Figure S1

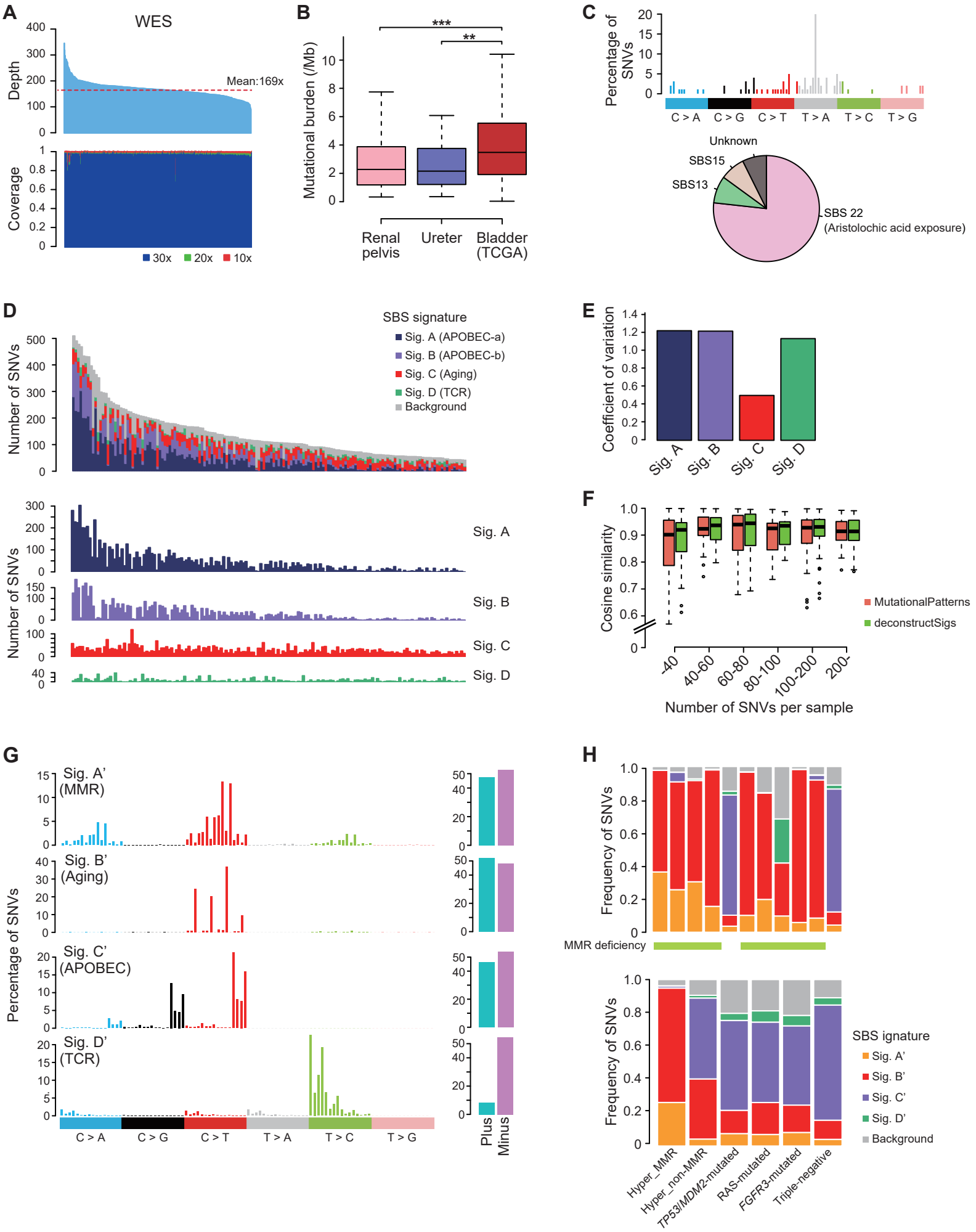

A
WES
Depth
Mean:169x
Coverage
30x
20x
10x
B
Mutational burden (/Mb)
***
**
Renal pelvis
Ureter
Bladder (TCGA)
C
Percentage of SNVs
C > A
C > G
C > T
T > A
T > C
T > G
Unknown
SBS15
SBS13
SBS 22
(Aristolochic acid exposure)
D
SBS signature
Sig. A (APOBEC-a)
Sig. B (APOBEC-b)
Sig. C (Aging)
Sig. D (TCR)
Background
Number of SNVs
Sig. A
Sig. B
Sig. C
Sig. D
E
Coefficient of variation
Sig. A
Sig. B
Sig. C
Sig. D
F
Cosine similarity
MutationalPatterns
deconstructSigs
-40
40-60
60-80
80-100
100-200
200-
Number of SNVs per sample
G
Sig. A' (MMR)
Sig. B' (Aging)
Sig. C' (APOBEC)
Sig. D' (TCR)
Percentage of SNVs
C > A
C > G
C > T
T > A
T > C
T > G
Plus
Minus
H
Frequency of SNVs
MMR deficiency
SBS ignature
Sig. A'
Sig. B'
Sig. C'
Sig. D'
Background
Hyper_MMR
Hyper_non-MMR
TP53/MDM2-mutated
RAS-mutated
FGFR3-mutated
Triple-negative

Figure S1, Related to Figure 1. Mutational burden and single base substitution signatures in UTUC.
(A) Bar plots showing depth (top) and coverage (bottom) of whole exome sequencing (WES).
(B) Box plots showing mutational burdens of cancers in renal pelvis (pink), ureter (purple), and bladder (red, TCGA cohort). *P* values calculated by two-sided Mann-Whitney *U* test with Bonferroni correction. The median, first and third quartiles, as well as outliners, are indicated with whiskers extending to the furthest value within 1.5 of the interquartile range in box plots. $^{**}P < 0.01$, $^{***}P < 0.001$.
(C) The mutational signature in UTUC98T, which is associated with SBS22.
(D) Bar plots showing the numbers of SNVs affected by Sig. A-D.
(E) Bar plots showing the coefficients of variations in each SBS signature.
(F) Box plots showing cosine similarity of four SBS signatures contributions between 'pmsignature' and 'MutationalPatterns' (orange) or 'deconstractSigs' (green). The median, first and third quartiles, as well as outliners, are indicated with whiskers extending to the furthest value within 1.5 of the interquartile range in box plots.
(G) Four predominant SBS signatures (Sig. A' -D' , left) and the strand bias for each signature (right) extracted from all UTUC samples, including hypermutated ones. TCR, transcriptional-coupled repair.
(H) Bar plots showing the frequency of each SBS signature in hypermutated samples (top) and in each mutational subtype (bottom). Hypermutated samples are further divided into two subgroups according to the status of defect in MMR genes in the bottom panel.
SNVs, single nucleotide variants; SBS, single base substitution; MMR, mismatch repair.

Table S2, related to Figure 1. Mismatch repair gene alterations in hypermutated samples.

Sample ID	SNV	Indel	Mismatch repair gene alterations	
			Germline	Somatic
UTUC185T	9038	183	*MSH6* _p.M1326fs	*MSH6* _p.P831fs
UTUC221T	3933	46	*MSH6* _p.Y214X	*MSH6* _p.T1085fs
UTUC252T	2194	573	*MSH2* _p.L173P	*MSH2* _ p.L187P
UTUC62T	2129	22	*MSH6* _p.T1085fs	*MSH6* _p.E434X
UTUC204T	1522	34	–	–
UTUC229rt	1474	35	*MSH6* _p.S602fs	*MSH6* _p.T336fs
UTUC282T	1073	328	–	*MLH1* _p.I330fs *MLH1* _p.S372fs
UTUC89T	1179	147	–	*MSH2* _p.Q76X
UTUC166T	1182	140	*MSH2* _p.G220fs	*MSH2* _c.645+1G>A
UTUC157T	1293	10	*MSH6* _c.3646+1G>C	*MSH6* _p.C687fs
UTUC175T	771	35	–	–

SNV, Single nucleotide variant.

Table S3, related to Figure 2. Driver genes identified by MutSigCV and dNdScv analysis.

MutSigCV

Gene	Type of mutations (Number)		*P*-value	*q* value
	Nonsilent	Silent		
ARID1A	31	5	2.20E-16	2.20E-16
CDKN1A	14	0	2.20E-16	2.20E-16
ELF3	27	0	2.20E-16	2.20E-16
HRAS	27	0	2.20E-16	2.20E-16
KDM6A	47	0	2.20E-16	2.20E-16
STAG2	24	3	2.20E-16	2.20E-16
TP53	73	1	2.20E-16	2.20E-16
KMT2D	132	5	1.55E-15	3.66E-12
ZFP36L1	11	1	5.65E-13	1.18E-09
CREBBP	21	1	8.85E-13	1.67E-09
TSC1	14	0	2.45E-06	4.19E-03
AMMECR1	6	0	6.36E-06	1.00E-02
FOXQ1	6	0	1.70E-05	2.32E-02
EP300	26	1	1.72E-05	2.32E-02
KMT2C	33	1	2.68E-05	3.31E-02
RHOB	9	0	2.81E-05	3.31E-02
CUL3	10	0	4.11E-05	0.045606
KRAS	7	0	6.95E-05	7.29E-02

dNdScv

Gene	Type of mutations (Number)					*P*-value	*q* value
	Synonymous	Missense	Nonsense	Splicing	Indel		
TP53	1	55	9	4	4	2.20E-16	2.20E-16
FGFR3	4	70	0	0	0	2.20E-16	2.20E-16
HRAS	0	25	0	0	3	2.20E-16	2.20E-16
KMT2D	5	22	49	5	54	2.20E-16	2.20E-16
KDM6A	0	3	20	4	19	2.20E-16	2.20E-16
STAG2	3	2	12	1	9	2.20E-16	2.20E-16
CDKN1A	0	0	0	0	13	6.66E-16	1.91E-12
PIK3CA	0	27	0	0	0	8.22E-15	2.06E-11
ARID1A	5	8	8	2	12	1.32E-14	2.95E-11
ELF3	1	5	1	1	19	1.75E-14	3.52E-11
ZFP36L1	1	1	0	0	10	1.56E-10	2.85E-07
CREBBP	2	7	8	0	7	1.97E-10	3.30E-07
KMT2C	1	8	7	1	17	7.55E-09	1.17E-05
RHOB	0	7	0	0	2	1.19E-08	1.71E-05
EP300	1	12	2	1	11	5.39E-08	7.22E-05
KMT2A	1	3	6	0	5	3.02E-07	3.79E-04
TSC1	1	6	4	0	4	6.33E-07	7.48E-04
RHOA	0	7	0	0	0	2.61E-06	2.91E-03
FBXW7	0	9	1	0	1	4.15E-06	4.39E-03
KRAS	0	7	0	0	0	7.94E-06	7.98E-03
ERCC2	1	11	1	0	0	1.69E-05	1.62E-02
SPTAN1	0	4	5	0	3	4.12E-05	3.76E-02
CUL3	0	7	0	2	1	5.37E-05	4.69E-02
DDX17	2	8	1	0	2	5.65E-05	4.73E-02

Table S4, related to Figure 2. *TERT* promoter mutation detection.

Sample ID	Chr	Start	End	Ref	Alt	Method	Variant allele frequency	
							Tumour	Normal
UTUC101T	5	1295228	1295228	G	A	PCR	0.065	0.000
UTUC112T	5	1295228	1295228	G	A	PCR	0.527	0.011
UTUC113T	5	1295228	1295228	G	A	PCR	0.344	0.000
UTUC115T	5	1295228	1295228	G	A	PCR	0.593	0.000
UTUC122T	5	1295228	1295228	G	A	PCR	0.314	0.000
UTUC123T	5	1295242	1295243	GG	AA	PCR	0.199	0.000
UTUC127T	5	1295250	1295250	G	A	PCR	0.588	0.124
UTUC12T	5	1295228	1295228	G	A	PCR	0.448	0.000
UTUC132T	5	1295228	1295228	G	A	PCR	0.475	0.000
UTUC133T	5	1295250	1295250	G	A	PCR	0.455	0.000
UTUC136T	5	1295228	1295228	G	A	PCR	0.333	0.000
UTUC13T	5	1295228	1295228	G	A	PCR	0.284	0.000
UTUC141T	5	1295228	1295228	G	A	PCR	0.170	0.000
UTUC142T	5	1295228	1295228	G	A	PCR	0.366	0.003
UTUC146T	5	1295228	1295228	G	A	PCR	0.408	0.000
UTUC148T	5	1295228	1295228	G	A	PCR	0.663	0.000
UTUC152T	5	1295228	1295228	G	A	PCR	0.656	0.000
UTUC153T	5	1295228	1295228	G	T	PCR	0.351	0.000
UTUC159T	5	1295228	1295228	G	A	PCR	0.359	0.010
UTUC160T	5	1295228	1295228	G	A	PCR	0.377	0.000
UTUC161T	5	1295228	1295228	G	A	PCR	0.422	0.022
UTUC162T	5	1295228	1295228	G	A	PCR	0.470	0.000
UTUC163T	5	1295228	1295228	G	A	PCR	0.497	0.000
UTUC171T	5	1295228	1295228	G	A	PCR	0.593	0.000
UTUC172T	5	1295228	1295228	G	A	PCR	0.117	0.000
UTUC173T	5	1295228	1295228	G	A	PCR	0.452	0.000
UTUC178T	5	1295250	1295250	G	A	PCR	0.086	0.000
UTUC179T	5	1295228	1295228	G	A	PCR	0.353	0.000
UTUC182T	5	1295250	1295250	G	A	PCR	0.500	0.011
UTUC183T	5	1295228	1295228	G	A	PCR	0.217	0.015
UTUC189T	5	1295228	1295228	G	A	PCR	0.552	0.014
UTUC18T	5	1295228	1295228	G	A	PCR	0.502	0.009
UTUC190T	5	1295228	1295228	G	A	PCR	0.211	0.000
UTUC191T	5	1295228	1295228	G	A	PCR	0.383	0.000
UTUC192T	5	1295228	1295228	G	A	PCR	0.302	0.000
UTUC205T	5	1295228	1295228	G	A	PCR	0.231	0.000
UTUC206T	5	1295228	1295228	G	A	PCR	0.546	0.000
UTUC207T	5	1295228	1295228	G	A	PCR	0.529	0.000
UTUC208T	5	1295228	1295228	G	A	PCR	0.357	0.000
UTUC211T	5	1295228	1295228	G	A	PCR	0.563	0.041
UTUC212T	5	1295228	1295228	G	A	PCR	0.542	0.000
UTUC213T	5	1295228	1295228	G	A	PCR	0.524	0.029
UTUC215T	5	1295228	1295228	G	A	PCR	0.520	0.000
UTUC217T	5	1295228	1295228	G	A	PCR	0.557	0.000
UTUC218T	5	1295228	1295228	G	A	PCR	0.729	0.000
UTUC222T	5	1295228	1295228	G	A	PCR	0.201	0.000
UTUC224T	5	1295228	1295228	G	A	PCR	0.437	0.000
UTUC227T	5	1295228	1295228	G	A	PCR	0.559	0.000

UTUC231T	5	1295228	1295228	G	A	PCR	0.651	0.000
UTUC233T	5	1295250	1295250	G	A	PCR	0.434	0.000
UTUC234T	5	1295228	1295228	G	A	PCR	0.570	0.000
UTUC236T	5	1295228	1295228	G	A	PCR	0.786	0.000
UTUC239T	5	1295228	1295228	G	A	PCR	0.530	0.000
UTUC247T	5	1295228	1295228	G	A	PCR/Sanger	0.378	0.000
UTUC248T	5	1295228	1295228	G	A	PCR	0.551	0.000
UTUC24T	5	1295228	1295228	G	A	PCR	0.523	0.000
UTUC251T	5	1295228	1295228	G	A	PCR	0.482	0.000
UTUC254T	5	1295228	1295228	G	A	PCR	0.094	0.000
UTUC255T	5	1295228	1295228	G	A	PCR	0.616	0.000
UTUC262T	5	1295228	1295228	G	A	PCR	0.495	0.000
UTUC265T	5	1295228	1295228	G	A	PCR	0.598	0.000
UTUC266T	5	1295228	1295228	G	A	PCR	0.661	0.000
UTUC268T	5	1295228	1295228	G	A	PCR	0.572	0.000
UTUC269T	5	1295228	1295228	G	A	PCR	0.547	0.000
UTUC270T	5	1295228	1295228	G	A	PCR	0.559	0.000
UTUC277T	5	1295228	1295228	G	A	PCR/Sanger	0.694	0.000
UTUC27T	5	1295228	1295228	G	A	PCR	0.514	0.000
UTUC280T	5	1295228	1295228	G	A	PCR/Sanger	0.599	0.000
UTUC282T	5	1295228	1295228	G	A	PCR/Sanger	0.573	0.000
UTUC284T	5	1295228	1295228	G	A	PCR/Sanger	0.585	0.000
UTUC285T	5	1295228	1295228	G	A	PCR/Sanger	0.657	0.000
UTUC286T	5	1295228	1295228	G	A	Sanger	–	–
UTUC287T	5	1295228	1295228	G	A	Sanger	–	–
UTUC288T	5	1295228	1295228	G	A	PCR/Sanger	0.594	0.000
UTUC289T	5	1295228	1295228	G	A	PCR/Sanger	0.600	0.000
UTUC291T	5	1295228	1295228	G	A	PCR/Sanger	0.616	0.000
UTUC29T	5	1295228	1295228	G	A	PCR	0.474	0.023
UTUC30T	5	1295228	1295228	G	A	PCR	0.331	0.000
UTUC31T	5	1295228	1295228	G	A	PCR	0.433	0.002
UTUC42T	5	1295250	1295250	G	A	PCR	0.411	0.000
UTUC43T	5	1295228	1295228	G	A	PCR	0.595	0.010
UTUC48T	5	1295228	1295228	G	A	PCR	0.193	0.000
UTUC4T	5	1295228	1295228	G	A	PCR	0.612	0.367
UTUC51T	5	1295228	1295228	G	A	PCR	0.485	0.027
UTUC55T	5	1295228	1295228	G	A	PCR	0.408	0.000
UTUC56T	5	1295242	1295243	GG	AA	PCR	0.415	0.039
UTUC59T	5	1295250	1295250	G	A	PCR	0.162	0.000
UTUC5T	5	1295250	1295250	G	A	PCR	0.535	0.000
UTUC60T	5	1295250	1295250	G	A	PCR	0.199	0.000
UTUC61T	5	1295228	1295228	G	A	PCR	0.563	0.000
UTUC66T	5	1295228	1295228	G	A	PCR	0.225	0.000
UTUC68T	5	1295228	1295228	G	A	PCR	0.427	0.000
UTUC71T	5	1295228	1295228	G	A	PCR	0.173	0.000
UTUC73T	5	1295250	1295250	G	A	PCR	0.524	0.000
UTUC78T	5	1295250	1295250	G	A	PCR	0.599	0.000
UTUC7T	5	1295250	1295250	G	A	PCR	0.408	0.000
UTUC80T	5	1295228	1295228	G	A	PCR	0.548	0.000
UTUC83T	5	1295250	1295250	G	A	PCR	0.307	0.000
UTUC8T	5	1295228	1295228	G	A	PCR	0.569	0.000
UTUC90T	5	1295228	1295228	G	A	PCR	0.364	0.012

UTUC92T	5	1295228	1295228	G	A	PCR	0.208	0.000
UTUC95T	5	1295228	1295228	G	A	PCR	0.131	0.000

Chr, chromosome; PCR, PCR–based deep sequencing; Sanger, Sanger sequencing.

Figure S2

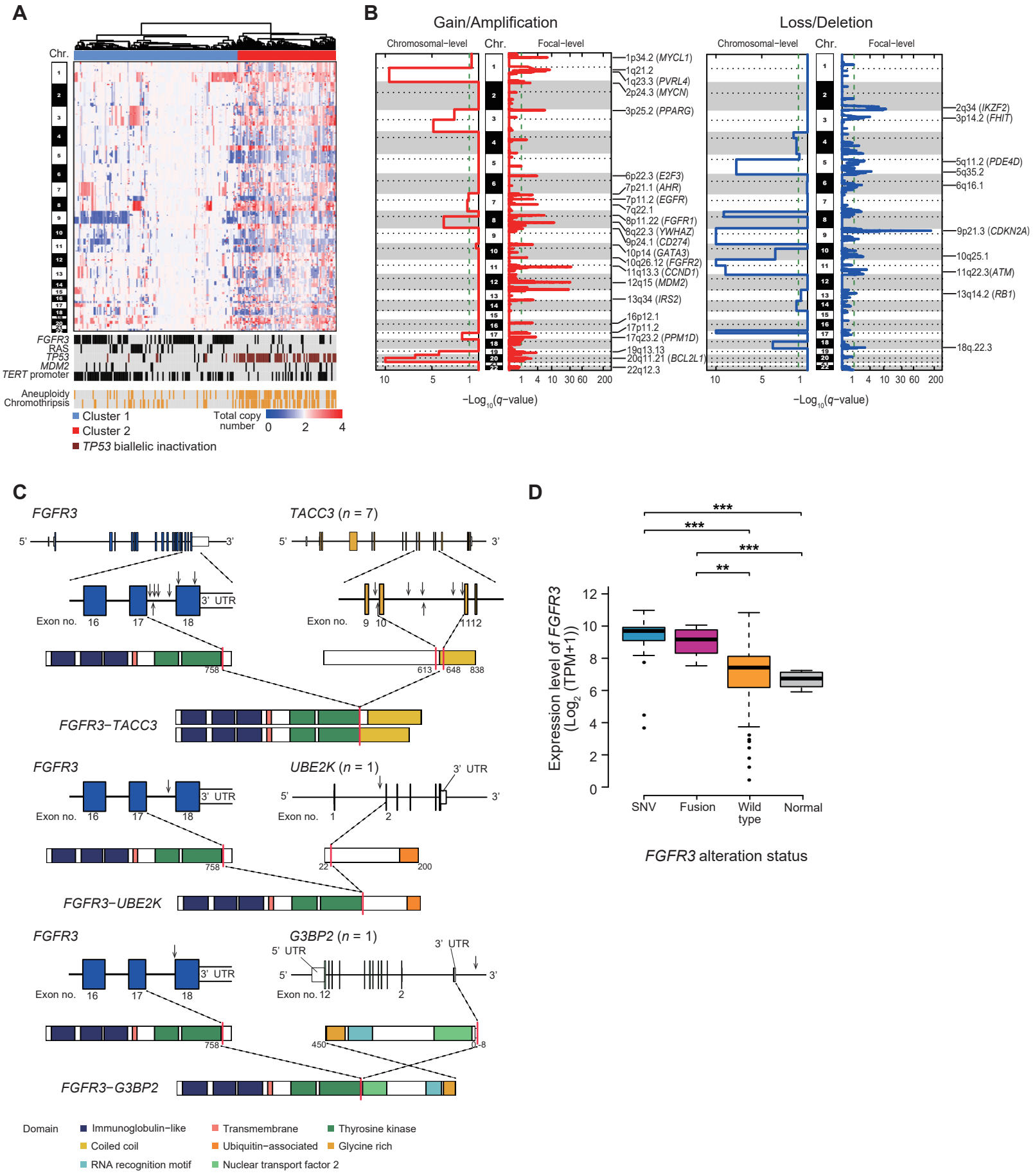

A
Chr.
FGFR3
RAS
TP53
MDM2
TERT promoter
Aneuploidy
Chromothripsis
Cluster 1
Cluster 2
TP53 biallelic inactivation
Total copy number
0 2 4
B
Gain/Amplification
Loss/Deletion
Chromosomal−level
Focal−level
1p34.2 (MYCL1)
1q21.2
1q23.3 (PVRL4)
2p24.3 (MYCN)
3p25.2 (PPARG)
6p22.3 (E2F3)
7p21.1 (AHR)
7p11.2 (EGFR)
7q22.1
8p11.22 (FGFR1)
8q22.3 (YWHAZ)
9p24.1 (CD274)
10p14 (GATA3)
10q26.12 (FGFR2)
11q13.3 (CCND1)
12q15 (MDM2)
13q34 (IRS2)
16p12.1
17p11.2
17q23.2 (PPM1D)
19q13.13
20q11.21 (BCL2L1)
22q12.3
2q34 (IKZF2)
3p14.2 (FHIT)
5q11.2 (PDE4D)
5q35.2
6q16.1
9p21.3 (CDKN2A)
10q25.1
11q22.3(ATM)
13q14.2 (RB1)
18q.22.3
−Log10(q−value)
C
FGFR3
TACC3 (n = 7)
5' 3'
3' UTR
Exon no. 16 17 18
Exon no. 9 10 11 12
758
613 648 838
FGFR3−TACC3
UBE2K (n = 1)
Exon no. 1 2
22 200
FGFR3−UBE2K
G3BP2 (n = 1)
5' UTR
Exon no. 12 2
450 0-8
FGFR3−G3BP2
Domain
Immunoglobulin−like
Transmembrane
Thyrosine kinase
Coiled coil
Ubiquitin−associated
Glycine rich
RNA recognition motif
Nuclear transport factor 2
D
Expression level of FGFR3 (Log2 (TPM+1))
***
**
SNV
Fusion
Wild type
Normal
FGFR3 alteration status

Figure S2, related to Figure 1. Copy number alterations and fusion genes in UTUC.

(A) Color-gradient maps of copy number alterations determined by SNP array karyotyping for 199 UTUC samples.

(B) Significant chromosomal- (left) and focal-level (right) copy number gain/amplification (red) and loss/deletion (blue) lesions detected by GISTIC 2.0. Putative target genes within the corresponding loci are indicated.

(C) Breakpoints in fusion genes of *FGFR3* and its partner genes and the resulting schematics of proteins. Arrows on the introns and exons of each gene represent breakpoints.

(D) Gene expression levels (RNA sequencing) of *FGFR3* in tumor samples with various status of alterations and in normal urothelial epithelia samples. *P* values calculated by two-sided Mann Whitney *U* test and Bonferroni corrected. The median, first and third quartiles, as well as outliers, are indicated with whiskers extending to the furthest value within 1.5 of the interquartile range. SNV, single nucleotide variant. $^{**}P < 0.01$, $^{***}P < 0.001$.

Figure S3

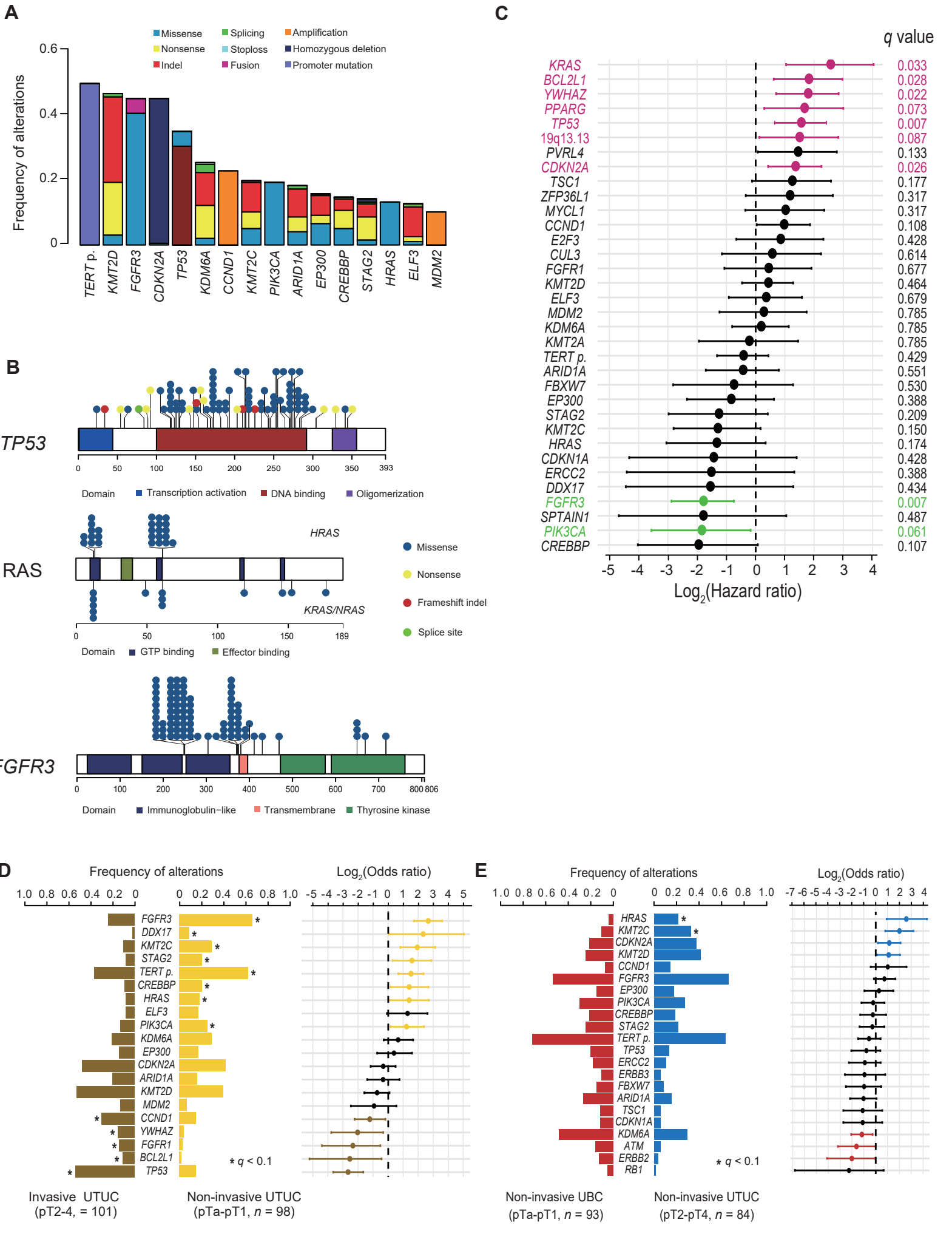

A
Missense
Nonsense
Indel
Splicing
Stoploss
Fusion
Amplification
Homozygous deletion
Promoter mutation
Frequency of alterations
TERT p.
KMT2D
FGFR3
CDKN2A
TP53
KDM6A
CCND1
KMT2C
PIK3CA
ARID1A
EP300
CREBBP
STAG2
HRAS
ELF3
MDM2
B
TP53
Domain
Transcription activation
DNA binding
Oligomerization
RAS
HRAS
KRAS/NRAS
Domain
GTP binding
Effector binding
Missense
Nonsense
Frameshift indel
Splice site
FGFR3
Domain
Immunoglobulin−like
Transmembrane
Thyrosine kinase
C
q value
KRAS 0.033
BCL2L1 0.028
YWHAZ 0.022
PPARG 0.073
TP53 0.007
19q13.13 0.087
PVRL4 0.133
CDKN2A 0.026
TSC1 0.177
ZFP36L1 0.317
MYCL1 0.317
CCND1 0.108
E2F3 0.428
CUL3 0.614
FGFR1 0.677
KMT2D 0.464
ELF3 0.679
MDM2 0.785
KDM6A 0.785
KMT2A 0.785
TERT p. 0.429
ARID1A 0.551
FBXW7 0.530
EP300 0.388
STAG2 0.209
KMT2C 0.150
HRAS 0.174
CDKN1A 0.428
ERCC2 0.388
DDX17 0.434
FGFR3 0.007
SPTAIN1 0.487
PIK3CA 0.061
CREBBP 0.107
Log2(Hazard ratio)
D
Frequency of alterations
Log2(Odds ratio)
FGFR3
DDX17
KMT2C
STAG2
TERT p.
CREBBP
HRAS
ELF3
PIK3CA
KDM6A
EP300
CDKN2A
ARID1A
KMT2D
MDM2
CCND1
YWHAZ
FGFR1
BCL2L1
TP53
* q < 0.1
Invasive UTUC (pT2-4, = 101)
Non-invasive UTUC (pTa-pT1, n = 98)
E
Frequency of alterations
Log2(Odds ratio)
HRAS
KMT2C
CDKN2A
KMT2D
CCND1
FGFR3
EP300
PIK3CA
CREBBP
STAG2
TERT p.
TP53
ERCC2
ERBB3
FBXW7
ARID1A
TSC1
CDKN1A
KDM6A
ATM
ERBB2
RB1
* q < 0.1
Non-invasive UBC (pTa-pT1, n = 93)
Non-invasive UTUC (pT2-pT4, n = 84)

Figure S3, related to Figure 1. Frequently altered driver genes in UTUC.

(A) Bar plots showing frequently altered genes (in > 10% of the entire cohort) in this study. Non-pathogenic mutations are also included.

(B) Mutation distributions for *TP53* (top), *HRAS/KRAS/NRAS* (middle), and *FGFR3* (bottom).

(C) Forest plots showing hazard ratios and corresponding *q* values (Exact Log-rank Test with Benjamini-Hochberg adjustment) by univariate analysis of driver gene alterations on disease−specific survival. Positively and negatively significant genes ($q < 0.1$) are shown in purple and green, respectively.

(D and E) Comparisons of frequency of common genetic lesions (left) between non-invasive ($n = 98$) vs. invasive UTUC ($n = 101$) **(D)** and non-invasive UBC ($n = 93$) vs. UTUC ($n = 84$) **(E)**. Log_2-transformed odds ratios are also shown with 95% confidence intervals (right). Significant enrichment is shown by asterisks ($q < 0.1$, two-sided Fisher' s exact test with Benjamini-Hochberg adjustment for multiple testing) in **(D)** and **(E)**.

TERT p., *TERT* promoter.

Figure S4

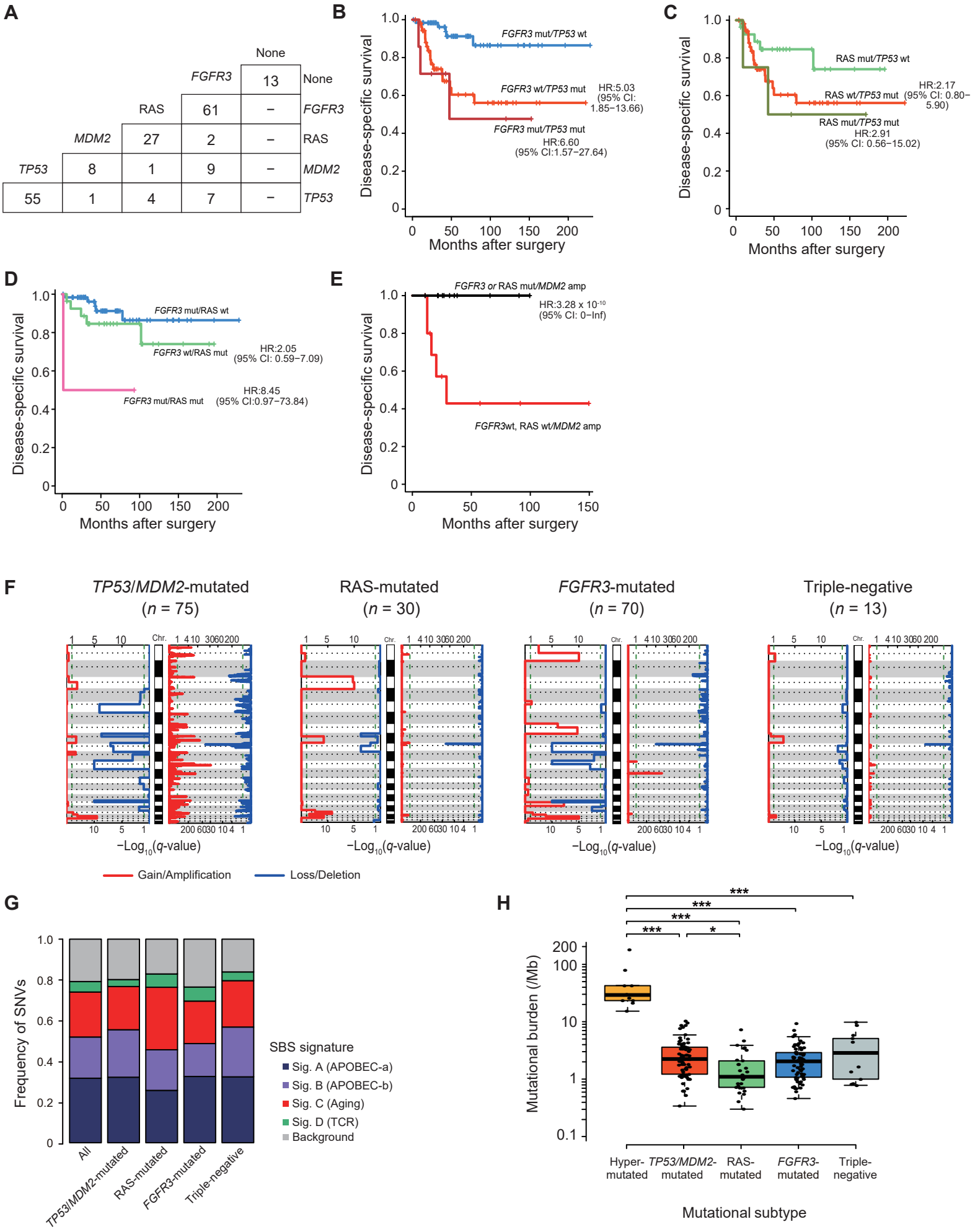

A
None
FGFR3
13
None
RAS
61
–
FGFR3
MDM2
27
2
–
RAS
TP53
8
1
9
–
MDM2
55
1
4
7
–
TP53
B
Disease-specific survival
FGFR3 mut/TP53 wt
FGFR3 wt/TP53 mut
HR:5.03
(95% CI:
1.85–13.66)
FGFR3 mut/TP53 mut
HR:6.60
(95% CI:1.57–27.64)
Months after surgery
C
Disease-specific survival
RAS mut/TP53 wt
RAS wt/TP53 mut
HR:2.17
(95% CI: 0.80–
5.90)
RAS mut/TP53 mut
HR:2.91
(95% CI: 0.56–15.02)
Months after surgery
D
Disease-specific survival
FGFR3 mut/RAS wt
FGFR3 wt/RAS mut
HR:2.05
(95% CI: 0.59–7.09)
HR:8.45
(95% CI:0.97–73.84)
FGFR3 mut/RAS mut
Months after surgery
E
Disease-specific survival
FGFR3 or RAS mut/MDM2 amp
HR:3.28 x 10-10
(95% CI: 0–Inf)
FGFR3wt, RAS wt/MDM2 amp
Months after surgery
F
TP53/MDM2-mutated
(n = 75)
RAS-mutated
(n = 30)
FGFR3-mutated
(n = 70)
Triple-negative
(n = 13)
Chr.
−Log10(q-value)
Gain/Amplification
Loss/Deletion
G
Frequency of SNVs
SBS signature
Sig. A (APOBEC-a)
Sig. B (APOBEC-b)
Sig. C (Aging)
Sig. D (TCR)
Background
All
TP53/MDM2-mutated
RAS-mutated
FGFR3-mutated
Triple-negative
H
Mutational burden (/Mb)
Hyper-mutated
TP53/MDM2-mutated
RAS-mutated
FGFR3-mutated
Triple-negative
Mutational subtype

Figure S4, related to Figures 2 and 3. Clinical effects of combinations of subtype-defining genetic alterations.

(A) Genetic status of subtype-defining genes; *TP53*, *MDM2*, RAS, and *FGFR3*. Numbers of samples with alterations in these genes are shown in cross section.

(B−D) Disease-specific survival for samples harboring at least one mutation in *TP53* and *FGFR3* **(B)**, *TP53* and RAS **(C)**, and RAS and *FGFR3* **(D)**.

(E) Disease-specific survival for samples harboring *MDM2* amplification with and without *FGFR3*/RAS mutation.

(F) Significant arm- and focal-level CNAs in the *TP53/MDM2-,* RAS-*, FGFR3*-mutated, and triple negative subtypes.

(G) Bar plots showing the frequency of each single base substitution (SBS) signature in each mutational subtype. SNV, single nucleotide variant.

(H) Log-scaled mutational burden of each mutational subtype of UTUC. *P* values calculated by two-sided Mann-Whitney *U* test with Bonferroni correction. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Hazard ratios (HRs) and 95% confidence intervals (CIs) calculated by Cox proportional-hazards model in **(B)−(E)**.

Figure S5

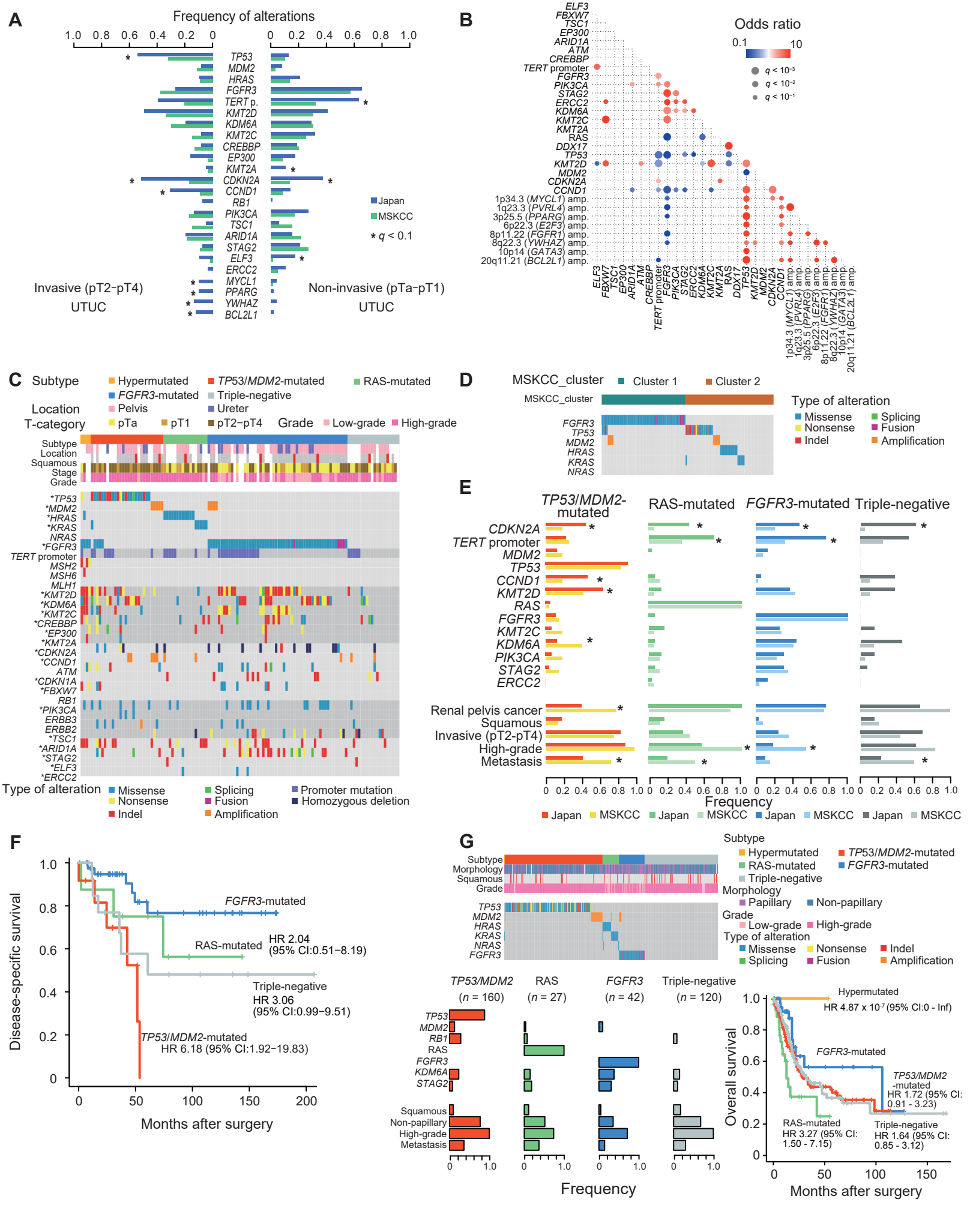

A
Frequency of alterations
TP53
MDM2
HRAS
FGFR3
TERT p.
KMT2D
KDM6A
KMT2C
CREBBP
EP300
KMT2A
CDKN2A
CCND1
RB1
PIK3CA
TSC1
ARID1A
STAG2
ELF3
ERCC2
MYCL1
PPARG
YWHAZ
BCL2L1
Japan
MSKCC
* q < 0.1
Invasive (pT2–pT4) UTUC
Non-invasive (pTa–pT1) UTUC
B
Odds ratio
0.1
10
q < 10⁻³
q < 10⁻²
q < 10⁻¹
ELF3
FBXW7
TSC1
EP300
ARID1A
ATM
CREBBP
TERT promoter
FGFR3
PIK3CA
STAG2
ERCC2
KDM6A
KMT2C
KMT2A
RAS
DDX17
TP53
KMT2D
MDM2
CDKN2A
CCND1
1p34.3 (MYCL1) amp.
1q23.3 (PVRL4) amp.
3p25.5 (PPARG) amp.
6p22.3 (E2F3) amp.
8p11.22 (FGFR1) amp.
8q22.3 (YWHAZ) amp.
10p14 (GATA3) amp.
20q11.21 (BCL2L1) amp.
C
Subtype
Hypermutated
TP53/MDM2-mutated
RAS-mutated
FGFR3-mutated
Triple-negative
Location
Pelvis
Ureter
T-category
pTa
pT1
pT2–pT4
Grade
Low-grade
High-grade
Subtype
Location
Squamous
Stage
Grade
*TP53
*MDM2
*HRAS
*KRAS
NRAS
*FGFR3
TERT promoter
MSH2
MSH6
MLH1
*KMT2D
*KDM6A
*KMT2C
*CREBBP
*EP300
*KMT2A
*CDKN2A
*CCND1
ATM
*CDKN1A
*FBXW7
RB1
*PIK3CA
ERBB3
ERBB2
*TSC1
*ARID1A
*STAG2
*ELF3
*ERCC2
Type of alteration
Missense
Splicing
Promoter mutation
Nonsense
Fusion
Homozygous deletion
Indel
Amplification
D
MSKCC_cluster
Cluster 1
Cluster 2
MSKCC_cluster
FGFR3
TP53
MDM2
HRAS
KRAS
NRAS
Type of alteration
Missense
Splicing
Nonsense
Fusion
Indel
Amplification
E
TP53/MDM2-mutated
RAS-mutated
FGFR3-mutated
Triple-negative
CDKN2A
TERT promoter
MDM2
TP53
CCND1
KMT2D
RAS
FGFR3
KMT2C
KDM6A
PIK3CA
STAG2
ERCC2
Renal pelvis cancer
Squamous
Invasive (pT2-pT4)
High-grade
Metastasis
0 0.2 0.4 0.6 0.8 1.0
Frequency
Japan MSKCC Japan MSKCC Japan MSKCC Japan MSKCC
F
Disease-specific survival
1.0
0.8
0.6
0.4
0.2
0
FGFR3-mutated
RAS-mutated
HR 2.04
(95% CI:0.51–8.19)
Triple-negative
HR 3.06
(95% CI:0.99–9.51)
TP53/MDM2-mutated
HR 6.18 (95% CI:1.92–19.83)
0
50
100
150
200
Months after surgery
G
Subtype
Morphology
Squamous
Grade
TP53
MDM2
HRAS
KRAS
NRAS
FGFR3
Subtype
Hypermutated
TP53/MDM2-mutated
RAS-mutated
FGFR3-mutated
Triple-negative
Morphology
Papillary
Non-papillary
Grade
Low-grade
High-grade
Type of alteration
Missense
Nonsense
Indel
Splicing
Fusion
Amplification
TP53/MDM2 (n = 160)
RAS (n = 27)
FGFR3 (n = 42)
Triple-negative (n = 120)
TP53
MDM2
RB1
RAS
FGFR3
KDM6A
STAG2
Squamous
Non-papillary
High-grade
Metastasis
0 1.0
Frequency
Overall survival
1.0
0.8
0.6
0.4
0.2
0
Hypermutated
HR 4.87 x 10⁻⁷ (95% CI:0 - Inf)
FGFR3-mutated
TP53/MDM2 -mutated
HR 1.72 (95% CI: 0.91 - 3.23)
RAS-mutated
HR 3.27 (95% CI: 1.50 - 7.15)
Triple-negative
HR 1.64 (95% CI: 0.85 - 3.12)
0
50
100
150
Months after surgery

Figure S5, related to Figure 3. Validation study in the MSKCC UTUC and TCGA UBC cohort.
(A) Comparisons of frequency of common genetic lesions in invasive (left) and non-invasive (right) UTUC between our (blue) and MSKCC (green) cohort. Significant enrichment is shown by asterisks (q < 0.1, two-sided Fisher' s exact test with Benjamini-Hochberg adjustment for multiple testing). *TERT* p., *TERT* promoter.
(B) Statistically significant (q < 0.1, pairwise Fisher' s exact test with Benjamini-Hochberg adjustment) positive (red) and negative (blue) correlations of major genetic events in the combined data of our and MSKCC UTUC cohort.
(C) Landscape of somatic alterations in MSKCC UTUC cohort, which are grouped into five discrete subtypes when they are analyzed with the Japanese cohort. Significantly altered and/or positively selected driver genes in the Japanese cohort are indicated by asterisks.
(D) Clustering analysis of the non-hypermutated MSKCC UTUC cases alone.
(E) Frequencies of common genetic lesions and clinical/histological features across four non-hypermutated subtypes in our and MSKCC UTUC cohort. Significant enrichment is shown by asterisks (P < 0.05, two-sided Fisher' s exact test).
(F) Disease-specific survival of four subtypes in MSKCC UTUC cohort.
(G) Landscape of somatic alterations (top), frequencies of common genetic lesions and clinical/histological features across four non-hypermutated subtypes (bottom left), and overall survival of five subtypes (bottom right) in the TCGA UBC cohort.
Hazard ratios (HRs) and 95% confidence intervals (CIs) calculated by Cox proportional-hazards model in **(F)** and **(G)**.
MSKCC, Memorial Sloan Kettering Cancer Center.

Figure S6

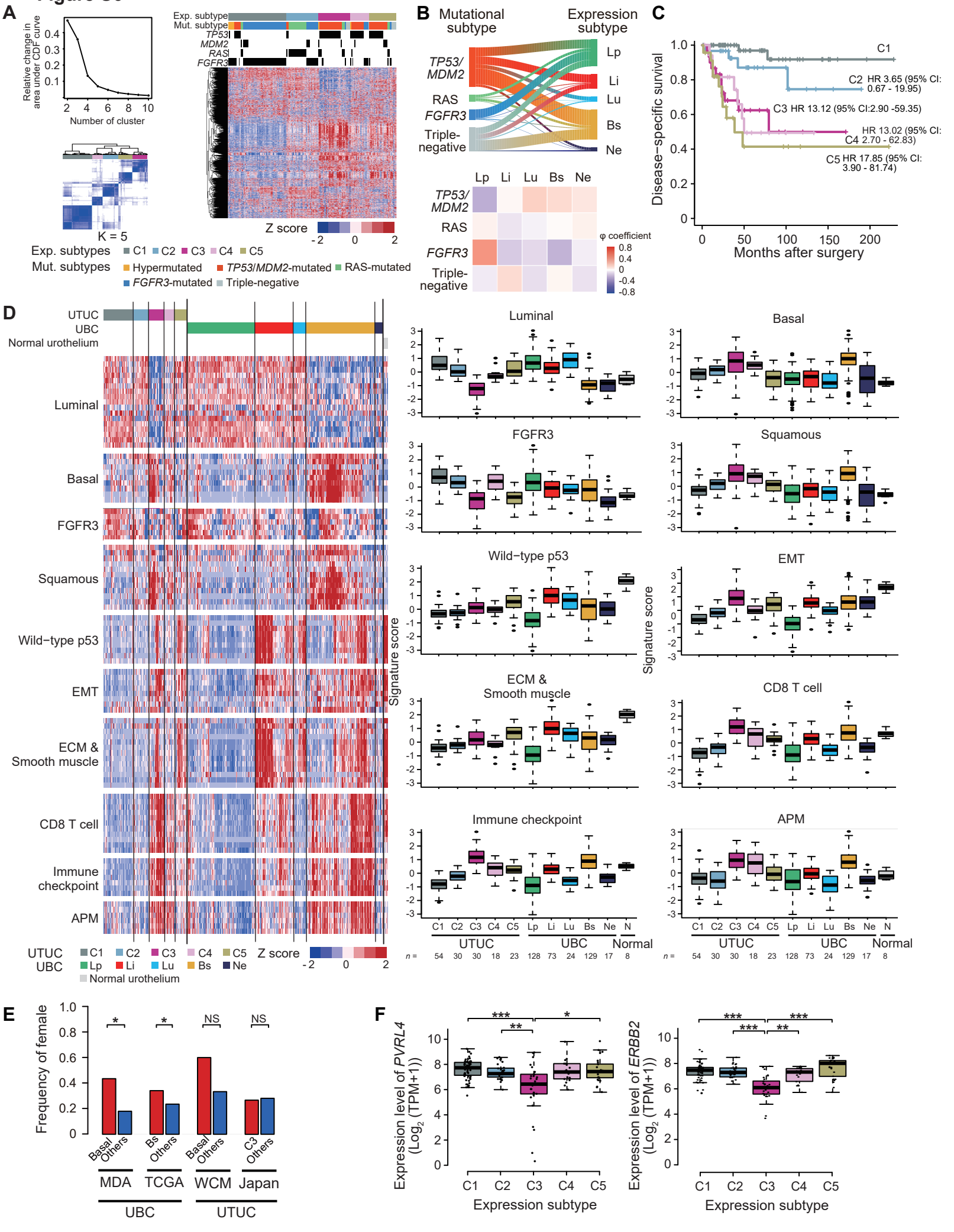

A
Relative change in area under CDF curve
Number of cluster
K = 5
Exp. subtype
Mut. subtype
TP53
MDM2
RAS
FGFR3
Z score
Exp. subtypes C1 C2 C3 C4 C5
Mut. subtypes Hypermutated TP53/MDM2-mutated RAS-mutated
FGFR3-mutated Triple-negative
B
Mutational subtype
Expression subtype
TP53/ MDM2
RAS
FGFR3
Triple-negative
Lp Li Lu Bs Ne
φ coefficient
C
Disease-specific survival
Months after surgery
C1
C2 HR 3.65 (95% CI: 0.67 - 19.95)
C3 HR 13.12 (95% CI:2.90 -59.35)
C4 HR 13.02 (95% CI: 2.70 - 62.83)
C5 HR 17.85 (95% CI: 3.90 - 81.74)
D
UTUC
UBC
Normal urothelium
Luminal
Basal
FGFR3
Squamous
Wild−type p53
EMT
ECM & Smooth muscle
CD8 T cell
Immune checkpoint
APM
Signature score
UTUC C1 C2 C3 C4 C5
UBC Lp Li Lu Bs Ne
Normal urothelium
Z score
n = 54 30 30 18 23 128 73 24 129 17 8
E
Frequency of female
Basal Others
Bs Others
Basal Others
C3 Others
MDA TCGA WCM Japan
UBC UTUC
F
Expression level of PVRL4 (Log2 (TPM+1))
Expression level of ERBB2 (Log2 (TPM+1))
Expression subtype

Figure S6, related to Figures 4 and 5. Expression subtypes in UTUC.

(A) Consensus clustering of gene expression data of in 158 UTUC samples using 2,000 most variably expressed genes. Delta area plot and consensus matrix of K = 5 (left) and integrated view of expression clustering combined with mutational subtypes, *TP53*, *MDM2*, RAS, and *FGFR3* mutations (right) are shown.

(B) Correlations between mutational and expression subtypes (left) and heatmap of Phi coefficients (right) in the TCGA UBC cohort.

(C) Disease-specific survival for different expression subtypes. Hazard ratios (HRs) and 95% confidence intervals (CIs) calculated by Cox proportional-hazards model.

(D) Heatmap showing expression level of biologically important signatures in UTUC expression subgroup, UBC subgroup (TCGA), and normal urothelial epithelia samples (left) and box plots of the normalized expression scores for each signature (right). Genes in each signature are in the same order as **Figure 4C**. The median, first and third quartiles, as well as outliers, are indicated with whiskers extending to the furthest value within 1.5 of the interquartile range in box plots. EMT, epithelial mesenchymal transition; ECM, extracellular matrix; APM, antigen presentation machinery.

(E) Bar plots showing the frequency of female in basal-like and other subtypes in the previous UBC and UTUC studies. MDA, MD Anderson Cancer Center; WCM, Weill Cornell Medicine. *P* values calculated by two-sided Fisher' s exact test. $*P < 0.05$; NS, not significant.

(F) Gene expression levels (RNA sequencing) of *PVRL4* and *ERBB2* in each expression subtype. *P* values calculated by two-sided Mann Whitney *U* test and Bonferroni corrected. The median, first and third quartiles, as well as outliers, are indicated with whiskers extending to the furthest value within 1.5 of the interquartile range. $*P < 0.05$,$**P < 0.01$, $***P < 0.001$.

Lp, luminal-papillary; Li, luminal-infiltrated; Lu, luminal; Bs, basal-squamous; Ne, neuronal.

Table S5, related to Figure 4. Genes involved in each expression signature.

Luminal	Basal	FGFR3	Squamous	Wild-type p53	EMT	ECM & smooth muscle	CD 8 T cell	Immune checkpoint	APM	F-TBRS
FGFR3	*KRT1*	*FGFR3*	*DSC3*	*ACTG2*	*ZEB1*	*PGM5*	*CD8A*	*CD274*	*HLA-A*	*ACTA2*
FOXA1	*CD44*	*TP63*	*GSDMC*	*CNN1*	*ZEB2*	*DES*	*GZMA*	*PDCD1LG2*	*B2M*	*COL4A1*
GPX2	*CDH3*	*IRS1*	*TGM1*	*MYH11*	*SNAI1*	*C7*	*GZMB*	*CTLA4*	*HLA-C*	*TAGLN*
CYP2J2	*KRT14*	*SEMA4B*	*PI3*	*MFAP4*	*TWIST1*	*SFRP4*	*GZMH*	*PDCD1*	*HLA-B*	*SH3PXD2A*
PPARG	*KRT16*	*PTPN13*	*TP63*	*PGM5*	*CDH2*	*COMP*	*GZMK*	*LAG3*	*TAP1*	
KRT7	*KRT5*	*TMPRSS4*	*DSC1*	*FLNC*	*VIM*	*SGCD*	*GZMM*	*HAVCR2*	*TAP2*	
FABP4	*KRT6A*		*DSC2*	*ACTC1*	*FOXC2*	*CNN1*	*PRF1*	*TIGIT*		
KRT18	*KRT6B*		*DSG1*	*DES*	*SNAI2*	*SYNM*	*CXCL9*			
KRT19	*KRT6C*		*DSG2*	*PCP4*		*ACTG2*	*CXCL10*			
KRT20			*DSG3*			*ACTC1*	*TBX21*			
KRT8			*S100A7*			*ACTA2*	*IFNG*			
XBP1			*S100A8*			*MYL9*				
UPK1A						*MYLH11*				
UPK2						*TAGLN*				
ERBB2										
ERBB3										
GATA3										

EMT, epithelial-mesenchymal transition; ECM, extracellular matrix; APM, antigen presentation machinery; F-TBRS, pan-fibroblast TGFβ response signature.

Figure S7

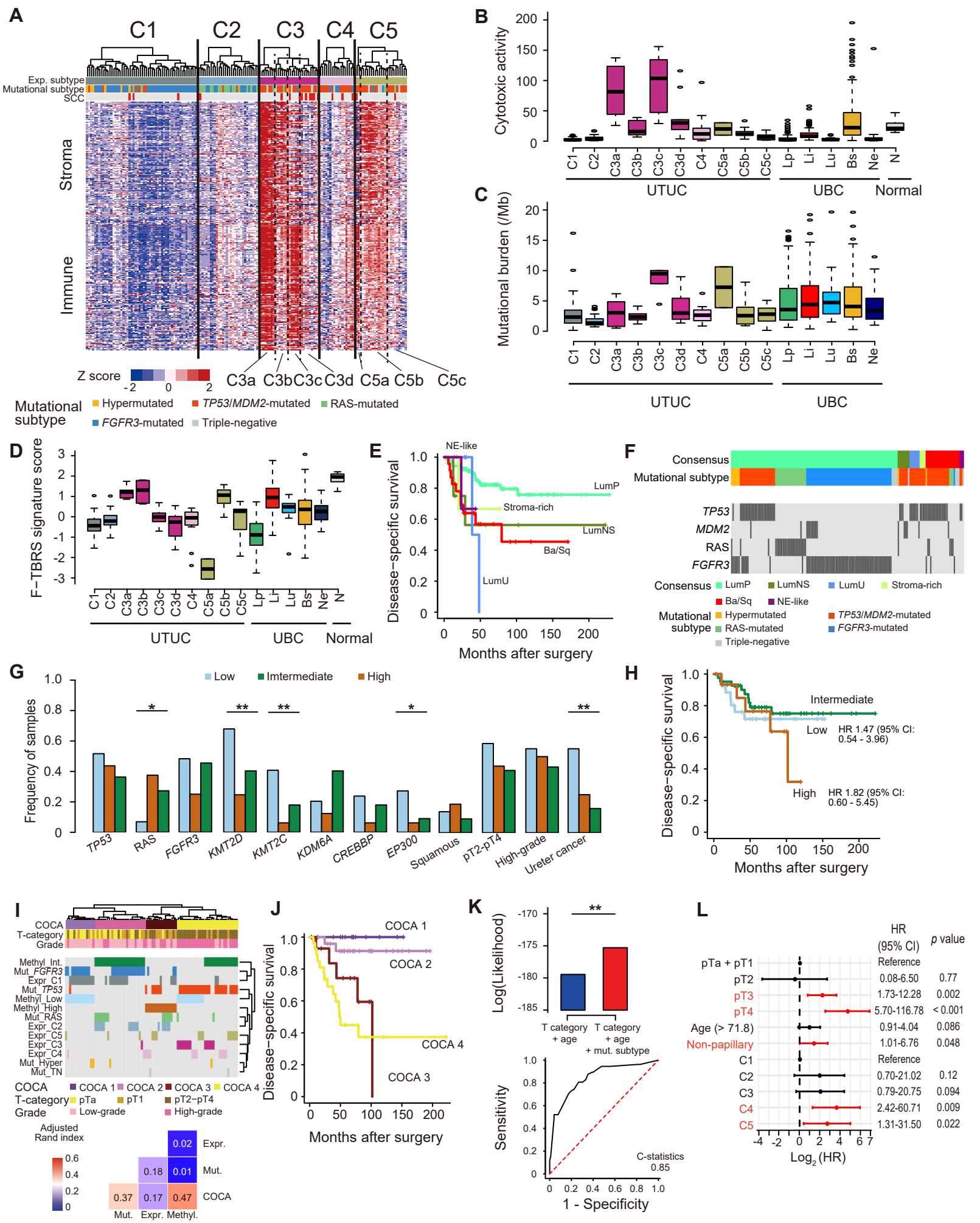

A
C1
C2
C3
C4
C5
Exp. subtype
Mutational subtype
SCC
Stroma
Immune
Z score
-2
0
2
C3a
C3b
C3c
C3d
C5a
C5b
C5c
Mutational subtype
Hypermutated
TP53/MDM2-mutated
RAS-mutated
FGFR3-mutated
Triple-negative
B
Cytotoxic activity
200
150
100
50
0
C1
C2
C3a
C3b
C3c
C3d
C4
C5a
C5b
C5c
Lp
Li
Lu
Bs
Ne
N
UTUC
UBC
Normal
C
Mutational burden (/Mb)
20
15
10
5
0
UTUC
UBC
D
F−TBRS signature score
3
2
1
0
-1
-2
-3
UTUC
UBC
Normal
E
NE-like
LumP
Stroma-rich
LumNS
Ba/Sq
LumU
Disease−specific survival
1.0
0.8
0.6
0.4
0.2
0
0
50
100
150
200
Months after surgery
F
Consensus
Mutational subtype
TP53
MDM2
RAS
FGFR3
Consensus
LumP
LumNS
LumU
Stroma-rich
Ba/Sq
NE-like
Mutational subtype
Hypermutated
TP53/MDM2-mutated
RAS-mutated
FGFR3-mutated
Triple-negative
G
Low
Imtermediate
High
Frequency of samples
1.0
0.8
0.6
0.4
0.2
0
*
**
**
*
**
TP53
RAS
FGFR3
KMT2D
KMT2C
KDM6A
CREBBP
EP300
Squamous
pT2-pT4
High-grade
Ureter cancer
H
Disease−specific survival
1.0
0.8
0.6
0.4
0.2
0
Intermediate
Low
HR 1.47 (95% CI: 0.54 - 3.96)
High
HR 1.82 (95% CI: 0.60 - 5.45)
0
50
100
150
200
Months after surgery
I
COCA
T-category
Grade
Methyl_Int.
Mut_FGFR3
Expr_C1
Mut_TP53
Methyl_Low
Methyl_High
Mut_RAS
Expr_C2
Expr_C5
Expr_C3
Expr_C4
Mut_Hyper
Mut_TN
COCA
COCA 1
COCA 2
COCA 3
COCA 4
T-category
pTa
pT1
pT2−pT4
Grade
Low-grade
High-grade
Adjusted Rand index
0.6
0.4
0.2
0
0.02
Expr.
0.18
0.01
Mut.
0.37
0.17
0.47
COCA
Mut.
Expr.
Methyl.
J
COCA 1
COCA 2
COCA 4
COCA 3
Disease−specific survival
1.0
0.8
0.6
0.4
0.2
0
0
50
100
150
200
Months after surgery
K
**
Log(Likelihood)
-170
-175
-180
-185
T category + age
T category + age + mut. subtype
Sensitivity
1.0
0.8
0.6
0.4
0.2
0
C-statistics
0.85
0
0.2
0.4
0.6
0.8
1.0
1 - Specificity
L
HR (95% CI)
p value
pTa + pT1
Reference
pT2
0.08-6.50
0.77
pT3
1.73-12.28
0.002
pT4
5.70-116.78
< 0.001
Age (> 71.8)
0.91-4.04
0.086
Non-papillary
1.01-6.76
0.048
C1
Reference
C2
0.70-21.02
0.12
C3
0.79-20.75
0.094
C4
2.42-60.71
0.009
C5
1.31-31.50
0.022
-4
-2
0
2
4
6
7
Log2 (HR)

Figure S7, related to Figures 4 and 5. Tumor microenvironment and methylation profile in UTUC.

(A) Heatmap showing immune and stromal expression profiles in five UTUC expression subgroups. Samples with C3 and C5 are further clustered into four and three small subclusters (C3a-d and C5a-c), respectively.

(B-D) Box plots showing the cytotoxic activity **(B)**, mutational burden **(C)**, and F-TBRS score **(D)** for each UTUC subgroup, UBC subgroup, and normal urothelial epithelia samples. The median, first and third quartiles, as well as outliers, are indicated with whiskers extending to the furthest value within 1.5 of the interquartile range. Lp, luminal-papillary; Li, luminal-infiltrated; Lu, luminal; Bs, basal-squamous; Ne, neuronal; UBC, urothelial bladder carcinoma; F-TBRS, pan fibroblast TGFβ response signature.

(E) Disease-specific survival for patients in each consensus subtype.

(F) Mutational subtypes and alteration of subtype-defying genes in each consensus subtype. LumP, Luminal Papillary; LumNS, Luminal Non-specified; LumU, Luminal Unstable; Ba/Sq, basal/Squamous; NE-like, neuroendocrine-like in **(E)** and **(F)**.

(G) Bar plots showing the frequencies of mutations in *TP53*, RAS, *FGFR3*, and chromatin modification genes as well as histopathological features in each methylation subtype. The frequencies were compared between CIMP and non-CIMP subtypes by two-sided Fisher' s exact test. **P* < 0.05, ***P* < 0.01.

(H) Disease-specific survival for patients in each methylation subtype. Hazard ratios (HRs) and 95% confidence intervals (CIs) calculated by Cox proportional-hazards model.

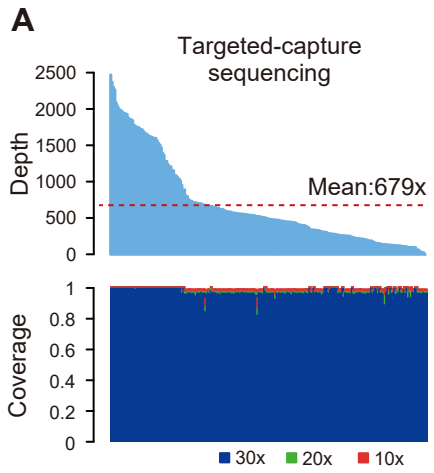
(I) Four clusters identified by cluster of cluster assignments analysis (top) and heatmap of adjusted Rand index between COCA, mutational, expression, and methylation subtypes (bottom). Mut., mutational subtypes; Expr., expression subtypes; Methyl., methylation subtypes.

(J) Disease-specific survival for patients in each COCA cluster.

(K) Bar plots showing the log likelihood values of two survival models (top) and ROC analysis for the prediction of disease-specific survival in the MSKCC UTUC cohort (bottom). ***P* < 0.01. Mut. subtype, mutational subtype.

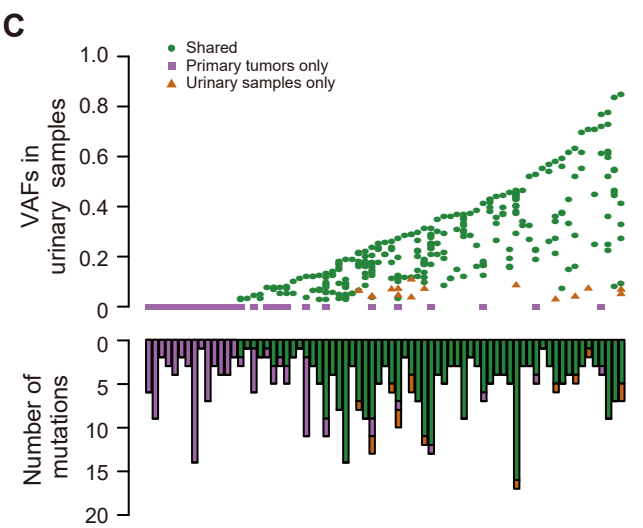
(L) HRs with 95% CIs for covariates significantly included in Cox proportional-hazards modeling of disease-specific survival. Factors that were significantly associated with disease-specific survival are

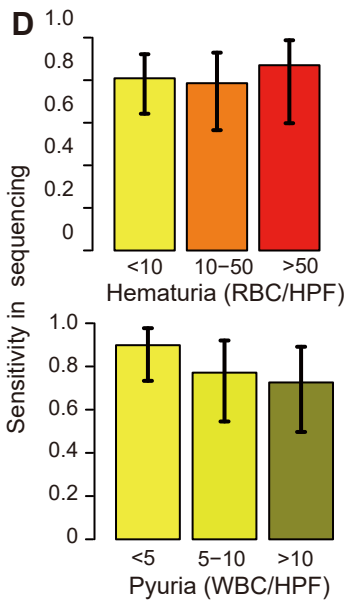
Figure S8

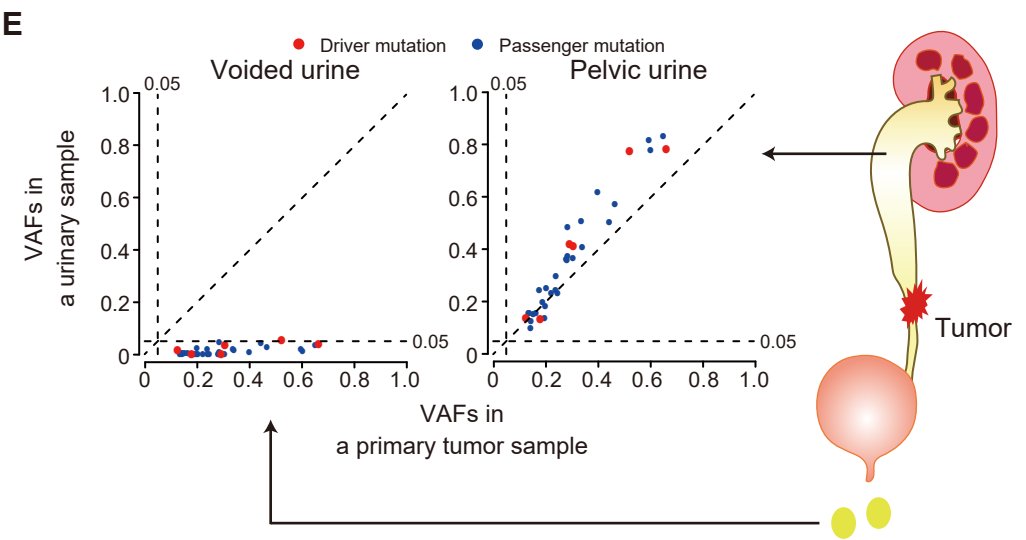


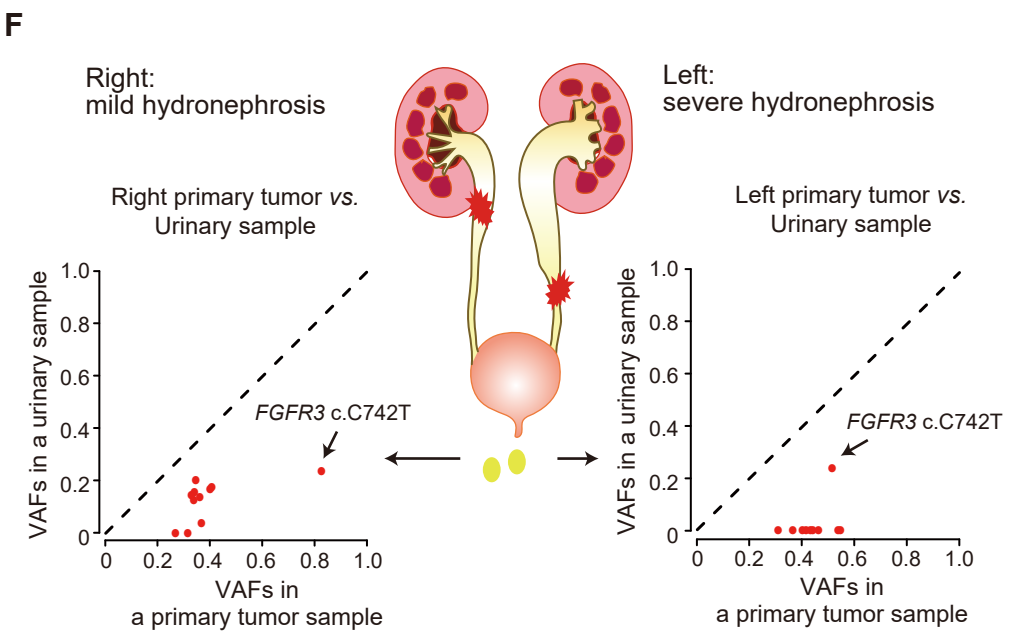
B

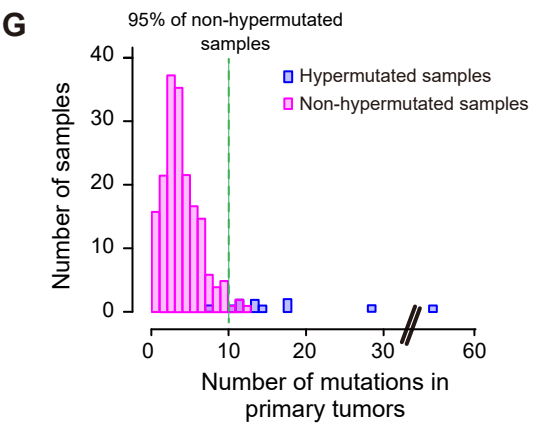
Our classification	The Paris System
Class 1: No atypical cell	Adequacy
Class 2: Reactive atypia	Negative for HGUC
Class 3: Atypical cells	Atypical urothelial cells
Class 4: Suspicious for malignancy	Suspicious for HGUC
Class 5: Malignant	HGUC/LGUN/ Other malignancies











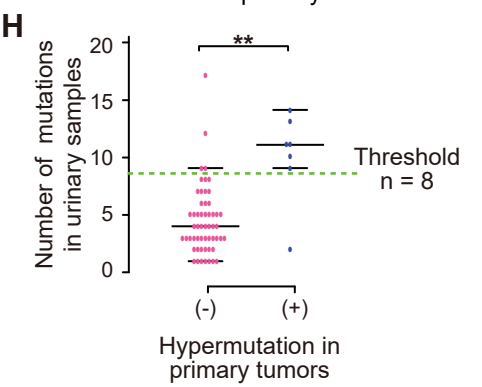


Figure S8, related to Figure 6. Urinary sediment-derived DNA sequencing.

(A) Bar plots showing depth (top) and coverage (bottom) of targeted-capture sequencing.

(B) A table showing correspondence of urinary cytology criteria in our institution classification and the Paris System. HGUC, high grade urothelial carcinoma; LGUN, low grade urothelial neoplasia.

(C) VAFs of mutations in urinary samples (top) and numbers of mutations (bottom) in primary tumors and/or urinary samples. Mutations shared by primary tumors and urinary samples, detected only in primary tumors, and only in urinary samples are shown in different colors.

(D) Bar plots with 95% confidence interval error bars indicating sensitivity of sequencing by the severity of hematuria (top) and pyuria (bottom).

(E) Diagonal plots of variant allele frequencies between mutations detected in the primary tumor and corresponding preoperative urinary sediment samples obtained from voided urine (left) or pre-obstructive pelvic urine (middle) in a case with severe hydronephrosis (UTUC203), as shown in the right schema.

(F) Comparisons of VAFs between primary tumors and preoperative urinary sediment in a case with bilateral tumor-induced hydronephrosis with different severities as shown in the schema (UTUC260). Note that left and right tumor shared no mutation except for *FGFR3* c.C742T.

(G) A histogram of numbers of mutations in the 30 panel genes detected by whole exome sequencing in non−hypermutated (pink) and hypermutated (blue) samples.

(H) Numbers of mutations detected in urinary sediment samples in non-hypermutated (pink) and hypermutated (blue) samples. A threshold of hyper- and non-hypermutated (green) and the median, first and third quartile (black) are indicated. *P* value calculated by two-sided Mann-Whitney *U* test. $^{**}P < 0.01$.

VAF, variant allele frequency.

Table S6, related to Figure 6. Patients' backgrounds of urinary sediment-derived DNA sequencing.

Characteristics	Entire cohort (*n* = 199)	Preoperative urine samples				Postoperative urine samples			
		Primary cohort (*n* = 41)	*P*-value (vs. entire cohort)	Validation cohort (*n* = 32)	*P*-value (vs. entire cohort)	Primary cohort (*n* = 25)	*P*-value (vs. entire cohort)	Validation cohort (*n* = 12)	*P*-value (vs. entire cohort)
Sex									
Male	139	31		22		16		8	
Female	60	10	0.57	10	1.00	9	0.65	4	0.76
T-category									
Ta	50	9		8		4		3	
T1	49	11		8		8		1	
T2	13	3		2		1		1	
T3	80	17		12		11		7	
T4	7	1	0.98	2	0.93	1	0.93	0	0.63
Grade									
Low	90	13		9		10		5	
High	109	28	0.12	23	0.08	15	0.67	7	1.00
Location									
Pelvis	123	24		19		15		5	
Ureter	65	16		13		9		7	
Unknown	11	1	0.67	0	0.40	1	0.93	0	0.22
Squamous									
positive	23	3		1		1		1	
negative	176	38	0.58	31	0.21	24	0.49	11	1.00
Mutational subtype									
Hypermutated	11	2		5		1		1	
TP53/MDM2-mutated	75	17		13		11		5	
RAS-mutated	30	3		3		0		3	
FGFR3-mutated	70	14		11		10		2	
Triple-negative	13	5	0.53	0	0.18	3	0.16	1	0.50

P values calculated by two-sided Fisher's exact test.

Table S8, related to Figure 6. Summary of urinary sediment-derived DNA sequencing and urinary cytology.

Cohort	Cancer detection				Mutation detection		Focal CNA detection	
	Sequencing		Cytology					
	Number of positive samples	Sensitivity (95% CI)	Number of positive samples	Sensitivity (95% CI)	Number of mutations (Urine/Tumor)	Sensitivity (95% CI)	Number of mutations (Urine/Tumor)	Sensitivity (95% CI)
Primary cohort	32/41	78.0% (62.4-89.4%)	12/41	29.3% (16.1-45.5%)	136/203	67.0% (60.1-73.4%)	26/27	96.3% (81.0-99.9%)
Validation cohort	28/32	87.5% (71.0-96.5%)	12/32	37.5% (21.1-56.3%)	147/177	83.1% (76.7-88.3%)	23/31	74.2% (55.4-88.1%)
All	60/73	82.2% (71.5-90.2%)	24/73	32.9% (22.3-44.9%)	283/380	74.5% (69.8-78.8%)	49/58	84.5% (72.6-92.7%)
Non-urothelial cancer patients	0/18	0% (0-18.5%)	2/18	11.1% (1.4%-34.7%)	—	—	—	—

CNA, copy number alteration; CI, confidence interval.