Supplementary

Table S1 Clinical characteristics of enrolled patients in each dataset

Characteristics	Training cohort	Validation cohorts		
	TCGA	Validation I (GSE30219)	Validation II (GSE31210)	Validation III (GSE37745)
Patients, n	500	83	226	106
Age (years)				
Mean	65.3	61.1	59.6	63.0
Gender				
Male	230	65	105	46
Female	270	18	121	60
Stage				
I-II	387	82	226	89
III-IV	105	1	0	17
NA	8	0	0	0
Survival				
Alive	318	40	191	29
Dead	182	43	35	77

Table S2 Patient characteristics in NCC cohort (N = 142)

Characteristics	Number of patients
Gender	
Male	87
Female	55
Age (years)	
≤60	72
>60	70
T stage	
I	51
II	62
III	16
IV	13
Lymph node metastasis	
No	52
Yes	90
TNM stage	
I	40
II	35
III	67
IV	0

https://dx.doi.org/10.21037/tlcr-21-715

Table S3 Pyroptosis-related genes

Gene Symbol	Ensembl ID	Entrez ID
AIM2	ENSG00000163568	9447
CARD8	ENSG00000105483	22900
CASP1	ENSG00000137752	834
CASP3	ENSG00000164305	836
CASP4	ENSG00000196954	837
CASP5	ENSG00000137757	838
CASP6	ENSG00000138794	839
CASP8	ENSG00000064012	841
DDX3X	ENSG00000215301	1654
GBP1	ENSG00000117228	2633
GBP2	ENSG00000162645	2634
GBP5	ENSG00000154451	115362
GSDMA	ENSG00000167914	284110
GSDMB	ENSG00000073605	55876
GSDMC	ENSG00000147697	56169
GSDMD	ENSG00000104518	79792
GSDME	ENSG00000105928	1687
GZMA	ENSG00000145649	3001
GZMB	ENSG00000100453	3002
HMGB1	ENSG00000189403	3146
IFI16	ENSG00000163565	3428
IL18	ENSG00000150782	3606
IL1B	ENSG00000125538	3553
IRF1	ENSG00000125347	3659
IRF2	ENSG00000168310	3660
IRF8	ENSG00000140968	3394
MEFV	ENSG00000103313	4210
NAIP	ENSG00000249437	4671
NEK7	ENSG00000151414	140609
NLRC3	ENSG00000167984	197358
NLRC4	ENSG00000091106	58484
NLRC5	ENSG00000140853	84166
NLRP1	ENSG00000091592	22861
NLRP12	ENSG00000142405	91662
NLRP2	ENSG00000022556	55655
NLRP3	ENSG00000162711	114548
NLRP6	ENSG00000174885	171389
NLRP7	ENSG00000167634	199713
NLRP9	ENSG00000185792	338321
NOD1	ENSG00000106100	10392
NOD2	ENSG00000167207	64127
NR2C2	ENSG00000177463	7182
P2RX7	ENSG00000089041	5027
PKN1	ENSG00000123143	5585
PKN2	ENSG00000065243	5586
PYCARD	ENSG00000103490	29108
TNF	ENSG00000232810	7124
ZBP1	ENSG00000124256	81030

https://dx.doi.org/10.21037/tlcr-21-715

Table S4 Prognostic analysis of pyroptosis-related genes using univariate Cox regression analysis

Gene Symbol	HR	HR.95L	HR.95H	P
NLRP1	0.782949744	0.66762702	0.918192767	0.002614241
NLRC3	0.771720026	0.634021223	0.939324704	0.009761218
IRF8	0.850075609	0.74458821	0.970507633	0.016269948
NOD1	0.802728159	0.669894701	0.961901172	0.017273893
NAIP	0.561459675	0.339698962	0.927989199	0.024355497
HMGB1	1.3578057	1.016876961	1.813037751	0.038133776
CASP6	1.262364165	0.999708309	1.594028248	0.050287199
NLRC4	0.817672575	0.668400122	1.000281746	0.050321034
GBP2	1.153303081	0.993796232	1.33841119	0.060377891
GBP1	1.108175468	0.989278108	1.241362625	0.076095307
CASP3	1.298411041	0.972617811	1.733333704	0.076459381
NLRP3	0.856083313	0.719881432	1.018054649	0.078820448
IL18	1.128606139	0.978595933	1.301611598	0.096387862
PYCARD	1.133072361	0.973641851	1.318609069	0.106370897
DDX3X	1.2195353	0.953374257	1.56000263	0.11413479
P2RX7	0.886237469	0.761837475	1.030950665	0.117588153
TNF	0.895586656	0.774465941	1.03564975	0.1368902
IRF1	1.128283179	0.958117031	1.328671644	0.14789195
PKN2	1.186319673	0.940062885	1.497085343	0.150068611
NLRP2	0.946593804	0.875945904	1.022939689	0.165483629
CASP4	1.144999495	0.935698621	1.401117641	0.188620623
ZBP1	0.904773197	0.76554901	1.069316957	0.240465503
GSDMC	1.062966096	0.95849913	1.178818934	0.24731095
NOD2	0.884214276	0.716311167	1.091473821	0.252080367
PKN1	1.135137151	0.901339356	1.429579595	0.281389125
CARD8	0.868879591	0.667170852	1.131571833	0.297023423
CASP8	1.133400238	0.879916026	1.459907607	0.332293649
GSDME	1.071477303	0.923788367	1.242777731	0.361579231
CASP1	0.931935016	0.800409166	1.085073623	0.363812216
GZMA	0.949643644	0.848736855	1.062547295	0.367341469
IRF2	1.146328186	0.846748112	1.551899899	0.376905266
GSDMA	0.935565431	0.778830378	1.123842495	0.476494961
NLRP12	1.059818633	0.888704088	1.263880239	0.517851905
NR2C2	0.934829375	0.760376133	1.149307458	0.522512859
IFI16	1.036309545	0.912141297	1.177380606	0.583882017
CASP5	1.065349346	0.837132249	1.35578247	0.606793584
NLRP9	1.266065987	0.314871355	5.090723739	0.739670056
IL1B	0.979734576	0.867907052	1.105970779	0.740574307
GBP5	0.983563445	0.879205206	1.100308601	0.772121565
GSDMD	0.972190891	0.801471967	1.179274096	0.774680432
MEFV	1.036063431	0.802397857	1.337774552	0.78586025
AIM2	1.011975458	0.928368043	1.103112429	0.786718011
GZMB	1.013897187	0.914597778	1.123977699	0.792981234
GSDMB	0.982799218	0.862229129	1.120229265	0.795002641
NEK7	1.01985776	0.824803011	1.261040317	0.855933014
NLRC5	0.989093367	0.848500971	1.152981224	0.888507528
NLRP7	1.017471289	0.749787388	1.380721844	0.911458801
NLRP6	0.985544494	0.692387268	1.402824683	0.935572403

https://dx.doi.org/10.21037/tlcr-21-715

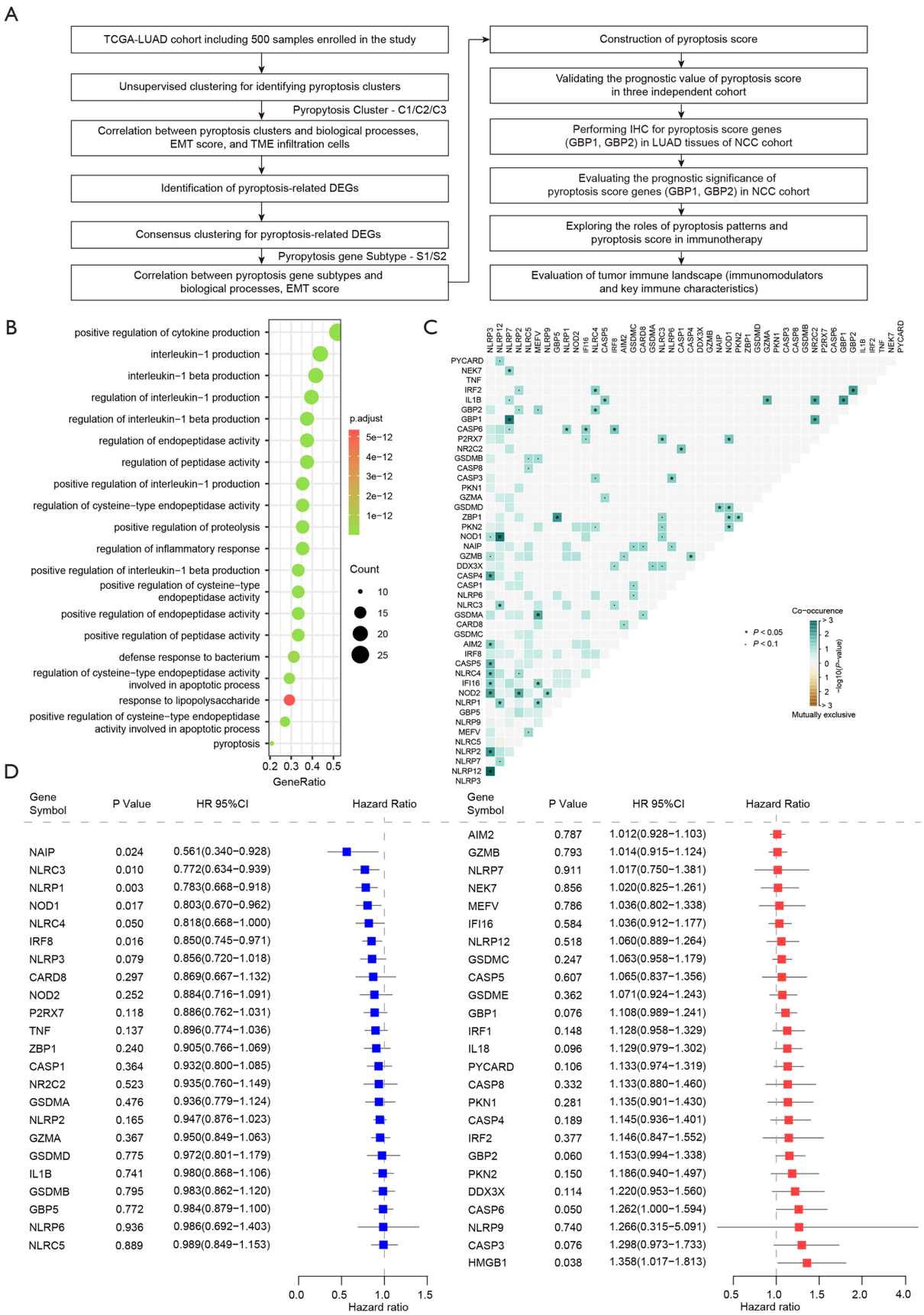

Figure S1 Correlation and prognostic analysis of 48 pyroptosis-related genes (PRGs). (A) Schematic overview of the workflow employed in this study. (B) Functional annotation for pyroptosis related genes by GO enrichment analysis. The dot size and color intensity represent the gene count and adjust P value, respectively; (C) The mutation co-occurrence and exclusion analysis for PRGs. Co-occurrence marked with green; Exclusion marked with yellow. (D) Forrest plot of the univariate Cox regression analysis showed that effect of pyroptosis related genes on clinical prognosis in LUAD. The vertical dotted line represented the hazard ratio (HR). The length of the horizontal line represented the 95% confidence interval, HR >1 indicated risk factor for survival, whereas HR <1 indicated protective factor for survival.

https://dx.doi.org/10.21037/tlcr-21-715

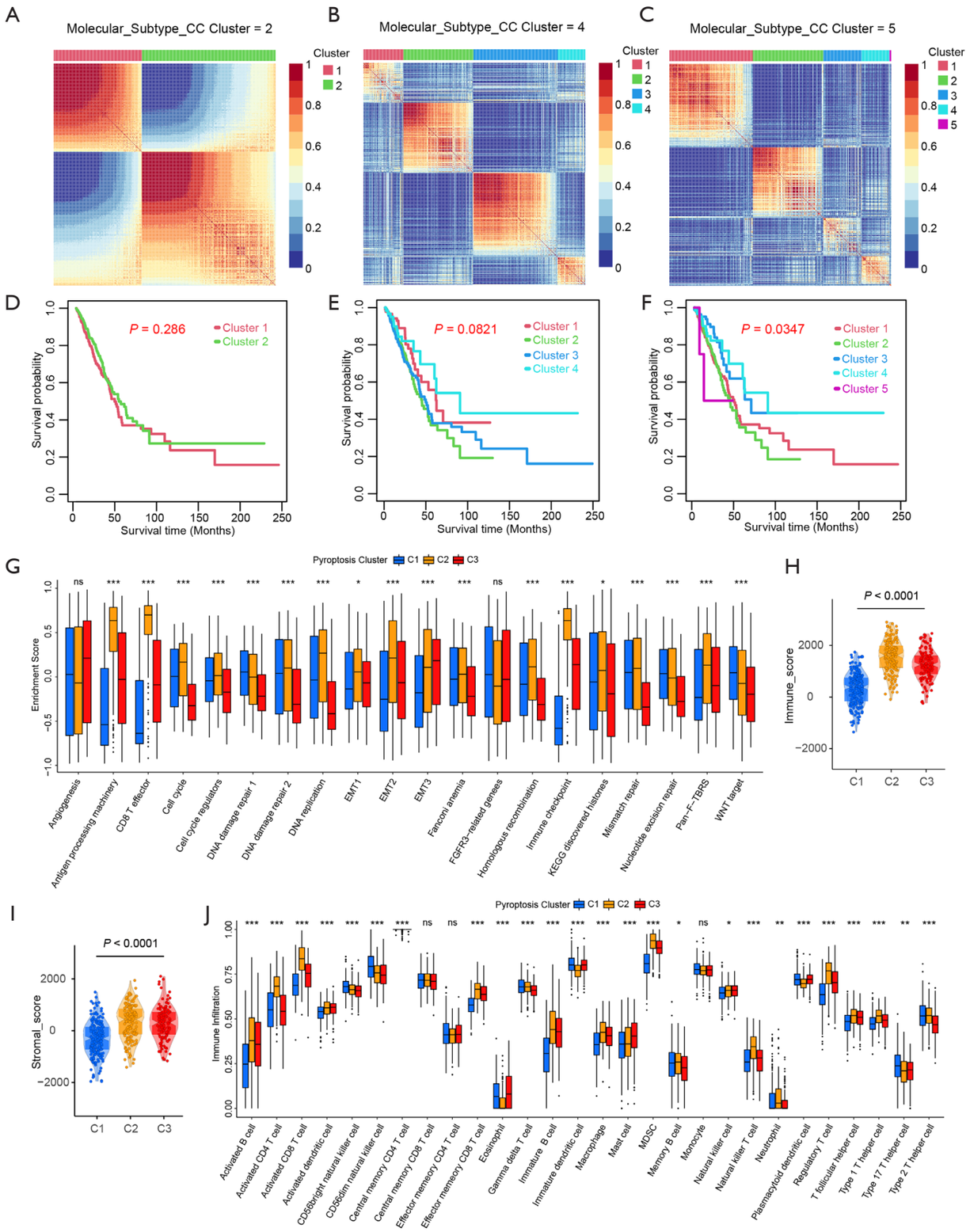

Figure S2 Pyroptosis-related molecular patterns correlated with EMT and immune cell infiltration characteristics in LUAD. (A-C) Unsupervised clustering of pyroptosis-related genes in TCGA cohort and consensus matrices for k =2,4,5. (D-F) Kaplan-Meier curves of overall survival (OS) for 500 LUAD patients in TCGA cohort with different pyroptosis cluster numbers with k=2,4,5. (G) Differences of the known biological gene signatures among three distinct pyroptosis clusters. Immune score (H) and stromal score (I) were compared among three pyroptosis clusters. (J)The fraction of tumor-infiltrating immune cells in three pyroptosis clusters of TCGA cohort by ssGSEA. The statistical difference of three clusters were compared through the Kruskal-Wallis test. ns: not significant; *P <0.05; **P <0.01; ***P <0.001.

https://dx.doi.org/10.21037/tlcr-21-715

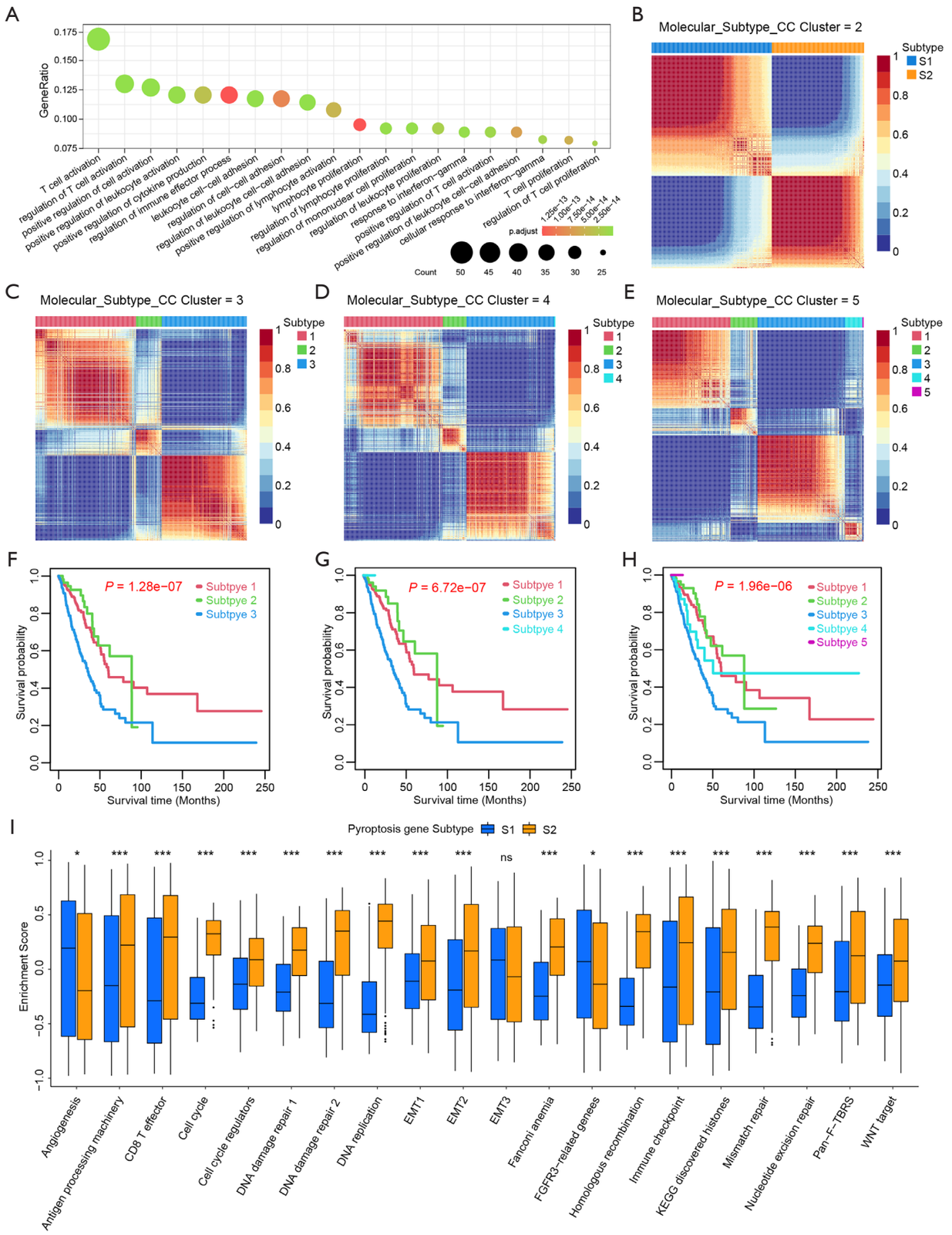

Figure S3 Pyroptosis gene subtypes with distinct prognosis and biological characteristics in LUAD. (A) GO enrichment analysis of the 470 pyroptosis-related differentially expressed genes (DEGs). The y-axis indicated gene ratio within each GO term. (B-E) Unsupervised clustering of prognostic pyroptosis-related DEGs in TCGA cohort and consensus matrices for k =2-5. (F-H) Kaplan-Meier curves of overall survival (OS) for 500 LUAD patients in TCGA cohort with different pyroptosis cluster numbers with k =3-5. (I) Differences in the known biological gene signatures among two distinct pyroptosis gene subtypes. The statistical differences between pyroptosis gene subtypes were analyzed through the Wilcoxon test. The asterisks represented the statistical P value (ns: not significant; *P <0.05; **P <0.01; ***P <0.001).

https://dx.doi.org/10.21037/tlcr-21-715

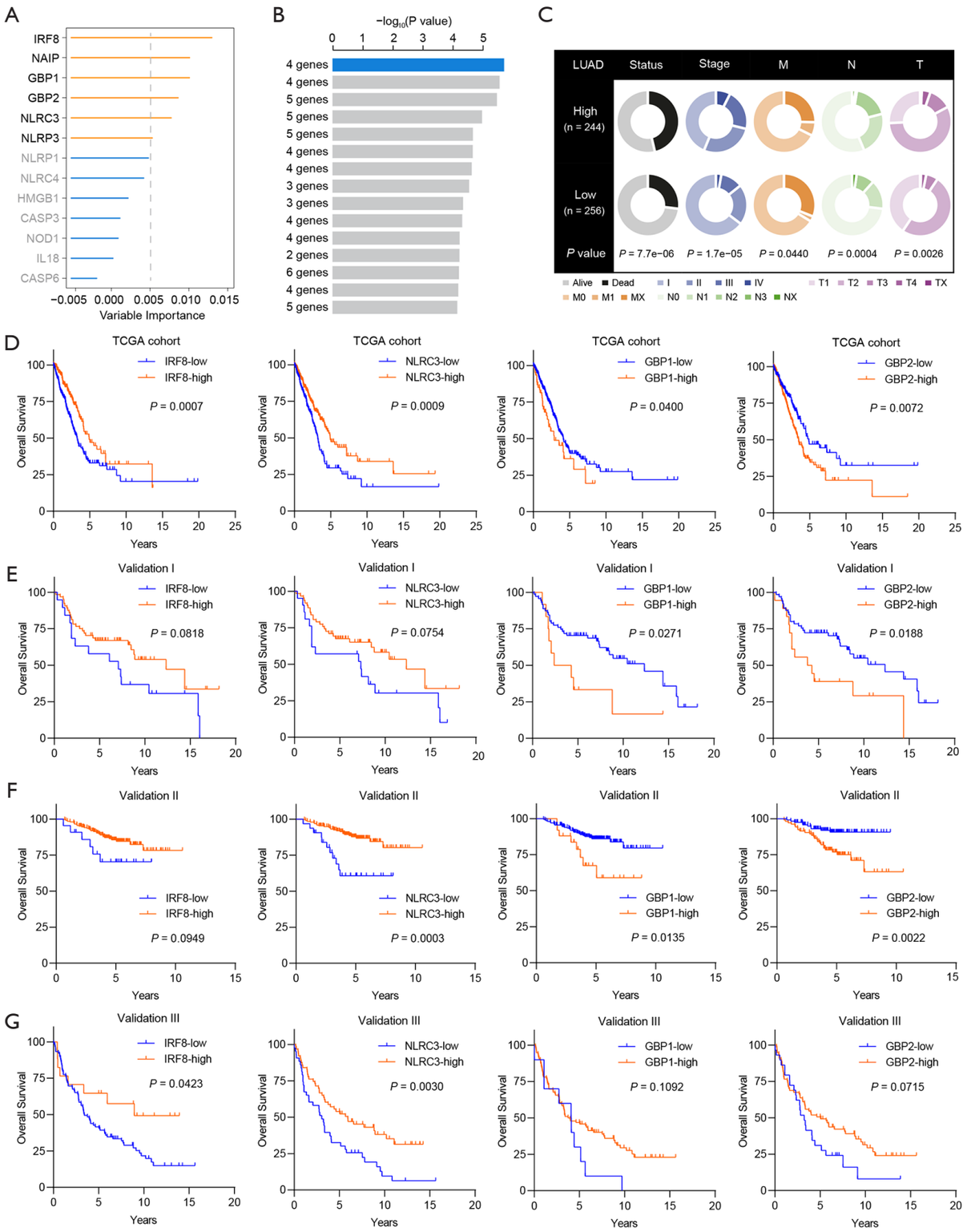

Figure S4 The impact of pyroptosis score and pyroptosis score genes on clinical outcome for LUAD patients. (A) Random survival forest analysis screened 13 genes ranked by importance. (B) The Log-rank P-values of the top 15 combinations were displayed. The signature including four genes was chosen, which had the biggest -log10 P-value and relatively small number of genes. (C) Comparisons of clinical features between the high and low pyroptosis score subgroups. (D-G) Kaplan-Meier curves for patients with high and low expression of 4 pyroptosis score genes (IRF8, NLRC3, GBP1, GBP2) in the TCGA cohort (D), Validation I (E), Validation II (F) and Validation III (G) cohorts, respectively.

https://dx.doi.org/10.21037/tlcr-21-715

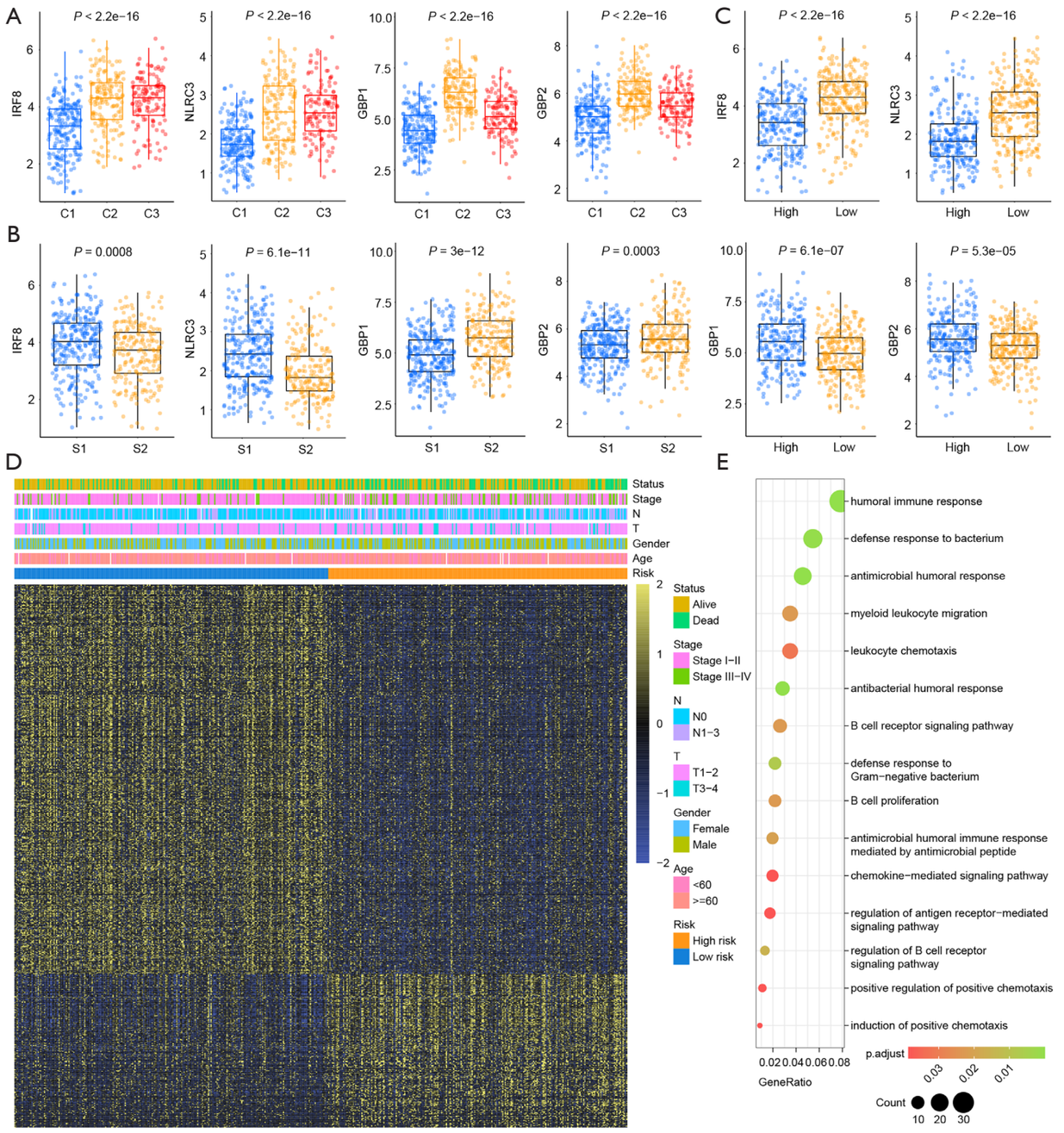

Figure S5 Expression of pyroptosis score genes in distinct subgroups and the biological processes associated with pyroptosis score. (A) Box plots showed the expression distribution of 4 genes (IRF8, NLRC3, GBP1, GBP2) across three pyroptosis clusters. (B) Comparisons of the expression of 4 genes (IRF8, NLRC3, GBP1, GBP2) between two pyroptosis gene subtypes. (C) Comparisons of the expression of 4 genes (IRF8, NLRC3, GBP1, GBP2) between high and low pyroptosis score groups. (D) Heatmap visualized the differences of clinicopathological features and DEGs between the high and low pyroptosis score groups. (E) GO enrichment analysis of DEGs between two groups.

https://dx.doi.org/10.21037/tlcr-21-715

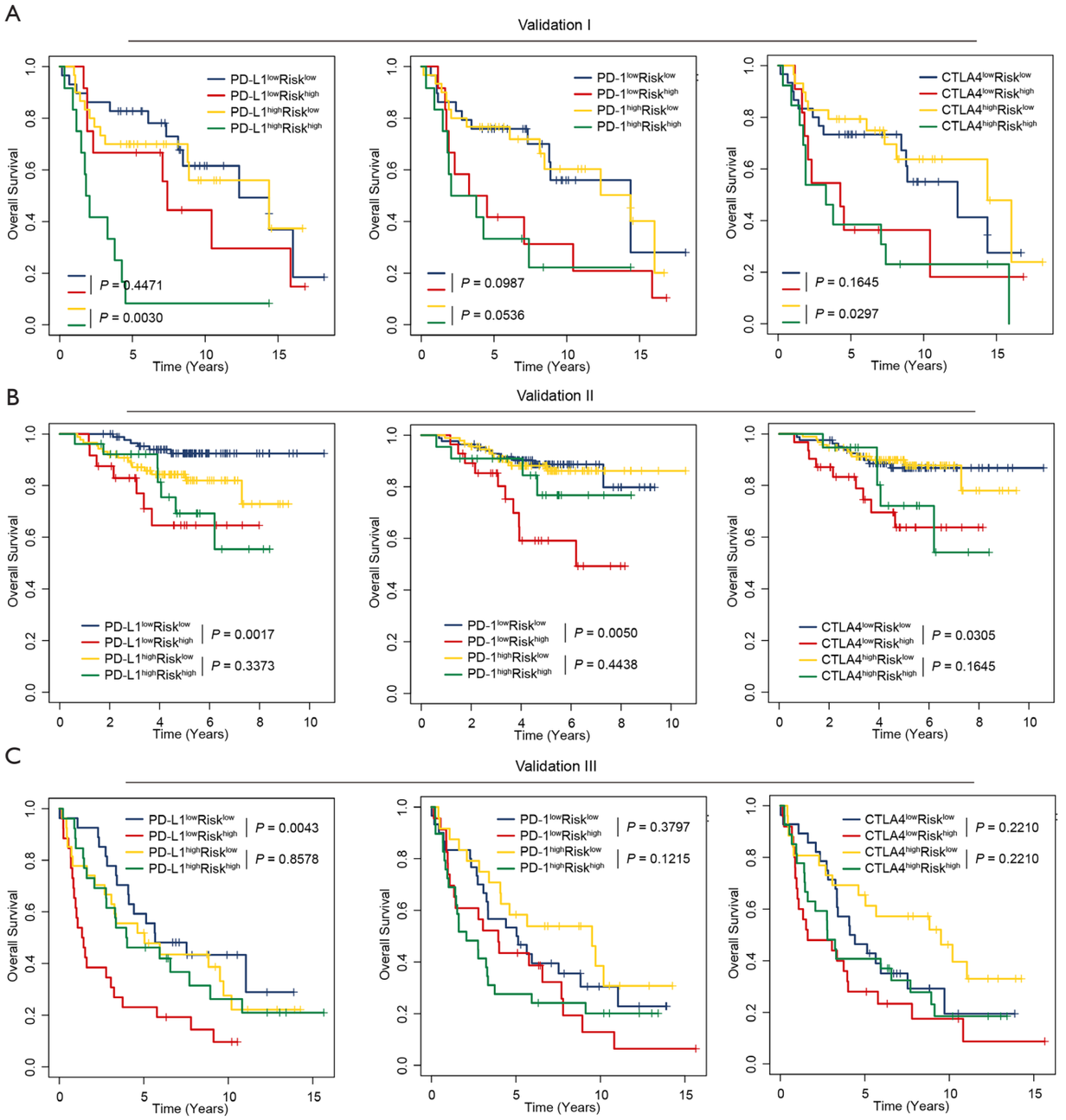

Figure S6 Impact of pyroptosis score and immune checkpoint genes expression on survival outcome in external validation cohorts. (A-C) Kaplan-Meier survival curves of OS among four patient groups divided by the pyroptosis score and PD-L1, PD-1 or CTLA-4 in Validation I (A), Validation II (B), Validation III (C) cohorts, respectively.

https://dx.doi.org/10.21037/tlcr-21-715

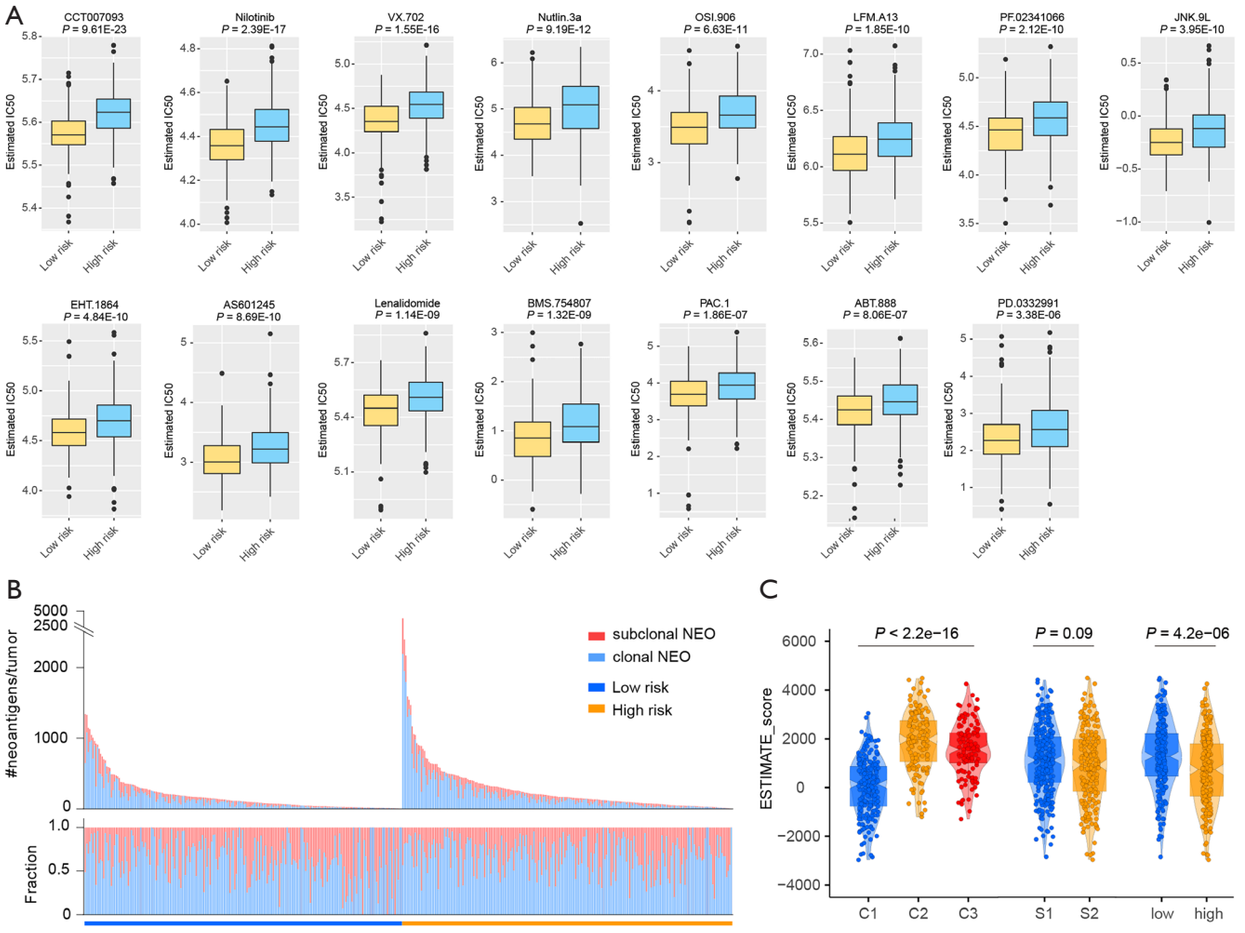

Figure S7 Potential small molecule compounds based on pyroptosis score. (A) The boxplots of the estimated IC50 for the top 15 significant compounds. (B) Proportion of neoantigens arising from clonal (blue) or subclonal (red) mutations was shown. Samples were grouped according to risk groups, with low risk (low pyroptosis score) on left and high risk (high pyroptosis score) on right. (C) The relative distributions of ESTIMATE score were compared among three pyroptosis clusters, pyroptosis gene subtypes and between pyroptosis score high versus low groups in TCGA cohort, respectively.

https://dx.doi.org/10.21037/tlcr-21-715