Supplementary Methods and Materials

Meta-analysis

Literature Search

Search terms used were as follows: (metastatic OR metastasis OR advanced) AND (urothelial carcinoma OR urothelial cancer OR bladder cancer OR upper tract urothelial carcinoma OR transitional cell carcinoma OR ureter OR pelvis) AND (Fibroblast Growth Factor Receptor 3 OR FGFR3 OR Fibroblast Growth Factor Receptor OR FGFR) AND (alteration OR alterations OR mutation OR mutations OR mutant). Detailed search strategies were shown in **Supplementary Table 1**. In addition, reference search was conducted to find additional recruited publications through screening reference lists.

Inclusion and exclusion criteria

Studies were included if they explored the survival and immunotherapy response of metastatic UC patients with FGFR3 mutation compared with those without FGFR3 mutation. PECOS/PICOS (Population/Problem, Exposure/Intervention, Comparator/Control, Outcome, and Study design) method was used to include and exclude the publications. Inclusive criteria were as follows: (P) Population/Problem: metastatic UC patients received ICB; (E/I) Exposure/Intervention: patients with FGFR3 mutation (FGFR3-mutated); (C) Comparator/Control: patients without FGFR3 mutation (FGFR3-wildtype); (O) Outcome: OS, progression free survival (PFS), disease specific survival (DSS), complete response rate, objective response rate, and disease control rate; and (S) Study design: cohort studies. We excluded case reports, letters, comments and reviews.

Quality assessment of studies

All enrolled studies were deemed as cohort studies based on the exposure of interest (FGFR3 mutation). Newcastle-Ottawa Scale [1] was used to evaluate the quality of cohort studies by two authors (YS and YD) independently. Any disagreement would be addressed through discussion of two sides or consultation of the third author (TX). Each study was evaluated according to selection, comparability and outcomes. A study scored 7-9 points was deemed as a high-quality study.

Data extraction

Two independent authors extracted the following data of each study: first author, trial name or cohort name, publication year, method for detection of FGFR3 mutation, sample capacity of two groups, details of ICB, follow-up period, survival outcomes (OS, PFS, and DSS), and pathological responses (complete response rate, objective response rate, and disease control rate). HR and 95%CI of survival outcomes were extracted from time-to-event data. If the study didn't report HR and 95%CI directly, the widely recognized method by Tierney et al [2] was used to calculate the HR as well as 95%CI from the published Kaplan-Meier curves. For Memorial Sloan Kettering Cancer Center cohort [3], we used the individual patient data downloaded from cBioportal [4] database (http://www.cbioportal.org/) to calculate the HR and 95%CI.

Single-cell RNA sequencing analysis

Sample collection

After tumor resection, tumor tissues were cut into two pieces along the long axis: One half was used for single-cell RNA sequencing and FGFR3 mutation status test, and the other half was used for pathological diagnosis.

Preparation of single-cell suspensions

Primary tumor tissue samples were transported in MACS Tissue Storage Solution

(cat#130-100-008, Miltenyi Biotec) at 4℃ immediately after surgical resection. The primary tumor tissues were rinsed with phosphate-buffered saline (PBS; Thermo Fisher Scientific), minced into ~1-mm cubic piece, and ground with a UTTD (ULTRA-TURRAX® Tube Drive) disperser (IKA, Germany). The ground tumor tissue samples were digested by 0.25% trypsin (Gibco, Life Technologies), terminated by H1640 supplemented with 10% fetal bovine serum (Gibco, Life Technologies), and then transferred to 10 ml of digestion medium containing collagenase IV (100 U/ml; Gibco, Life Technologies) and dispase (0.6 U/ml; Gibco, Life Technologies). The digested samples were filtered through a 70-μm nylon mesh. After centrifuging, the pelleted cells were suspended with ice-cold red blood cell lysis buffer (Solarbio) and filtered with a 40-μm nylon mesh. Last, the pelleted cells were suspended with 1 ml of Dulbecco's PBS (Solarbio), and the concentrations of live cells and clumped cells were determined using an automated cell counter (Countstar).

Droplet-based single-cell sequencing

Using the Single Cell 3′ Library and Gel Bead Kit V3 (10X Genomics) and Chromium Single Cell A Chip Kit (10X Genomics), the cell suspension was loaded onto the Chromium single-cell controller (10X Genomics) to generate single-cell gel beads in the emulsion (GEMs) according to the manufacturer's protocol. Briefly, single cells were suspended in PBS containing 0.04% bovine serum albumin. Approximately 10,000 cells were added to each channel. The captured cells were lysed, and the released RNA was barcoded via reverse transcription in individual GEMs. Reverse transcription was performed at 53°C for 45 min, followed by 85°C for 5 min, after which the temperature was held at 4°C. Complementary DNA was generated and amplified, after which its quality was assessed using an Agilent 4200 according to the manufacturer's instructions. Single-cell RNA sequencing libraries

were constructed using the Single Cell 3′ Library Gel Bead Kit V3. Sequencing was performed with Illumina (HiSeq PE150) according to the manufacturer's instructions (Illumina).

Single-cell RNA sequencing data processing

Quality check was performed by fastp (version 0.20.0). Raw gene expression matrices were generated for each sample using the Cell Ranger (version 7.1.2) Pipeline coupled with human reference version GRCh38. The filtered gene expression matrices were analyzed by R software (version 4.1.2) with the Seurat [5] package (version 4.0.0). DoubletFinder R package (version 2.0) was used to computationally infer and remove doublets in each sample individually, with default parameters. In brief, low-quality cells were removed if they met the following conditions: (1) >10000 or <200 genes (2) >50000 UMIs; and (3) $>25\%$ mitochondrial gene counts. Then, the filtered expression matrix was then normalized with the function NormalizeData, followed by the identification of 2000 genes of high cell-to-cell variation by using the function FindVariableFeatures. For multi-sample integration, we employed the function FindIntegrationAnchors to obtain anchors across individual samples. By inputting the anchors into the function IntegrateData and regressing out the influence of library size and percentage of mitochondria genes, we created a batch-corrected expression matrix of all cells on the 2000 highly-variable genes. Based on the batch-corrected data, we performed Principal Component Analysis (PCA) with top 2000 variable features by using the function RunPCA. Cells were then clustered using the functions FindNeighbors and FindClusters with the first 50 principal components (PCs). Finally, we conducted nonlinear dimensional reduction for data visualization. In brief, UMAP was performed on the top 50 PCs by using the function RunUMAP.

Cell type annotation and cluster marker identification

After nonlinear dimensional reduction and projection of all cells into two-dimensional space by UMAP, cells clustered together according to common features. The FindAllMarkers function in Seurat was used to find markers for each of the identified clusters. Clusters were then classified and annotated on the basis of expression of canonical markers of particular cell types. Clusters that expressed two or more canonical cell type markers were classified as doublet cells, and clusters that expressed no canonical cell type markers were classified as low-quality cells. Both doublet cell clusters and low-quality cell clusters were excluded from further analyses.

Subclustering of major cell types

For each major cell type, cells were extracted from the overview integrated dataset first. Next, these major cell types were integrated for further subclustering. After integration, genes were scaled to unit variance. Scaling, PCA, and clustering were performed as described above.

CNV estimation and identification of malignant cells

To infer CNVs from the single-cell RNA sequencing data, we used infercnv R package (version 1.14.2) with the default parameters. Immune cells and stromal cells were considered as putative nonmalignant cells, and their CNV estimates were used to define a baseline [6]. The calculated CNV score was defined as the mean square of the CNV estimates across all genomic locations. All epithelial cells were inputted.

Gene set variation analysis

Pathway analyses were performed on the hallmark gene sets (H) and KEGG gene sets (C2) described in the molecular signature database [7]. To assign pathway activity estimates to individual cells, we applied GSVA using standard settings, as

implemented in the GSVA R package (version 1.48.1) as described previously.

Defining cell scores

We used cell scores to evaluate the degree to which individual cells expressed a certain predefined expression gene set. The cell scores were initially based on the average expression of the genes from the predefined gene set in the respective cell. The AddModuleScore function in Seurat was used to implement the method with default settings.

The scores for evaluating epithelial cells and T cells were from Molecular Signatures Database (MSigDB) (http://www.gsea-msigdb.org/gsea/index.jsp). The progenitor exhausted score and terminally exhausted score for evaluating T cells were from Long et al [8].

Pseudotime trajectory analysis

To characterize the potential process of T functional changes and determine the potential lineage differentiation, we applied the Monocle2 (version 2.16.0) R package [9].

Cell-cell interaction analysis

To enable a systematic analysis of cell-cell communication, we reclustered each cell type. The CellPhoneDB (version 3.0) [10] was then used to explore the interactions of ligand-receptor pairs between cell subtypes. The interactions between distinct cell subpopulations via putative ligand-receptor pairs were visualized using the ggplot2 R package.

Bulk RNA sequencing and microarray data analysis

Bulk RNA-seq data and processing

Bulk RNA-seq expression data of IMvigor210 cohort were extracted from the R

package IMvigor210CoreBiologies. The R package edgeR was used to perform differential gene expression. Differential expressed genes with |logFC|>1 were enrolled into functional enrichment analysis. The relative proportions of 28 immune cell types were calculated based on the immune cell signature reported by Charoentong et al [11]. The gene set scores were initially based on the average expression of the genes from the predefined gene set in the respective sample.

Microarray data and processing

To study the role of FGFR3 in UC, microarray data from the RT-112 bladder cancer cell lines, with or without short hairpin RNA (shRNA) knockdown of FGFR3, were obtained from Gene Expression Omnibus database (accession number: GSE41035) [12]. The R package limma was used to perform differential gene expression. The R packages clusterProfiler and GSVA were used to perform gene set enrichment analysis and gene set variation analysis, respectively.

References:

1. Wells G, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for Assessing The Quality of Nonrandomised Studies in Meta-analyses. Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. *Ottawa Health Research Institute Web site: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp* 2009;7
2. Tierney JF, Stewart LA, Ghersi D, Burdett S, Sydes MR. Practical methods for incorporating summary time-to-event data into meta-analysis. *Trials* 2007;8:16-16.
3. Samstein RM, Lee C, Shoushtari AN, et al. Tumor mutational load predicts

survival after immunotherapy across multiple cancer types. *Nat Genet* 2019;51:202-206.

4. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2012; 2:401-404.
5. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* 2018;36:411-420.
6. Patel AP, Tirosh I, Trombetta JJ, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* 2014;344:1396-1401.
7. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *P Natl Acad Sci Usa* 2005;102:15545-15550.
8. Long Z, Sun C, Tang M, et al. Single-cell multiomics analysis reveals regulatory programs in clear cell renal cell carcinoma. *Cell Discov* 2022;8:68.
9. Trapnell C, Cacchiarelli D, Grimsby J, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* 2014;32:381-386.
10. Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat Protoc* 2020;15:1484-1506.
11. Charoentong P, Finotello F, Angelova M, et al. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell Rep* 2017;18:248-262.
12. Du X, Wang Q, Chan E, et al. FGFR3 stimulates stearoyl CoA desaturase 1

activity to promote bladder tumor growth. *Cancer Res* 2012;72:5843-5855.