Cell Reports Medicine, Volume *5*

Supplemental information

Proteomic landscape profiling of primary prostate cancer reveals a 16-protein panel for prognosis prediction

Rui Sun, Jun A, Haolan Yu, Yan Wang, Miaoxia He, Lingling Tan, Honghan Cheng, Jili Zhang, Yingrui Wang, Xiaochen Sun, Mengge Lyu, Min Qu, Lingling Huang, Zijian Li, Wenhui Zhang, Kunpeng Ma, Zhenyang Dong, Weigang Ge, Yun Zhang, Xuan Ding, Bo Yang, Jianguo Hou, Chuanliang Xu, Linhui Wang, Yi Zhu, Tiannan Guo, Xu Gao, and Chenghua Yang

Supplemental information

Proteomic landscape profiling of primary prostate cancer reveals a 16-protein panel for prognosis prediction

Rui Sun, Jun A, Haolan Yu, Yan Wang, Miaoxia He, Lingling Tan, Honghan Cheng, Jili Zhang, Yingrui Wang, Xiaochen Sun, Mengge Lyu, Min Qu, Lingling Huang, Zijian Li, Wenhui Zhang, Kunpeng Ma, Zhenyang Dong, Weigang Ge, Yun Zhang, Xuan Ding, Bo Yang, Jianguo Hou, Chuanliang Xu, Linhui Wang, Yi Zhu, Tiannan Guo, Xu Gao, Chenghua Yang

Figure S1

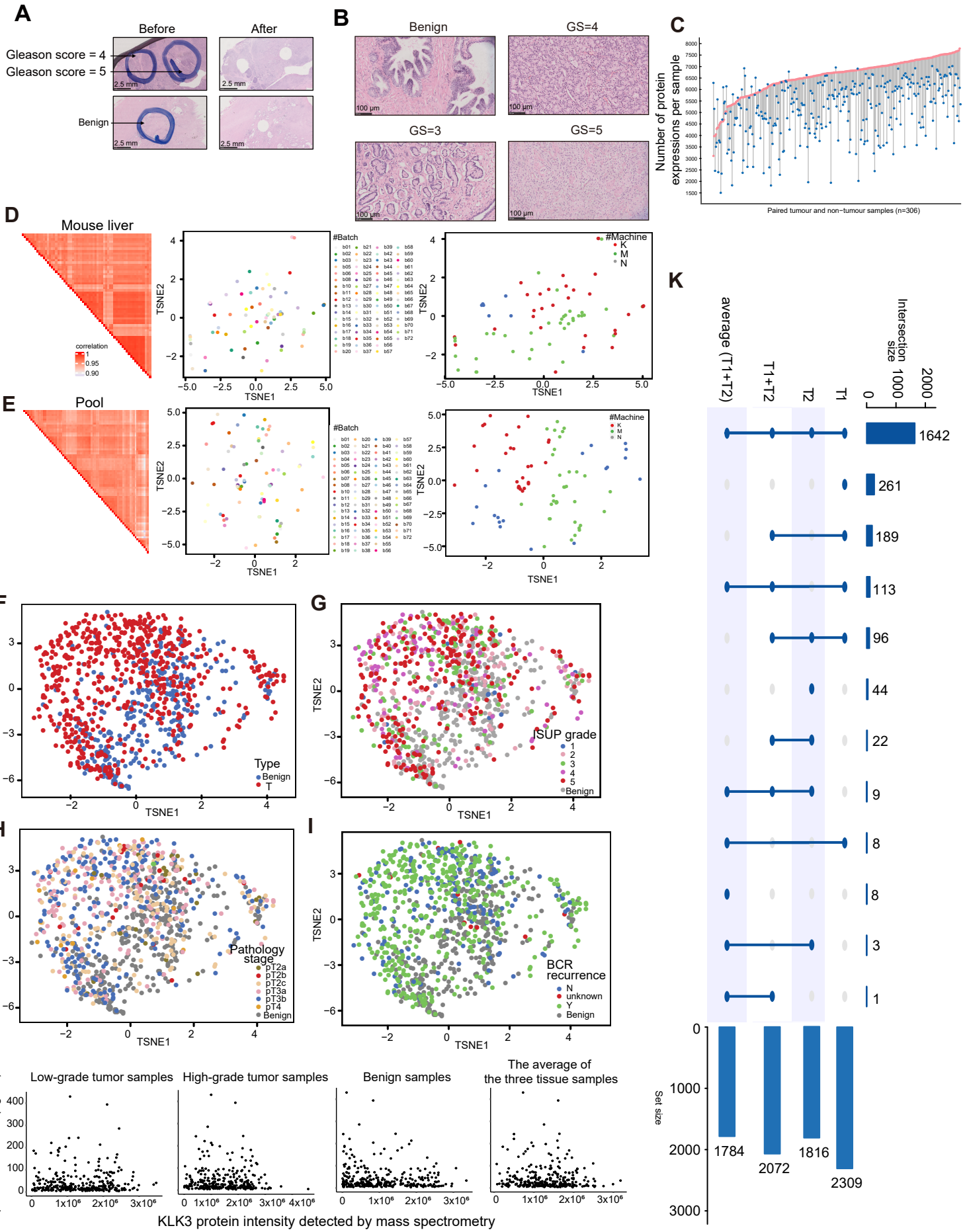
A
Before
After
Gleason score = 4
Gleason score = 5
Benign
2.5 mm
B
Benign
GS=4
GS=3
GS=5
100 μm
C
Number of protein expressions per sample
Paired tumour and non−tumour samples (n=306)
D
Mouse liver
correlation
#Batch
#Machine
TSNE1
TSNE2
E
Pool
F
Type
Benign
T
G
ISUP grade
Benign
H
Pathology stage
pT2a
pT2b
pT2c
pT3a
pT3b
pT4
Benign
I
BCR recurrence
N
unknown
Y
Benign
J
Low-grade tumor samples
High-grade tumor samples
Benign samples
The average of the three tissue samples
prePSA in serum (ng/mL)
KLK3 protein intensity detected by mass spectrometry
K
average (T1+T2)
T1+T2
T2
T1
Intersection size
Set size
1642
261
189
113
96
44
22
9
8
8
3
1
1784
2072
1816
2309

Figure S1 Representative HE staining showing the FFPE sampling of PCa in this study. Related to Figure 1.
(A) HE staining shows the FFPE sample before and after the punch.
(B) HE staining shows the different Gleason score criteria used by pathologists in our study.
(C) The quantification of protein numbers for each patient is shown. Each line represents a patient, with the red dot representing the average protein number identified in two tumor samples, and the blue dot representing the protein number identified in the adjacent benign sample.
(D-E) Left panel: Correlation matrix for quantality samples, including 72 mouse liver samples (D) for sample preparation QC and 72 pool samples for MS QC (E). Middle panel: tSNE plots showing the proteomic data distribution of 72 batches of mixed PCa QC samples collected. Right panel: tSNE plots showing the distribution of proteomic data of mixed PCa QC samples collected by three mass spectrometers (K,N,M).
(F-I) tSNE plots showing the proteomic data distribution of different sample type (F), ISUP grade (G), pathology stage (H) and BCR recurrence (I).
(J) The correlation of between serum PSA and tissue KLK3 expression.
(K) Comparison of four methods for differentially expressed proteins between tumor and benign samples. “T1”: tumor samples from primary Gleason score (GS1) were compared with benign samples; “T2”: tumor samples from secondary GS (GS2) were compared with benign samples. “T1+T2”: Two tumor samples from the same patient, considered as individual entities, were compared with benign samples; “average (T1+T2)”: The average levels of the two tumor samples from the same patient were compared with benign samples.

Figure S2

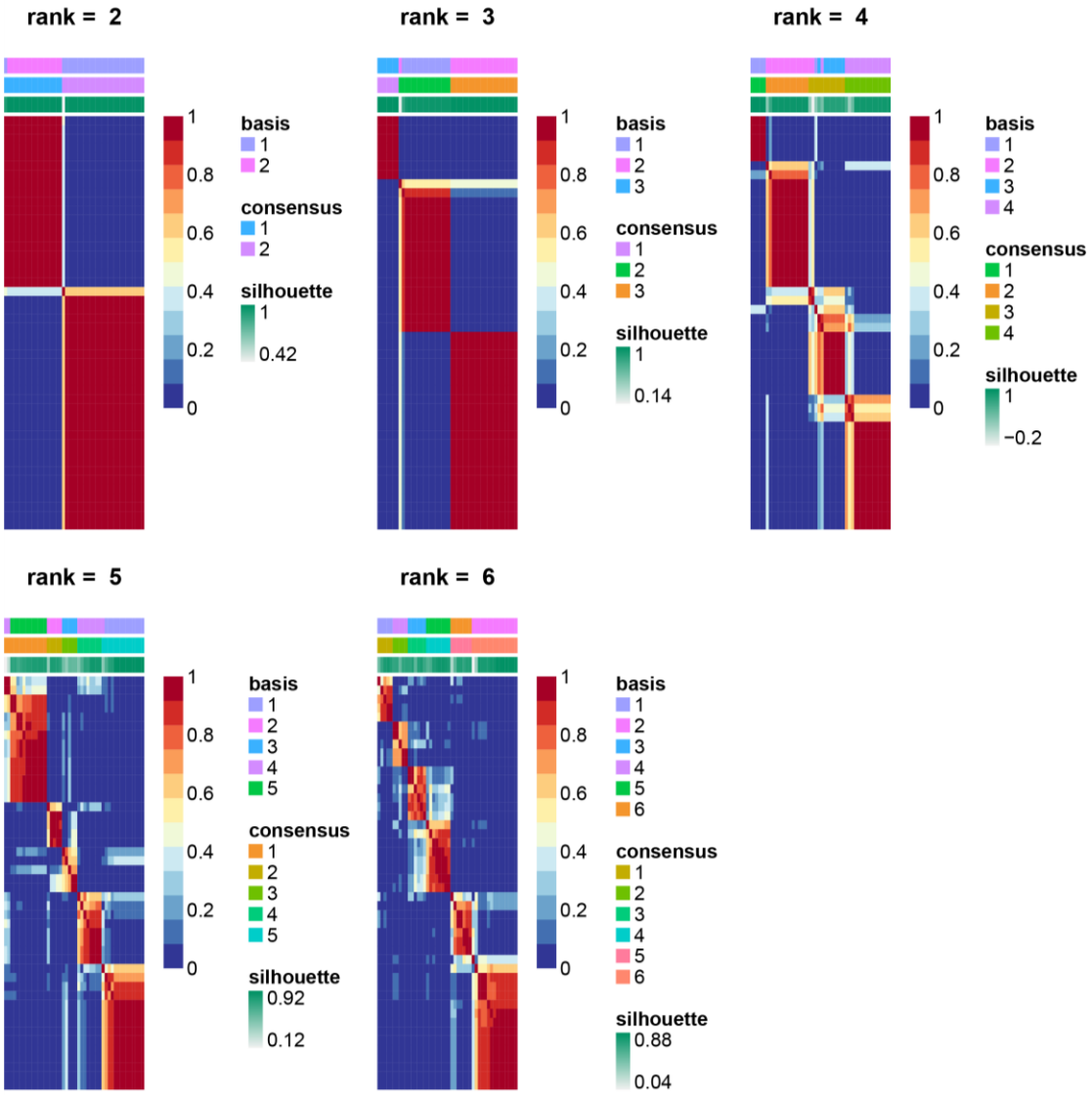

Figure S2. Consensus clustering of the proteomic data. Related to Figure 3. Subgroups are identified based on proteomic data by K-means consensus clustering of their abundances.

Figure S3

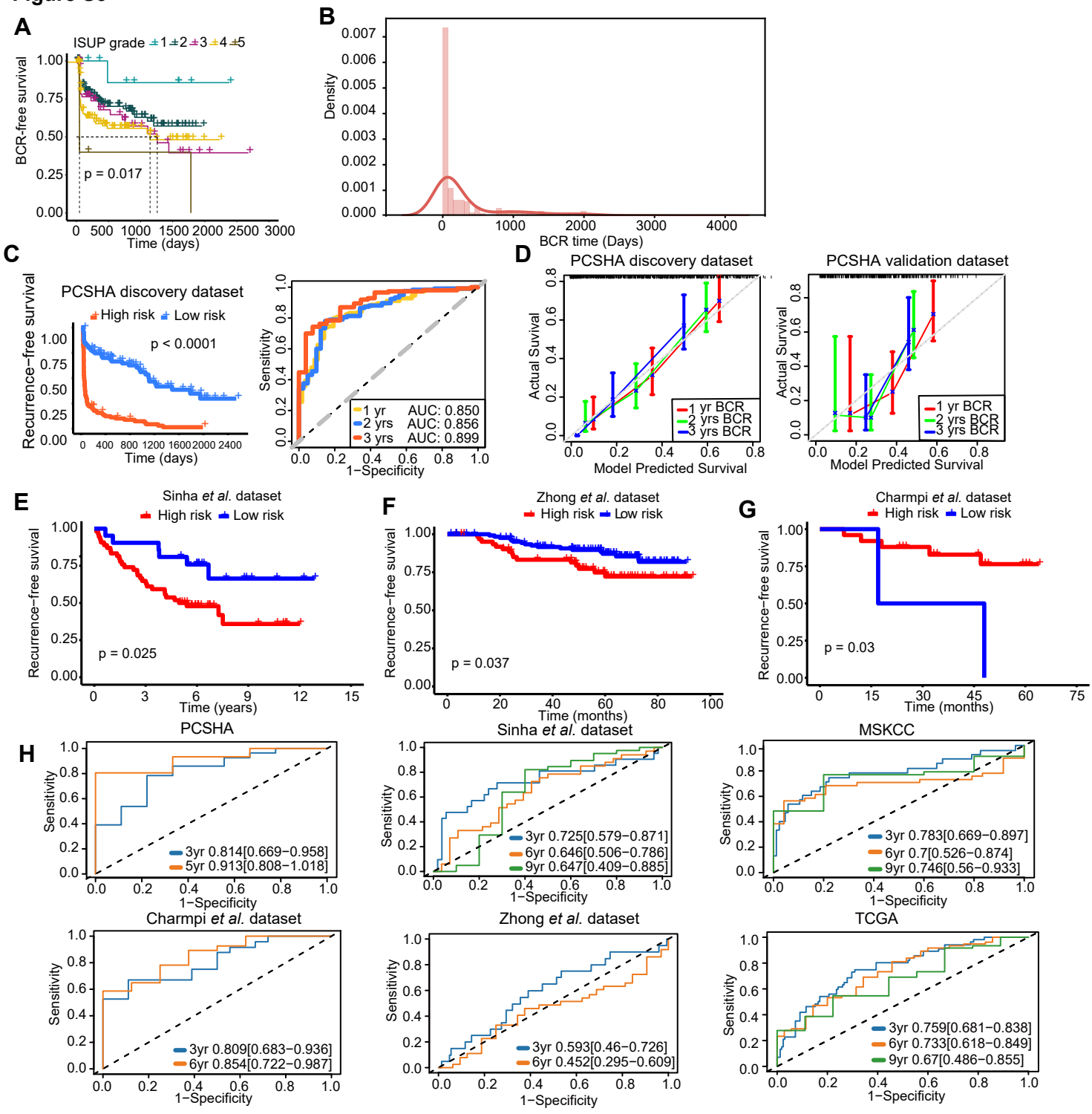

Figure S3 Identification and validation of a panel of 16 protein biomarkers for prognostic prediction Related to Figure 4.

(A) Kaplan-Meier analysis of BCR-free survival for PCa across different ISUP grade (log rank p = 0.017).
(B) BCR time distribution of all 306 patients in our study.
(C) Kaplan-Meier analysis (right) and ROC curves (left) of the PCSHA discovery dataset for BCR-free survival prediction based on our 16-protein model, with 1-year predictive power (yellow), 2-year predictive power (blue), and 3-year predictive power (red).
(D) Calibration curves of the 16-protein panel in PCSHA discovery dataset (left panel) and validation dataset (right panel).
(E-G) Kaplan-Meier analysis of BCR-free survival based on our 16-protein panel of the dataset Zhong et al. dataset (E), the Sinha et al. dataset (F), and the Charmpi et al. dataset (G).
(H) ROC curves of the six datasets for BCR-free survival prediction based on our 16-protein model, with various years' predictive power.

Figure S4

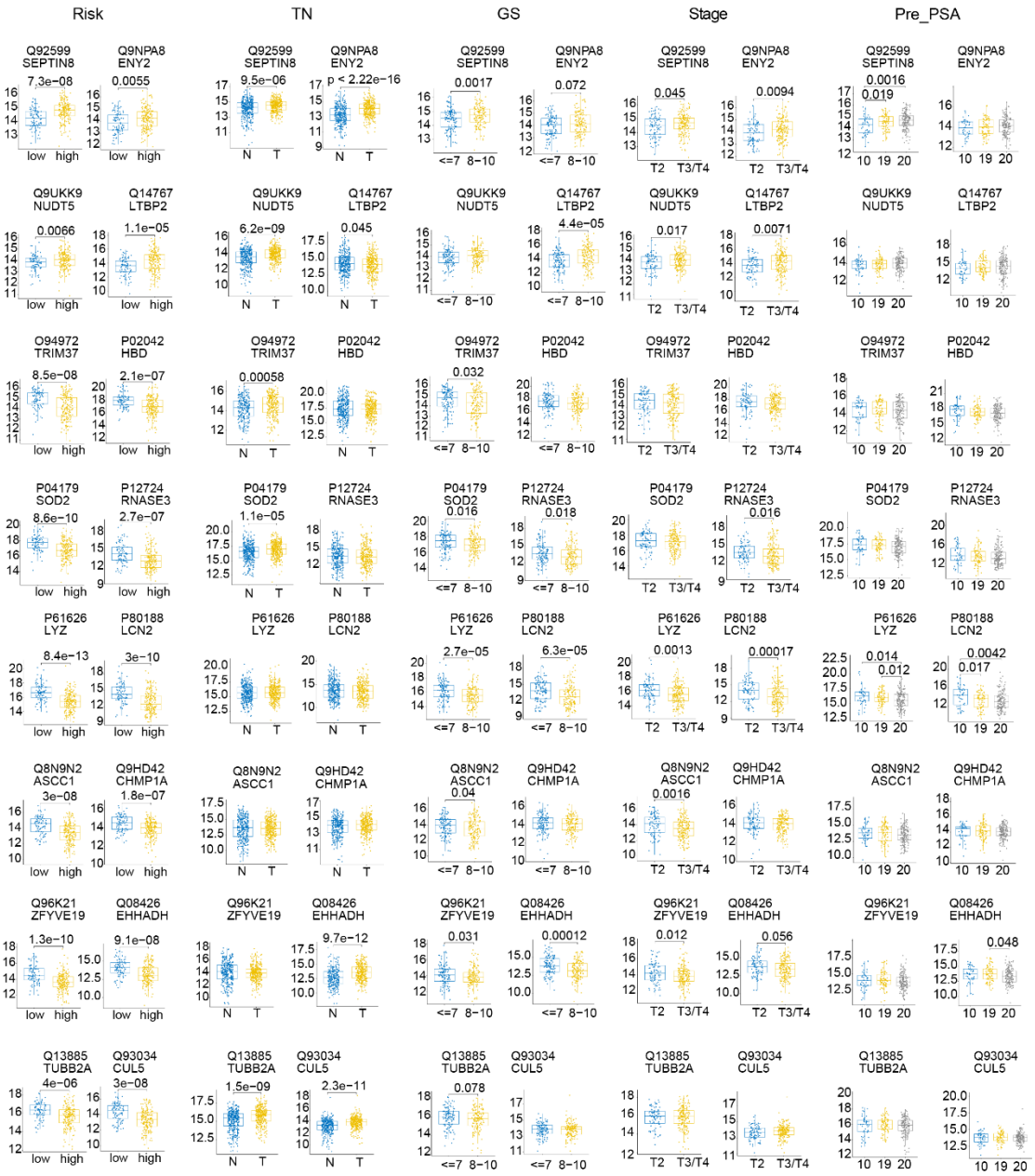

Figure S4 Relative protein expression of the 16 proteins in different groups. Related to Figure 4.

Figure S5

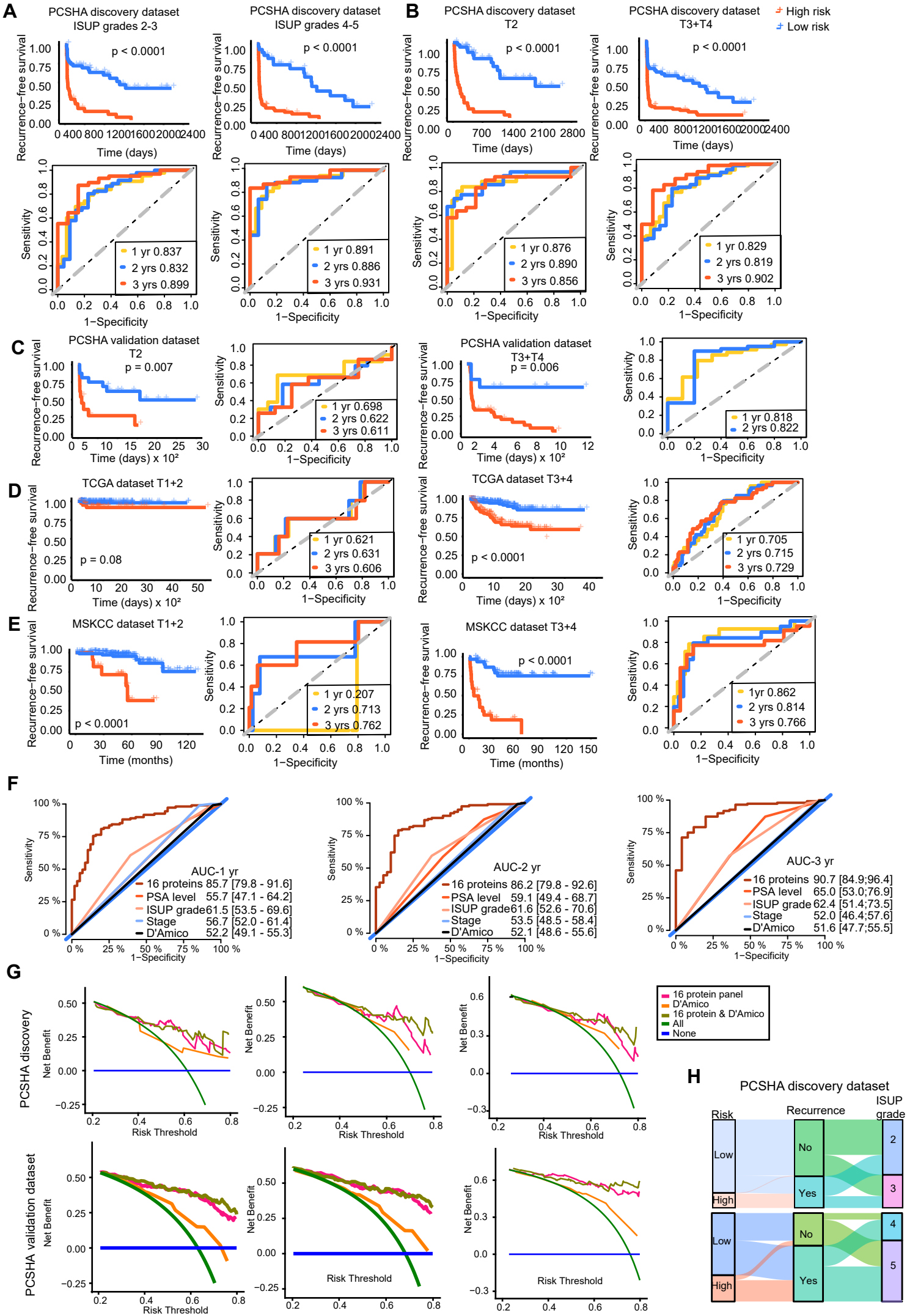
A
PCSHA discovery dataset
ISUP grades 2-3
p < 0.0001
Recurrence−free survival
Time (days)
Sensitivity
1−Specificity
1 yr 0.837
2 yrs 0.832
3 yrs 0.899
PCSHA discovery dataset
ISUP grades 4-5
p < 0.0001
1 yr 0.891
2 yrs 0.886
3 yrs 0.931
B
PCSHA discovery dataset
T2
p < 0.0001
1 yr 0.876
2 yrs 0.890
3 yrs 0.856
PCSHA discovery dataset
T3+T4
p < 0.0001
1 yr 0.829
2 yrs 0.819
3 yrs 0.902
High risk
Low risk
C
PCSHA validation dataset
T2
p = 0.007
Time (days) x 10²
1 yr 0.698
2 yrs 0.622
3 yrs 0.611
PCSHA validation dataset
T3+T4
p = 0.006
1 yr 0.818
2 yrs 0.822
D
TCGA dataset T1+2
p = 0.08
1 yr 0.621
2 yrs 0.631
3 yrs 0.606
TCGA dataset T3+4
p < 0.0001
1 yr 0.705
2 yrs 0.715
3 yrs 0.729
E
MSKCC dataset T1+2
p < 0.0001
Time (months)
1 yr 0.207
2 yrs 0.713
3 yrs 0.762
MSKCC dataset T3+4
p < 0.0001
1yr 0.862
2 yrs 0.814
3 yrs 0.766
F
Sensitivity
1−Specificity
AUC-1 yr
16 proteins 85.7 [79.8 - 91.6]
PSA level 55.7 [47.1 - 64.2]
ISUP grade 61.5 [53.5 - 69.6]
Stage 56.7 [52.0 - 61.4]
D'Amico 52.2 [49.1 - 55.3]
AUC-2 yr
16 proteins 86.2 [79.8 - 92.6]
PSA level 59.1 [49.4 - 68.7]
ISUP grade 61.6 [52.6 - 70.6]
Stage 53.5 [48.5 - 58.4]
D'Amico 52.1 [48.6 - 55.6]
AUC-3 yr
16 proteins 90.7 [84.9;96.4]
PSA level 65.0 [53.0;76.9]
ISUP grade 62.4 [51.4;73.5]
Stage 52.0 [46.4;57.6]
D'Amico 51.6 [47.7;55.5]
G
PCSHA discovery
PCSHA validation dataset
Net Benefit
Risk Threshold
16 protein panel
D'Amico
16 protein & D'Amico
All
None
H
PCSHA discovery dataset
Risk
Recurrence
ISUP grade
Low
High
No
Yes
2
3
4
5

Figure S5 16-protein panel to prediction BCR in different subgroups. Related to Figures 4-5.
(A-B) Kaplan–Meier plots and ROC curves of BCR-free survival based on our 16-protein model in different ISUP grades (A) and pathological stage (B) using the PCSHA discovery dataset with a 1-year predictive power (yellow), a 2-year predictive power (blue), and a 3-year predictive power (red).
(C-E) Kaplan–Meier plots and ROC curves of BCR-free survival based on our 16-protein prediction model in different pathological stage using the PCSHA validation set (C), the TCGA dataset (D), and MSKCC dataset (E).
(F)ROC curves of the 16-protein panel and clinicopathological characteristics (PSA level, Gleason score, pathology stage, and D'Amico) using the PCSHA discovery dataset at 1 (upper panel), 2 (middle panel), and 3 (lower panel) years.
(G)Decision curve analysis (DCA) showing the clinical judgement of the benefits in the PCSHA discovery and validation sets.
(H)Sankey plots showing the PCa patients overlapping among the 16-protein panel prediction system, the recurrence status, and the ISUP grade using the PCSHA discovery set.

Figure S6

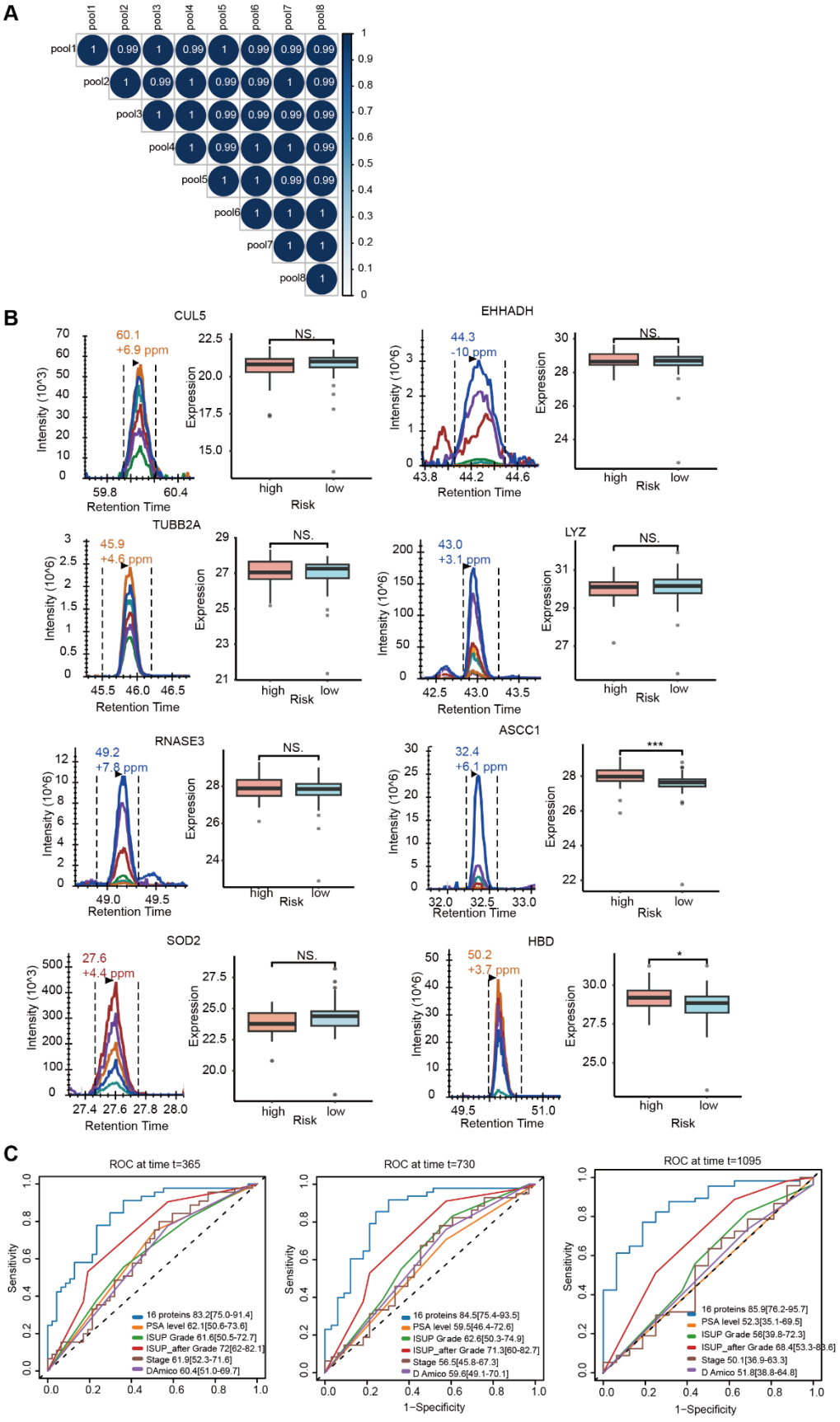
A
pool1
pool2
pool3
pool4
pool5
pool6
pool7
pool8
B
CUL5
EHHADH
TUBB2A
LYZ
RNASE3
ASCC1
SOD2
HBD
Intensity (10^3)
Intensity (10^6)
Retention Time
Expression
Risk
high
low
NS.
C
ROC at time t=365
ROC at time t=730
ROC at time t=1095
Sensitivity
1−Specificity
16 proteins 83.2[75.0-91.4]
PSA level 62.1[50.6-73.6]
ISUP Grade 61.6[50.5-72.7]
ISUP_after Grade 72[62-82.1]
Stage 61.9[52.3-71.6]
DAmico 60.4[51.0-69.7]
16 proteins 84.5[75.4-93.5]
PSA level 59.5[46.4-72.6]
ISUP Grade 62.6[50.3-74.9]
ISUP_after Grade 71.3[60-82.7]
Stage 56.5[45.8-67.3]
D Amico 59.6[49.1-70.1]
16 proteins 85.9[76.2-95.7]
PSA level 52.3[35.1-69.5]
ISUP Grade 56[39.8-72.3]
ISUP_after Grade 68.4[53.3-83.6]
Stage 50.1[36.9-63.3]
D Amico 51.8[38.8-64.8]

Figure S6. Independent biopsy samples for validation. Related to Figure 6.
(A) Correlation matrix for quality control samples, including eight pool samples for QC.
(B) The peak groups for the unique peptides and the protein abundance between high- and low-risk patients.
(C) ROC curves of the 16-protein model and clinicopathological characteristics for BCR prediction (PSA level, ISUP grade based on biopsy and surgical samples, pathological stage, and D'Amico) in the PCSHA biopsy test set with three years.

Figure S7

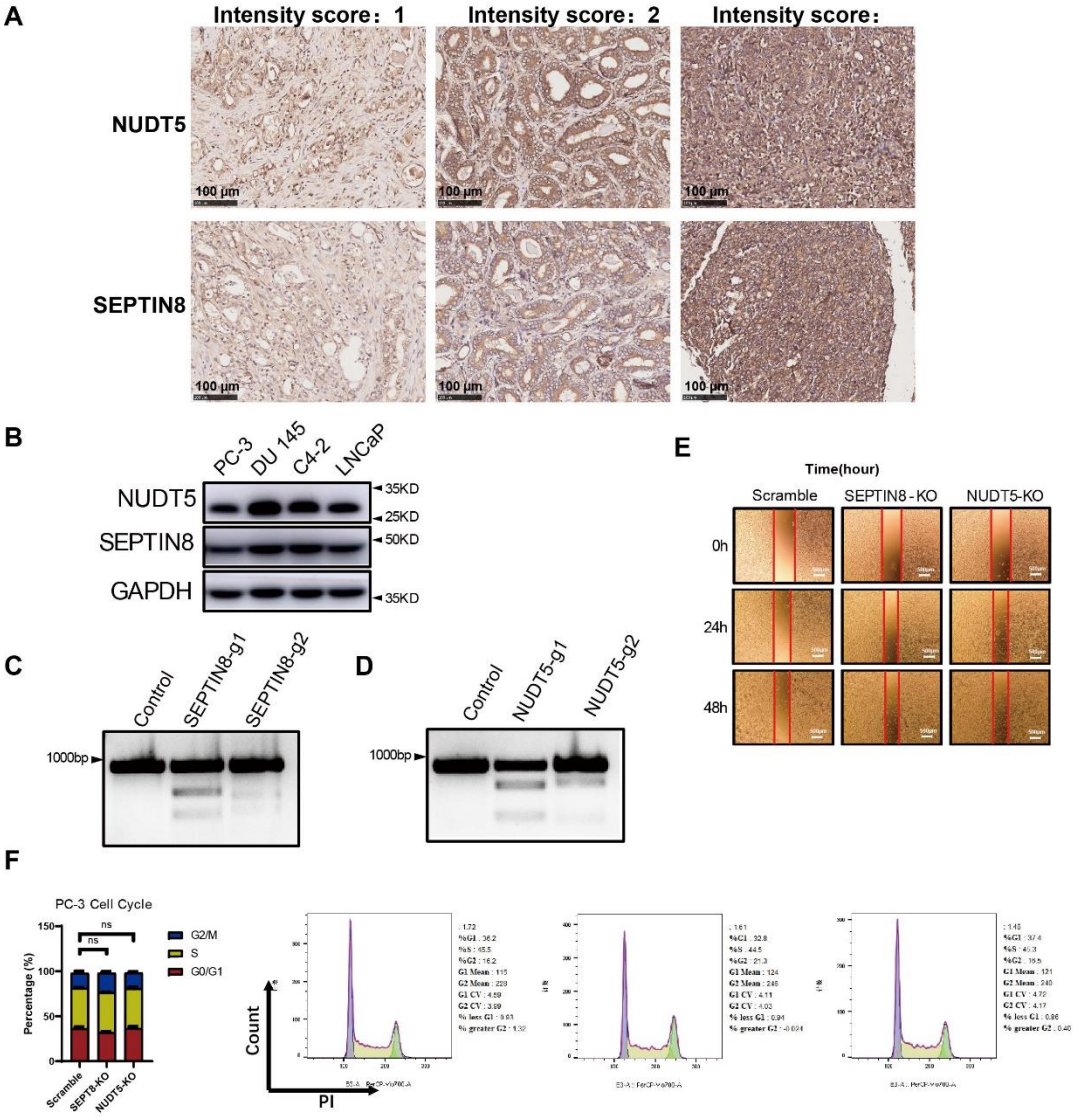

Figure S7. Validation of NUDT5 and SEPTIN8 single knockout. Related to Figure 7.
(A) Representative Immunohistochemical staining intensity score showing the protein expression of SEPTIN8 and NUDT5 in different PCa tissue samples.
(B) Expression level of NUDT5 and SEPTIN8 in different prostate cancer cell lines measured by western blot.
(C-D) Surveyor assay showing the deletion efficiency of SEPTIN8 (C) and NUDT5 (D) genes by the CRISPR/Cas9 system.
(E-F) Cell functional assay of PC-3 cells with SEPTIN8 or NUDT5 knockout: wound healing assay (E) and cell cycle analysis (F).