Supplementary material and methods

Data collection and processing

We downloaded normalized gene expression data based on fragments per kilobase per million (FPKM) of patient tumor samples and paired adjacent normal tissues across 31 cancer types from TCGA data portal (http://gdac.broadinstitute.org/). We obtained the APA usage profiles for 10,303 tumors across 31 cancer types from The Cancer 3′ UTR Atlas (TC3A)[1], 553 paired adjacent normal and tumor samples across 12 cancer types from Synapse (https://www.synapse.org/, syn7888354)[2], and 1172 samples across 10 tissues of Genotype-Tissue Expression project (GTEx) from APAatlas (https://hanlab.uth.edu/apa/)[3]. Protein level data was obtained from ProteomeXchange with accessible ID PXD006003 (http://www.proteomexchange.org). We download raw data of PolyA RNA-seq of melanoma patients with ICB treatment from Gene Expression Omnibus (GEO) under accession no. GSE91061[4] and m6A-seq data of HeLa cell line under accession no. GSE102493[5].

Tumor immune infiltration fraction (TIF) of TCGA 31 cancer types obtained from previous study[6]. We calculated the cytolytic activity (CYT) score by the geometric mean of gene expression of cytolytic markers, including GZMA, GZMB, PRF1, and GNLY[7]. We used Gene Set Variation Analysis (GSVA)[8] to estimate the T cell–inflamed gene expression profile (GEP) level[9] in each sample of TCGA 31 cancer types. A list of 22-tumor-infiltrating immune cell type markers was derived from CIBERSORT[10].

qRT-PCR to quantify the usage of *COL1A1* 3'UTR

Total RNAs were extracted using TRIpure (Bioteke Corporation, China), and reverse transcription reactions were performed using HiScript II Q RT SuperMix for qPCR (Vazyme, China) according to the manufacturer's instructions. Then, 40 cycles of quantitative reverse-transcription PCR (qRT-PCR) were conducted in 96-well plates using Ultra SYBR Mixture (CWBIO, China) on the QuantStudio3 Real-Time PCR System.

To detect dPAS usage, two pairs of primers were designed, with one targeting the ORF to represent the total transcript level and the other targeting sequences just before the dPAS to detect long transcripts that used the dPAS[11]. The percentage of dPAS usage was calculated as follows: $_{\Delta}CT$ = CT long − CT total. The $_{\Delta\Delta}CT$ ($_{\Delta}CT$ target − $_{\Delta}CT$ control) represents the percentage of long transcripts in target/percentage of long transcripts in Control. For tumor vs control,

$_{\Delta\Delta}CT=_{\Delta}CT_{tumor}-{}_{\Delta}CT_{control}$. The fold change of gene expression was calculated by $2^{-\Delta\Delta CT}$. The sequences of primers were as follows: GAPDH: forward, 5'-GGAGCGAGATCCCTCCAAAAT-3', reverse, 5'-GGCTGTTGTCATACTTCTCATGG-3'; COL1A1 total: forward 5'-CAGACTGGCAACCTCAAGAA-3', reverse, 5'-CAGTGACGCTGTAGGTGAAG-3'; COL1A1 long: forward 5'-GACAACTTCCCAAAGCACAAA-3', reverse 5'-GGGTCATTTCCACATGCTTTATT-3'.

Statistical analysis

Gene Set Enrichment Analysis (GSEA) involved in ImmAPAScore algorithm was performed by R package *clusterProfiler*. We used t-distributed stochastic neighbor embedding (t-SNE) to cluster ImmAPAs via R package *Rtsne*. The Kaplan-Meier survival analysis was performed via R package *survival* using log-rank test. Maxstat package was used to determine the optimal cutoff point between PDUI or gene expression and overall survival for each data set. Multivariable cox proportional hazards regression model was used to assess the association of variables with survival through the *coxph* function. The R package *survivalROC* was used to calculate time-dependent receiver-operating characteristic (ROC) curves and the area under the ROC curve (AUC) for ICB_APASig score. All the statistical analyses were performed in R version 3.6.

Supplementary Reference

1. Feng, X., Li, L., Wagner, E. J. & Li, W. TC3A: The Cancer 3′ UTR Atlas. *Nucleic Acids Research* **46**, D1027–D1030 (2018).

2. Xiang, Y., Ye, Y., Lou, Y., Yang, Y., Cai, C., Zhang, Z., Mills, T., Chen, N., Kim, Y., Ozguc, F. M., Diao, L., Karmouty-quintana, H., Xia, Y., Kellems, R., Chen, Z., Yoo, S., Shyu, A., Mills, G. B. & Han, L. Comprehensive characterization of alternative polyadenylation in human cancer. *Journal of the National Cancer Institute* **110**, 379–389 (2018).

3. Hong, W., Ruan, H., Zhang, Z., Ye, Y., Liu, Y., Li, S., Jing, Y., Zhang, H., Diao, L., Liang, H. & Han, L. APAatlas: Decoding alternative polyadenylation across human tissues. *Nucleic Acids Research* **48**, D34–D39 (2020).

4. Riaz, N., Havel, J. J., Makarov, V., Desrichard, A., Urba, W. J., Sims, J. S., Hodi, F. S., Martín-Algarra, S., Mandal, R., Sharfman, W. H., Bhatia, S., Hwu, W. J., Gajewski, T. F., Slingluff, C. L., Chowell, D., Kendall, S. M., Chang, H., Shah, R., Kuo, F., *et al.* Tumor and Microenvironment Evolution during Immunotherapy with Nivolumab. *Cell* **171**, 934–949 (2017).

5. Liu, J., Yue, Y., Liu, J., Cui, X., Cao, J., Luo, G., Zhang, Z., Cheng, T., Gao, M., Shu, X., Ma, H., Wang, F., Wang, X., Shen, B., Wang, Y., Feng, X. & He, C. VIRMA mediates preferential m6A mRNA methylation in 3′UTR and near stop codon and associates with alternative polyadenylation. *Cell Discovery* **4**, (2018).

6. Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Ou Yang, T.-H., Porta-Pardo, E., Gao, G. F., Plaisier, C. L., Eddy, J. A., Ziv, E., Culhane, A. C., Paull, E. O., Sivakumar, I. K. A., Gentles, A. J., Malhotra, R., Farshidfar, F., Colaprico, A., Parker, J. S., *et al.* The Immune Landscape of Cancer. *Immunity* 812–830 (2018).

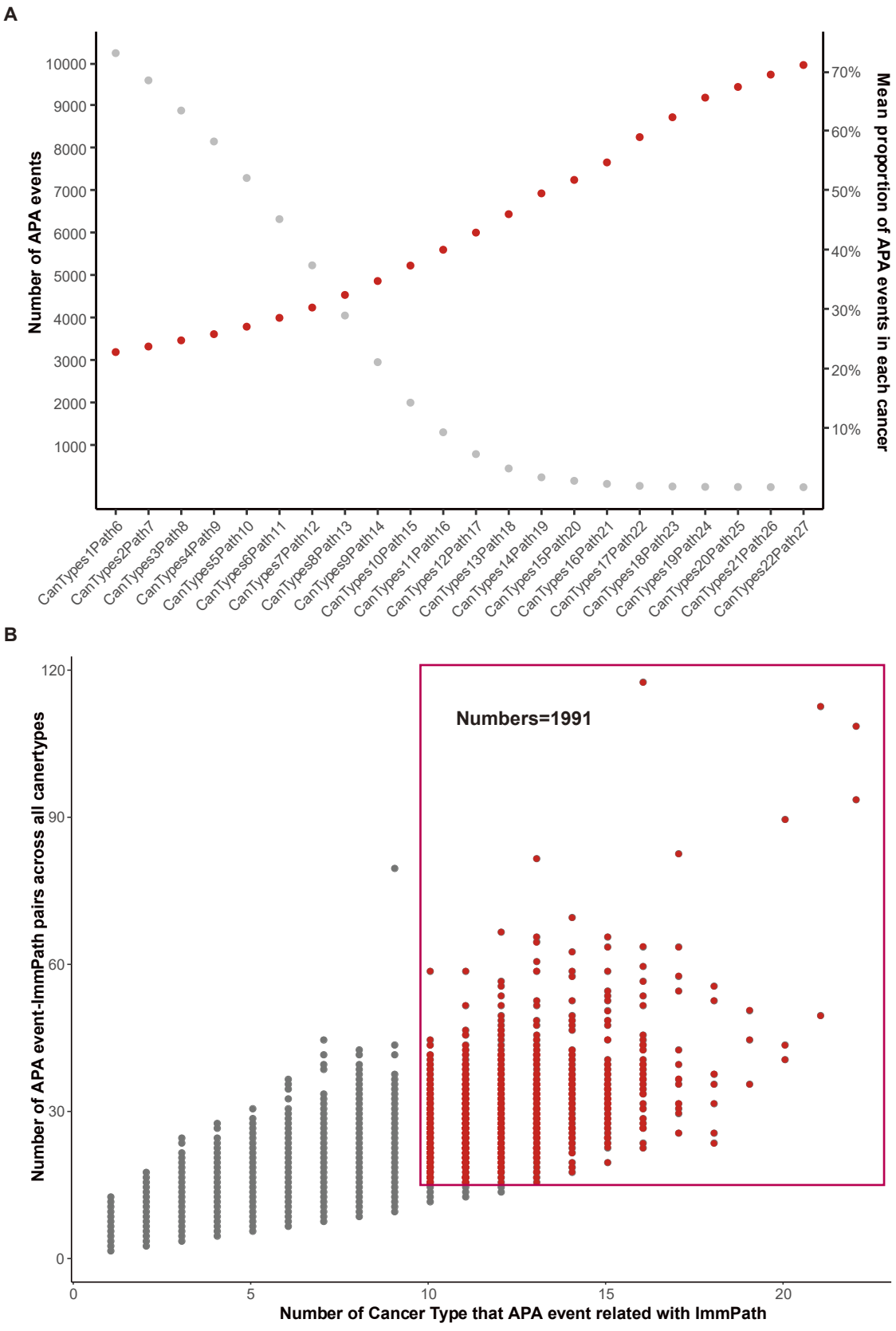
7. Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G. & Hacohen, N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* **160**, 48–61 (2015).

8. Hänzelmann, S., Castelo, R. & Guinney, J. GSVA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics* **14**, (2013).

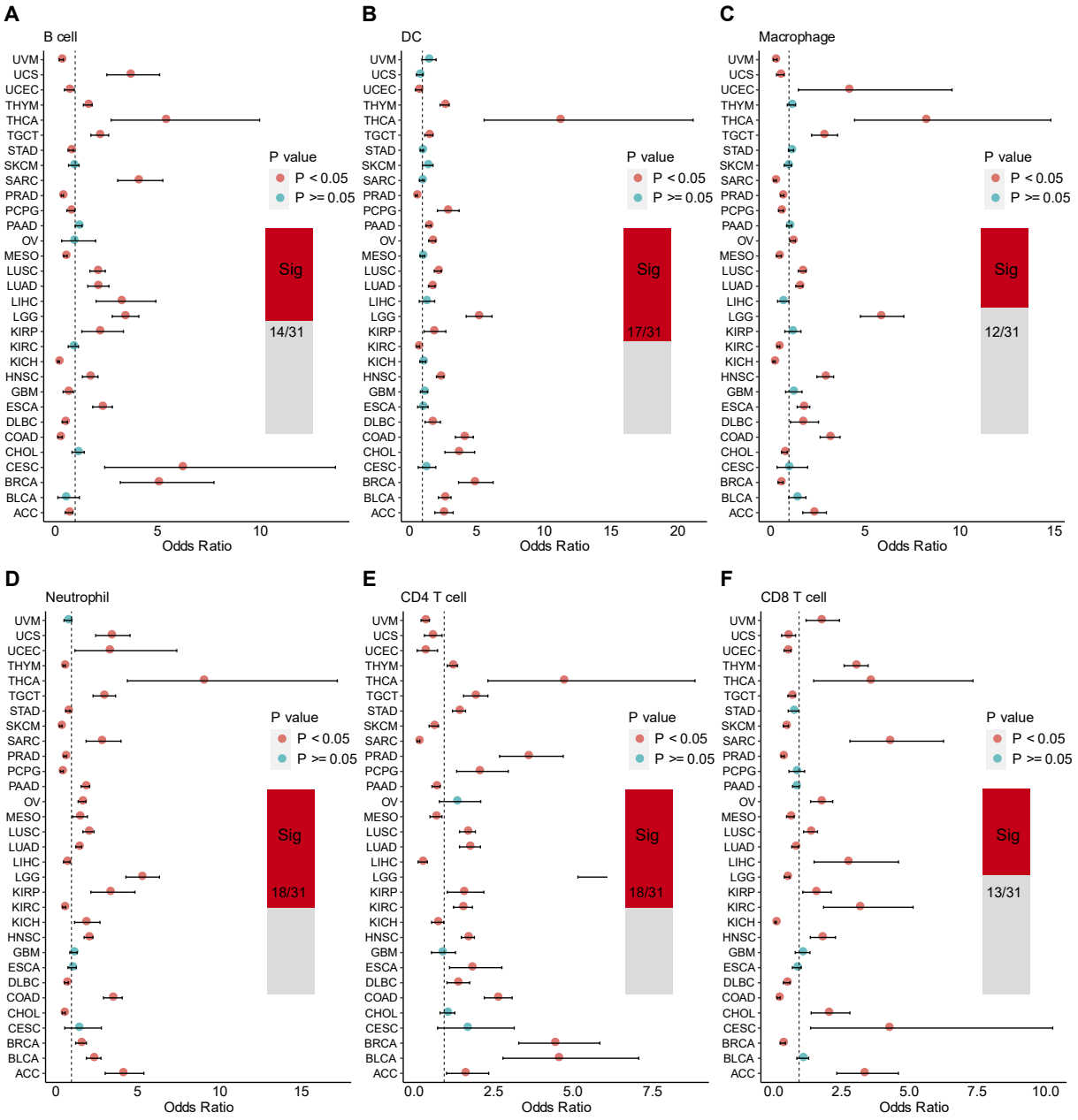
9. Cristescu, R., Mogg, R., Ayers, M., Albright, A., Murphy, E., Yearley, J., Sher, X., Liu, X. Q., Lu, H., Nebozhyn, M., Zhang, C., Lunceford, J. K., Joe, A., Cheng, J., Webber, A. L., Ibrahim, N., Plimack, E. R., Ott, P. A., Seiwert, T. Y., *et al.* Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. *Science (New York, N.Y.)* **362**, eaar3593 (2018).

10. Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., Hoang, C. D., Diehn, M. & Alizadeh, A. A. Robust enumeration of cell subsets from tissue expression profiles. *Nature Methods* **12**, 453–457 (2015).

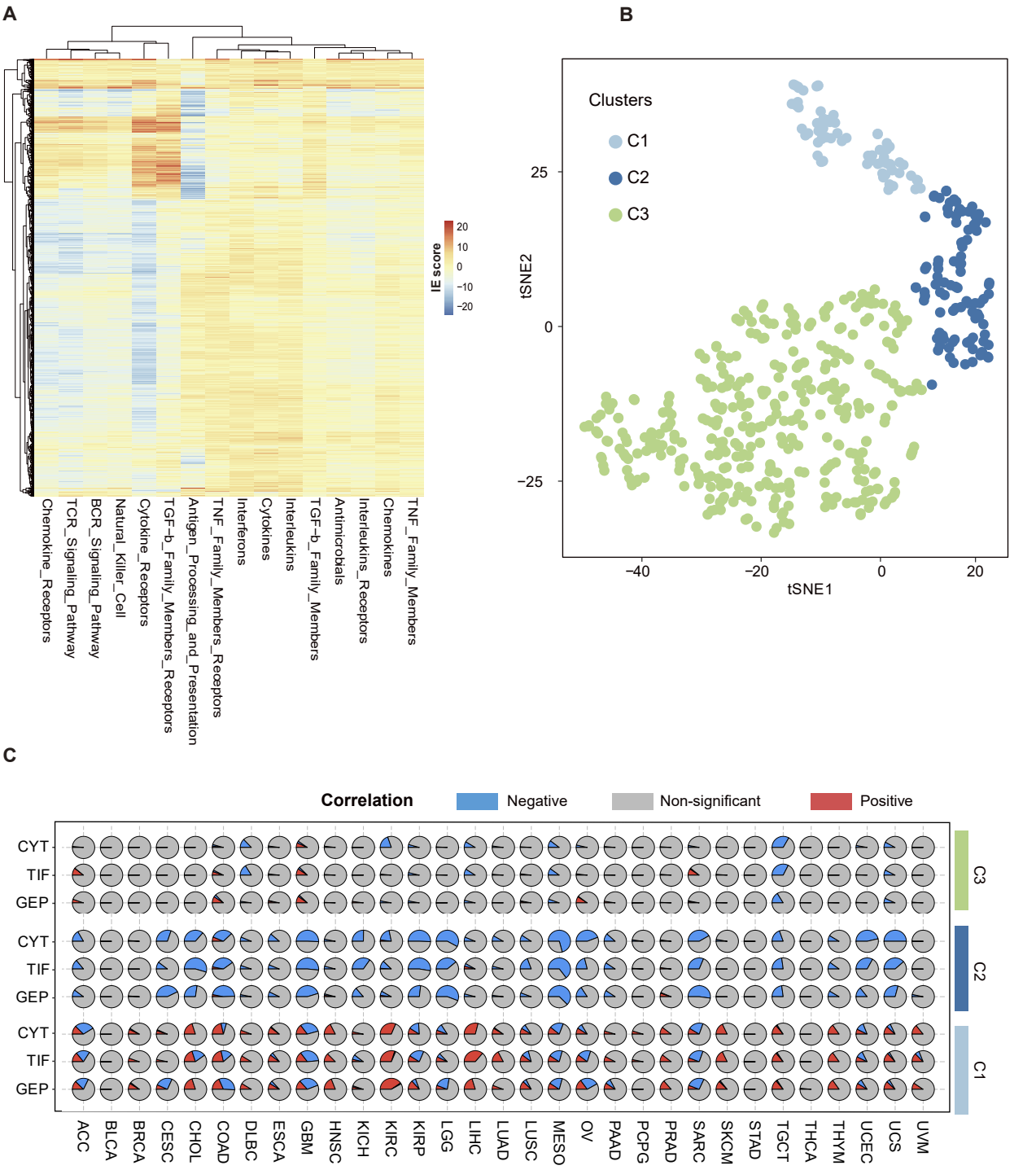
11. Weng, T., Ko, J., Masamha, C. P., Xia, Z., Xiang, Y., Chen, N. yuan, Molina, J. G., Collum, S., Mertens, T. C., Luo, F., Philip, K., Davies, J., Huang, J., Wilson, C., Thandavarayan, R. A., Bruckner, B. A., Jyothula, S. S. K., Volcik, K. A., Li, L., *et al.* Cleavage factor 25 deregulation contributes to pulmonary fibrosis through alternative polyadenylation. *Journal of Clinical Investigation* **129**, 1984–1999 (2019).



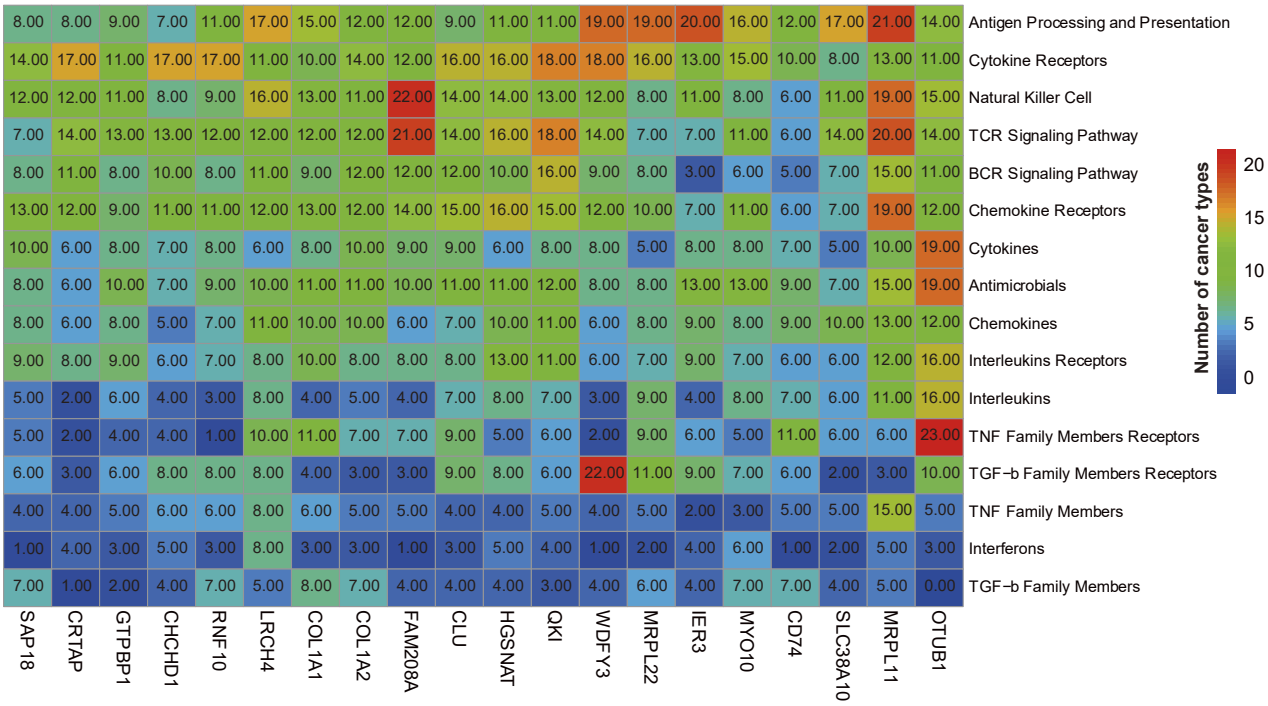
Supplementary Figure 1. Define the ImmAPAs based on the APA event-ImmPath pairs. A. Number of APA events filtered by different cutoffs. X-axis presents the different cutoff, which defined as the number of cancer types of each APA events that shown related with immune pathway, and number of total APA event-ImmPath pairs of each APA events across all cancer types. Left y-axis indicates the number of APA events under different cutoffs, right y-axis indicates the mean proportion of APA events in each cancer types of the total APA events under different cutoffs. B. Number of cancer types in which the APA event significant related with immune pathway, and the number of APA events-ImmPath pairs of each APA events across all cancer types.



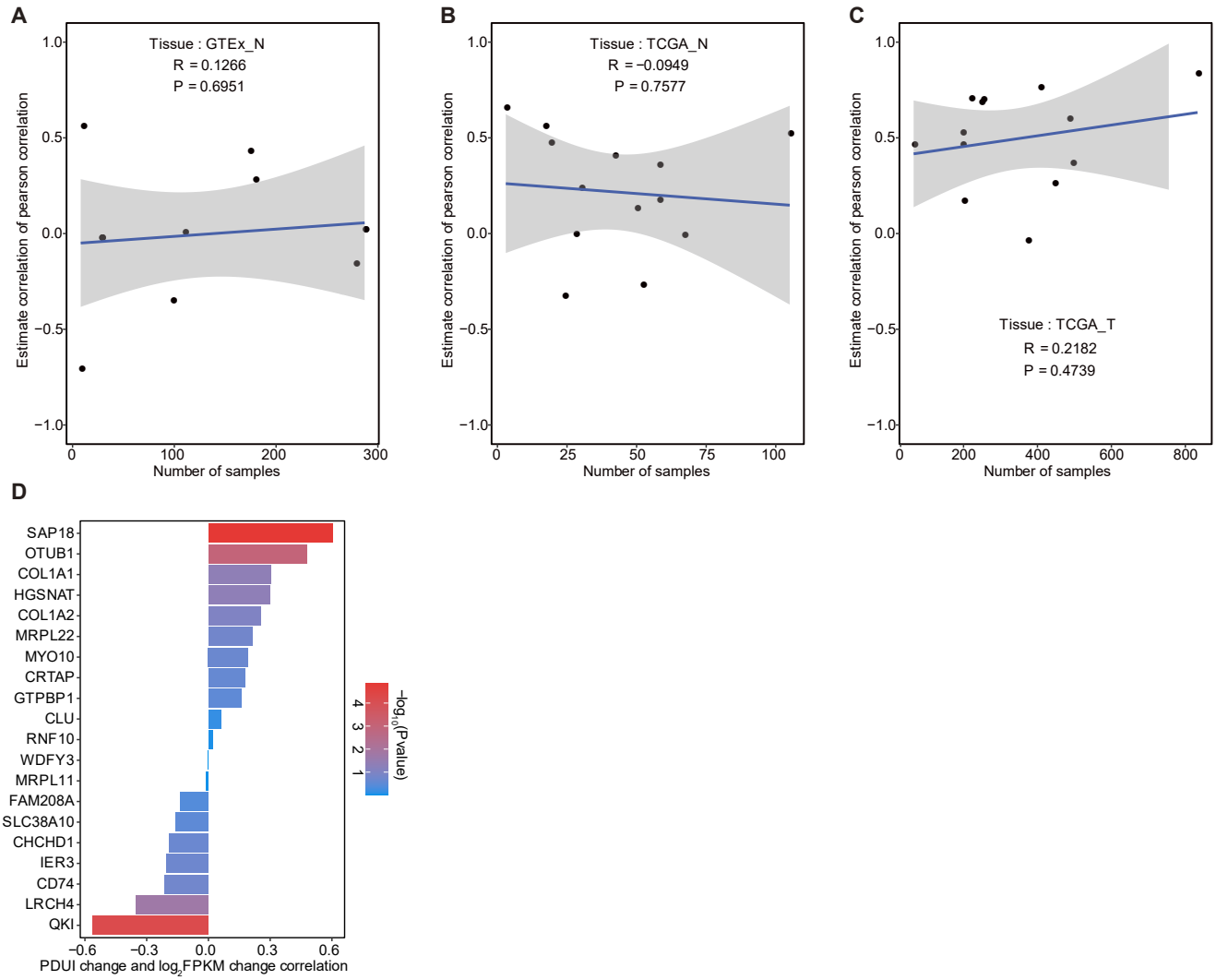
Supplementary Figure 2. The ImmAPAs are correlated with immune cell infiltration. The ImmAPAs were likely to be enriched in APA events that were correlated with immune cell infiltration, including B cells (A), dendritic cells (B), macrophages (C), neutrophils (D), CD4 T cells (E), and CD8 T cells (F). The dots represent the odds ratio (OR) of Fisher's exact test and the error bars show the 95% confidence intervals of the OR. The numbers inside bar plots show the proportion of cancers, where red indicates cancers with P < 0.05 and OR > 1 in two-sided Fisher's exact test.



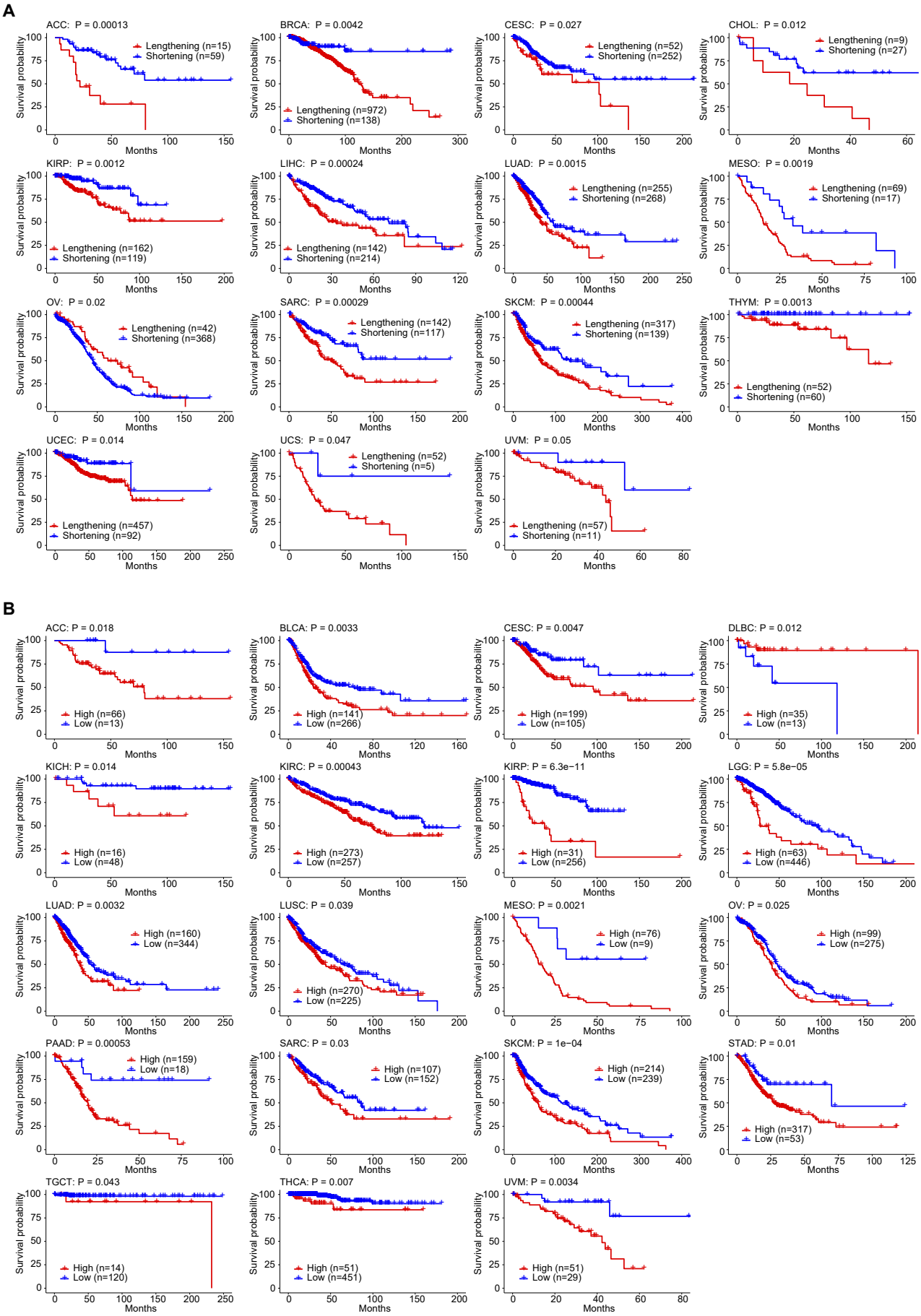
Supplementary Figure 3. ImmAPAs in different clusters correlated with immune features. A. Heatmap shows the IE (Immune Enrichment) score of all 1991 ImmAPAs. B. t-SNE embedding of 3 cluster ImmAPAs. C. The proportion of ImmAPA with positive (Rs > 0.2 and FDR < 0.05 in red) or negative (Rs < -0.2 and FDR< 0.05 in blue) correlation with CTY, GEP, and TIF in each cluster in 31 cancer types.



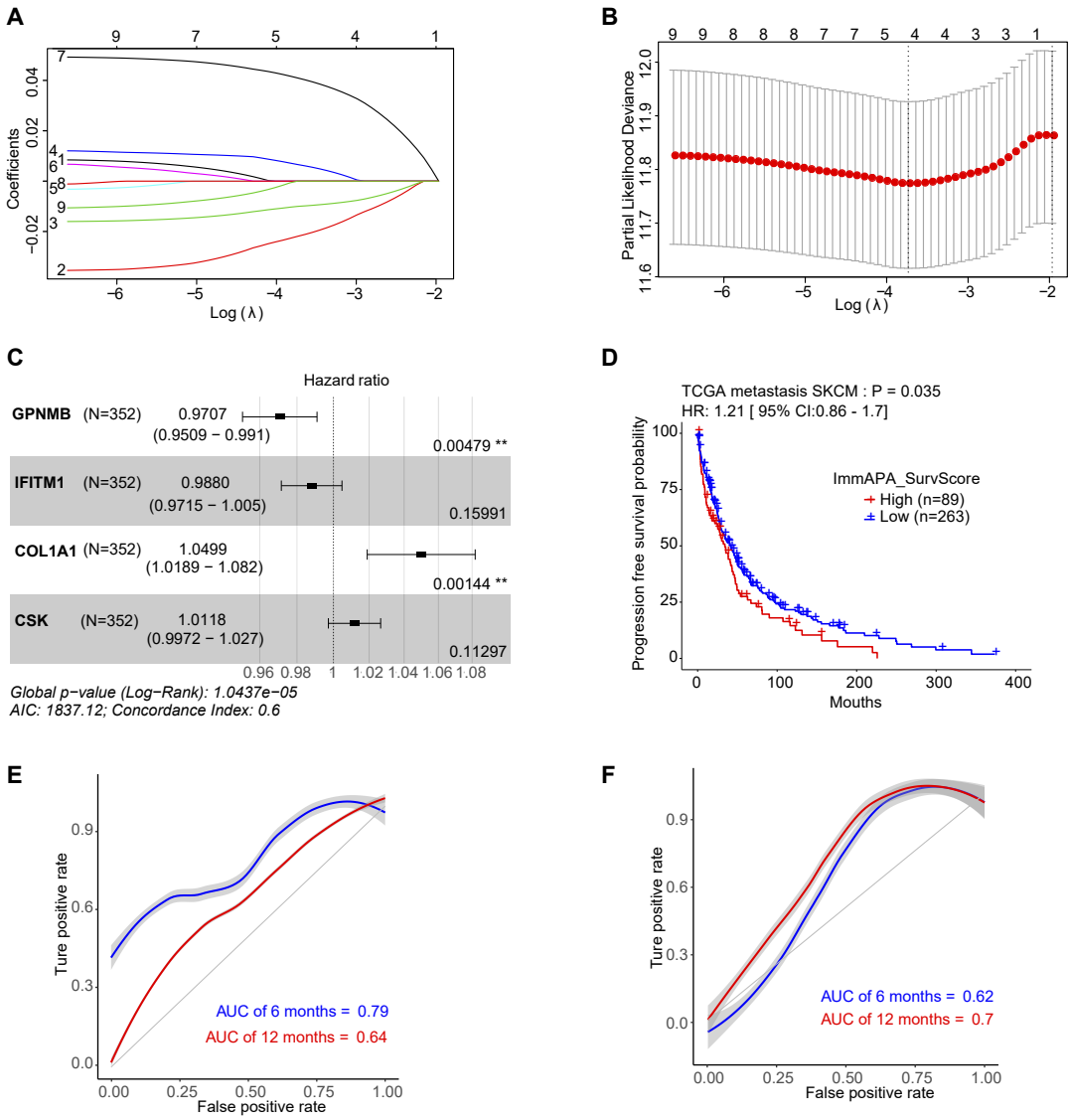
Supplementary Figure 4. Top 20 ImmAPAs related with 16 immune pathways in all 31 cancer types. A. Heatmap shows the number of cancer types in which the ImmAPA was significant correlated with specifical immune pathway.



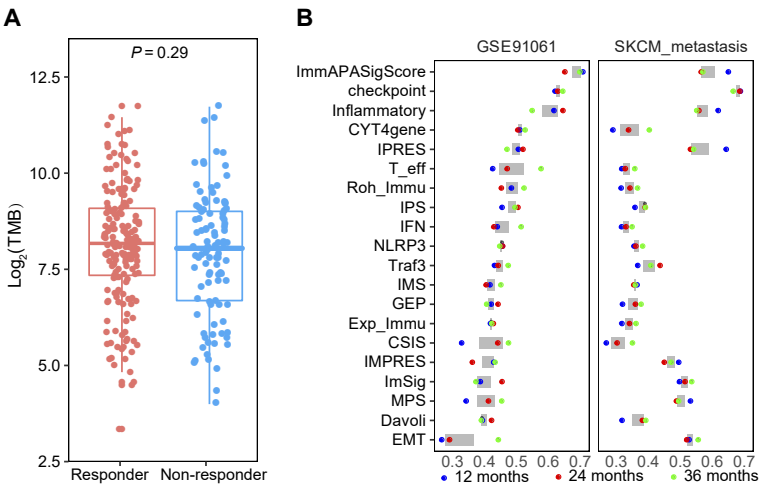
Supplementary Figure 5. Correlation of FPKM and PDUI of *COL1A1* was not related with number of samples. A-C. Number of samples and person correlation of FPKM and PDUI of *COL1A1* in different tissues in GTEx (A), adjacent normal tissues in TCGA (B) and tumor tissues in TCGA (C). (D) Bar plots showing PDUI change and the mRNA expression change of top 20 ranked ImmAPAs between melanoma and adjacent normal tissues detected by RNA-seq.



Supplementary Figure 6. The prognosis of *COL1A1* 3'UTR usage and mRNA expression, and the m6A binding in *COL1A1* 3'UTR. A-B. Kaplan-Meier curves show the significant association of PDUI (A) or FPKM (B) of *COL1A1* and Overall survival in different cancer types



Supplementary Figure 7. Identification of the ImmAPA_ICBSig. A. The partial likelihood deviance plot of LASSO Cox regression used to select candidate APA events. B. LASSO coefficient profiles of candidate ImmAPAs. The vertical dotted lines were the optimal values by using the minimum criteria. C. HR Forest plot of multivariate Cox model of four ICB-related ImmAPAs with overall survival. D. Kaplan–Meier survival curves of progressive free survival between high and low ICB_APASig score patients. E. Time-dependent ROC curve at 6 and 12-months of OS for the ICB_APASig score. F. Time-dependent ROC curve at 12 and 24-months of OS for the ICB_APASig score. A log-rank test was used in D. ROC, receiver operating characteristic curve; AUC, Area Under the ROC Curve.



Supplementary Figure 8. Comparisons of ICB_APASig score and other molecular biomarkers. A. Boxplot showing the log2-transformation TMB between responders and non-responders. The boxes indicate the median ± 1 quartile, with the whiskers extending from the hinge to the smallest or largest value within 1.5× IQR from the box boundaries. P value calculated by two-sided Wilcoxon rank-sum test. B. Performance of ICB_APASig score and other published signatures based on AUC of time-dependent ROC of 12, 24, and 36-months.