Cell Reports Medicine, Volume *4*

Supplemental information

Plasma proteomic profiling discovers molecular features associated with upper tract urothelial carcinoma

Yuanyuan Qu, Zhenmei Yao, Ning Xu, Guohai Shi, Jiaqi Su, Shiqi Ye, Kun Chang, Kai Li, Yunzhi Wang, Subei Tan, Xiaoru Pei, Yijiao Chen, Zhaoyu Qin, Jinwen Feng, Jiacheng Lv, Jiajun Zhu, Fahan Ma, Shaoshuai Tang, Wenhao Xu, Xi Tian, Aihetaimujiang Anwaier, Sha Tian, Wenbo Xu, Xinqiang Wu, Shuxuan Zhu, Yu Zhu, Dalong Cao, Menghong Sun, Hualei Gan, Jianyuan Zhao, Hailiang Zhang, Dingwei Ye, and Chen Ding

Supplementary Figure 1

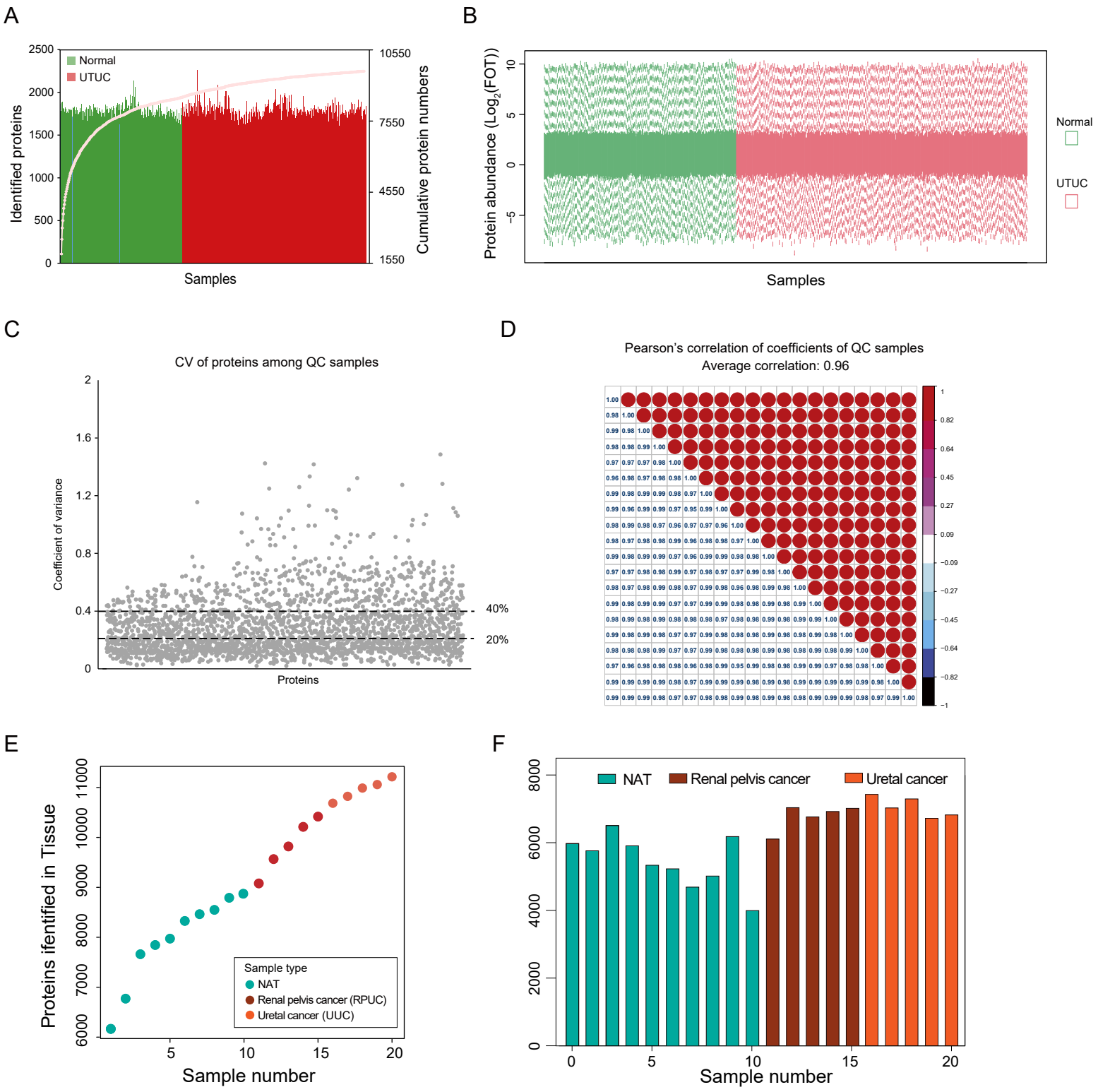


Figure S1. Overall synopsis of the plasma proteome profiling of patients with UTUC. Related to Figure 1.
(A) Proteins identified in each UTUC and normal plasma samples (down). Cumulative number of proteins Identified (up). **(B)** Distribution of log2-transformed FOT abundance of identified proteins in 601 plasma proteome samples that passed quality control. Green presents Normal plasma samples (n = 239), red denotes UTUC samples (n = 362). In the box plots, the middle bar represents the median, and the box represents the interquartile range; bars extend to 1.5× the interquartile range. **(C)** Coefficient of variance of proteins of plasma pooled QC samples. **(D)** Pearson correlation of QC samples. **(E)** Proteins identified and cumulative number of protein identifications in UTUC tissues. Green presents normal tissue samples (NAT); dark red presents RPUC; light red presents UUC. **(F)** Bar chart showing proteins identified in normal and UTUC tissue samples.

Suppermentary Figure 2

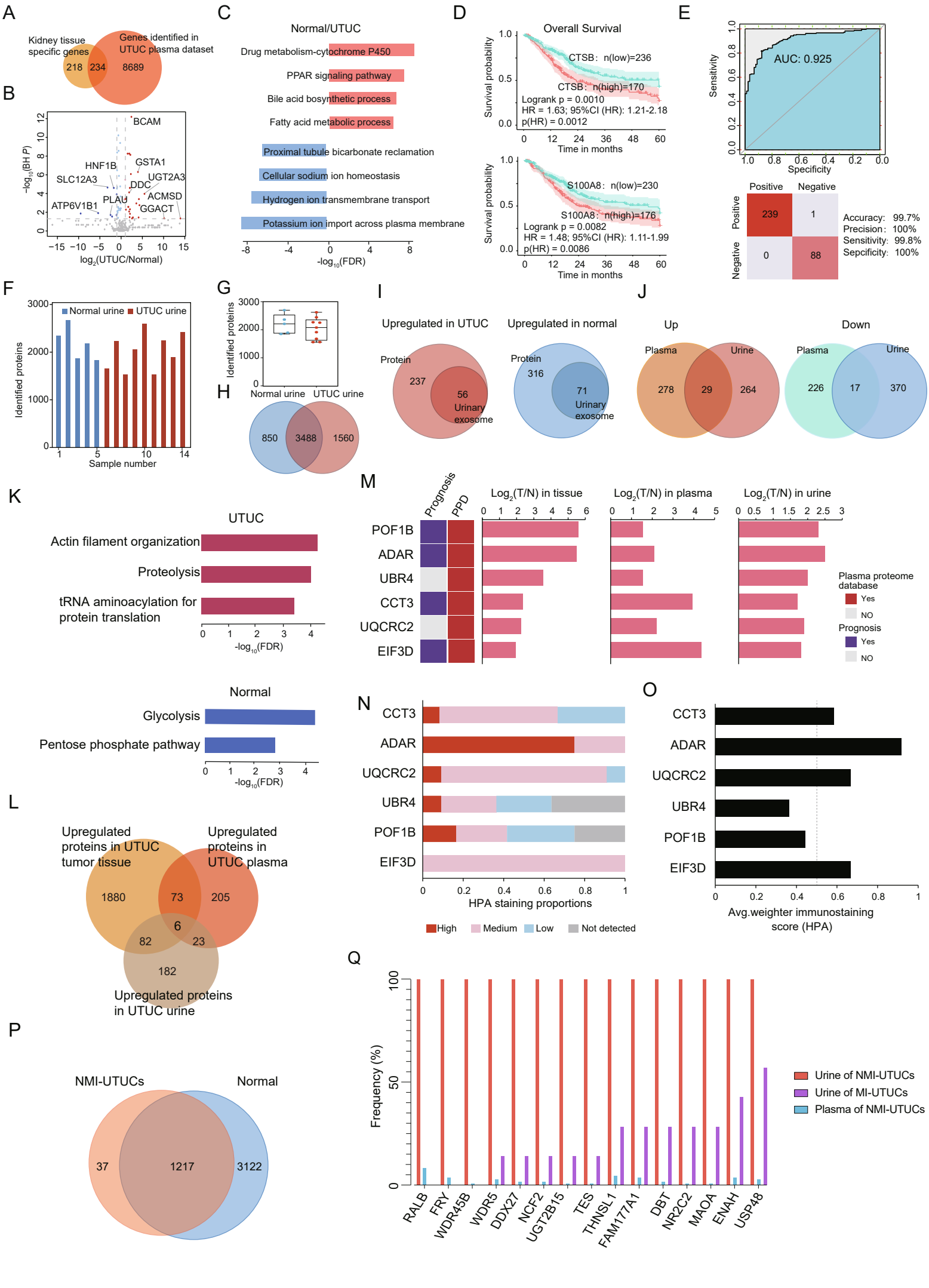


Figure S2. Plasma and urine proteome profiles differed between UTUC and normal Samples. Related to Figure 2.

(A) VENN plot of kidney tissue specific proteins identified between the Human Protein Atlas (HPA) dataset and UTUC plasma discovery cohort. **(B)** Volcano plot showing kidney tissue specific proteins differently expressed in UTUC and normal plasma samples (Wilcoxon rank-sum test). **(C)** Pathways enriched in normal and UTUC plasma samples. **(D)** Overall survival analyses of UC patients with high or low levels of CTSB in TCGA BLCA cohort, p value from log-rank test. 95% confidence interval and hazard ratios (HR) were also presented. **(E)** ROC of plasma biomarkers in the independent validation cohort. **(F)** Bar chart showing proteins identified in normal and UTUC urine samples. **(G)** Box plot showing the proteins identified in normal and UTUC urine samples. **(H)** VENN plot of proteins identified in normal and UTUC urine samples. **(I)** VENN plot showing that urinary exosome in upregulated proteins in normal and UTUC samples. **(J)** VENN plot showing the upregulated (up) and downregulated (down) proteins between plasma and urine of UTUC patients. **(K)** Pathway analysis of upregulated (red) or downregulated protein (blue). **(L)** Venn diagram showing the upregulated protein (Wilcoxon rank-sum test, BH $P < 0.05$, UTUC / Normal ratio > 1.5) overlap among UTUC tumor tissue, UTUC plasma, and UTUC urine. **(M)** Left, the annotation of 6 proteins from previously reported urothelial cancer cohort and Plasma Proteome Database (PPD, http://www.plasmaproteomedatabase.org/). Right, fold-changes of 6 proteins in UTUC (T) and Normal (N). **(N-O)** Proportions of urothelial tumors with high, medium, or low staining, or not detected (ND) as reported by the Human Protein Atlas (HPA) **(N)** and urothelial tumor-cell immunohistochemistry (IHC) staining scores defined by the HPA **(O)**. **(P)** VENN plot of proteins identified between normal samples and NMI-UTUC urine samples. **(Q)** Protein frequency in urine of NMI-UTUC compared with the plasma of NMI-UTUC.

Suppementary Figure 3

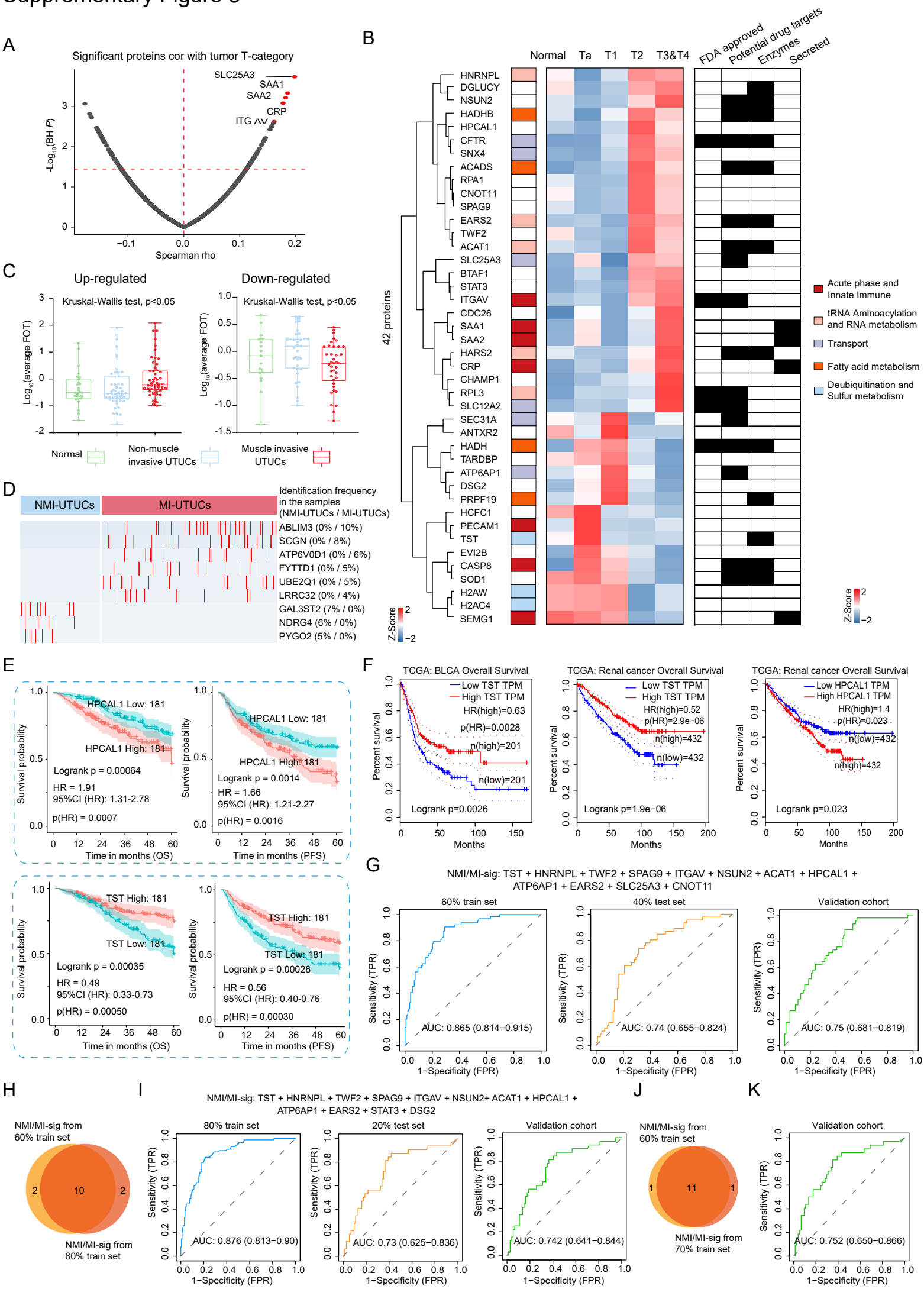
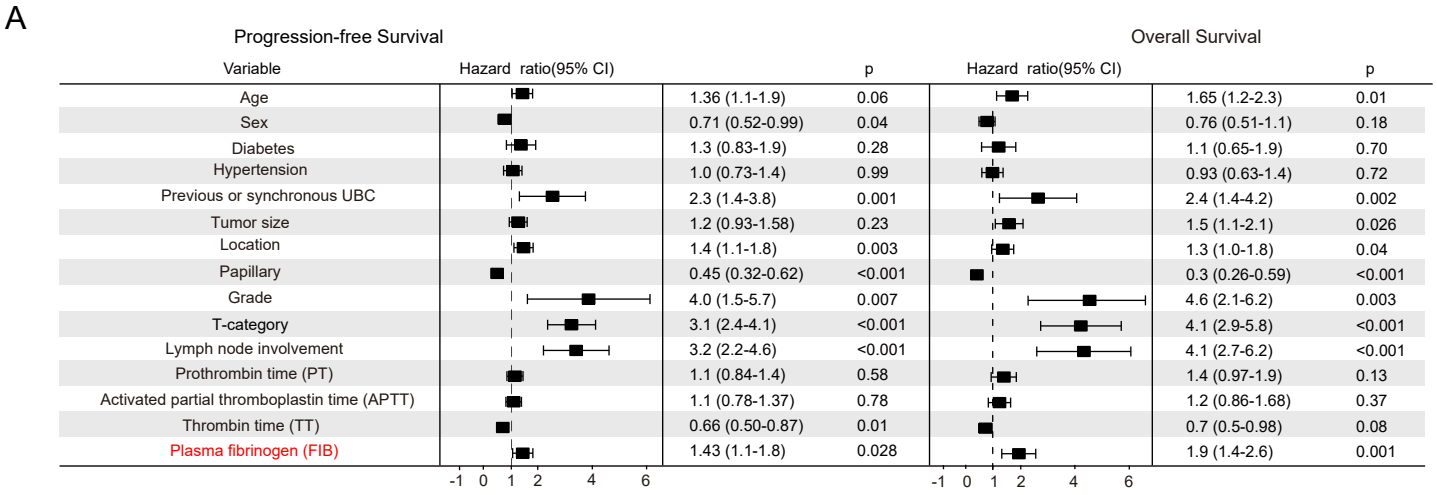
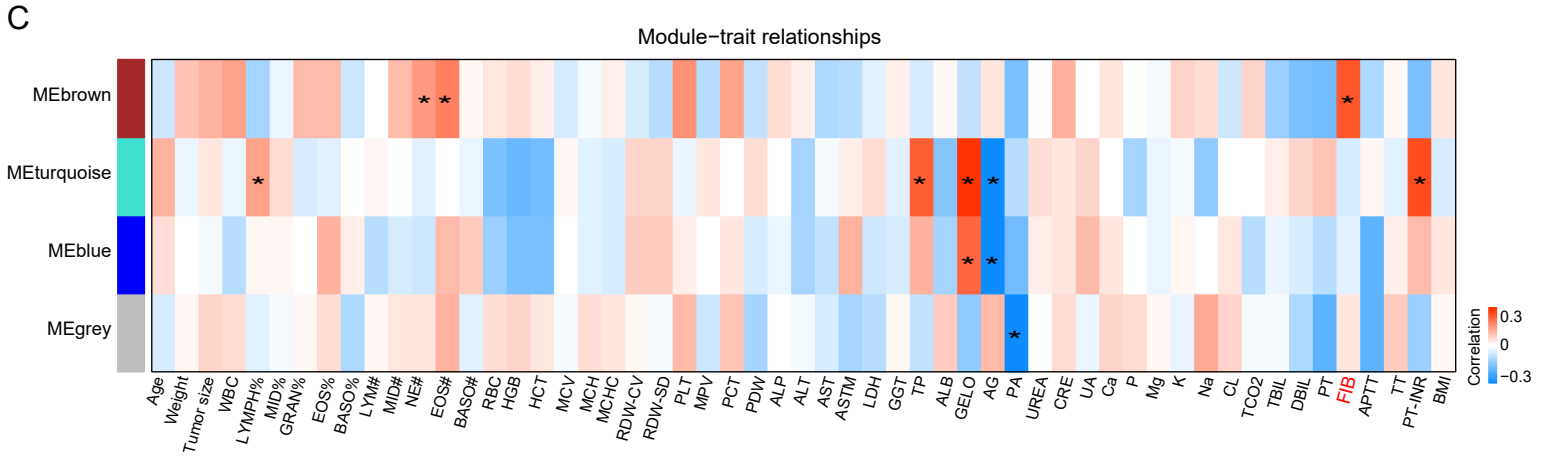


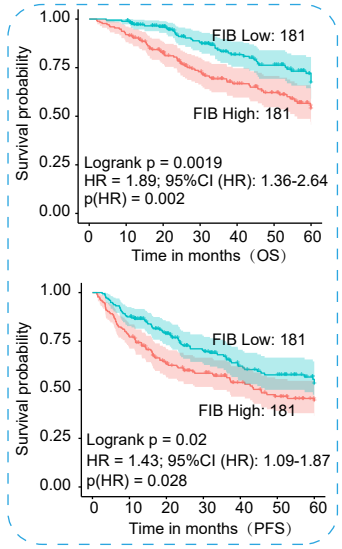
Figure S3. Plasma proteomic profiles identify patients with muscle-invasive. Related to Figure 5.

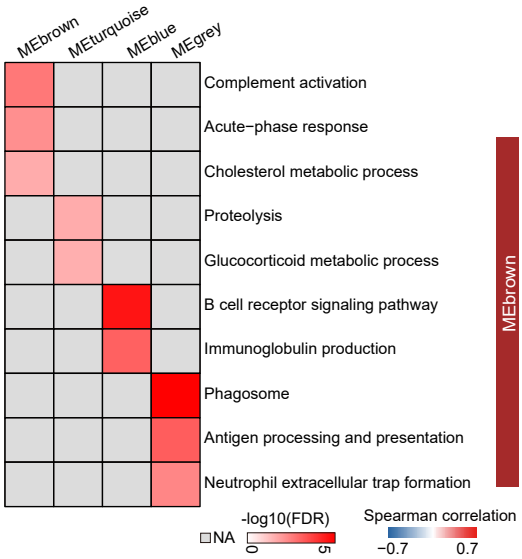
(A) Volcano plot showing the correlation between tumor T-category and protein abundance (Spearman's correlation test). **(B)** Left: the dendrogram shows the 42 differentially expressed proteins from the comparison of NMI-UTUCs and MI-UTUCs; Different color blocks represent the functional categorization of proteins. Middle: heatmap of 42 differentially expressed proteins abundance across histologic stages of normal, pTa, pT1, pT2, pT3 and pT4. Right: FDA approved drug targets, potential drug targets, enzymes, and secreted proteins annotated by the Human Protein Atlas (HPA). **(C)** Boxplots of differentially expressed proteins from Normal, NMI-UTUCs to MI-UTUCs (Kruskal-Wallis test). **(D)** Heatmap of exclusively identified plasma proteins in MI-UTUCs or NMI-UTUCs samples. **(E)** Overall survival and progression-free survival analyses of UTUC patients with high or low levels of TST (down) or HPCAL1 (up) protein abundance (p value from log-rank test) in UTUC plasma discovery cohort. 95% confidence interval and hazard ratios (HR) were also presented. **(F)** Overall survival analyses of BLCA TCGA and Renal cancer TCGA patients with high or low levels of TST (left) or HPCAL1(right) mRNA abundance (p value from log-rank test). **(G)** The ROC curves of the classifier model in predicting NMI-UTUCs and MI-UTUCs in the 60% train set, 40% test set, and validation cohort. **(H)** Overlap of NMI/MI-sigs from 80% training set and 60% training set. **(I)** The ROC curves of the classifier model in predicting NMI-UTUCs and MI-UTUCs in the 80% train set, 20% test set, and validation cohort. **(J)** Overlap of NMI/MI-sigs from 70% training set and 60% training set. **(K)** ROC curve and corresponding AUC statistics in the independent validation cohort (n=89) using PRM assays.

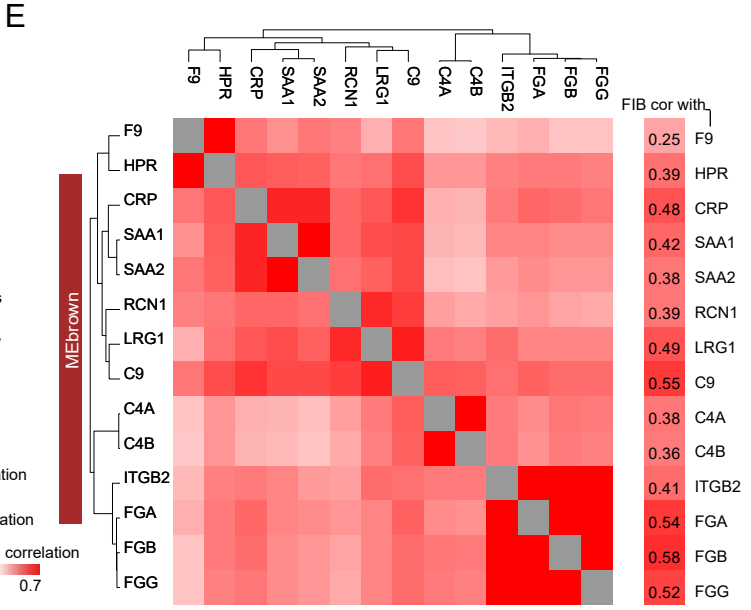
Suppermentary Figure 4











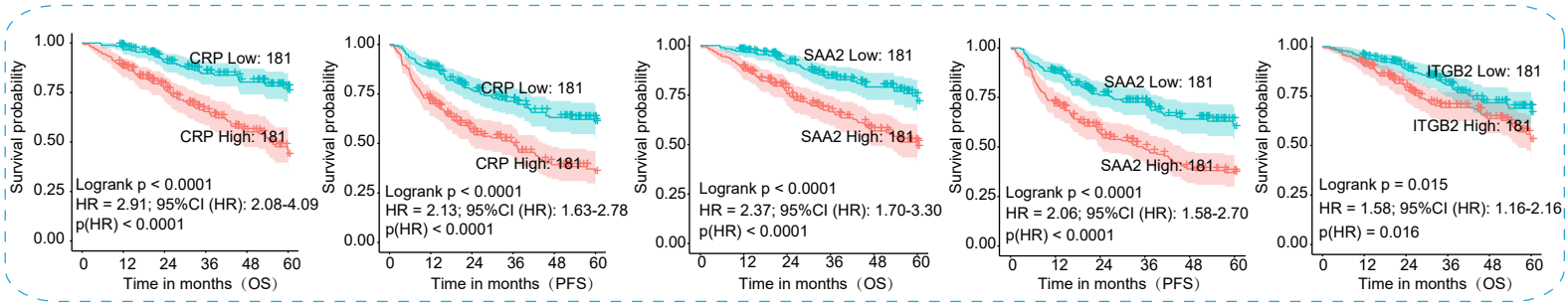
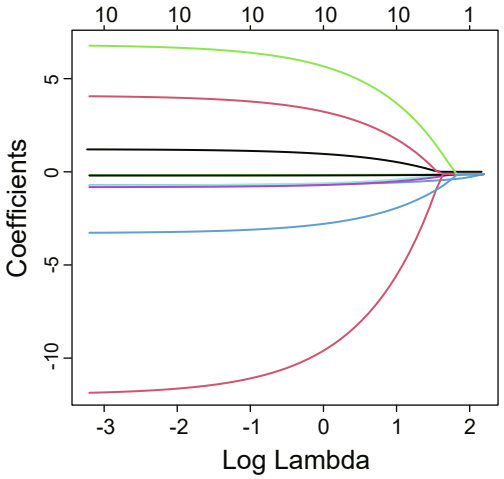


Figure S4. Clinical features associated with proteomic profiles. Related to Figure 6.

(A) Forest plot for univariate Cox regressions for age, gender, diabetes, hypertension, previous or synchronous urothelial bladder carcinoma (UBC), tumor size, location, grade, papillary, T-category, lymph node involvement, prothrombin time (PT), activated partial thromboplastin time (APTT), thrombin time (TT), and plasma fibrinogen (FIB). The main effects are presented as hazard ratios with 95% confidence intervals. **(B)** Overall survival and progression-free survival analyses of UTUC patients with high or low levels of plasma fibrinogen (p value from log-rank test) in UTUC plasma discovery cohort. 95% confidence interval and hazard ratios (HR) were also presented. **(C)** Heatmap showing the correlation between modules obtained from WGCNA analysis and clinical outcomes. **(D)** Enrichment pathway of different modules. **(E)** The correlation map showing the pertinence of plasma proteins correlated with FIB in the MEbrown module. **(F)** Overall survival and progression-free survival analyses of UTUC patients with high or low levels of CRP (left), SAA1, SAA2 or ITGB2 (right) protein abundance (p value from log-rank test) in UTUC plasma discovery cohort. 95% confidence interval and hazard ratios (HR) were also presented.

Supplementary Figure 5

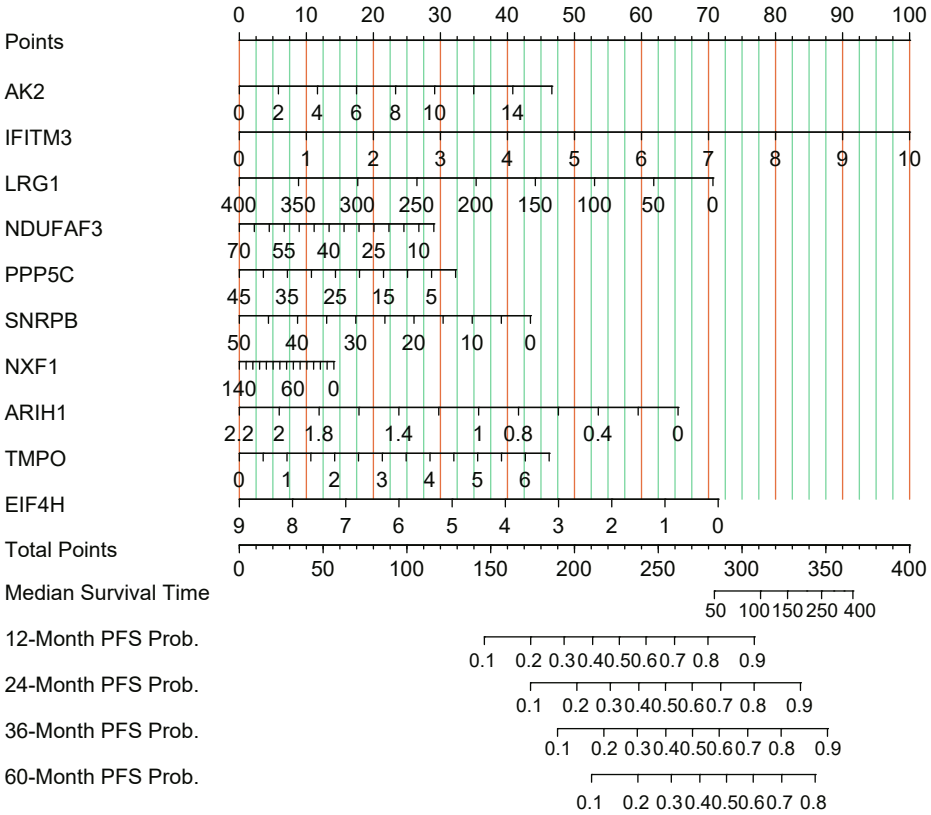
A



B

No.	Name	Coefficient
1	Intercept	53.557609
2	TMPO	1.5458147
3	IFITM3	0.1816506
4	AK2	0.0634408
5	LRG1	-0.020376
6	NXF1	-0.039298
7	PPP5C	-0.052722
8	SNRPB	-0.159858
9	NDUFAF3	-0.341377
10	ARIH1	-0.82372
11	EIF4H	-0.845486

C



D

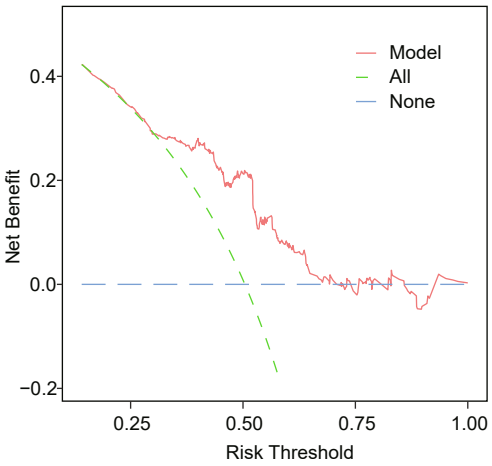


Figure S5. Progression-related proteins classifier of patients with UTUC. Related to Figure 7.
(A) Least absolute shrinkage and selection operator (LASSO) coefficient profiles of 10 proteins. **(B)** The coefficient using in Lasso Cox regression of 10 proteins. Ak2, IFITM3, LRG1, NDUFAF3, PPP5C, SNRPB, NXF1, ARIH1, TMPO, and EIF4H. **(C)** Nomogram for predicting 1-, 3-, and 5- years for patients with UTUC and the 10 proteins parameters. **(D)** Decision curves of the nomogram for survival prediction of UTUC patients.

Supplementary Figure 6

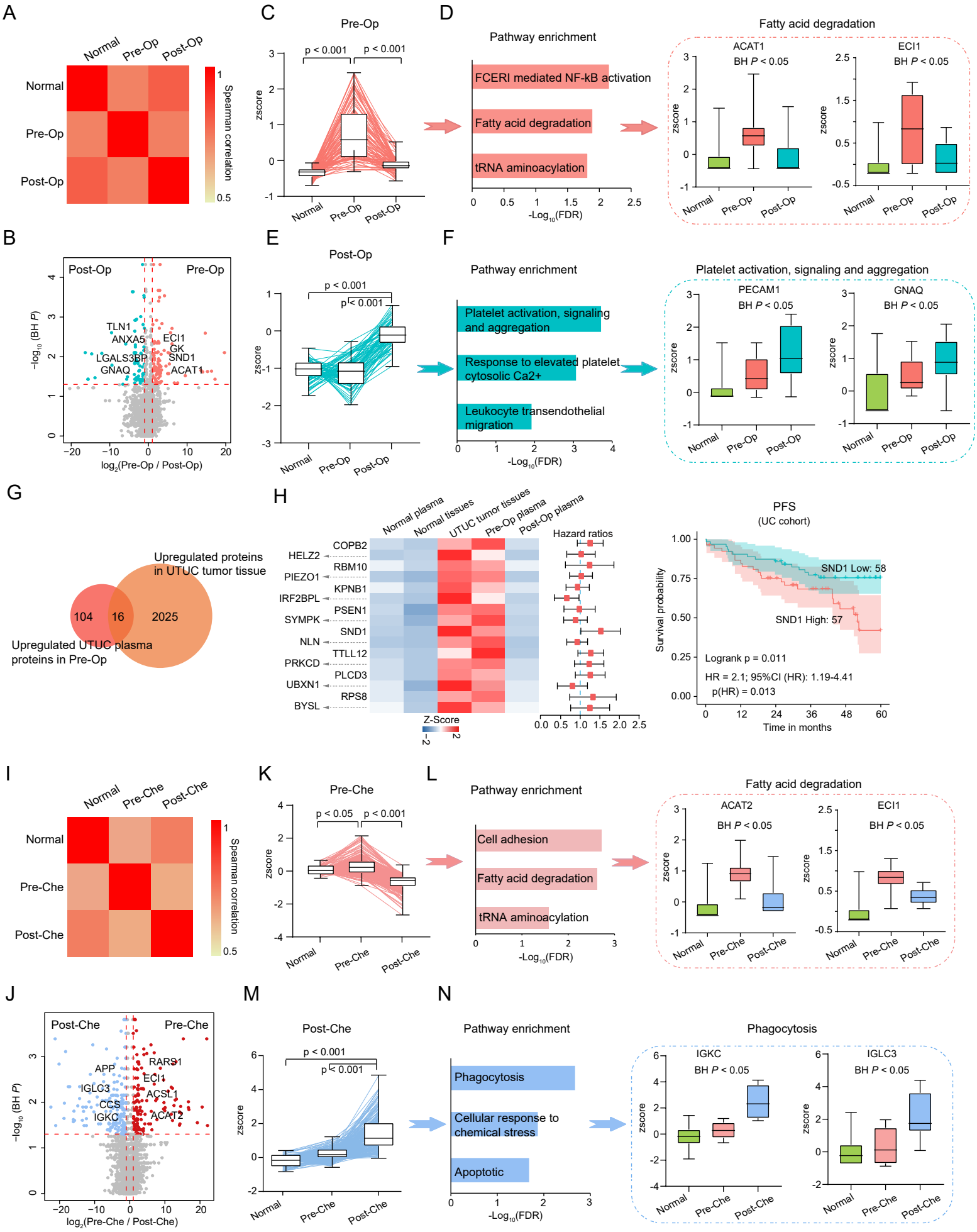


Figure S6. The dynamic changes in response to major changes in tumor burden. Related to Figure 2.

(A) Spearman correlation of plasma proteins among healthy controls (Normal), Pre-Op, and Post-Op. **(B)** Proteins abundance differences between Pre-Op and Post-Op (Wilcoxon rank-sum test). **(C)** Line plots and boxplots of DEPs that were upregulated in Pre-Op (Wilcoxon rank-sum test). Values are standardized by z-score. **(D) Left:** Pathways enrichment analysis based on the DEPs in Pre-Op. **Right:** The protein abundance of ACAT1 and ECI1 in Normal, Pre-Op, and Post-Op groups (Kruskal-Wallis test). Values are standardized by z-score. **(E)** Line plots and boxplots of DEPs that were upregulated in Post-Op (Wilcoxon rank-sum test). **(F) Left:** Pathways enrichment analysis based on the DEPs in Post-Op. **Right:** The protein abundance of PECAM1 and GNAQ in Normal, Pre-Op, and Post-Op groups (Kruskal-Wallis test). **(G)** The Venn diagram shows the overlap of upregulated proteins in Post-Op versus upregulated proteins in UTUC tumor tissues. **(H) Left:** Heatmap showing 16 protein abundance among five groups (Normal plasma, Normal tissues, UTUC tumor tissues, Pre-Op plasma, and Post-Op plasma). Values are standardized by z-score. **Middle,** prognostic risk scores (hazard ratio) of each protein. The middle red points indicate hazard ratio for each protein; endpoints represent lower or upper 95% confidence intervals. **Right:** Progression-free survival (PFS) analyses of patients with high or low levels of SND1 protein abundance in urothelial bladder cancer cohort. **(I)** Spearman correlation of plasma proteins among healthy controls (Normal), Pre-Che, and Post-Che. **(J)** Proteins abundance differences between Pre-Che and Post-Che (Wilcoxon rank-sum test). **(K)** Line plots and boxplots of DEPs that were upregulated in Pre-Che (Wilcoxon rank-sum test). Values are standardized by z-score. **(L) Left:** Pathways enrichment analysis based on the DEPs in Pre-Che. **Right:** The protein abundance of ACAT2 and ECI1 in Normal, Pre-Che, and Post-Che groups (Kruskal-Wallis test). Values are standardized by z-score. **(M)** Line plots and boxplots of DEPs that were upregulated in Post-Che (Wilcoxon rank-sum test). **(N) Left:** Pathways enrichment analysis based on the DEPs in Post-Che. **Right:** The protein abundance of IGKC and IGLC3 in Normal, Pre-Che, and Post-Che groups (Kruskal-Wallis test). Values are standardized by z-score.

Supplementary Figure 7

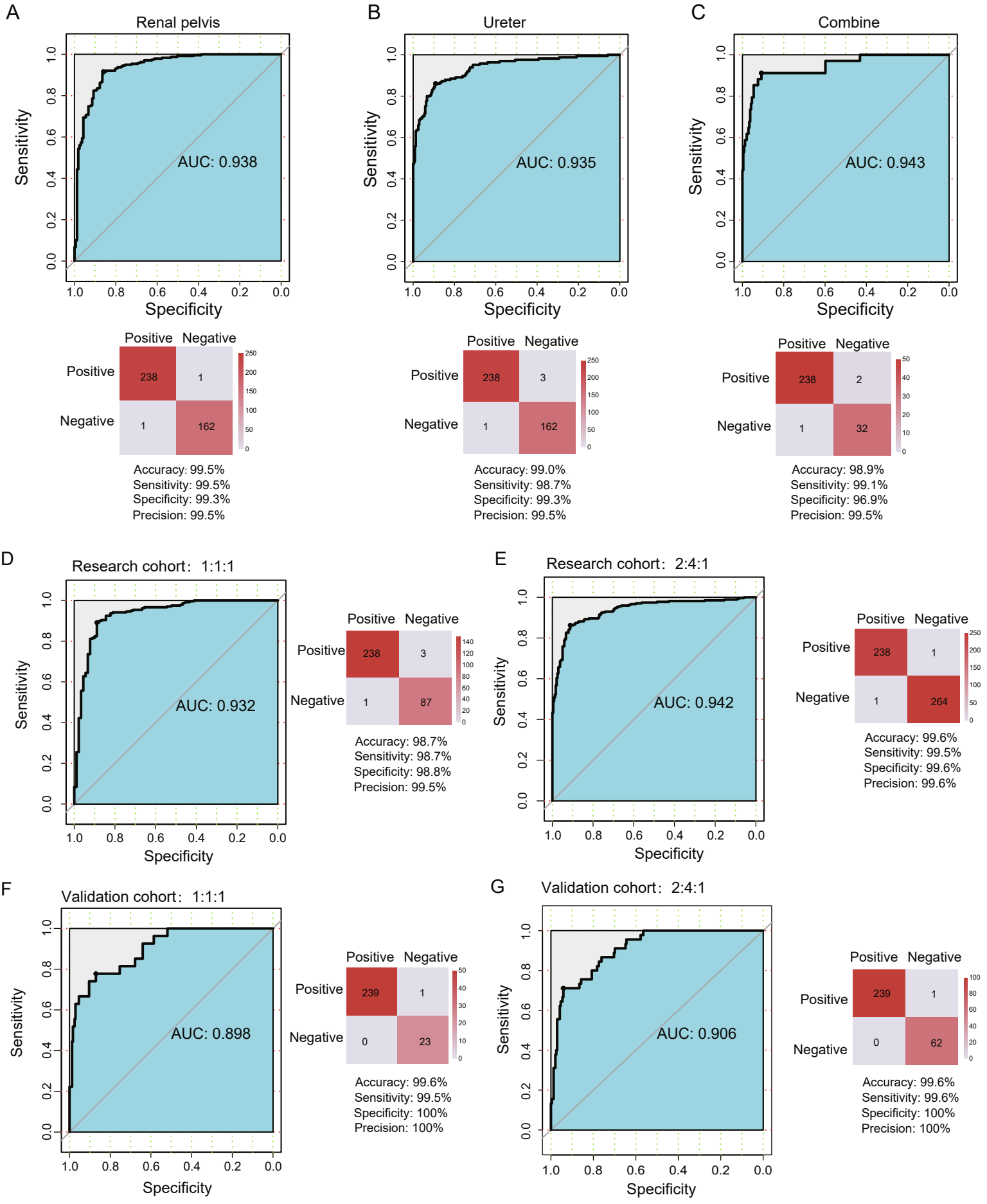


Figure S7. ROC of plasma biomarkers in the renal pelvis samples, ureter samples, and combined samples. Related to Figure 2. **(A)-(C)** ROC curve and corresponding AUC statistics in the renal pelvis samples **(A)**, ureter samples **(B)**, and combined samples **(C)**. **(D)-(E)** ROC curve of mixed the UTUC samples in a ratio of 1:1:1(D) and 2:4:1**(E)** in research cohort. **(F)-(G)** ROC curve of mixed the UTUC samples in a ratio of 1:1:1**(F)** and 2:4:1**(G)**.