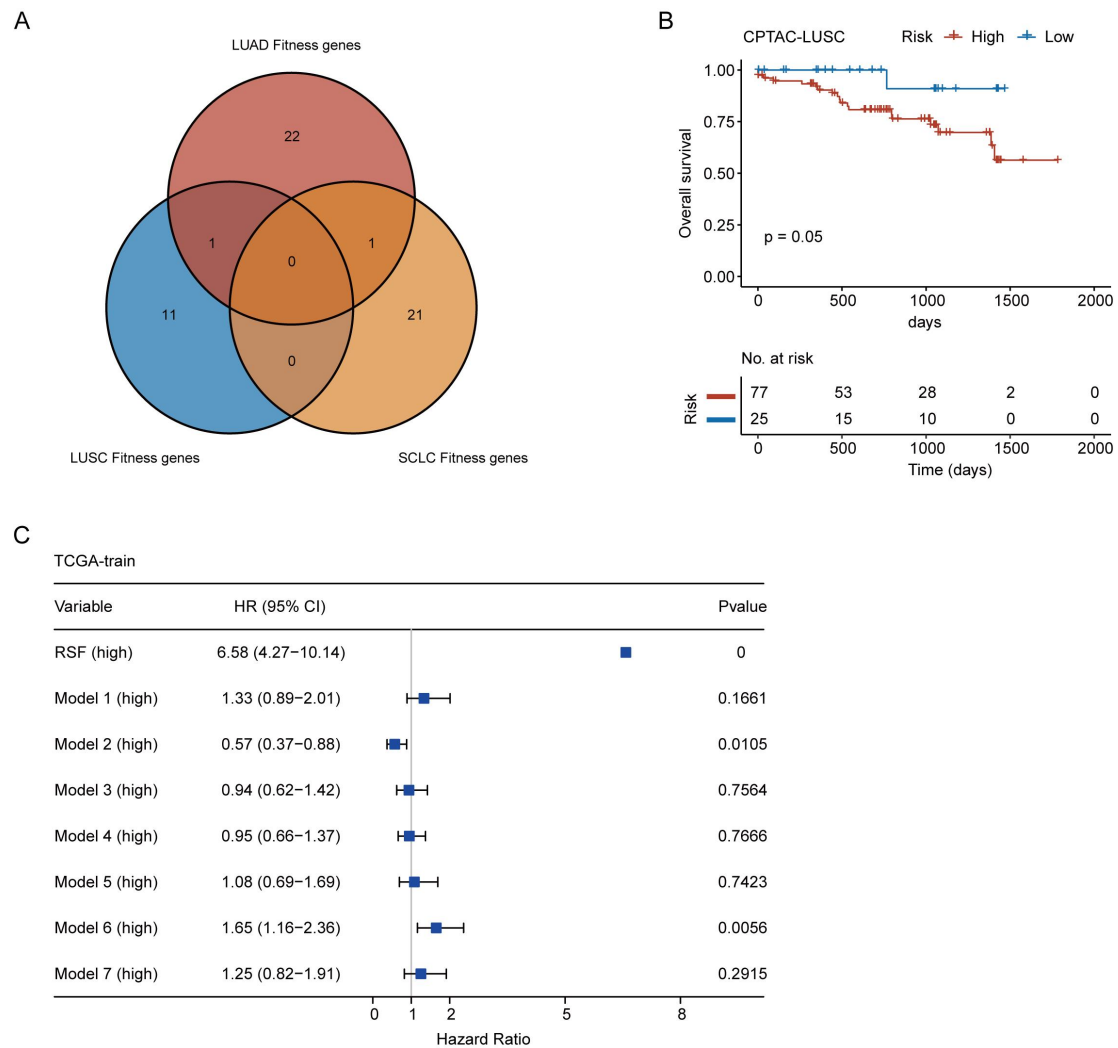
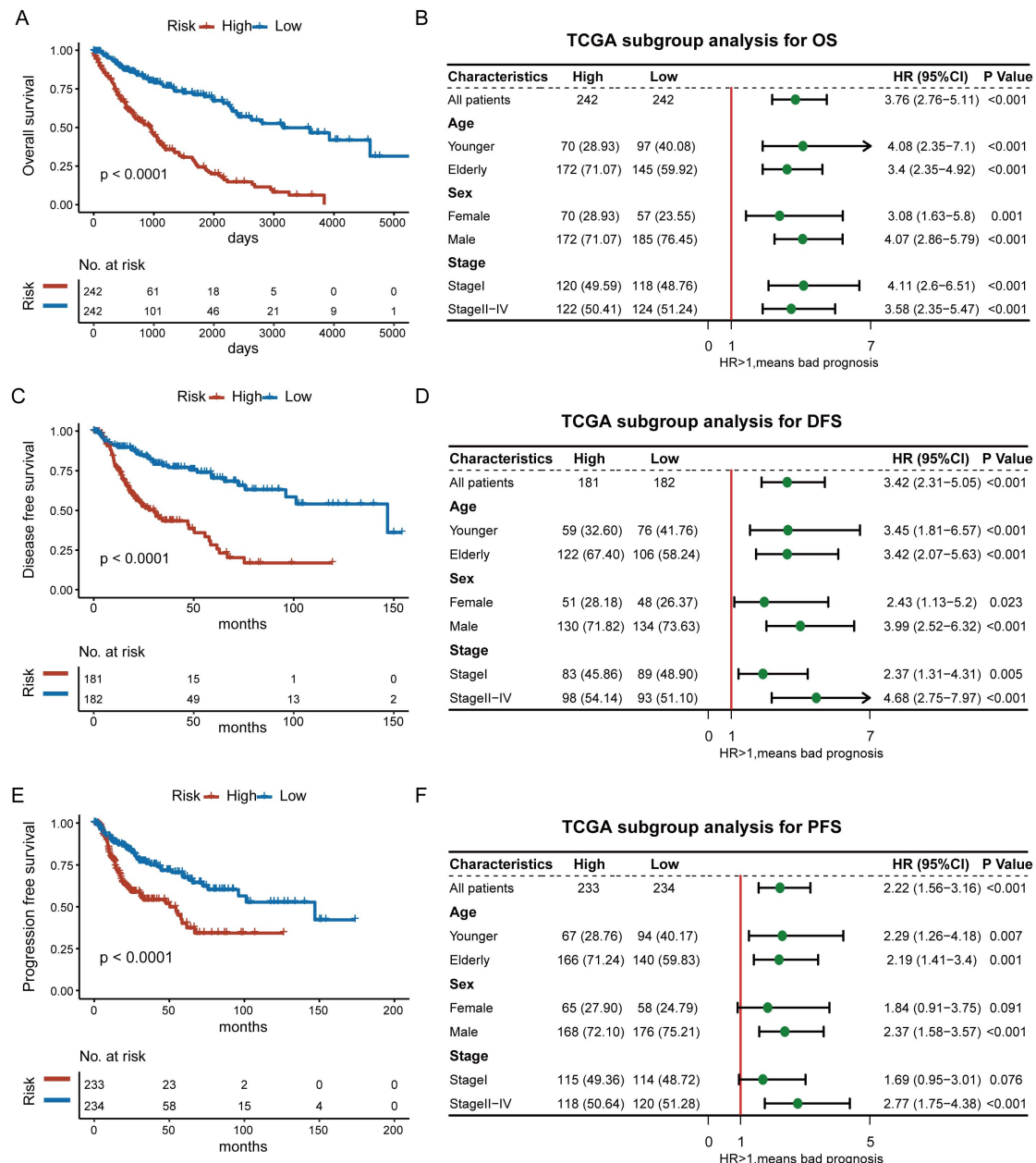


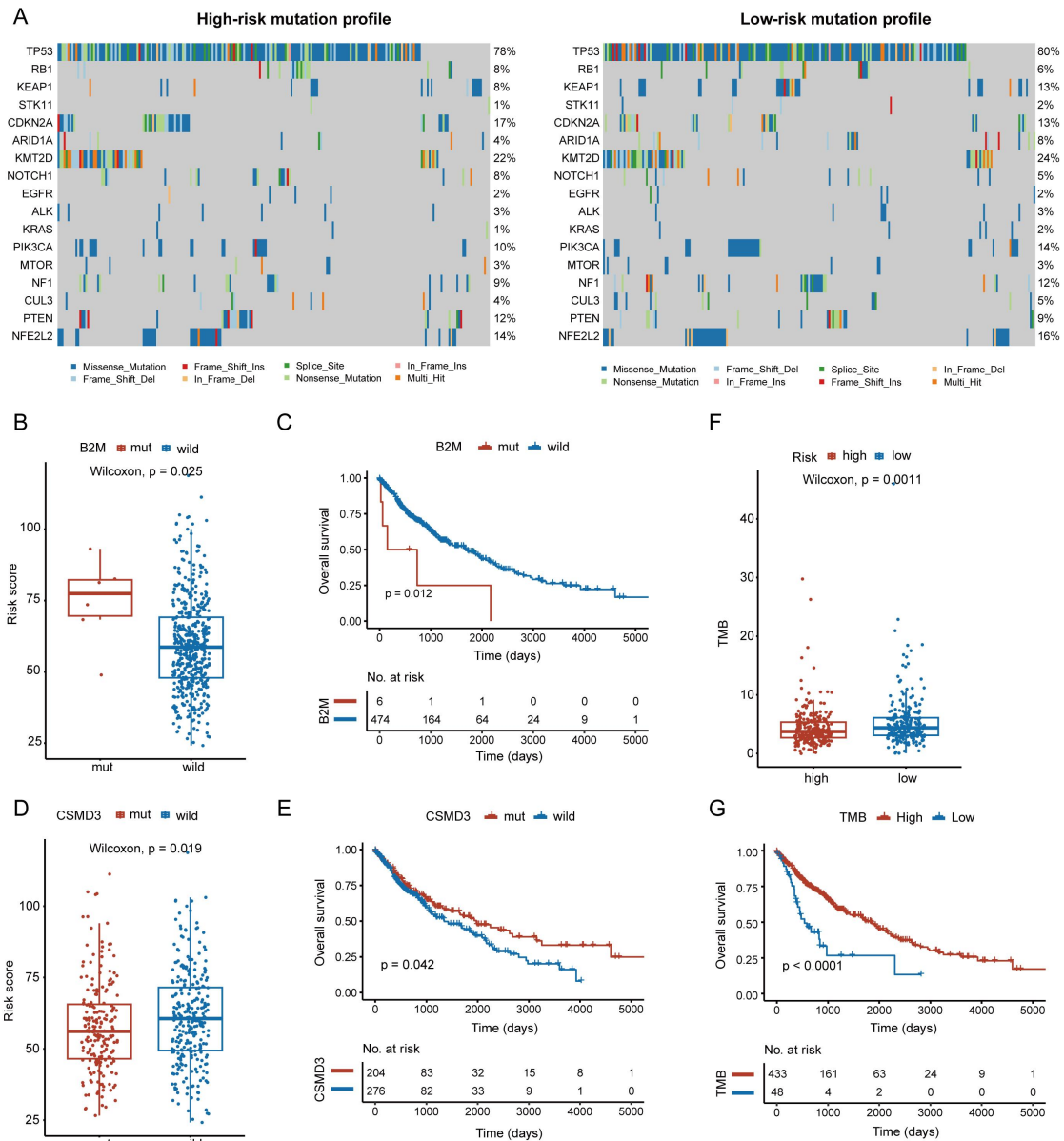
Supplementary Figures



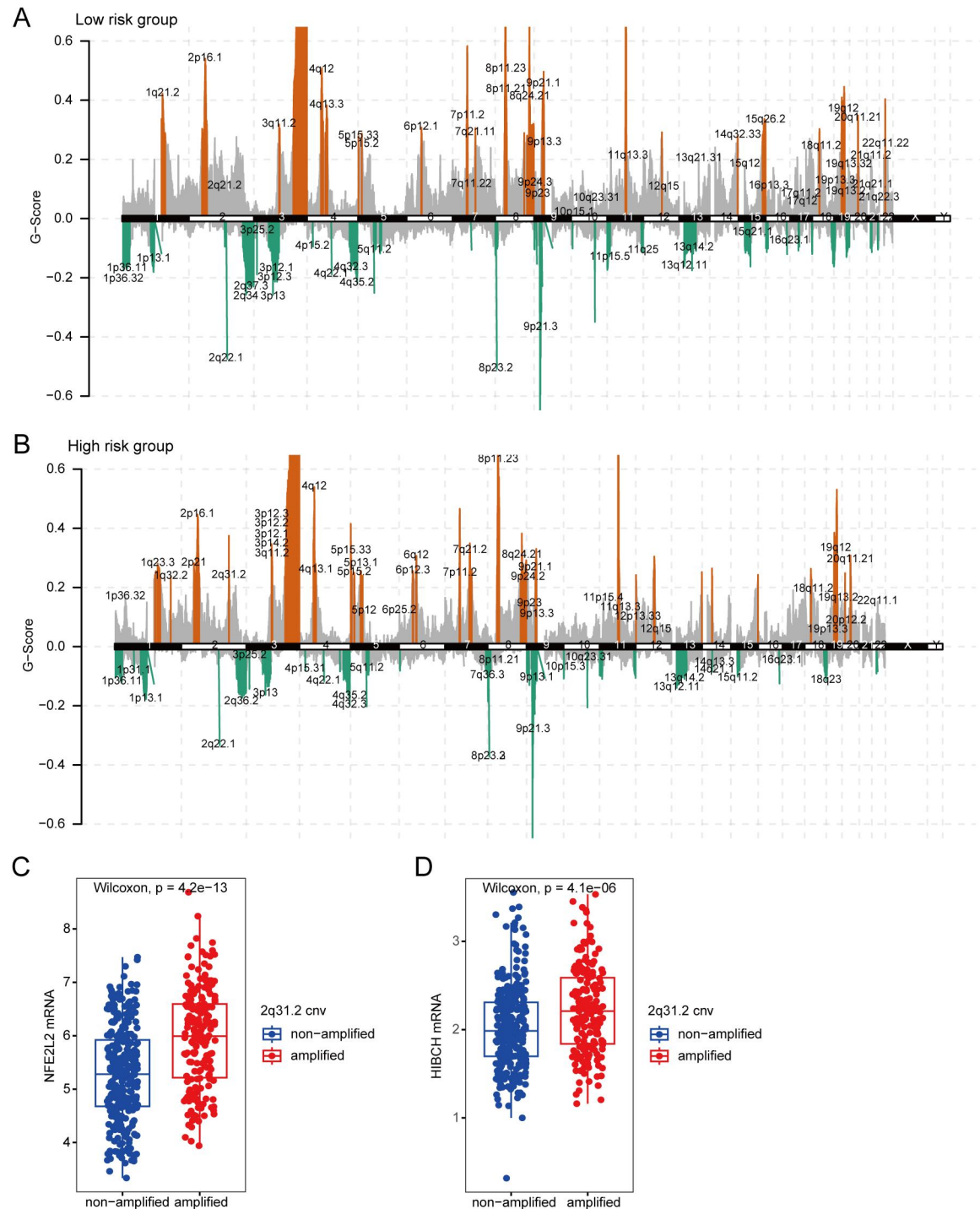
Supplementary Figure 1. Specificity of fitness genes in LUSC and RSF model validation. A. Venn diagram of the fitness genes of LUSC, LUAD, and SCLC. B. Kaplan-Meier estimates of overall survival according to the RSF risk score in the CPTAC-LUSC cohort (n=102) (Log-rank test). C. Multivariate Cox regression analysis for the RSF and seven prognostic models in the training cohort from the TCGA dataset (n=345).



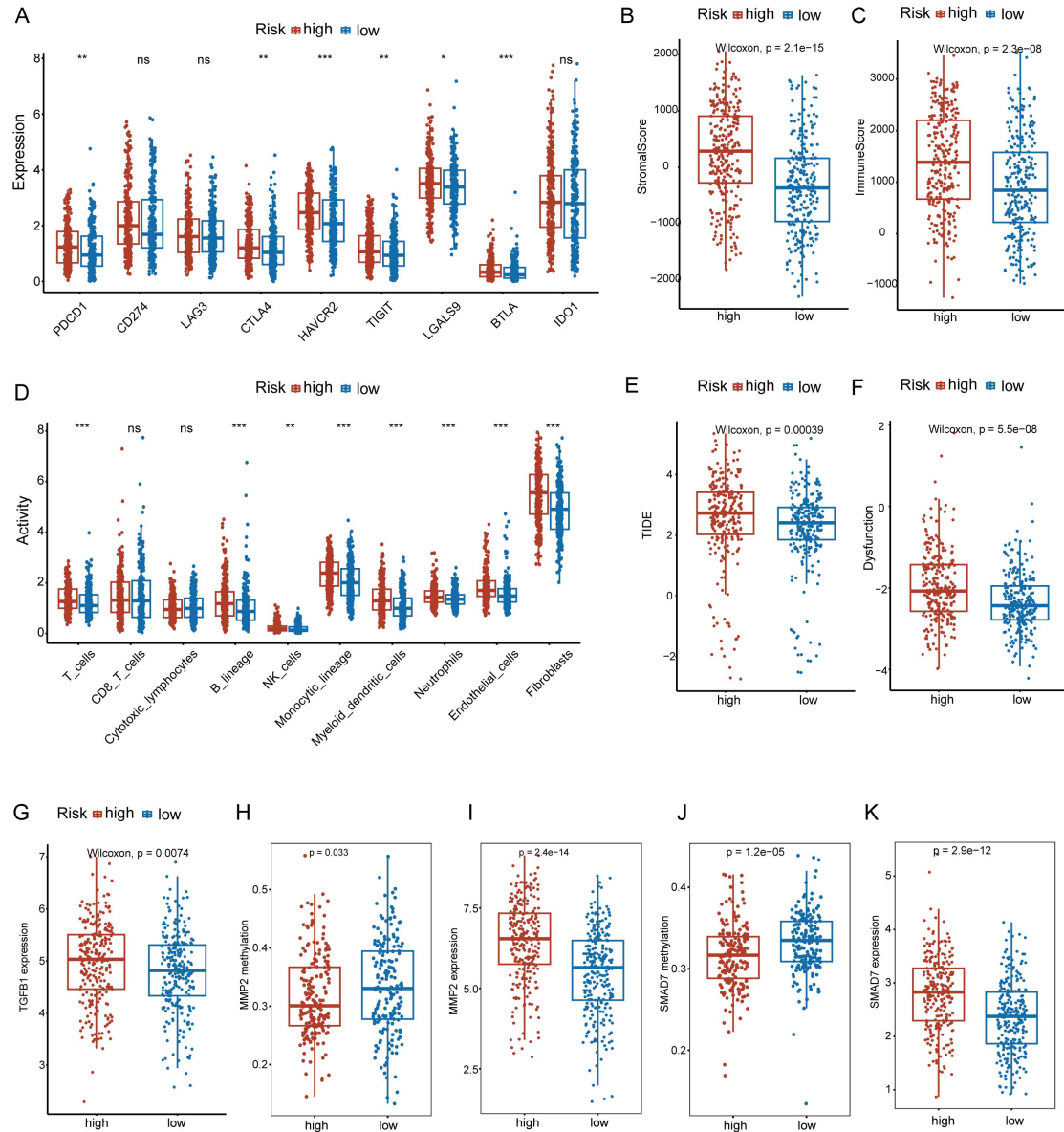
Supplementary Figure 2. Subgroup prognosis analysis for the RSF risk score in the overall TCGA-LUSC dataset. A. Kaplan-Meier estimates of OS according to the RSF risk score in the overall TCGA dataset (n=484) (Log-rank test). B. Subgroup analysis for OS in the overall TCGA dataset (n=484). C. Kaplan-Meier estimates of DFS according to the RSF risk score in the overall TCGA dataset (n=363) (Log-rank test). D. Subgroup analysis for DFS in the overall TCGA dataset (n=363). E. Kaplan-Meier estimates of PFS according to the RSF risk score in the overall TCGA dataset (n=467) (Log-rank test). F. Subgroup analysis for PFS in the overall TCGA dataset (n=467).



Supplementary Figure 3. Analysis of mutational profiles in the high-risk group and low-risk group in the TCGA-LUSC dataset. A. Oncoplot displaying the somatic landscape of the high-risk and low-risk groups (high=244, low=237). B. *B2M* mutation was associated with a higher risk score (mut=6, wild=474). C. *B2M* mutation was associated with worse OS (Log-rank test). D. *CSMD3* mutation was associated with a lower risk score (mut=204, wild=276). E. *B2M* mutation was associated with good OS (Log-rank test). F. The high-risk group exhibited a lower TMB (high=244, low=237). G. A high TMB was associated with good OS (high=433, low=48) (Log-rank test).

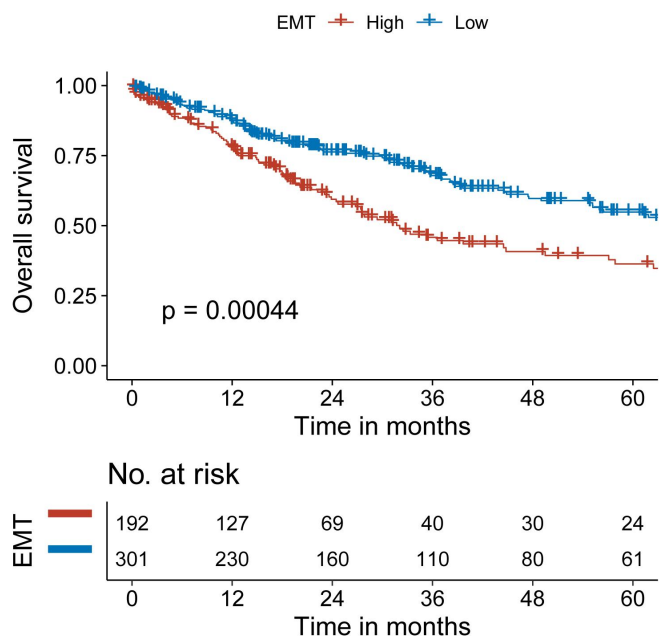


Supplementary Figure 4. Analysis of copy number variations in the low-risk group and high-risk group in the TCGA-LUSC dataset. A. Significantly altered CNV regions in the low-risk group (n=246). B. Significantly altered CNV regions in the high-risk group (n=246). C. Increased NFE2L2 mRNA expression in LUSC tumours with 2q31.2 amplifications in the TCGA-LUSC cohort (n=500). D. Increased HIBCH mRNA expression in LUSC tumours with 2q31.2 amplifications in the TCGA-LUSC cohort (n=500).

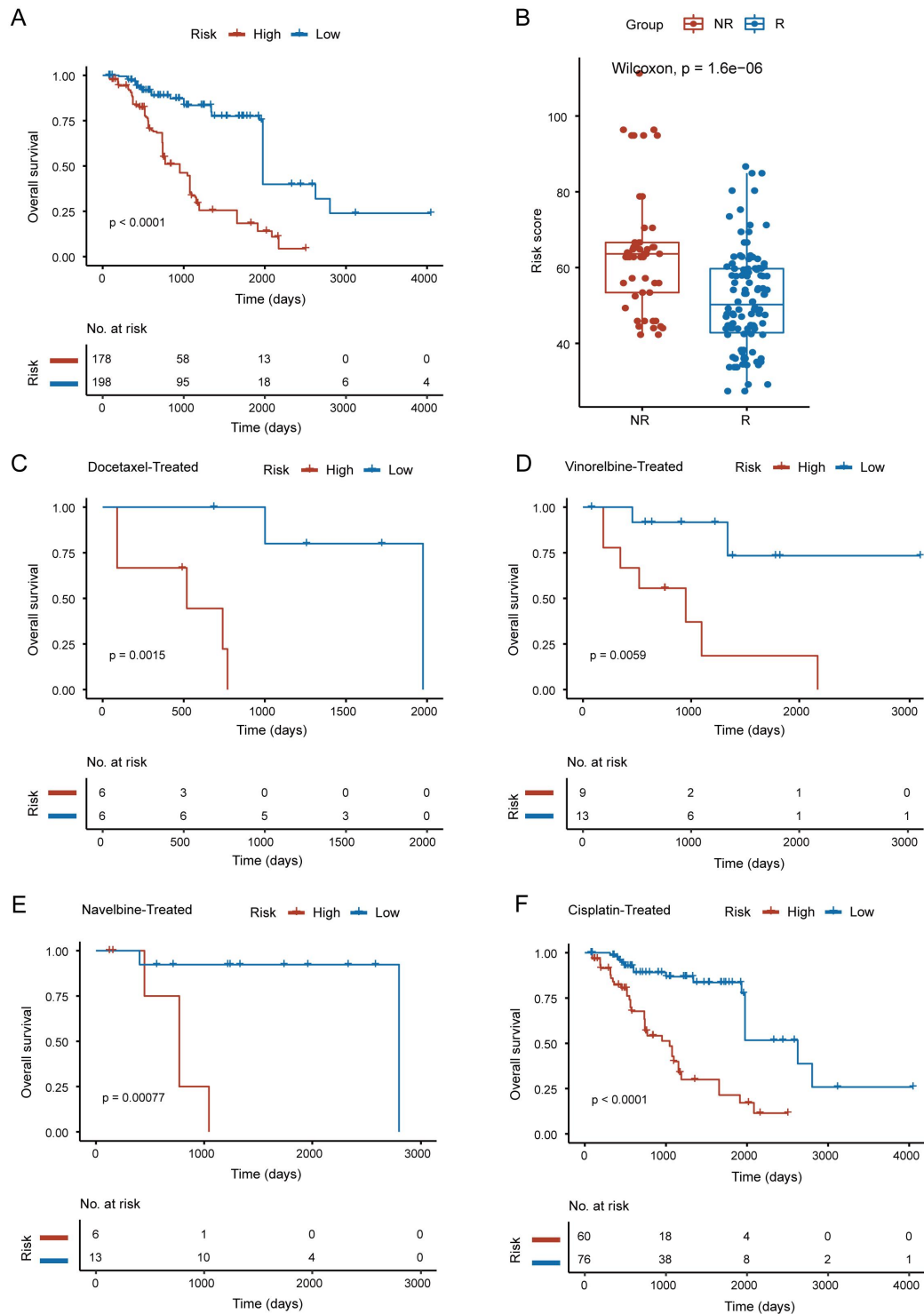


Supplementary Figure 5. TME analysis for the RSF model in the TCGA-LUSC dataset (high=246, low=247). A. Immune checkpoint differences between the high- and low-risk groups (ns, not significant; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$, Wilcoxon Test). B. Immune score differences between the high- and low-risk groups. C. Stromal score differences between the high- and low-risk groups. D. Immune cell differences between the high- and low-risk groups (ns, not significant; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$, Wilcoxon Test). E. TIDE score differences between the high- and low-risk groups. F. Dysfunction score differences between the high- and low-risk groups. G. *TGFBI* expression differences between the high- and low-risk groups. H. Boxplot displaying the difference in the methylation levels of *MMP2* between the high- and low-risk groups (high=188, low=174). I. Boxplot displaying the differences in the mRNA expression of *MMP2* between the high- and low-risk groups. J. Boxplot displaying the differences in the methylation levels of *SMAD7* between the high- and low-risk groups (high=188, low=174). K. Boxplot displaying the differences in *SMAD7* mRNA expression

between the high- and low-risk groups.



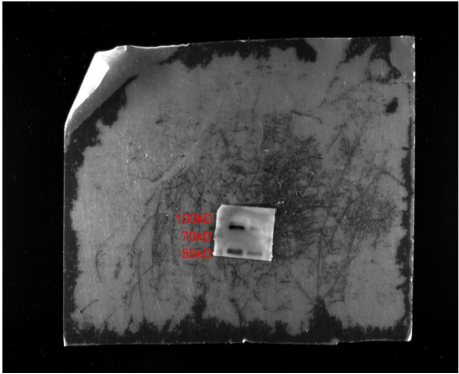
Supplementary Figure 6. Analysis of the prognostic value of the EMT signature in the TCGA-LUSC dataset (n=493) (Log-rank test).



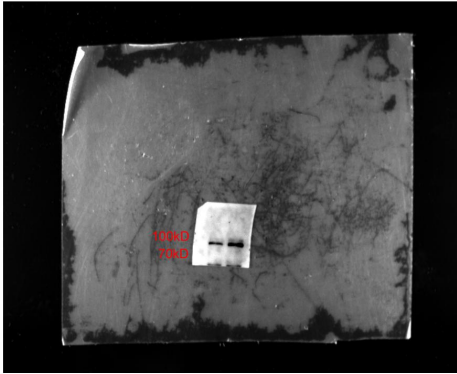
Supplementary Figure 7. The RSF model risk score can effectively predict the response to chemotherapy in the TCGA-LUSC dataset. A. The high-risk group was associated with worse survival in chemotherapy-treated patients (n=376) (Log-rank test). B. Patients who did not respond to chemotherapy exhibited higher RSF risk scores (NR=51, R=110). C-F. Kaplan-Meier estimates of overall survival according to the RSF risk score in patients who received docetaxel (C), vinorelbine (D), navelbine (E), and cisplatin (F) (Log-rank test). NR: nonresponse; R: response.

H520

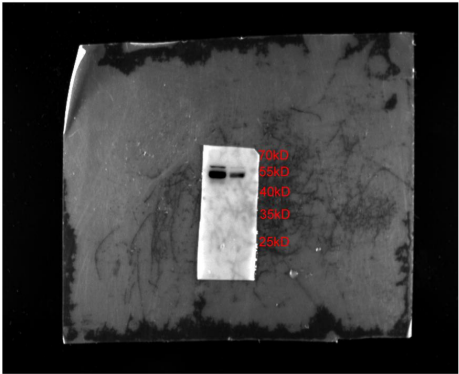
N-Cadherin



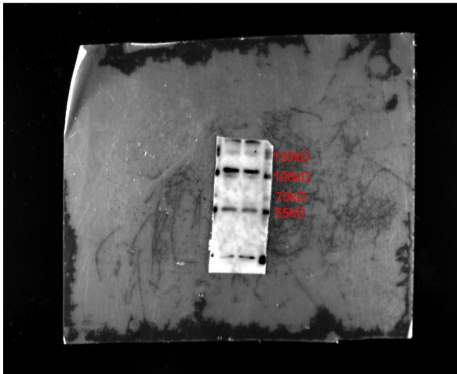
E-Cadherin



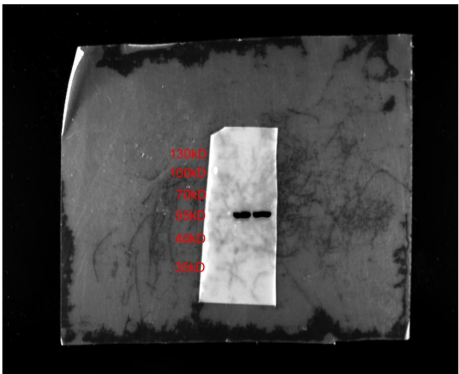
VIM



TGFβR1

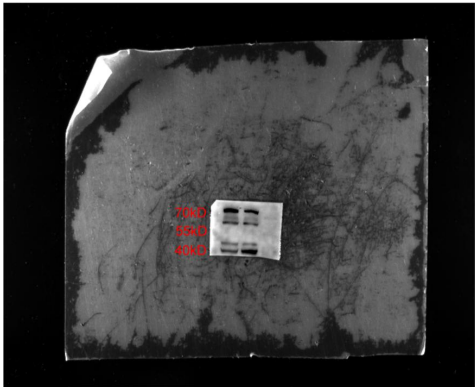


β-actin

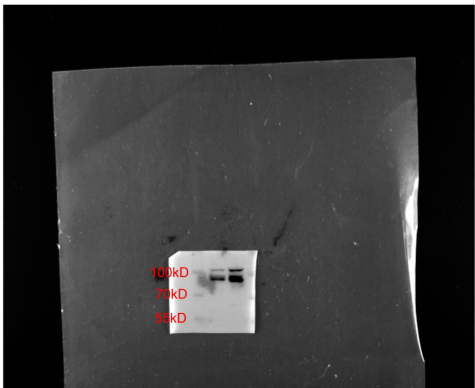


HCC15

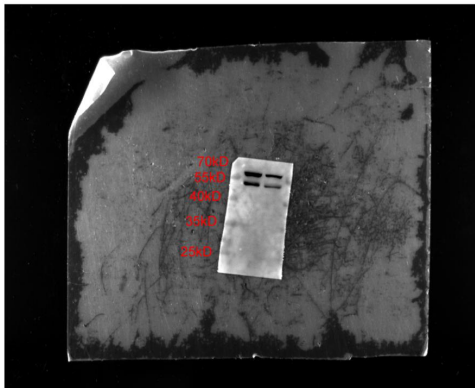
N-Cadherin



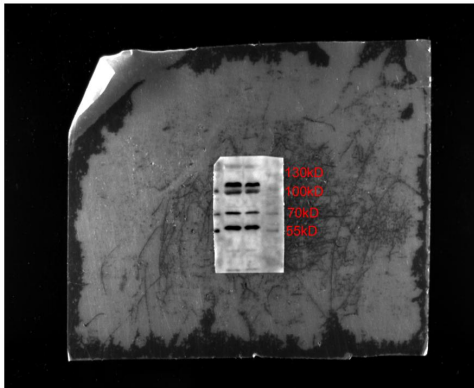
E-Cadherin



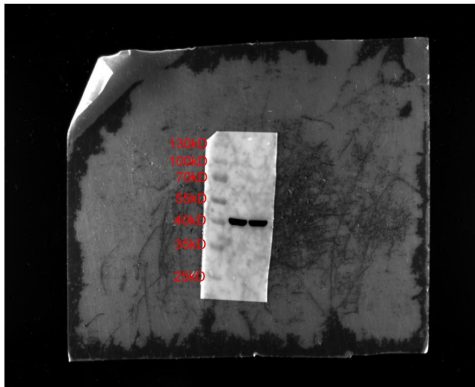
VIM



TGFβR1

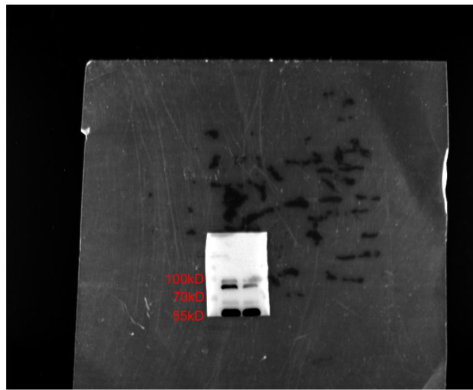


β-actin

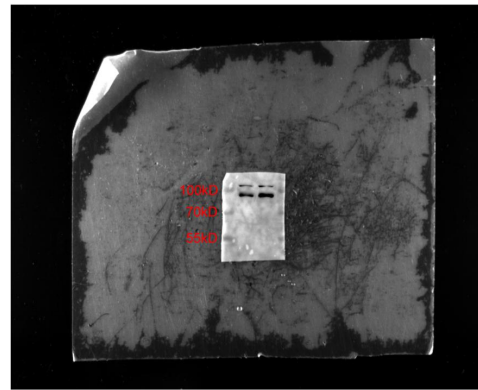


H226

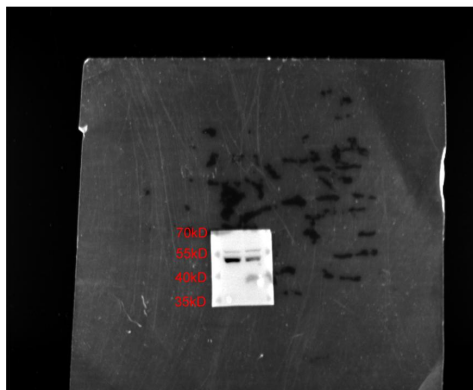
N-Cadherin



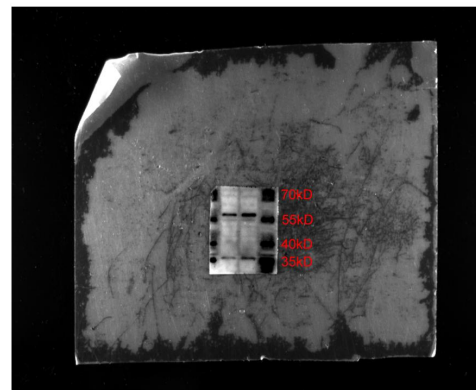
E-Cadherin



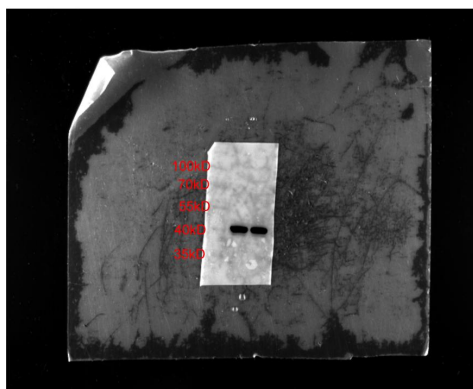
VIM



TGF β R1



β -actin



Supplementary Figure 8. Original Blots for Figure 5E.

Supplementary Tables

Supplementary Table 1. LUSC fitness genes (n=31).

Symbol
BRF2
CDC45
CDC7
ERH
LRR1
LSM7
MED6
NEDD8
PABPN1
PHB2
POLR3K
PSMA3
PSMC6
RPS11
RPS19
RPS29
SCFD1
SNW1
SRSF3
SUPT5H
USP39
LSG1
ARIH1
FARSA
SNRNP25
SS18L2
C1orf131
SNRPC
BCAS2
FTSJ3
PCBP2

Supplementary Table 2. 12 core fitness genes in the RSF model.

Symbol
SUPT5H
ARIH1
SNRPC
PHB2
PABPN1
FARSA

MED6
ERH
SNW1
RPS29
BCAS2
LSM7

Supplementary Table 3. Multivariate cox regression analysis in TCGA-train cohort.

	HR	95% CI	P
Risk (high)	5.84	3.89 - 8.78	0
Age	1.01	0.99 - 1.04	0.1844
Sex (male)	1.19	0.79 - 1.77	0.4052
Stage			
I	Reference		
II	1.11	0.76 - 1.62	0.5917
III	1.55	0.99 - 2.44	0.0577
IV	2.38	0.85 - 6.62	0.0979

Supplementary Table 4. Multivariate cox regression analysis in TCGA-test cohort.

	HR	95% CI	P
Risk (high)	1.78	1.04 - 3.06	0.0349
Age	1.00	0.97 - 1.03	0.9643
Sex (male)	1.43	0.79 - 2.57	0.2329
Stage			
I	Reference		
II	1.11	0.59 - 2.09	0.7439
III	1.37	0.71 - 2.67	0.3498
IV	3.24	0.42 - 24.70	0.2569

Supplementary Table 5. Multivariate cox regression analysis in GSE157011 cohort.

	HR	95% CI	P
Risk (high)	1.31	1.02 - 1.68	0.0316
Age	1.02	1 - 1.03	0.0099
Sex (male)	1.28	0.98 - 1.66	0.0667
Stage			
T1	Reference		
T2	1.46	1.11 - 1.92	0.0064
T3	2.52	1.7 - 3.73	0

Supplementary Table 6. The RSF risk score is independent of unbiased

prognostic signature in LUSC in the TCGA training dataset.

	Univariate Analysis			Multivariate Analysis		
	HR	95% CI	P	HR	95% CI	P
RSF risk score (high)	6.06	4.05 - 9.06	<2e-16	4.87	3.23 - 7.37	5.60e-14
Unbiased prognostic signature (high)	3.20	2.26 - 4.55	7.36e-11	2.18	1.52 - 3.12	2.35e-05

Supplementary Table 7. The RSF risk score is independent of unbiased prognostic signature in LUSC in the TCGA test dataset.

	Univariate Analysis			Multivariate Analysis		
	HR	95% CI	P	HR	95% CI	P
RSF risk score (high)	1.69	1.03 - 2.78	0.039	1.69	1.01 - 2.83	0.047
Unbiased prognostic signature (high)	1.15	0.70 - 1.89	0.57	1.00	0.60 - 1.67	0.998

Supplementary Table 8. The RSF risk score is independent of unbiased prognostic signature in LUSC in the GSE157011 dataset.

	Univariate Analysis			Multivariate Analysis		
	HR	95% CI	P	HR	95% CI	P
RSF risk score (high)	1.29	1.01 - 1.65	0.039	1.29	1.01 - 1.65	0.041
Unbiased prognostic signature (high)	1.03	0.81 - 1.31	0.81	1.01	0.79 - 1.29	0.948

Supplementary Table 9. The RSF risk score is independent of genetic abnormalities for prognosis.

	Univariate Analysis			Multivariate Analysis		
	HR	95% CI	P	HR	95% CI	P
Risk (high)	3.58	2.65 - 4.82	<2e-16	3.58	2.64 - 4.87	4.24e-16
B2M (mut)	2.94	1.21 - 7.16	0.02	1.91	0.77 - 4.71	0.16
CSMD3 (mut)	0.74	0.56 - 0.97	0.03	0.95	0.71 - 1.27	0.73
TMB (high)	0.42	0.28 - 0.62	<0.001	0.41	0.27 - 0.62	3.04e-05
22q11.22 (amp)	0.74	0.57 - 0.97	0.03	0.77	0.59 - 1.02	0.06
2q31.2 (amp)	0.99	0.75 - 1.30	0.93			
7q36.3 (del)	1.07	0.78 - 1.47	0.67			
15q21.1 (del)	1.12	0.84 - 1.46	0.45			

MMP2 meth (high)	0.95	0.69 - 1.30	0.74			
SMAD7 meth (high)	0.91	0.67 - 1.26	0.58			

Supplementary Table 10. LUSC cell lines and RSF risk score.

Cell line	Risk score	Group
LK2	32.58688158	low
NCIH1703	41.61586976	low
NCIH520	49.05845618	low
EBC1	58.26025507	low
HCC95	61.22053461	low
HCC15	69.03711116	high
RERFLCSQ1	71.49550792	high
NCIH226	86.38512239	high
CALU1	91.71089949	high