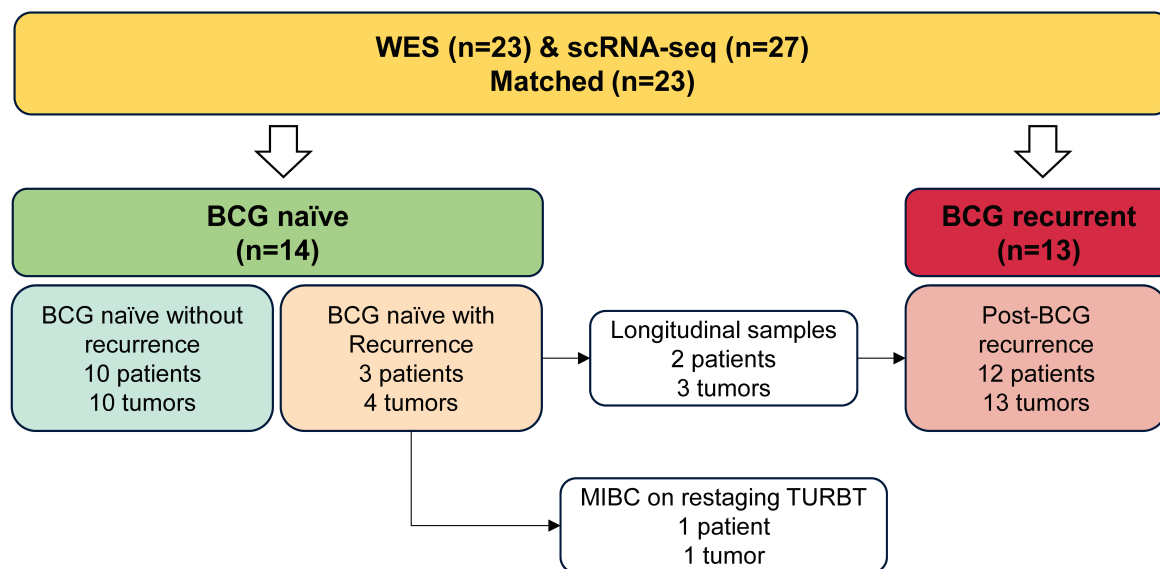


Figure S1



Supplementary Figure S1. Flow diagram of patients used for this study.

A Stroma

100
50
0

● Arterial endothelial
● Venous endothelial
● Interstitial fibroblasts
● Intradetrusor fibroblasts
● Myofibroblasts
● Periuothelial fibroblasts
● Smooth muscle cells

B Immune

100
50
0

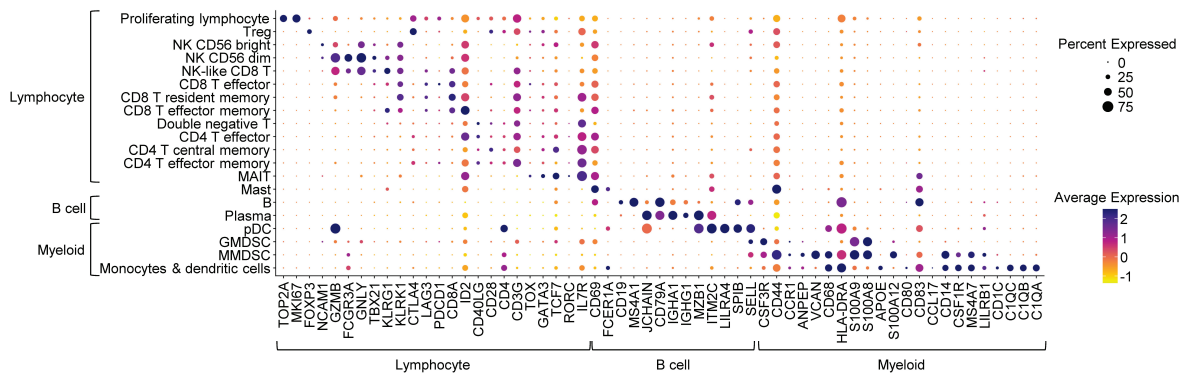
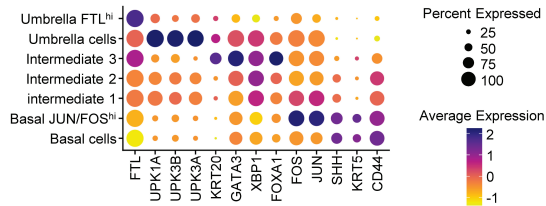
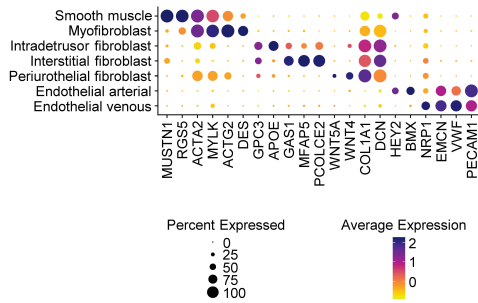
● B cells
● CD4 T central memory
● CD4 T effector
● CD4 effector memory
● CD8 T effector
● CD8 T effector memory
● CD8 T resident memory
● Double negative T
● GMDSC
● MAIT
● Mast cells
● MMDSC
● Monocytes & DC
● NK CD56 bright
● NK CD56 dim
● NK-like CD8 T
● pDC
● Plasma
● Proliferating lymphocytes
● Treg

C Urothelial

100
50
0

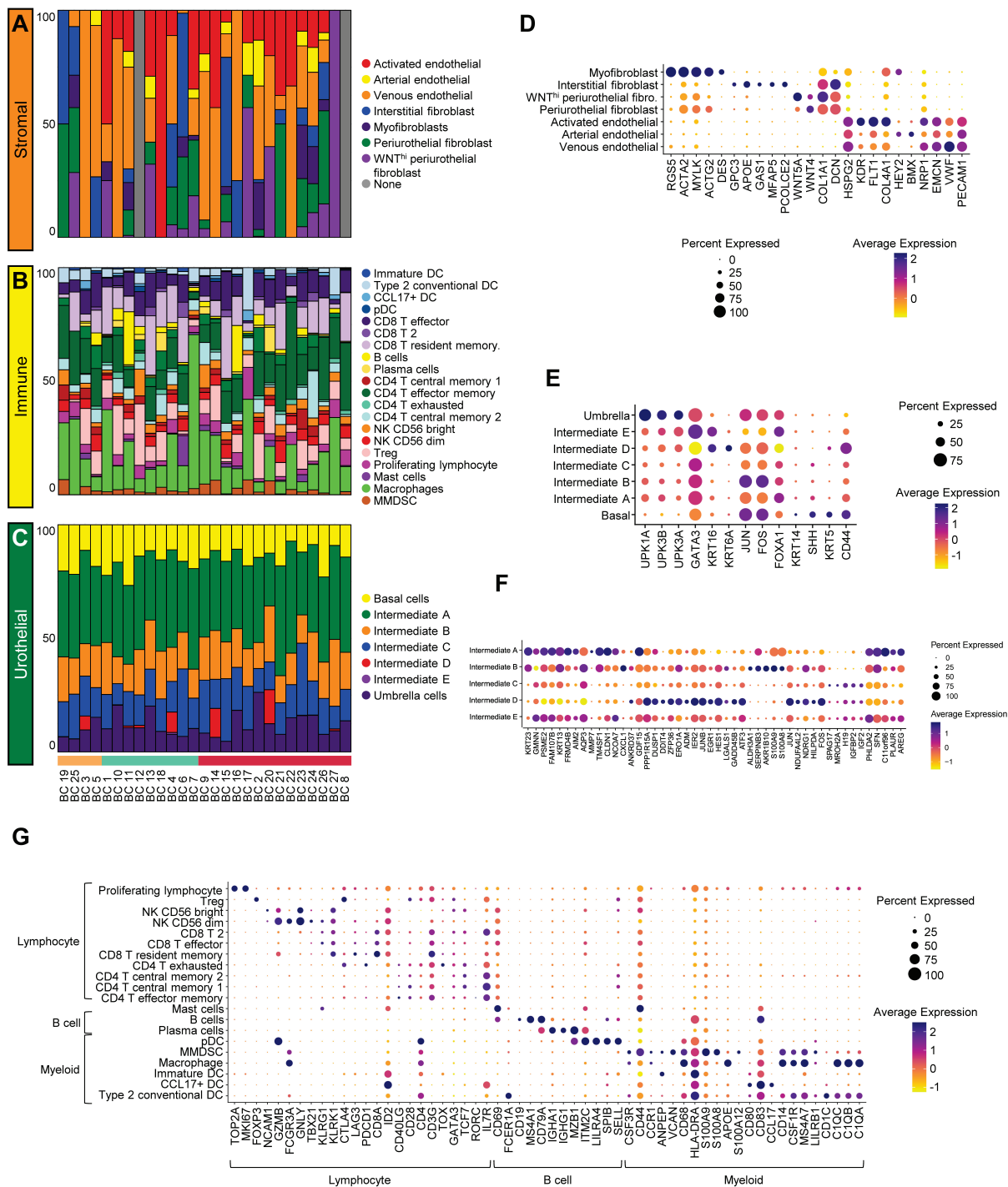
● Basal cells
● Basal JUN/FOS^{hi}
● Intermediate 1
● Intermediate 2
● Intermediate 3
● Umbrella cells
● Umbrella FTL^{hi}

BN 1 BN 2 BN 3 BN 4 BN 5 BN 6 BN 7 BN 8 BN 9 BN 10



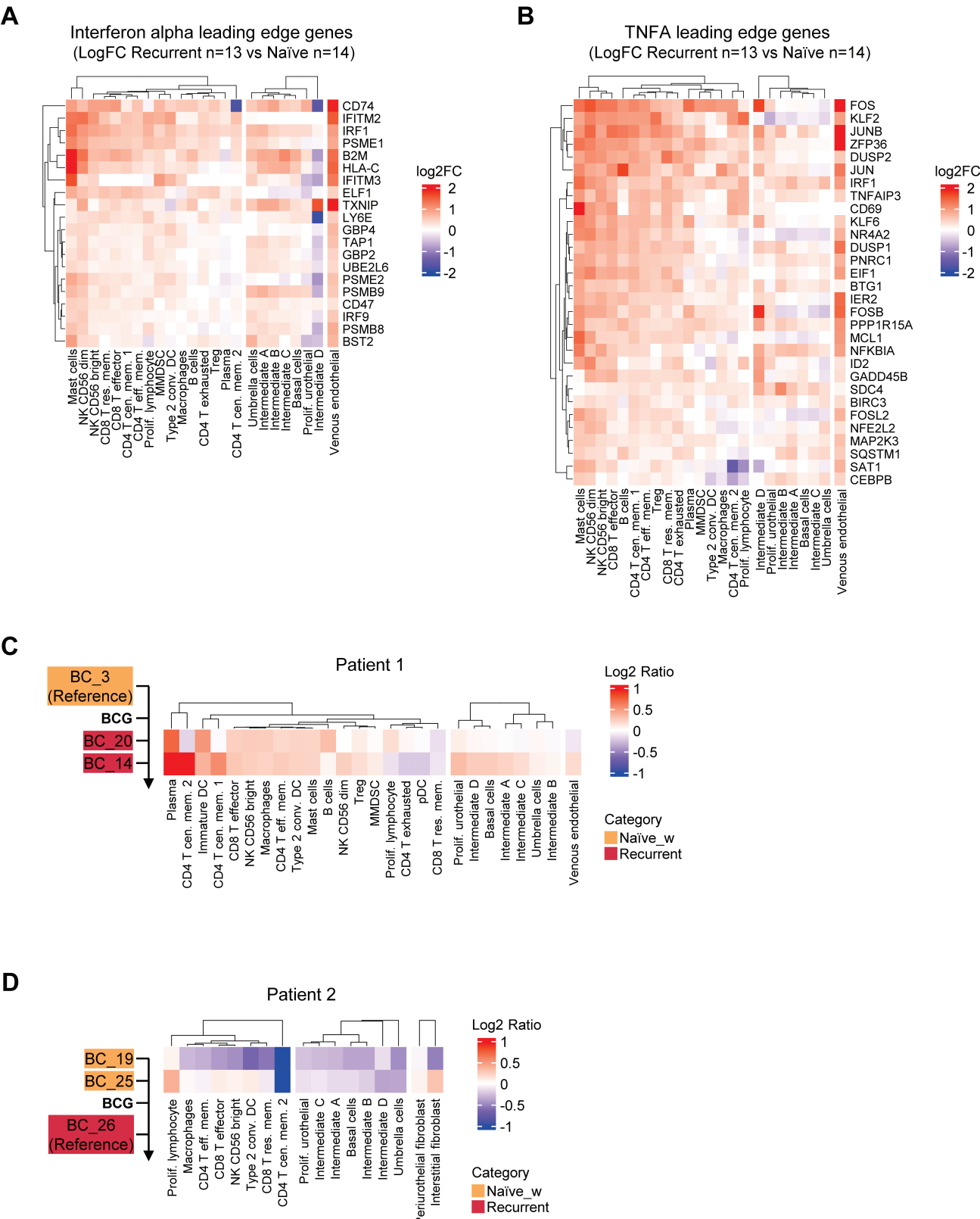
Supplementary Figure S2. Benign bladder scRNA-seq analysis. A-C) Detailed cell typing in stromal (A), immune (B), and urothelial (C) compartments was performed and the fraction of cells within each compartment per sample are represented as cell abundance/sample composition bar graphs. D-F) Detailed cell naming was performed by using AI based prediction tools in BioTuring software along with manually curated markers as represented in dot plots of key markers.

Figure S3



Supplementary Figure S3. NMIBC sample scRNA-seq analysis. A-C) Detailed cell typing in stromal (A), immune (B), and urothelial (C) compartments was performed and the fraction of cells within each compartment per sample are represented as cell abundance/sample composition bar graphs. D-F) Detailed cell naming was performed by using AI based prediction tools, mapping to benign bladder reference UMAP, and validated using manually curated markers as represented in dot plots of key markers.

Figure S4



Supplementary Figure S4. Inflammatory pathway enrichment analysis. A) Log2 fold change of leading-edge genes for the interferon alpha (IFNA) pathway (red= enriched in recurrent, blue= enriched in naïve). B) Log2 fold change of leading-edge genes for the TNF alpha pathway (red= enriched in recurrent, blue= enriched in naïve). C) Patient 1 had a single sample prior to BCG treatment (naïve, orange) and two samples after BCG recurrence (red). Expression of IFNG pathway genes were summarized as ssGSEA scores for each cell type with >100 cells. Using the BCG naïve sample (BC_3) as a reference, the log2 ratio change in IFNG ssGSEA score was plotted for each cell type for the BCG recurrence samples (BC_20 and BC_14). Positive log2 ratio (red) indicates increases in recurrence, and negative ratio (blue) indicates decrease in recurrence. D) Patient 2 had two samples prior to BCG treatment (naïve, orange) and a single sample after BCG recurrence (red). Expression of IFNG pathway genes were summarized as ssGSEA scores for each cell type with >100 cells. Using the BCG recurrence sample (BC_26) as a reference, the log2 ratio change in IFNG ssGSEA score was plotted for each cell type for the BCG naïve samples (BC_19 and BC_25). Positive log2 ratio (red) indicates decrease in recurrence, and negative ratio (blue) indicates increase in recurrence.

A

Naïve Tumors

Recurrent Tumors

T cells to Urothelial

T cells to T cells

Urothelial to T cells

Urothelial to Urothelial

Number of interactions

B

Recurrent
Naïve

CD4 T cen.
mem. 2

CD4 T cen.
mem. 1

CD4 T eff.
mem.

CD8 T eff.

Prolif. lymph.

CD8 T res. mem.

Average Expression

Percent Expressed

C

log₂FC

D

Recurrence-free Survival

Months

p = 0.014

N = 141 68 19 3 1 CD6^{high}
N = 141 88 30 5 0 CD6^{low}

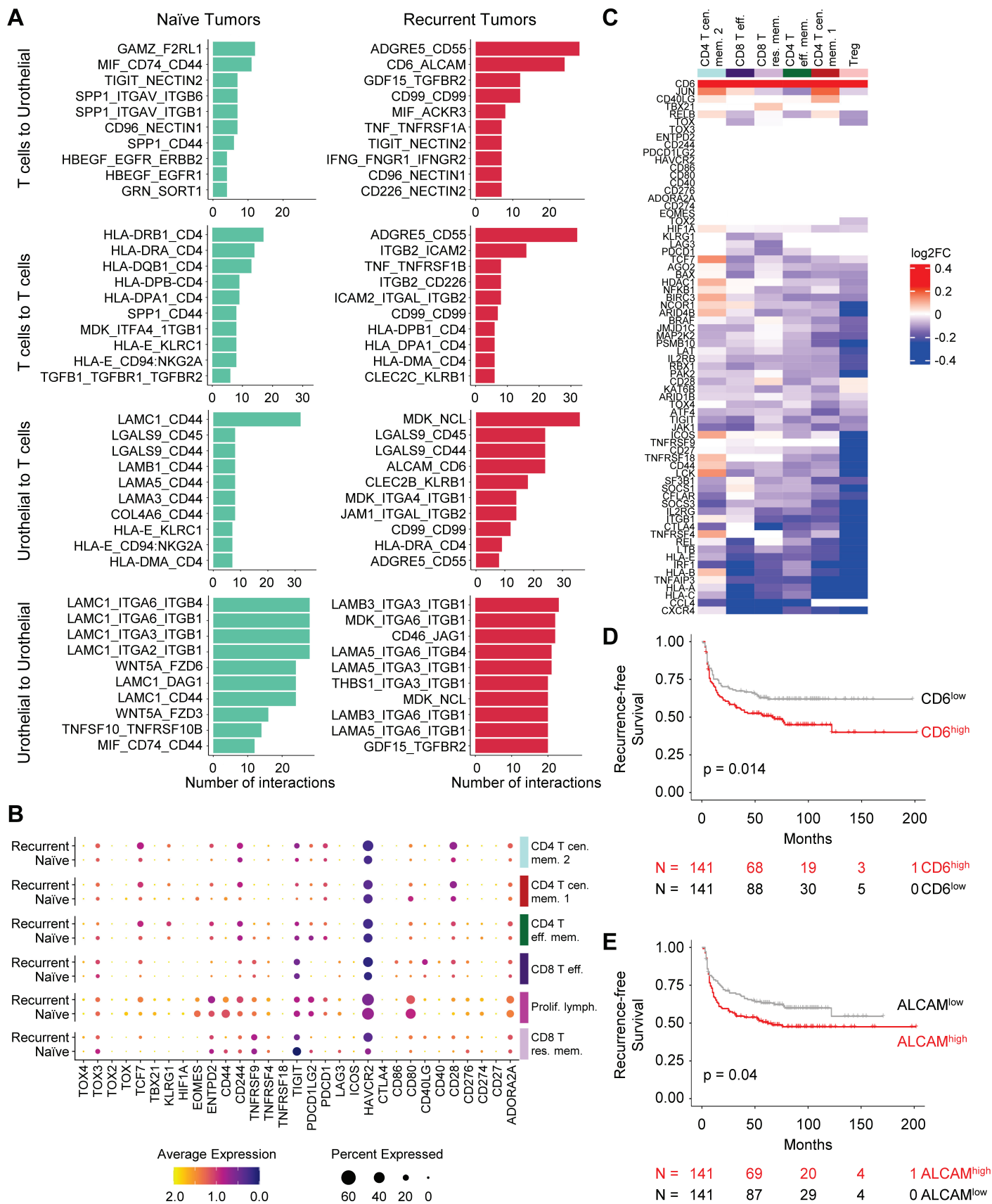
E

Recurrence-free Survival

Months

p = 0.04

N = 141 69 20 4 1 ALCAM^{high}
N = 141 87 29 4 0 ALCAM^{low}



Supplementary Figure S5. T cell characterization in BCG recurrence. A) Top enriched interactions between T cells-to-Urothelial, T cells-to-T cells, Urothelial-to-T cells, and Urothelial-to-Urothelial in naïve (green) and recurrent (red) samples. B) Expression of immune checkpoint related gene in T cell populations with >100 cells per sample in naïve and recurrent specimens. C) Heatmap of differential gene expression in curated T cell activation genes (Nanostring T cell activation gene set) between CD6^{hi} and CD6^{low} (stratified by median CD6 expression) T cell populations. D-E) Time to disease recurrence post-BCG treatment for Jong *et al.* NMIBC cohort (n=282), stratified by median CD6 (D) and ALCAM (E) gene expression. P-values were derived from log-rank test.

Descriptions of Supplementary Data files

Supplementary Data S1. Single cell RNA-seq quality control assessment using Cellranger metrics. Post alignment quality control metrics were collected using Cellranger 5.0.0 as described in the methods and recorded on a per sample basis.

Supplementary Data S2. Differential gene expression between naïve and recurrent samples for each cell type (each tab).

Supplementary Data S3. Hallmark pathway gene set enrichment analysis of differentially expressed genes between naïve and recurrent samples for each cell type (each tab).

Supplementary Data S4. Jong *et al.* bulk RNA-seq cohort data (from original publication supplementary materials) with ssGSEA scores for CD6, ALCAM, and CD6/ALCAM dual signatures.