Supplementary

Table S1 Clinical characteristics of the patients in GSE21034, GSE70770 and GSE116918 datasets

Characteristics	All patients	Patients with biochemical recurrence	Patients without biochemical recurrence	P value
Clinicopathological characteristics of the GSE21034				
Patients, n (%)	131 (100.0)	32 (24.4)	99 (75.6)	
PSA, n (%)				<0.0001
<10 ng/mL	100 (66.4)	15 (48.9)	83 (83.8)	
≥10 ng/mL	31 (33.6)	17 (51.1)	16 (16.2)	
AJCC T staging, n (%)				0.001
T2	85 (36.5)	13 (40.6)	72 (72.7)	
T3T4	46 (63.5)	19 (59.4)	27 (27.3)	
Gleason score, n (%)				<0.0001
≤7	106 (80.9)	20 (62.5)	96 (97.0)	
≥8	15 (19.1)	12 (37.5)	3 (3.0)	
Follow-up after surgery				
Median months (range)	47.5 (1.4–147.1)	30.5 (1.4–91.9)	53.0 (1.8–147.1)	<0.0001
Clinicopathological characteristics of the GSE70770				
Patients, n (%)	193 (100)	59 (30.1)	134 (69.4)	
PSA, n (%)				0.26
<10 ng/mL	132 (68.4)	37 (62.7)	95 (70.9)	
≥10 ng/mL	61 (31.6)	22 (27.3)	39 (29.1)	
AJCC T staging, n (%)				0.001
T2	78 (40.4)	15 (25.4)	71 (53.0)	
T3T4	115 (59.6)	44 (74.6)	63 (47.0)	
Gleason score, n (%)				<0.0001
≤7	173 (89.6)	46 (78.0)	127 (94.8)	
≥8	20 (10.4)	13 (22.0)	7 (5.2)	
Follow-up after surgery				
Median months (range)	54.1 (1.8–115.5)	67.6 (2.0–115.5)	48.2 (1.8–103.4)	<0.0001
Clinicopathological characteristics of the GSE116918				
Patients, n (%)	248 (100)	56 (22.6)	192 (77.4)	
PSA, n (%)				0.213
<10 ng/mL	50 (20.2)	8 (14.3)	42 (21.9)	
≥10 ng/mL	198 (79.8)	48 (85.7	150 (78.1)	
AJCC T staging, n (%)				0.005
T2	152 (61.3)	25 (44.6)	127 (66.1)	
T3T4	96 (38.7)	31 (55.4)	65 (38.9)	
Gleason score, n (%)				<0.0001
≤7	141 (56.9)	28 (50.0)	113 (58.9)	
≥8	107 (43.1)	28 (50.0)	79 (41.1)	
Follow-up after radical radiation therapy				
Median months (range)	78.3 (10.0–117)	58.8 (10.0–106.0)	83.9 (11.0–117)	<0.0001

PSA, prostate-specific antigen.

http://dx.doi.org/10.21037/tau-20-1019

Table S2 Details of the GEO datasets included in this study

Dataset ID	Prostate cancer sample size	Platform
GSE21034	131	[HuEx-1_0-st] Affymetrix Human Exon 1.0 ST Array [probe set (exon) version]
GSE70770	193	Illumina HumanHT-12 V4.0 expression beadchip
GSE116918	248	[ADXPCv1a520642] Almac Diagnostics Prostate Disease Specific Array (DSA)

GEO, the gene expression omnibus.

Table S3 Univariate Cox regression analysis of 9 mRNAs and BCR-free survival of PCa patients in four datasets

Genes	TCGA		GSE21034		GSE70770		GSE116918	
	HR (95% CI)	P value	HR (95% CI)	P value	HR (95% CI)	P value	HR (95% CI)	P value
ASPM	1.50 (1.22–1.84)	<0.0001	6.18 (2.11–18.14)	<0.0001	5.03 (2.60–9.74)	<0.0001	2.50 (1.51–4.12)	<0.0001
MMP11	1.43 (1.24–1.67)	<0.0001	3.31 (1.13–9.70)	0.02	1.48 (1.11–1.96)	0.007	4.10 (2.04–8.27)	<0.0001
CENPF	1.47 (1.23–1.75)	<0.0001	4.82 (2.28–10.21)	<0.0001	4.61 (2.50–8.49)	<0.0001	1.82 (1.89–2.79)	0.005
KIF4A	1.65 (1.35–2.00)	<0.0001	5.79 (1.68–20.00)	0.005	5.54 (2.50–12.30)	<0.0001	1.50 (1.06–2.11)	0.02
COL1A1	1.44 (1.20–1.72)	<0.0001	2.55 (1.57–4.12)	<0.0001	1.31 (1.04–1.67)	0.02	3.39 (1.84–6.26)	<0.0001
ASPN	1.44 (1.19–1.74)	<0.0001	2.91 (1.92–4.40)	<0.0001	6.211.05–36.76)	0.04	1.63 (1.24–2.14)	<0.0001
FANCI	1.78 (1.30–2.52)	0.001	3.77 (1.29–10.99)	0.01	9.90 (2.82–34.78)	0.0003	2.68 (1.07–6.74)	0.03
C9orf152	0.57 (0.44–0.75)	<0.0001	0.49 (0.30–0.80)	0.004	0.38 (0.24–0.60)	<0.0001	0.72 (0.53–1.00)	0.04
EPHX2	0.56 (0.43–0.72)	<0.0001	0.39 (0.23–0.66)	<0.0001	0.47 (0.30–0.75)	<0.0001	0.47 (0.25–0.87)	0.01

HR, hazard ratio; CI, confidence interval; C9orf152, chromosome 9 open reading frame 152; EPHX2, epoxide hydrolase 2; ASPM, assembly factor for spindle microtubules; MMP11, matrix metallopeptidase 11; CENPF, centromere protein F; KIF4A, kinesin family member 4A; COL1A1, collagen type I alpha 1 chain; ASPN, Aspirin; FANCI, FA complementation group I.

Table S4 Comparison of C-index between the 9-mRNA signature and D'Amico model

Dataset ID	D'Amico model (Reference models)	9-mRNA signature	P value
	C-index (95% CI)	C-index (95% CI)	
TCGA	0.71 (0.65–0.78)	0.72 (0.66–0.78)	0.68
GSE21034	0.84 (0.76–0.91)	0.79 (0.71–0.87)	0.86
GSE70770	0.72 (0.64–0.80)	0.74 (0.67–0.80)	0.32
GSE116918	0.64 (0.56–0.73)	0.71 (0.64–0.78)	0.066

C-index, concordance index; TCGA, The Cancer Genome Atlas.

Table S5 Comparison of C-index between the nomogram and D'Amico model

Dataset ID	D'Amico model (Reference models), C-index (95% CI)	Nomogram, C-index (95% CI)	P value
TCGA	0.71 (0.65–0.78)	0.74 (0.68–0.80)	0.68
GSE21034	0.84 (0.76–0.91)	0.85 (0.78–0.91)	0.20
GSE70770	0.72 (0.64–0.80)	0.76 (0.70–0.83)	0.055
GSE116918	0.64 (0.56–0.73)	0.71 (0.64–0.79)	0.026

C-index, concordance index; TCGA, The Cancer Genome Atlas.

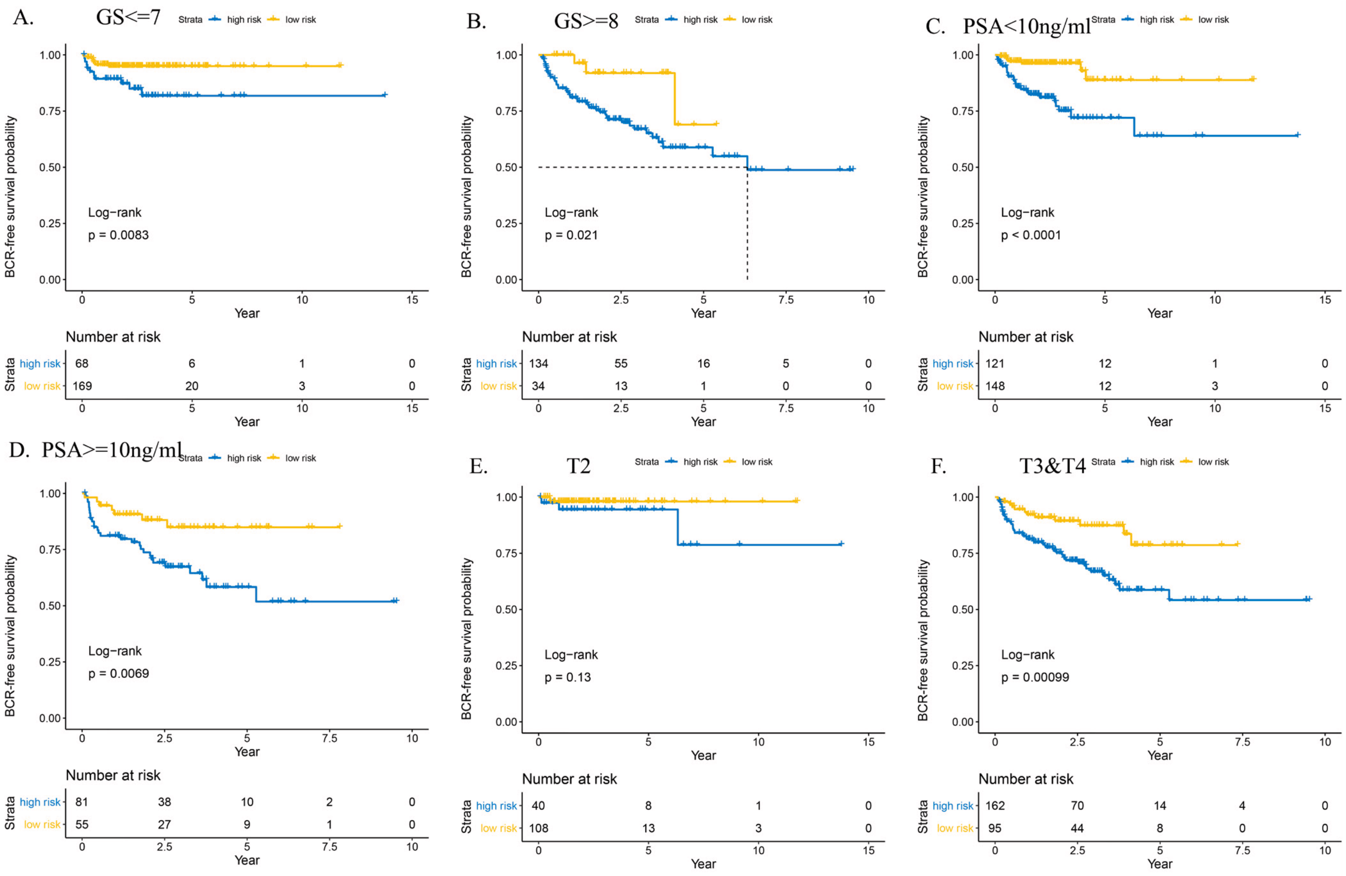

Figure S1 Kaplan-Meier survival analysis for PCa patients according to the 9-mRNA signature stratified by clinicopathological variables in TCGA dataset. (A, B) Gleason score (GS); (C, D) PSA; (E, F) AJCC T staging. BCR, biochemical recurrence; PSA, prostate-specific antigen; GS, Gleason score; TCGA, The Cancer Genome Atlas.

http://dx.doi.org/10.21037/tau-20-1019