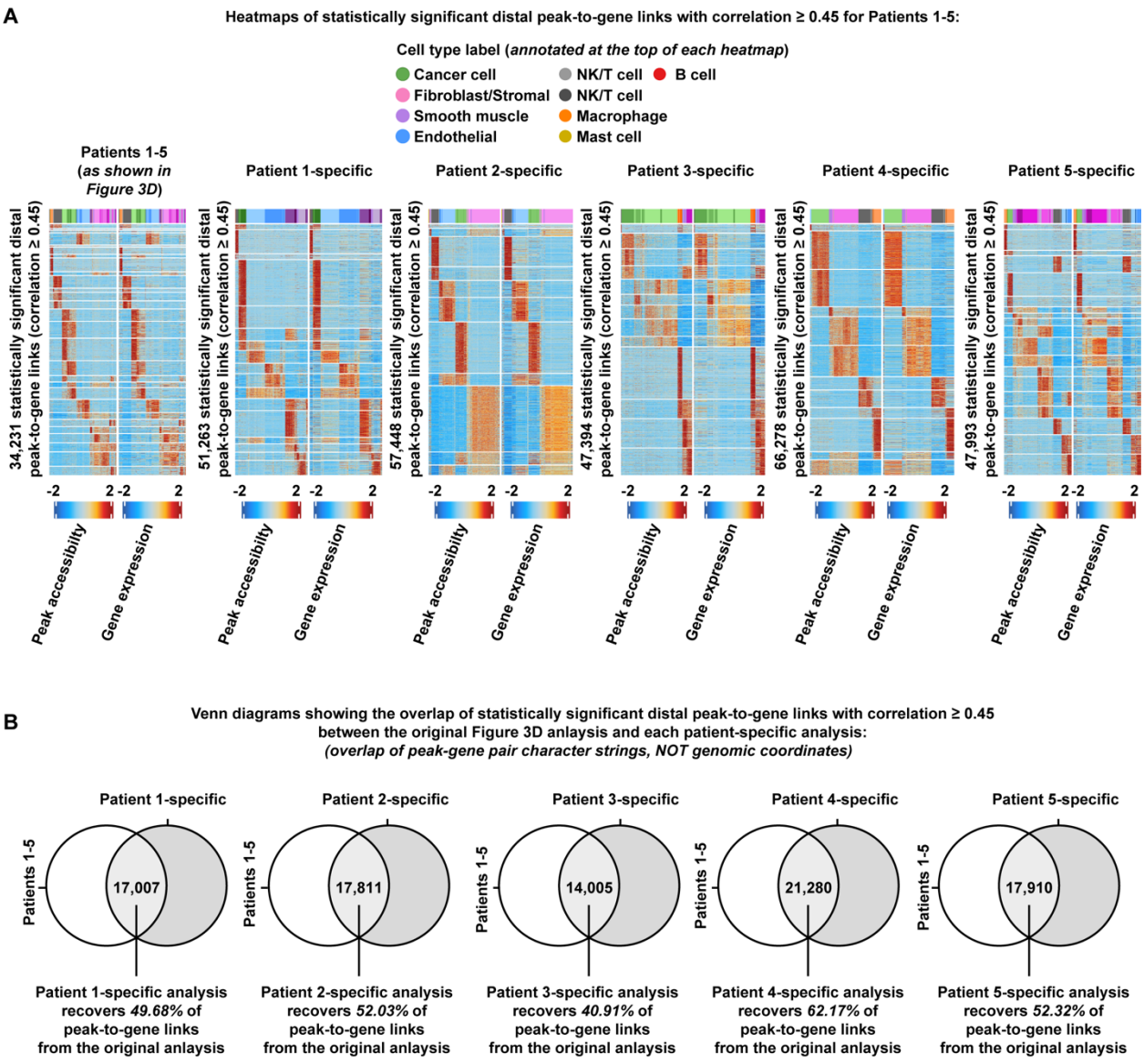
A
Heatmaps of statistically significant distal peak-to-gene links with correlation ≥ 0.45 for Patients 1-5:
Cell type label (annotated at the top of each heatmap)
Cancer cell
Fibroblast/Stromal
Smooth muscle
Endothelial
NK/T cell
NK/T cell
Macrophage
Mast cell
B cell
Patients 1-5 (as shown in Figure 3D)
Patient 1-specific
Patient 2-specific
Patient 3-specific
Patient 4-specific
Patient 5-specific
34,231 statistically significant distal peak-to-gene links (correlation ≥ 0.45)
51,263 statistically significant distal peak-to-gene links (correlation ≥ 0.45)
57,448 statistically significant distal peak-to-gene links (correlation ≥ 0.45)
47,394 statistically significant distal peak-to-gene links (correlation ≥ 0.45)
66,278 statistically significant distal peak-to-gene links (correlation ≥ 0.45)
47,993 statistically significant distal peak-to-gene links (correlation ≥ 0.45)
-2
2
Peak accessibility
Gene expression
B
Venn diagrams showing the overlap of statistically significant distal peak-to-gene links with correlation ≥ 0.45 between the original Figure 3D anlaysis and each patient-specific analysis:
(overlap of peak-gene pair character strings, NOT genomic coordinates)
Patient 1-specific
Patient 2-specific
Patient 3-specific
Patient 4-specific
Patient 5-specific
Patients 1-5
17,007
17,811
14,005
21,280
17,910
Patient 1-specific analysis recovers 49.68% of peak-to-gene links from the original anlaysis
Patient 2-specific analysis recovers 52.03% of peak-to-gene links from the original anlaysis
Patient 3-specific analysis recovers 40.91% of peak-to-gene links from the original anlaysis
Patient 4-specific analysis recovers 62.17% of peak-to-gene links from the original anlaysis
Patient 5-specific analysis recovers 52.32% of peak-to-gene links from the original anlaysis

Figure S12. Distal peak-to-gene links identified in the Endometrioid Endometrial Cancer patient cohort (Patients 1-5) are recovered by patient-specific peak-to-gene correlation analyses, Related to Figure 3.

A) Row-scaled heatmaps of statistically significant distal peak-to-gene links where each row represents the expression of a gene (*right*) correlated to the accessibility of a distal peak (*left*). The leftmost heatmap was presented in **Fig. 3D** and the remaining heatmaps represent peak-to-gene analyses using only cells from the indicated patients.

B) Venn diagrams showing the overlap of peak-gene pair terms between the original **Fig. 3D** analysis and each patient-specific analysis. The percentage of peak-to-gene links recovered from the original Fig. 3D analysis is stated below each Venn diagram.

Figure S13

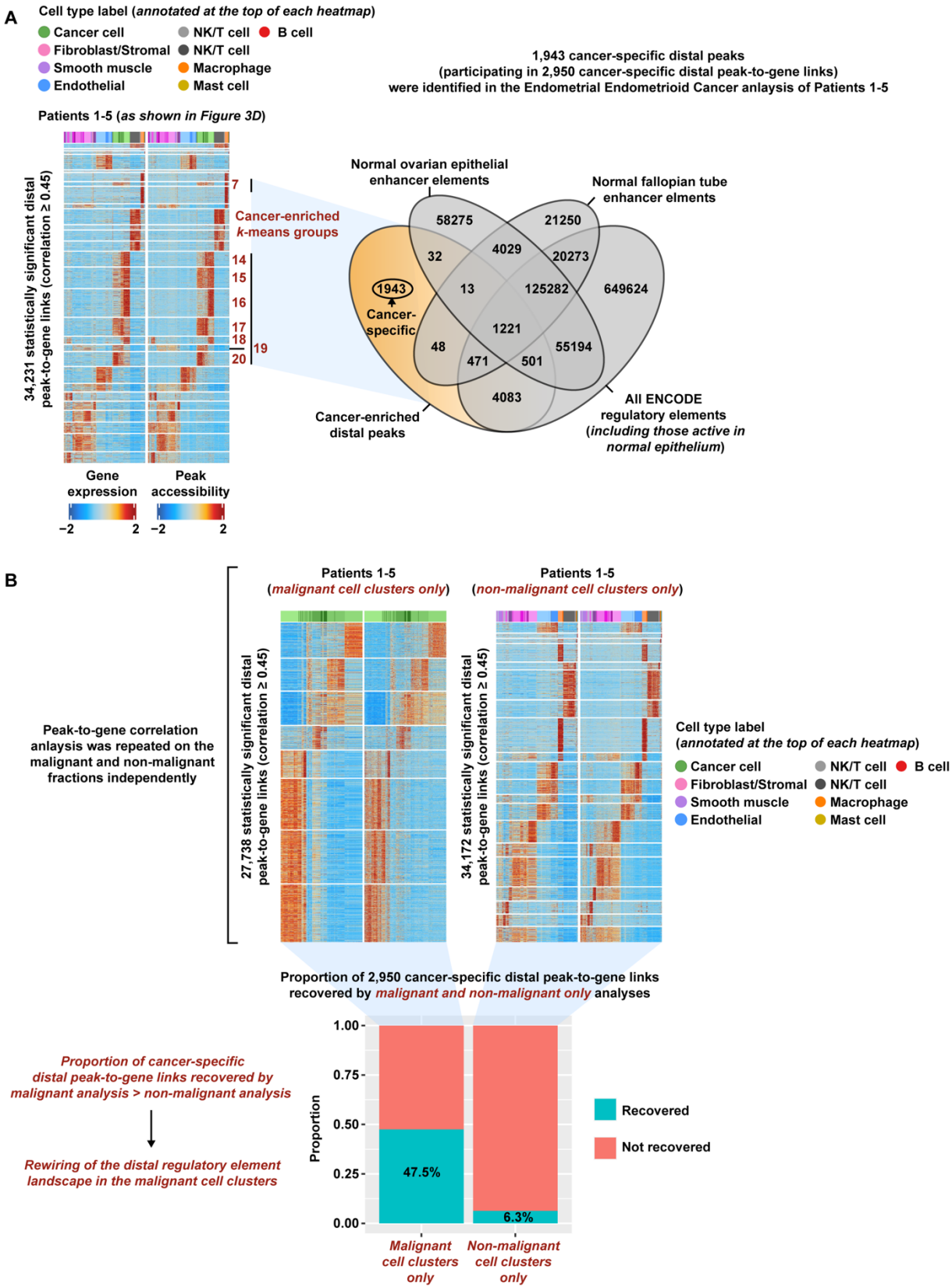
A
Cell type label (annotated at the top of each heatmap)
Cancer cell
Fibroblast/Stromal
Smooth muscle
Endothelial
NK/T cell
NK/T cell
Macrophage
Mast cell
B cell
1,943 cancer-specific distal peaks
(participating in 2,950 cancer-specific distal peak-to-gene links)
were identified in the Endometrial Endometrioid Cancer anlaysis of Patients 1-5
Patients 1-5 (as shown in Figure 3D)
34,231 statistically significant distal peak-to-gene links (correlation ≥ 0.45)
7
14
15
16
17
18
19
20
Cancer-enriched k-means groups
Gene expression
Peak accessibility
−2
2
−2
2
Normal ovarian epithelial enhancer elements
Normal fallopian tube enhancer elments
58275
21250
32
4029
20273
1943
Cancer-specific
13
125282
649624
1221
48
55194
471
501
4083
Cancer-enriched distal peaks
All ENCODE regulatory elements (including those active in normal epithelium)
B
Peak-to-gene correlation anlaysis was repeated on the malignant and non-malignant fractions independently
Patients 1-5 (malignant cell clusters only)
Patients 1-5 (non-malignant cell clusters only)
27,738 statistically significant distal peak-to-gene links (correlation ≥ 0.45)
34,172 statistically significant distal peak-to-gene links (correlation ≥ 0.45)
Cell type label (annotated at the top of each heatmap)
Cancer cell
Fibroblast/Stromal
Smooth muscle
Endothelial
NK/T cell
NK/T cell
Macrophage
Mast cell
B cell
Proportion of 2,950 cancer-specific distal peak-to-gene links recovered by malignant and non-malignant only analyses
Proportion of cancer-specific distal peak-to-gene links recovered by malignant analysis > non-malignant analysis
Rewiring of the distal regulatory element landscape in the malignant cell clusters
Proportion
1.00
0.75
0.50
0.25
0.00
47.5%
6.3%
Recovered
Not recovered
Malignant cell clusters only
Non-malignant cell clusters only

Figure S13. Cancer-specific distal peak-to-gene links identified in the Endometrioid Endometrial Cancer patient cohort (Patients 1-5) are recovered by a malignant-specific peak-to-gene correlation analysis, Related to Figure 3.

A) Row-scaled heatmaps of 34,231 statistically significant distal peak-to-gene links where each row represents the expression of a gene (*left*) correlated to the accessibility of a distal peak (*right*). Peak-to-gene links are grouped into *k*-means clusters and cancer-enriched *k*-means clusters are marked in red text. Cancer-enriched peaks that participate in cancer-enriched *k*-means groups are used as input into the genomic coordinate overlap analysis shown in the Venn diagram (*far right*). The Venn diagram shows the number of cancer-specific distal peaks (*orange slice*) after overlapping the genomic coordinates of cancer-enriched distal peaks with the genomic coordinates of normal ovarian surface epithelium enhancer elements, normal fallopian tube enhancer elements, and all ENCODE regulatory element annotations (*gray circles*).

B) Row-scaled heatmaps of statistically significant distal peak-to-gene links, as in **A**, but using only cells from malignant cell type subclusters (*left*) and non-malignant cell type subclusters (*right*). The proportion of cancer-specific distal peak-to-gene links recovered by each malignant and non-malignant-specific analysis is plotted in the stacked bar chart (*bottom*).

Figure S14

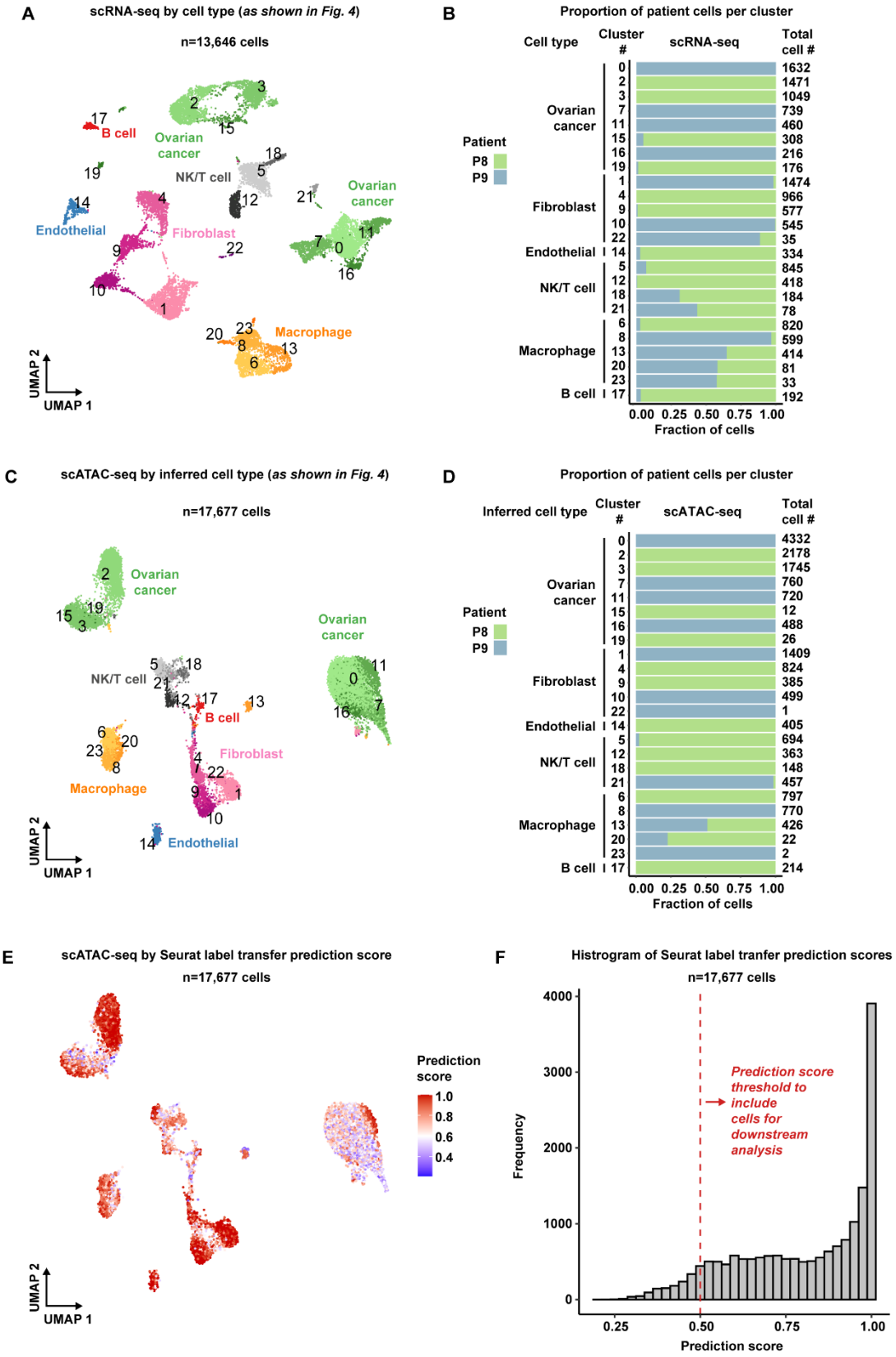
A
scRNA-seq by cell type (as shown in Fig. 4)
n=13,646 cells
B cell
Ovarian cancer
NK/T cell
Endothelial
Fibroblast
Ovarian cancer
Macrophage
UMAP 2
UMAP 1
B
Proportion of patient cells per cluster
Cell type
Cluster #
scRNA-seq
Total cell #
Patient
P8
P9
Ovarian cancer
Fibroblast
Endothelial
NK/T cell
Macrophage
B cell
0.00 0.25 0.50 0.75 1.00
Fraction of cells
C
scATAC-seq by inferred cell type (as shown in Fig. 4)
n=17,677 cells
Ovarian cancer
NK/T cell
B cell
Fibroblast
Macrophage
Endothelial
UMAP 2
UMAP 1
D
Proportion of patient cells per cluster
Inferred cell type
Cluster #
scATAC-seq
Total cell #
Patient
P8
P9
Ovarian cancer
Fibroblast
Endothelial
NK/T cell
Macrophage
B cell
0.00 0.25 0.50 0.75 1.00
Fraction of cells
E
scATAC-seq by Seurat label transfer prediction score
n=17,677 cells
Prediction score
1.0
0.8
0.6
0.4
UMAP 2
UMAP 1
F
Histrogram of Seurat label tranfer prediction scores
n=17,677 cells
Frequency
Prediction score threshold to include cells for downstream analysis
Prediction score

Figure S14. Expanded view of matched scRNA-seq and scATAC-seq for the High-Grade Serous Ovarian Cancer patient cohort (Patient 8 & 9), Related to Figure 4.
A) UMAP plot of 13,646 scRNA-seq cells color-coded by cell type across the **HGSOC patient cohort**. Cell type subclusters as determined by graph-based Louvain clustering are labeled on the UMAP.
B) Stacked bar chart showing the contribution of each patient to each cell type subcluster in scRNA-seq.
C) UMAP plot of 17,677 scATAC-seq cells color-coded by inferred cell type across the **HGSOC patient cohort**. Inferred cell type subclusters are labeled on the UMAP.
D) Stacked bar chart showing the contribution of each patient to each inferred cell type subcluster in scATAC-seq.
E) UMAP plot of 17,677 scATAC-seq cells, as in **C**, but colored by prediction score.
F) Histogram of scATAC-seq inferred cell type subcluster prediction scores. The dashed vertical line in red at 0.5 represents a threshold cutoff for including cells in downstream analysis.

Figure S15

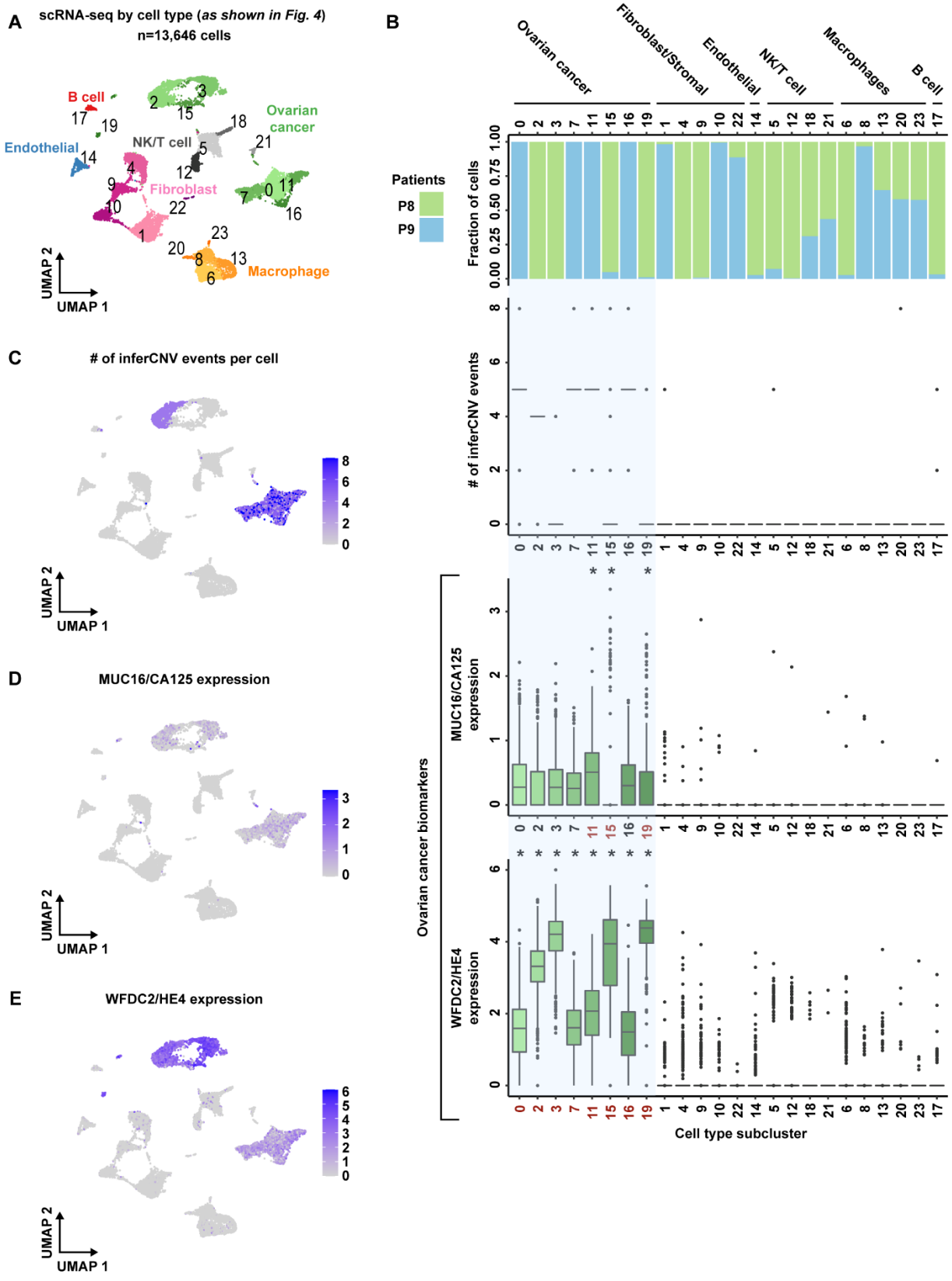
A
scRNA-seq by cell type (as shown in Fig. 4)
n=13,646 cells
B cell
Endothelial
NK/T cell
Ovarian cancer
Fibroblast
Macrophage
UMAP 1
UMAP 2
B
Ovarian cancer
Fibroblast/Stromal
Endothelial
NK/T cell
Macrophages
B cell
Patients
P8
P9
Fraction of cells
# of inferCNV events
Ovarian cancer biomarkers
MUC16/CA125 expression
WFDC2/HE4 expression
Cell type subcluster
C
# of inferCNV events per cell
D
MUC16/CA125 expression
E
WFDC2/HE4 expression

Figure S15. Identification of malignant cell type subclusters in scRNA-seq across the High-Grade Serous Ovarian Cancer patient cohort (Patient 8 & 9), Related to Figure 4.
A) UMAP plot of 13,646 scRNA-seq cells color-coded by cell type and cell type subcluster number.
B) Stacked bar chart showing the contribution of each patient to each cell type subcluster in scRNA-seq (*top*). Boxplot showing the number of inferCNV events in each cell type subcluster (*second from top*). The blue shadow highlights cell type subclusters that are predicted to be malignant. Boxplot showing the expression level of ovarian cancer biomarker *MUC16*/CA125 across all cell type subclusters. The blue shadow highlights cell type subclusters that are predicted to be malignant. Asterisks denote a statistically significant difference in gene expression (Wilcoxon Rank Sum test, Bonferroni-corrected *p*-value <0.01) between the marked cell type subcluster and the remaining cells outside of the blue shadow (*second from bottom*). Boxplot showing the expression level of ovarian cancer biomarker *WFDC2*/HE4 across all cell type subclusters. The blue shadow highlights cell type subclusters that are predicted to be malignant. Asterisks denote a statistically significant difference in gene expression (Wilcoxon Rank Sum test, Bonferroni-corrected *p*-value <0.01) between the marked cell type subcluster and the remaining cells outside of the blue shadow (*bottom*).
C) UMAP plot of 13,646 scRNA-seq cells, as in **A**, but colored by total number of inferCNV events per cell.
D) UMAP plot of 13,646 scRNA-seq cells, as in **A**, but colored by the normalized expression level of ovarian cancer biomarker *MUC16*/CA125.
E) UMAP plot of 13,646 scRNA-seq cells, as in **A**, but colored by the normalized expression level of ovarian cancer biomarker *WFDC2*/HE4.

Figure S16

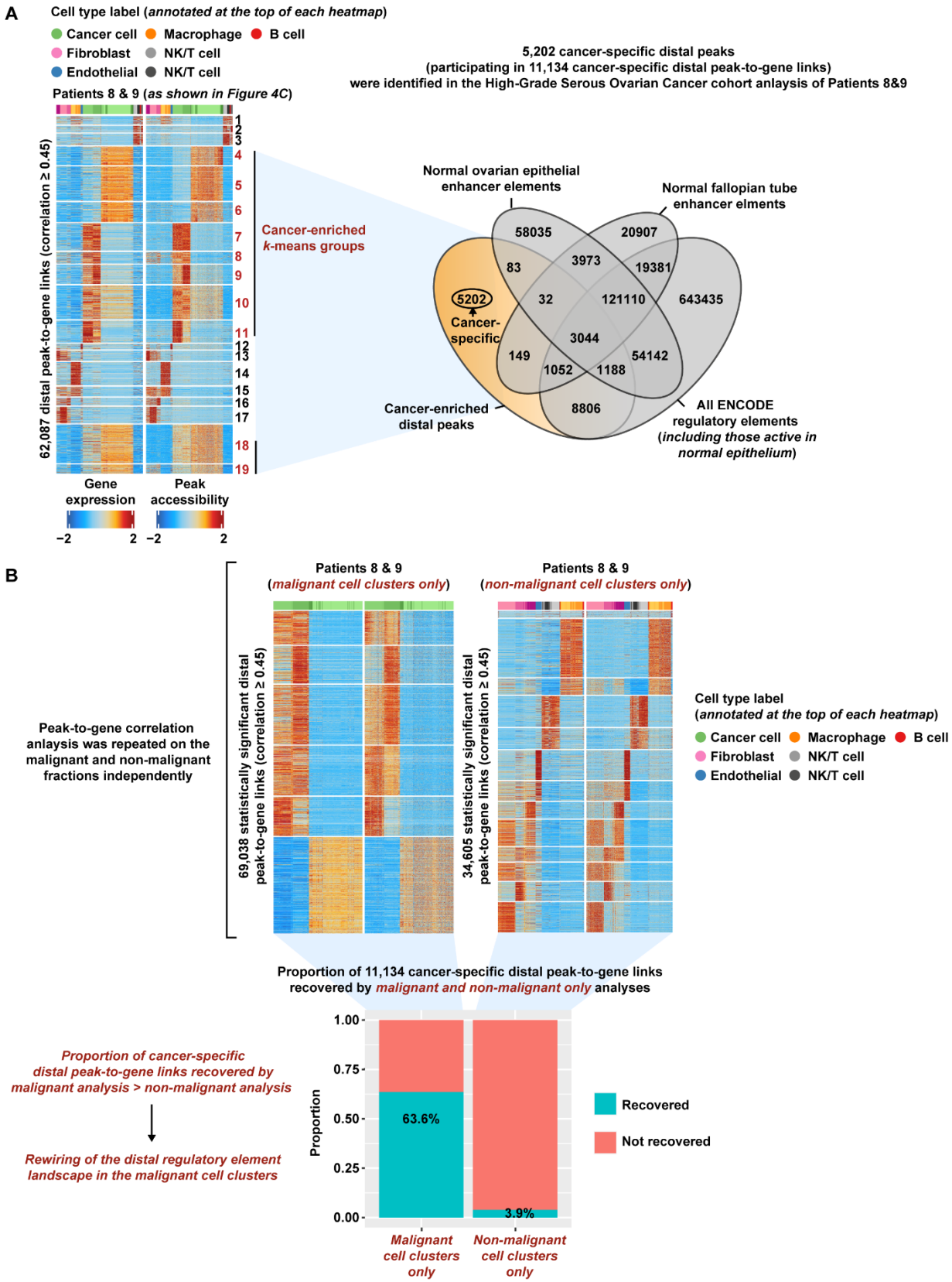
A
Cell type label (annotated at the top of each heatmap)
Cancer cell
Macrophage
B cell
Fibroblast
NK/T cell
Endothelial
NK/T cell
Patients 8 & 9 (as shown in Figure 4C)
62,087 distal peak-to-gene links (correlation ≥ 0.45)
Cancer-enriched k-means groups
Gene expression
Peak accessibility
−2
2
−2
2
5,202 cancer-specific distal peaks
(participating in 11,134 cancer-specific distal peak-to-gene links)
were identified in the High-Grade Serous Ovarian Cancer cohort anlaysis of Patients 8&9
Normal ovarian epithelial enhancer elements
Normal fallopian tube enhancer elments
58035
20907
83
3973
19381
5202
Cancer-specific
32
121110
643435
3044
149
54142
1052
1188
8806
Cancer-enriched distal peaks
All ENCODE regulatory elements (including those active in normal epithelium)
B
Patients 8 & 9 (malignant cell clusters only)
Patients 8 & 9 (non-malignant cell clusters only)
69,038 statistically significant distal peak-to-gene links (correlation ≥ 0.45)
34,605 statistically significant distal peak-to-gene links (correlation ≥ 0.45)
Peak-to-gene correlation anlaysis was repeated on the malignant and non-malignant fractions independently
Cell type label (annotated at the top of each heatmap)
Cancer cell
Macrophage
B cell
Fibroblast
NK/T cell
Endothelial
NK/T cell
Proportion of 11,134 cancer-specific distal peak-to-gene links recovered by malignant and non-malignant only analyses
Proportion of cancer-specific distal peak-to-gene links recovered by malignant analysis > non-malignant analysis
Rewiring of the distal regulatory element landscape in the malignant cell clusters
Proportion
1.00
0.75
0.50
0.25
0.00
63.6%
3.9%
Recovered
Not recovered
Malignant cell clusters only
Non-malignant cell clusters only

Figure S16. Cancer-specific distal peak-to-gene links identified in the High-Grade Serous Ovarian Cancer patient cohort (Patients 8 & 9) are recovered by a malignant-specific peak-to-gene correlation analysis, Related to Figure 4.

A) Row-scaled heatmaps of 62,087 statistically significant distal peak-to-gene links where each row represents the expression of a gene (*left*) correlated to the accessibility of a distal peak (*right*). Peak-to-gene links are grouped into *k*-means clusters and cancer-enriched *k*-means clusters are marked in red text. Cancer-enriched peaks that participate in cancer-enriched *k*-means groups are used as input into the genomic coordinate overlap analysis shown in the Venn diagram (*far right*). The Venn diagram shows the number of cancer-specific distal peaks (*orange slice*) after overlapping the genomic coordinates of cancer-enriched distal peaks with the genomic coordinates of normal ovarian surface epithelium enhancer elements, normal fallopian tube enhancer elements, and all ENCODE regulatory element annotations (*gray circles*).

B) Row-scaled heatmaps of statistically significant distal peak-to-gene links, as in **A**, but using only cells from malignant cell type subclusters (*left*) and non-malignant cell type subclusters (*right*). The proportion of cancer-specific distal peak-to-gene links recovered by each malignant and non-malignant-specific analysis is plotted in the stacked bar chart (*bottom*).

Figure S17

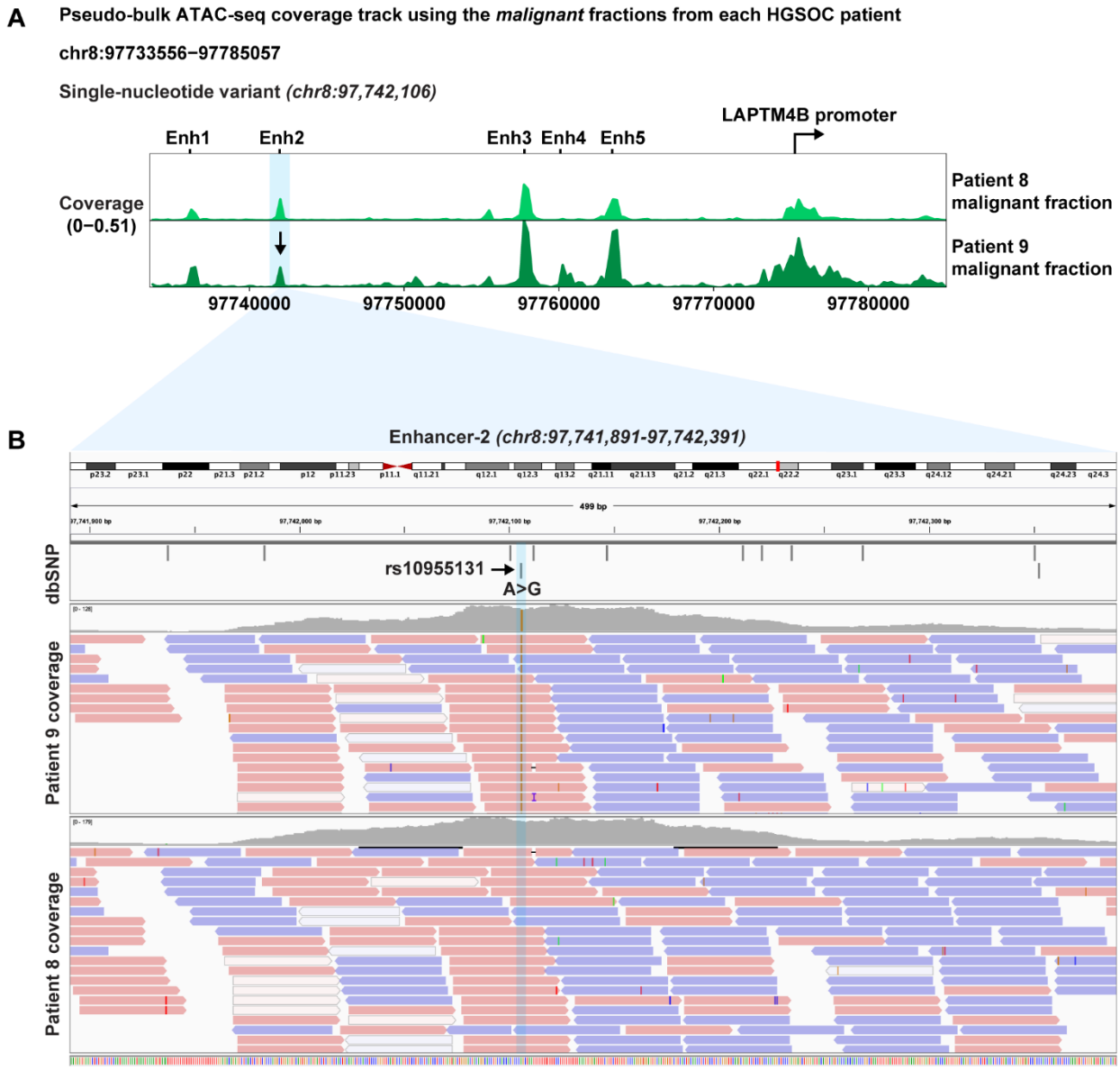
A
Pseudo-bulk ATAC-seq coverage track using the malignant fractions from each HGSOC patient
chr8:97733556−97785057
Single-nucleotide variant (chr8:97,742,106)
LAPTM4B promoter
Enh1
Enh2
Enh3
Enh4
Enh5
Coverage (0−0.51)
Patient 8 malignant fraction
Patient 9 malignant fraction
97740000
97750000
97760000
97770000
97780000
B
Enhancer-2 (chr8:97,741,891-97,742,391)
499 bp
dbSNP
rs10955131
A>G
Patient 9 coverage
Patient 8 coverage

Figure S17. Patient 9 malignant cells harbor an annotated SNP within the *LAPTM4B* Enhancer-2 region, Related to Figure 4.

A) Browser track showing the accessibility of the *LAPTM4B* locus in malignant cell barcodes for the Patient 8 and Patient 9 tumor samples. The blue shadow denotes the Enhancer-2 region while the black arrow denotes the location of a single-nucleotide variant in the Patient 9 malignant cells.

B) Integrated Genomics Viewer (IGV) screenshot showing the pileup of scATAC-seq reads from malignant cell barcodes in the Patient 8 tumor sample (*bottom*) and in the Patient 9 tumor sample (*top*). The blue shadow and black arrow denote the location of the annotated variant present in the scATAC-seq reads from the Patient 9 tumor sample.

Figure S18

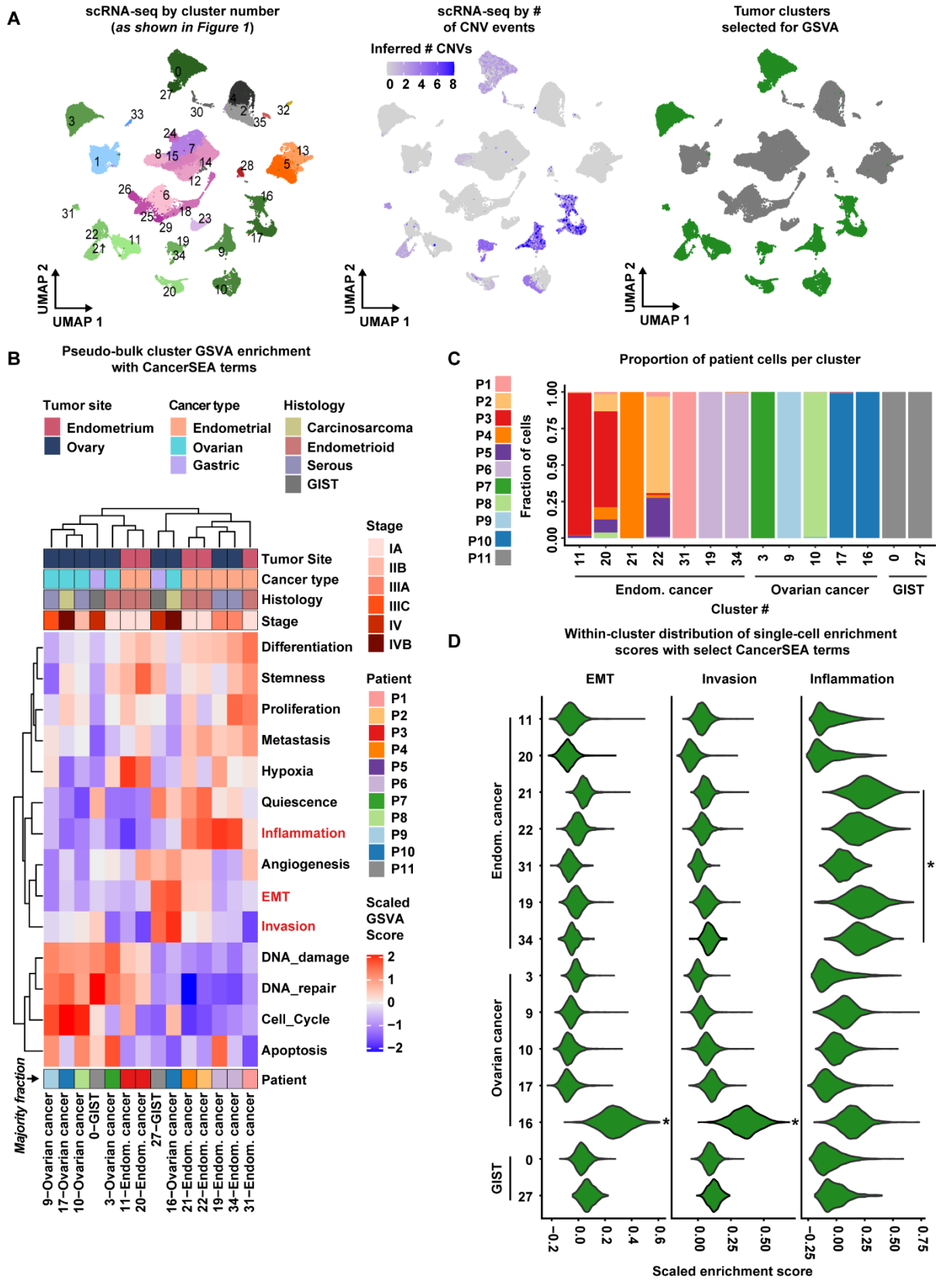
A
scRNA-seq by cluster number
(as shown in Figure 1)
scRNA-seq by #
of CNV events
Inferred # CNVs
0 2 4 6 8
Tumor clusters
selected for GSVA
UMAP 2
UMAP 1
B
Pseudo-bulk cluster GSVA enrichment
with CancerSEA terms
Tumor site
Endometrium
Ovary
Cancer type
Endometrial
Ovarian
Gastric
Histology
Carcinosarcoma
Endometrioid
Serous
GIST
Stage
IA
IIB
IIIA
IIIC
IV
IVB
Patient
P1
P2
P3
P4
P5
P6
P7
P8
P9
P10
P11
Tumor Site
Cancer type
Histology
Stage
Differentiation
Stemness
Proliferation
Metastasis
Hypoxia
Quiescence
Inflammation
Angiogenesis
EMT
Invasion
DNA_damage
DNA_repair
Cell_Cycle
Apoptosis
Patient
Scaled
GSVA
Score
2
1
0
−1
−2
Majority fraction
9−Ovarian cancer
17−Ovarian cancer
10−Ovarian cancer
0−GIST
3−Ovarian cancer
11−Endom. cancer
20−Endom. cancer
27−GIST
16−Ovarian cancer
21−Endom. cancer
22−Endom. cancer
19−Endom. cancer
34−Endom. cancer
31−Endom. cancer
C
Proportion of patient cells per cluster
Fraction of cells
0.00
0.25
0.50
0.75
1.00
11
20
21
22
31
19
34
3
9
10
17
16
0
27
Endom. cancer
Ovarian cancer
GIST
Cluster #
D
Within-cluster distribution of single-cell enrichment
scores with select CancerSEA terms
EMT
Invasion
Inflammation
Endom. cancer
Ovarian cancer
GIST
−0.2
0.0
0.2
0.4
0.6
0.00
0.25
0.50
−0.25
0.00
0.25
0.50
0.75
Scaled enrichment score

Figure S18. Profiling malignant cell clusters in scRNA-seq with CancerSEA gene signatures representing 14 functional states of cancer, Related to Figures 1 & 6.
A) UMAP plot of 75,523 scRNA-seq cells color-coded by cell type, by number of inferred CNVs and by selection for GSVA (*left, middle, right, respectively*). Note that the CNV events were inferred for each patient tumor individually and each CNV region has a posterior probably of being a normal diploid state <0.05 as determined by inferCNV's Bayesian Network Latent Mixture Model.
B) Unsupervised hierarchical clustering heatmap of GSVA enrichment scores for 14 malignant cell type pseudo-bulk profiles. The patient identity for each pseudo-bulk cell type profile was determined based on the majority fraction of patient cells. Notable CancerSEA enrichment terms are marked in red text.
C) Stacked bar chart showing the contribution of each patient to each cell type subcluster in scRNA-seq.
D) Violin plots showing the within cluster distribution of single cell enrichment scores for three CancerSEA terms as calculated by Seurat's *AddModuleScore()*. Asterisks denote statistically significant differences in gene signature enrichment across all subclusters (Kruskal-Wallis test, *p*-value <0.01).

Figure S19

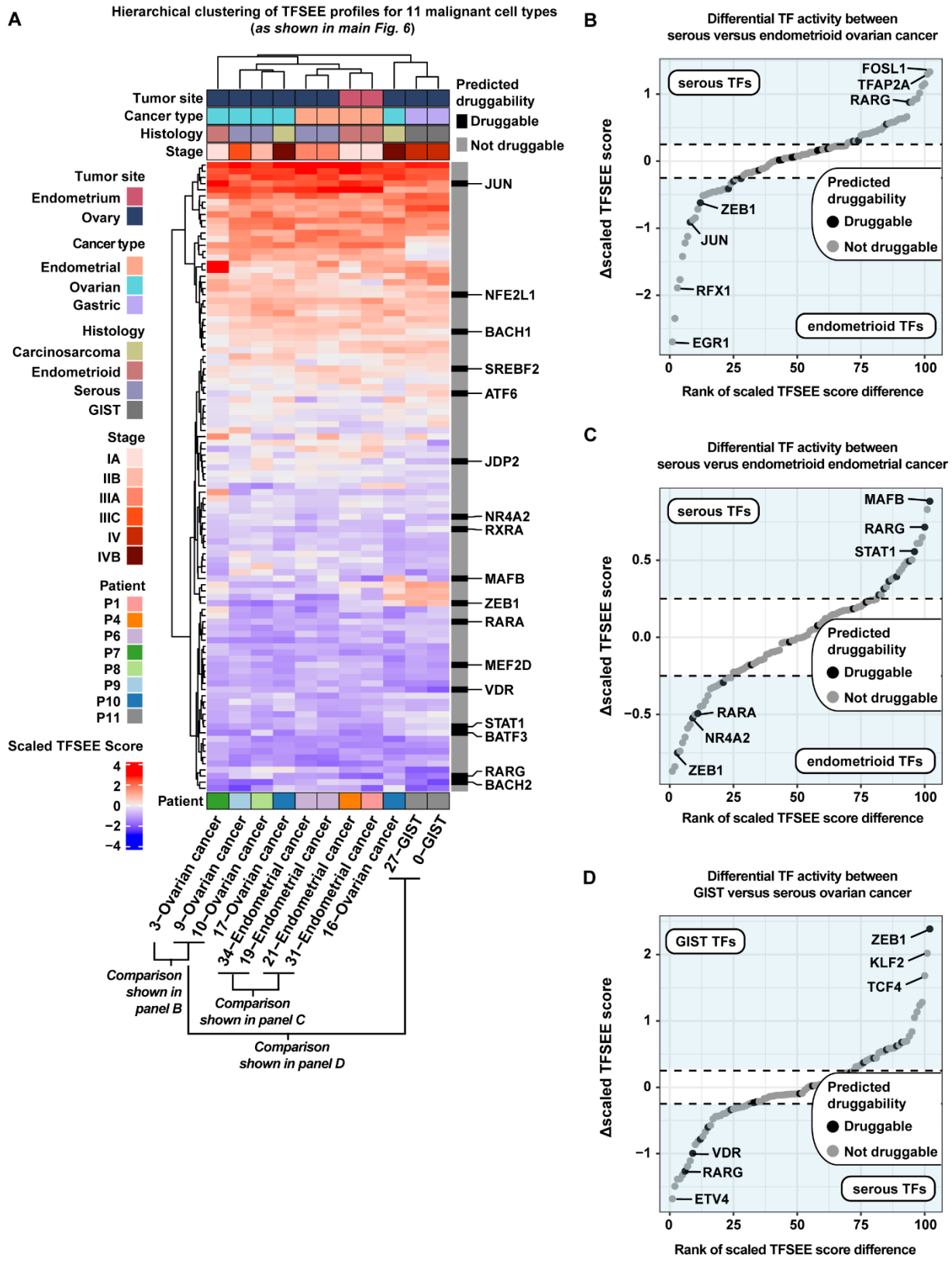
A
Hierarchical clustering of TFSEE profiles for 11 malignant cell types
(as shown in main Fig. 6)
Tumor site
Cancer type
Histology
Stage
Predicted druggability
Druggable
Not druggable
Tumor site
Endometrium
Ovary
Cancer type
Endometrial
Ovarian
Gastric
Histology
Carcinosarcoma
Endometrioid
Serous
GIST
Stage
IA
IIB
IIIA
IIIC
IV
IVB
Patient
P1
P4
P6
P7
P8
P9
P10
P11
Scaled TFSEE Score
4
2
0
−2
−4
JUN
NFE2L1
BACH1
SREBF2
ATF6
JDP2
NR4A2
RXRA
MAFB
ZEB1
RARA
MEF2D
VDR
STAT1
BATF3
RARG
BACH2
Patient
3–Ovarian cancer
9–Ovarian cancer
10–Ovarian cancer
17–Ovarian cancer
34–Endometrial cancer
19–Endometrial cancer
21–Endometrial cancer
31–Endometrial cancer
16–Ovarian cancer
27–GIST
0–GIST
Comparison shown in panel B
Comparison shown in panel C
Comparison shown in panel D
B
Differential TF activity between serous versus endometrioid ovarian cancer
serous TFs
endometrioid TFs
FOSL1
TFAP2A
RARG
ZEB1
JUN
RFX1
EGR1
Predicted druggability
Druggable
Not druggable
Δscaled TFSEE score
Rank of scaled TFSEE score difference
C
Differential TF activity between serous verus endometrioid endometrial cancer
serous TFs
endometrioid TFs
MAFB
RARG
STAT1
RARA
NR4A2
ZEB1
Predicted druggability
Druggable
Not druggable
Δscaled TFSEE score
Rank of scaled TFSEE score difference
D
Differential TF activity between GIST versus serous ovarian cancer
GIST TFs
serous TFs
ZEB1
KLF2
TCF4
VDR
RARG
ETV4
Predicted druggability
Druggable
Not druggable
Δscaled TFSEE score
Rank of scaled TFSEE score difference

Figure S19. Additional TFSEE analyses reveal active enhancers and their cognate TFs enriched in endometrial, ovarian and GIST tumors, Related to Figure 6.
A) Unsupervised hierarchical clustering heatmap of cell type normalized TFSEE scores (n=102 TFs across active enhancers). Each row of the heatmap represents TF activity across cell type-specific enhancers enriched in each column. Theoretical druggability status for each TF is marked with druggable/not druggable according to the *canSAR* database. Select malignant cell type comparisons are annotated at the bottom for follow up in panels **B-D**.
B) Rank-ordered plot showing the difference in scaled TFSEE score for each TF between the mean of 9-,10-Ovarian cancer profiles and 3-Ovarian cancer (representing serous versus endometrioid ovarian cancer). Each point represents a TF and is colored by theoretical druggability status. Notable TFs enriched in either condition (serous/endometrioid) are labeled in the light blue regions of the plot.
C) Rank-ordered plot showing the difference in scaled TFSEE score for each TF between the mean of 19-,34-Endometrial cancer profiles and the mean of 21-,31-Endometrial cancer profiles (representing serous versus endometrioid endometrial cancer). Each point represents a TF and is colored by theoretical druggability status. Notable TFs enriched in either condition (serous/endometrioid) are labeled in the light blue regions of the plot.
D) Rank-ordered plot showing the difference in scaled TFSEE score for each TF between the mean of 0-,27-GIST profiles and the mean of 9-,10-Ovarian cancer profiles (representing GIST versus serous ovarian cancer). Each point represents a TF and is colored by theoretical druggability status. Notable TFs enriched in either condition (GIST/serous) are labeled in the light blue regions of the plot.

Figure S20

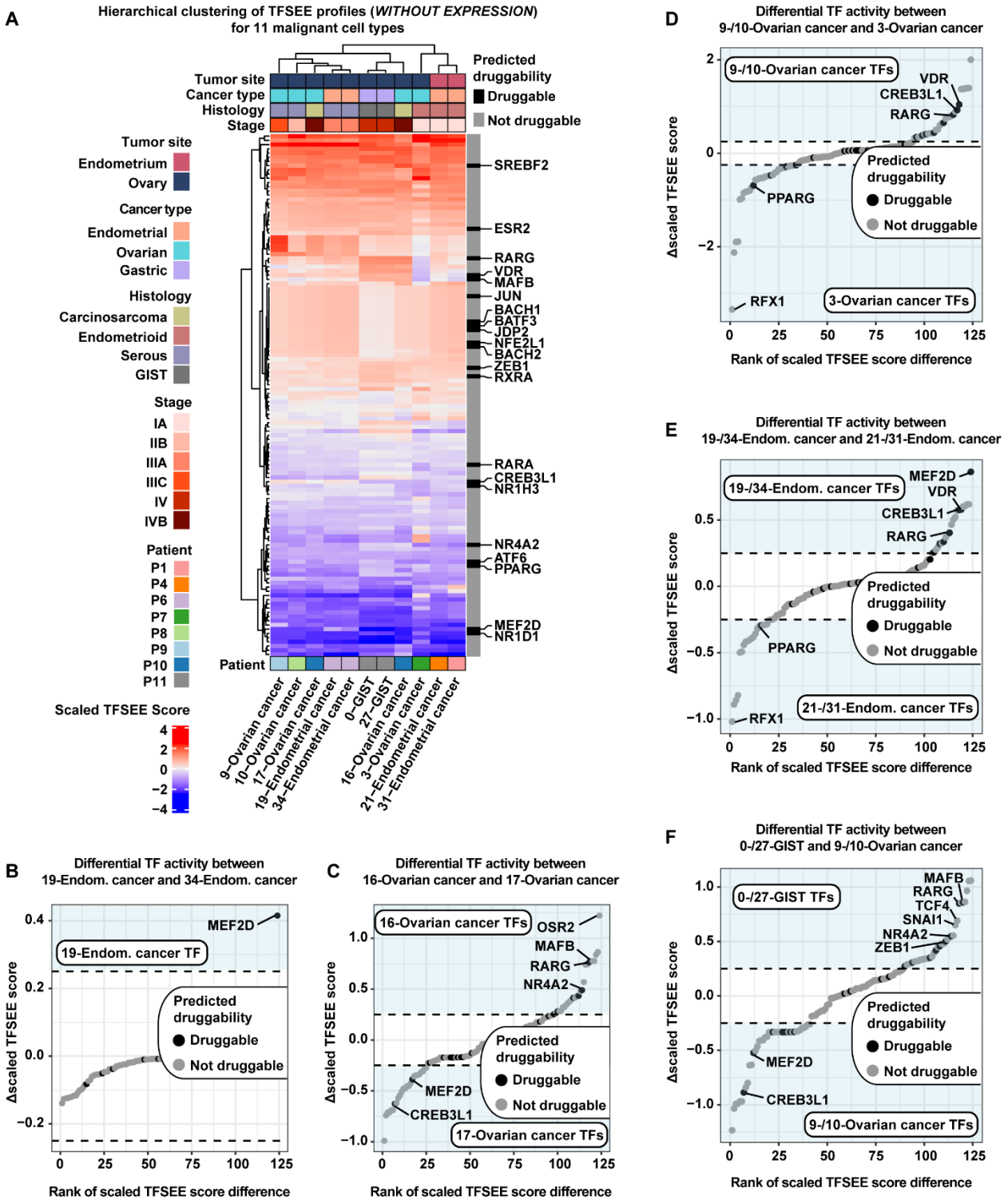
A
Hierarchical clustering of TFSEE profiles (WITHOUT EXPRESSION) for 11 malignant cell types
Predicted druggability
Druggable
Not druggable
Tumor site
Cancer type
Histology
Stage
Patient
SREBF2
ESR2
RARG
VDR
MAFB
JUN
BACH1
BATF3
JDP2
NFE2L1
BACH2
ZEB1
RXRA
RARA
CREB3L1
NR1H3
NR4A2
ATF6
PPARG
MEF2D
NR1D1
Endometrium
Ovary
Endometrial
Ovarian
Gastric
Carcinosarcoma
Endometrioid
Serous
GIST
IA
IIB
IIIA
IIIC
IV
IVB
P1
P4
P6
P7
P8
P9
P10
P11
Scaled TFSEE Score
9–Ovarian cancer
10–Ovarian cancer
17–Ovarian cancer
19–Endometrial cancer
34–Endometrial cancer
0–GIST
27–GIST
16–Ovarian cancer
3–Ovarian cancer
21–Endometrial cancer
31–Endometrial cancer
B
Differential TF activity between 19-Endom. cancer and 34-Endom. cancer
MEF2D
19-Endom. cancer TF
Predicted druggability
Druggable
Not druggable
Δscaled TFSEE score
Rank of scaled TFSEE score difference
C
Differential TF activity between 16-Ovarian cancer and 17-Ovarian cancer
16-Ovarian cancer TFs
OSR2
MAFB
RARG
NR4A2
MEF2D
CREB3L1
17-Ovarian cancer TFs
D
Differential TF activity between 9-/10-Ovarian cancer and 3-Ovarian cancer
9-/10-Ovarian cancer TFs
VDR
CREB3L1
RARG
PPARG
RFX1
3-Ovarian cancer TFs
E
Differential TF activity between 19-/34-Endom. cancer and 21-/31-Endom. cancer
19-/34-Endom. cancer TFs
MEF2D
VDR
CREB3L1
RARG
PPARG
RFX1
21-/31-Endom. cancer TFs
F
Differential TF activity between 0-/27-GIST and 9-/10-Ovarian cancer
0-/27-GIST TFs
MAFB
RARG
TCF4
SNAI1
NR4A2
ZEB1
MEF2D
CREB3L1
9-/10-Ovarian cancer TFs

Figure S20. TFSEE without expression provides alternate viewpoint of enriched enhancers and their cognate TFs in endometrial, ovarian and GIST tumors, Related to Figure 6.

A) Unsupervised hierarchical clustering heatmap of cell type normalized expression-agnostic TFSEE scores (n=124 TFs across active enhancers). Each row of the heatmap represents expression-agnostic TF activity across cell type-specific enhancers enriched in each column. Theoretical druggability status for each TF is marked with druggable/not druggable according to the *canSAR* database.

B) Rank-ordered plot showing the difference in scaled expression-agnostic TFSEE score for each TF between 19-Endometrial cancer and 34-Endometrial cancer (representing subclone 1 versus subclone 2). Each point represents a TF and is colored by theoretical druggability status. Notable TFs enriched in either condition (subclone 1/subclone 2) are labeled in the light blue regions of the plot.

C) Rank-ordered plot, as in **B**, but comparing 16-Ovarian cancer with 17-Ovarian cancer (representing sarcoma versus carcinoma).

D) Rank-ordered plot, as in **B**, but comparing the mean of 9-,10-Ovarian cancer profiles with 3-Ovarian cancer (representing serous versus endometrioid ovarian cancer).

E) Rank-ordered plot, as in **B**, but comparing the mean of 19-,34-Endometrial cancer profiles with the mean of 21-,31-Endometrial cancer profiles (representing serous versus endometrioid endometrial cancer).

F) Rank-ordered plot, as in **B**, but comparing the mean of 0-,27-GIST profiles with the mean of 9-,10-Ovarian cancer profiles (representing GIST versus serous ovarian cancer).

Table S6. sgRNA Oligonucleotides for CRISPRi (listed 5' to 3'), Related to STAR Methods.

Name	Sequence	Genomic Coordinates
sgScramble Fwd	CACCGCTGATCTATCGCGGTCGTC	n/a
sgScramble Rev	AAACGACGACCGCGATAGATCAGC	n/a
sgEnhancer2-1 Fwd	CACCGCAGCGTTACTGCCCTACTA	chr8:97743071-97743090
sgEnhancer2-1 Rev	AAACTAGTAGGGCAGTAACGCTGC	chr8:97743071-97743090
sgEnhancer2-2 Fwd	CACCGTGCGGTGCCTTTCAACTAG	chr8:97743164-97743182
sgEnhancer2-2 Rev	AAACCTAGTTGAAAGGCACCGCAC	chr8:97743164-97743182
sgEnhancer3-1 Fwd	CACCGGTGGATTGCGTTACTAATC	chr8:97758790-97758809
sgEnhancer3-1 Rev	AAACGATTAGTAACGCAATCCACC	chr8:97758790-97758809
sgEnhancer3-2 Fwd	CACCGGATTCCTTGGACATGCCAG	chr8:97759393-97759411
sgEnhancer3-2 Rev	AAACCTGGCATGTCCAAGGAATCC	chr8:97759393-97759411

Table S7. siRNA Sequences (listed 5' to 3', all except GAPDH were used as pools of 4 siRNAs), Related to STAR Methods.

siRNA Name	Sequence
GAPDH	UGGUUUACAUGUUCCAAUA
Non-Targeting #1	UAAGGCUAUGAAGAGAUAC
Non-Targeting #2	AUGUAUUGGCCUGUAUUAG
Non-Targeting #3	AUGAACGUGAAUUGCUCAA
Non-Targeting #4	UGGUUUACAUGUCGACUAA
YY1 #1	GGAUAACUCGGCCAUGAGA
YY1 #2	CAAGAAGAGUUACCUCAGC
YY1 #3	GAACUCACCUCCUGAUUAU
YY1 #4	GCUUAGUAAUGCUACGUGU
CEBPD #1	GGGAGAAGAGCGCCGGCAA
CEBPD #2	GAGAAGAGCGCCGGCAAGA
CEBPD #3	UGGUGGAGCUGUCGGCUGA
CEBPD #4	GCGCCUACAUCGACUCCAU
KLF6 #1	GCCUAGAGCUGGAACGUUA
KLF6 #2	GCAGGAAAGUUUACACCAA
KLF6 #3	UGCAAGAAGUGAUGAGUUA
KLF6 #4	AAAUUGAGCUCCUCUGUCA

Table S8. Oligonucleotides for RT-qPCR (listed 5' to 3'), Related to STAR Methods.

Oligonucleotide name	Sequence
ACTB Fwd	GACGACATGGAGAAAATCTG
ACTB Rev	ATGATCTGGGTCATCTTCTC
GAPDH Fwd	ACAGTTGCCATGTAGACC
GAPDH Rev	TTGAGCACAGGGTACTTTA
LAPTM4B Fwd	TGAACTGGGAGGTGACTTTGAG
LAPTM4B Rev	TTGCTTGTACGCTCCGTAAG
YY1 Fwd	ACAAGAAGTGGGAGCAGAAG
YY1 Rev	TTTTCATCTGAGGACCACATGG
CEBPD Fwd	GCCATGTACGACGACGAGAG
CEBPD Rev	TGTGATTGCTGTTGAAGAGGTC
KLF6 Fwd	GCAACAGACCTGCCTAGAG
KLF6 Rev	TTTTCTCCCGAGCCAGAATG