Supplementary

Table S1 Details of kits used in the present study

Name	Company	Country
QIAamp DNA FFPE Tissue Kit	Qiagen	USA
DNeasy Blood and tissue Kit	Qiagen	USA
dsDNA HS Assay Kit	ThermoFisher Scientific	USA
KAPA Hyper Prep Kit	KAPA Biosystems	USA
xGen Exome Research Panel and Hybridization and Wash Reagents Kit	Integrated DNA Technology	USA
RNeasy Plus Universal Kit	Qiagen	USA
Qubit™ RNA HS Assay Kit	ThermoFisher Scientific	USA
Take 3	BioTek	USA
RNA Cartridge kit of the Qseq100 Bio-Fragment Analyzer	Bioptic	China
VAHTS mRNA-seq V3 Library Prep Kit	Vazyme	China

Table S2 Details of software used in the present study

Name	Version
SOAPnuke	1.5.6
Burrows-Wheeler Alignment tool	0.7.12
SAMtools	1.3
SAMBLASTER	0.1.22
VarScan	2.4.1
SnpEff	4.3
CNVkit	0.8.1
ascatNgs	3.1.0
POLYSOLVER	1.0
Bwakit	0.7.11
trim galore	0.6.7
Kallisto	0.46.2
Gencode	38.0
MiXCR	2.1.10
VDJtools	1.2.1
PyClone	0.13.0

https://dx.doi.org/10.21037/tlcr-23-800

Table S3 Detailed information on hyperparameter combinations per model

Model	Hyperparameter combinations
Support vector machine	kernel: ['linear'], C: [0.1, 1, 10]
Random forest	n_estimators: randint(100, 1000), max_depth: randint(5, 20), min_samples_split: randint(2, 10), min_samples_leaf: randint(1, 10), max_features: ['auto', 'sqrt'], bootstrap: [True, False]
Gradient boosting classifier	n_estimators: [50, 100, 200], learning_rate: [0.1, 0.01, 0.001], max_depth: [4]
Decision tree classifier	max_depth: [None, 5, 10, 15], min_samples_split: [2, 5, 10], min_samples_leaf: [1, 2, 3]
Extra tree classifier	n_estimators: [100, 200, 300], max_depth: [None, 5, 10], min_samples_split: [2, 5, 10]
Gaussian process classifier	kernel = 1.0 *RBF(1.0, length_scale_bounds=(1e-3, 1e3)) n_restarts_optimizer=10, max_iter_predict=100
K-nearest neighbors	n_neighbors=9, weights='uniform'

For unspecified model parameters, default values are used.

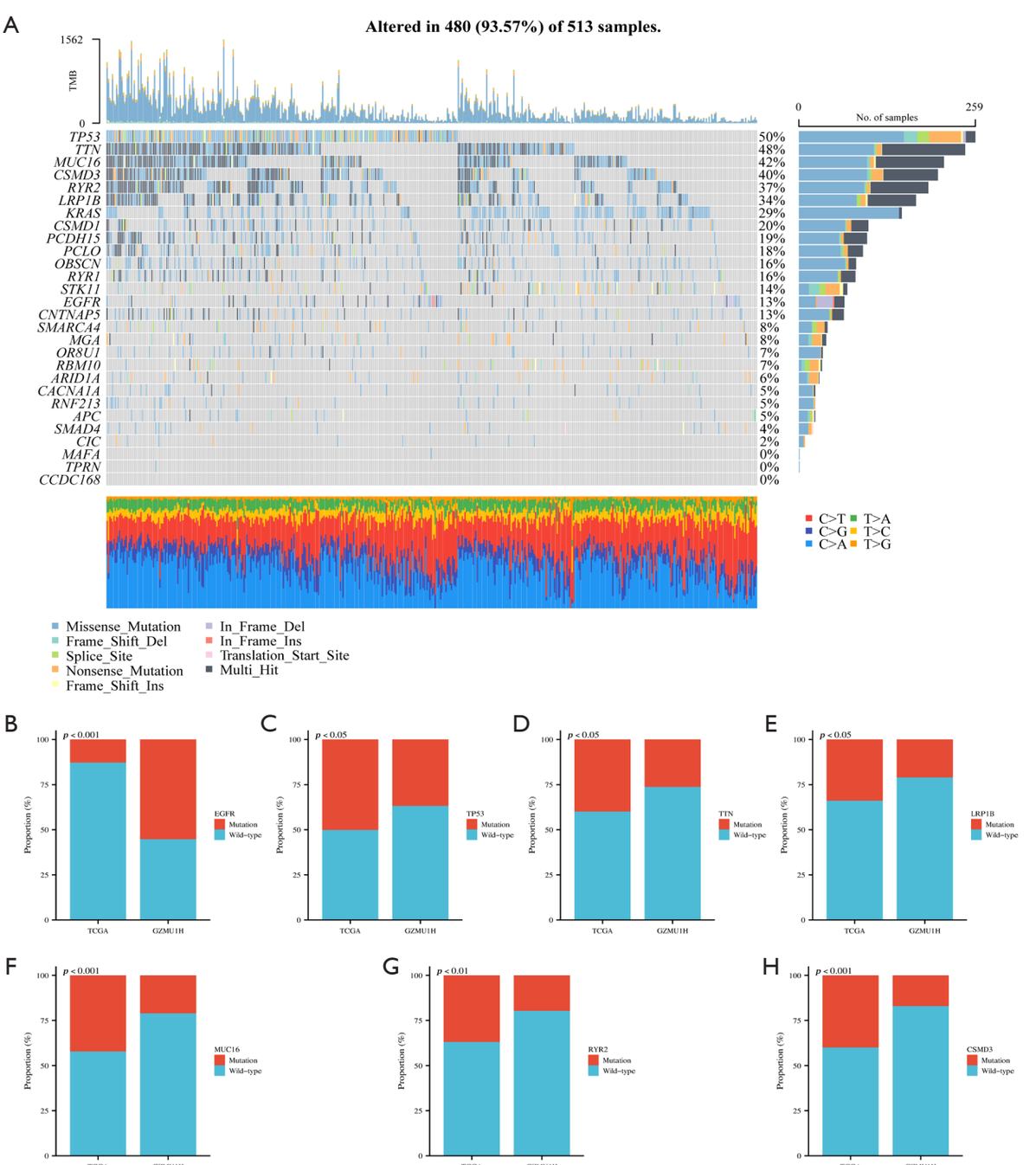

Figure S1 Mutation landscape of lung adenocarcinoma in the TCGA-LUAD cohort. Oncoplot demonstrating the highly mutant genes in the TCGA-LUAD dataset (A). Among the top ten mutated genes in the GZMU1H cohort, seven genes with significantly different mutant frequencies than the TCGA-LUAD cohort, including *EGFR* (B), *TP53* (C), *TTN* (D), *LRP1B* (E), *MUC16* (F), *RYR2* (G), and *CSMD3* (H), were found.

https://dx.doi.org/10.21037/tlcr-23-800

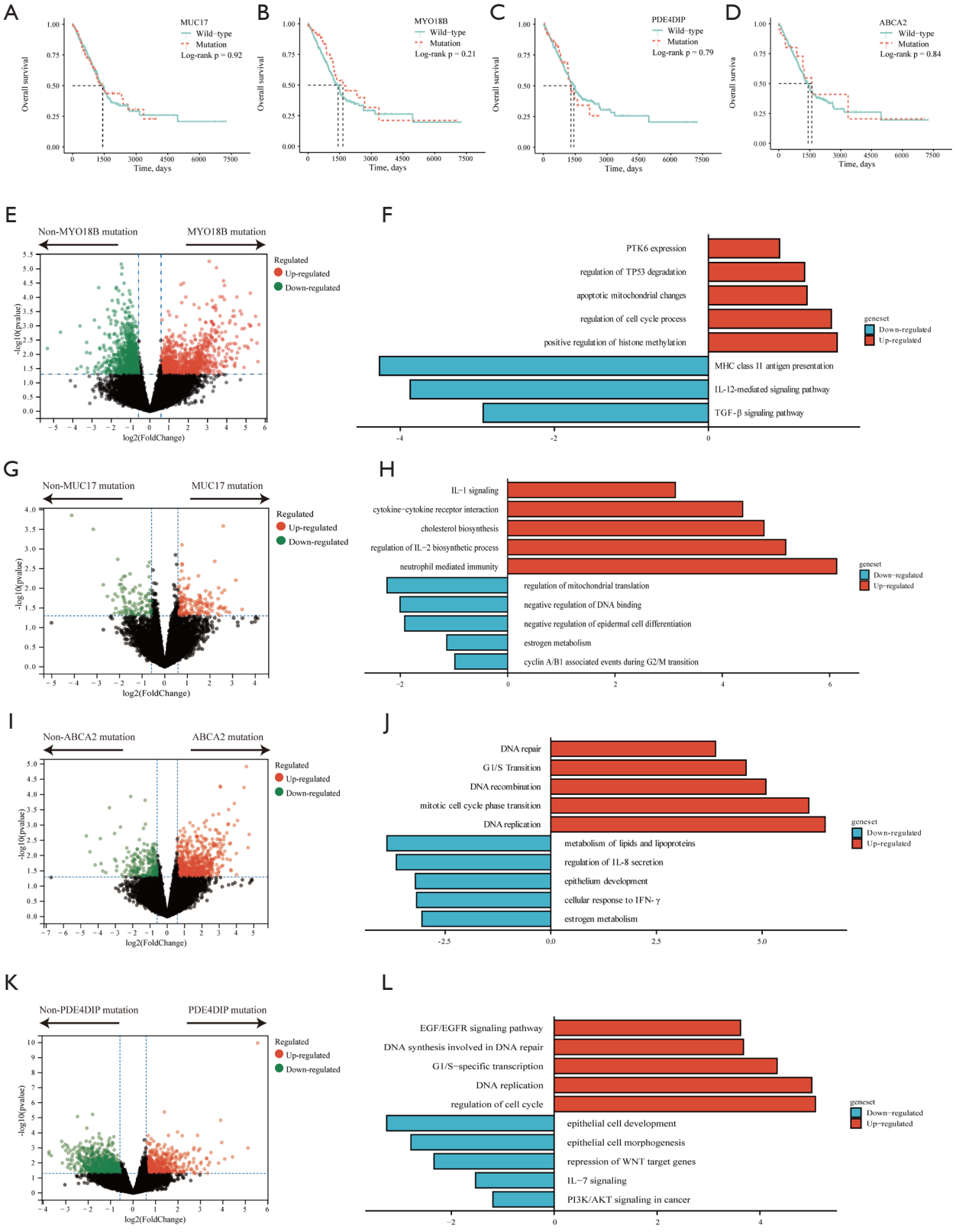

Figure S2 Transcriptomic spectrums and prognostic effects of specific gene mutations of early-stage non-squamous non-small cell lung cancer. Prognostic effects of monogenic mutation, including *MUC17* (A), *MYO18B* (B), *PDE4DIP* (C), and *ABCA2* (D), in the TCGA-LUAD cohort. Differentially expressed genes and corresponding enriched pathways of mutant versus wild-type *MYO18B* (E,F), *MUC17* (G,H), *ABCA2* (I,J), *PDE4DIP* (K,L) in the *GZMU1H* cohort. Red and green dots refer to significantly up-regulated and down-regulated genes, respectively. Black dots represent genes with insignificant changes in expression levels.

https://dx.doi.org/10.21037/tlcr-23-800

Table S4 Differentially expressed genes between tumor nest and paratumor tissues

A2M, ABCC4, ACKR1, ACKR2, ACKR3, ACKR4, ACO1, ACTG1, ACVR1B, ACVR2A, ACVR2B, ACVRL1, ADA2, ADAR, ADCYAP1R1, ADIPOR1, ADIPOR2, ADM2, ADRB1, ADRB2, ADRM1, AEN, AGER, AGRP, AGT, AGTR1, AGTR2, AHNAK, AKT2, AKT3, AMELX, AMHR2, ANGPT1, ANGPT4, ANGPTL1, ANGPTL3, ANGPTL4, ANGPTL5, ANGPTL6, ANGPTL7, ANOS1, ANXA6, AP3B1, APLN, APOBEC3A, APOH, APOM, AQP9, AR, ARG2, ARRB1, ARTN, AVPR1B, AVPR2, AZGP1, BACH2, BCL10, BCL3, BDNF, BECN1, BID, BIRC5, BLNK, BMP1, BMP2, BMP3, BMP5, BMP6, BMP8A, BMP8B, BMPR1A, BMPR2, BPHL, BPIFA1, BPIFA2, BPIFB2, BPIFB4, BRAF, BRD8, BST2, BTK, C5AR1, CACYBP, CALCRL, CALR, CAMP, CANX, CARD11, CASP3, CAT, CBL, CBLB, CBLC, CCK, CCL13, CCL14, CCL15, CCL16, CCL17, CCL18, CCL19, CCL20, CCL22, CCL23, CCL24, CCL25, CCL26, CCL28, CCL3, CCL3L1, CCL4, CCL5, CCL7, CCL8, CCR1, CCR10, CCR4, CCR5, CCR6, CCR7, CCR8, CCR9, CCRL2, CD14, CD19, CD1A, CD1B, CD1E, CD209, CD22, CD244, CD247, CD320, CD3D, CD3E, CD4, CD40LG, CD70, CD74, CD79A, CD79B, CD86, CDC42, CDH1, CDK4, CETP, CGB5, CHGB, CHUK, CIITA, CKLF, CLDN4, CLEC11A, CLEC4M, CMA1, CMKLR1, CMTM2, CMTM3, CMTM5, CMTM6, CMTM8, CNTF, CNTFR, COLEC10, COLEC12, CR2, CRABP1, CRABP2, CREB1, CRH, CRHR1, CRIM1, CRLF1, CRLF3, CSF2, CSF2RB, CSF3, CSF3R, CSH2, CSK, CSPG5, CSRP1, CST4, CTLA4, CTSB, CTSE, CTSG, CX3CL1, CX3CR1, CXCL1, CXCL10, CXCL11, CXCL12, CXCL13, CXCL14, CXCL16, CXCL17, CXCL2, CXCL3, CXCL5, CXCL6, CXCL8, CXCL9, CXCR1, CXCR2, CXCR3, CXCR4, CXCR5, CXCR6, CYBB, CYLD, CYSLTR2, DAXX, DCK, DDX17, DDX58, DEFA1, DEFA3, DEFB1, DEFB124, DEFB125, DEFB126, DEFB127, DEFB4A, DES, DKK1, DLL4, DUOX1, DUOX2, EBI3, ECD, EDN1, EDN2, EDN3, EDNRA, EDNRB, EGF, EGFR, EIF2AK2, ELAVL1, ELN, ENDOU, ENG, EPGN, EPO, EPOR, EPPIN, ERAP2, EREG, ESM1, ESR1, ESR2, ESRRA, ESRRB, F2R, F2RL1, FABP12, FABP4, FABP5, FABP7, FAM3B, FAM3C, FAM3D, FAS, FASLG, FCER1G, FCGR3A, FCGR3B, FCGRT, FGA, FGF1, FGF10, FGF11, FGF12, FGF13, FGF14, FGF17, FGF18, FGF19, FGF2, FGF21, FGF22, FGF23, FGF7, FGF7P6, FGF9, FGFR1, FGFR2, FGFR4, FGFRL1, FGR, FLT1, FLT3, FLT3LG, FLT4, FOS, FPR1, FPR2, FSHR, FURIN, FYN, GAL, GALR2, GCGR, GDF10, GDF11, GDF15, GDF3, GDF5, GDF6, GDF7, GDNF, GFAP, GHR, GHRHR, GHRL, GIPR, GLP2R, GMFB, GMFG, GNAI1, GNLY, GNRH1, GNRH2, GNRHR, GPER1, GPHA2, GPI, GPR17, GRAP2, GRB2, GREM1, GREM2, GRK2, GSK3B, GUCA2A, GZMB, HAMP, HBEGF, HCK, HDAC1, HDGF, HDGFL3, HFE, HGF, HMGB1, HNF4A, HNF4G, HRAS, HRG, HSP90AA1, HSP90AB1, HSPA1A, HSPA1B, HSPA1L, HSPA2, HSPA4, HSPA5, HSPA6, HSPA8, HTR3A, HTR3B, HTR3C, ICAM1, ICAM2, IDO1, IFITM1, IFNA1, IFNA13, IFNA21, IFNA5, IFNAR2, IFNB1, IFNE, IFNGR1, IFNGR2, IFNK, IFNL1, IFNL3, IGF1, IGF1R, IGHA1, IGHA2, IGHD, IGHE, IGHG1, IGHG2, IGHG3, IGHG4, IGHM, IGKC, IGLC1, IGLC2, IGLC3, IGLC6, IGLC7, IKBKB, IKBKE, IL10RA, IL10RB, IL11, IL11RA, IL12A, IL12B, IL12RB1, IL13, IL13RA2, IL15, IL15RA, IL16, IL17D, IL17RA, IL17RC, IL17RD, IL17RE, IL18, IL18R1, IL18RAP, IL1A, IL1B, IL1R1, IL1RL1, IL1RL2, IL2, IL20RA, IL20RB, IL21, IL21R, IL22RA2, IL23A, IL23R, IL27, IL27RA, IL2RA, IL31RA, IL33, IL34, IL36A, IL36G, IL36RN, IL37, IL3RA, IL4R, IL5, IL5RA, IL6, IL6ST, IL7R, ILK, INHBA, INHBB, INPP5D, INSL3, IREB2, IRF1, IRF3, IRF5, IRF7, ISG15, ISG20, ISG20L2, ITGAL, ITGAV, ITK, JAG1, JAK1, JAK2, JUN, JUND, KDR, KIR2DL1, KIR2DL3, KIR2DS4, KIR3DL1, KIR3DL2, KITLG, KL, KLKB1, KLRC1, KLRC2, KLRC3, KLRC4, KLRD1, KLRK1, LANCL1, LAT, LBP, LCN10, LCN12, LCN2, LCN6, LCNL1, LCP2, LEAP2, LEFTY2, LEPR, LGMN, LGR4, LGR5, LHCGR, LIF, LIFR, LILRB3, LMBR1, LMBR1L, LPA, LRP1, LRSAM1, LTA, LTB, LTB4R, LTB4R2, LTBP2, LTBP4, LTF, LYN, LYZ, MALT1, MANF, MAP2K1, MAP2K2, MAP3K14, MAP3K8, MAPK1, MAPK3, MAPK8, MAPT, MARCO, MASP1, MASP2, MAVS, MBL2, MC1R, MC2R, MC4R, MDK, MET, MIA, MICA, MIF, MMP12, MMP9, MPL, MPO, MSR1, MSTN, MTNR1A, MUC4, MUC5AC, MX1, MX2, MYDGF, NCK1, NCK2, NCR1, NCR2, NDP, NEDD4, NENF, NEO1, NFAT5, NFATC1, NFATC2, NFATC3, NFKBIA, NFKBIB, NFKBIE, NFKBIZ, NFYB, NGF, NLRX1, NMB, NMBR, NOD1, NOD2, NODAL, NOS1, NOS2, NOX1, NOX4, NPFF, NPPA, NPPC, NPR1, NPR3, NR1D1, NR1D2, NR1H2, NR1H3, NR1I3, NR2C1, NR2C2, NR2E1, NR2E3, NR2F1, NR2F2, NR2F6, NR3C1, NR3C2, NR4A1, NR4A2, NR4A3, NR5A2, NR6A1, NRG1, NRG2, NRG3, NRG4, NRP1, NRP2, NRTN, NTF3, NUDT6, OAS1, OASL, OBP2A, OGFR, OGN, OLR1, OPRD1, OPRK1, OPRL1, ORM1, ORM2, OSGIN1, OSM, OSMR, OSTN, OXT, OXTR, PAEP, PAK1, PAK2, PAK4, PAK6, PDF, PDGFA, PDGFB, PDGFC, PDGFD, PDGFRA, PDGFRB, PDGFRL, PDIA3, PDK1, PF4, PGC, PGF, PGLYRP1, PGLYRP2, PGLYRP3, PGLYRP4, PGR, PI15, PI3, PIK3CA, PIK3CB, PIK3CG, PIK3R1, PIK3R2, PIK3R5, PLA2G2A, PLAU, PLAUR, PLCG1, PLCG2, PLSCR1, PLTP, PLXNA2, PLXNA3, PLXNA4, PLXNB1, PLXNB2, PLXNB3, PLXNC1, PLXND1, PMCH, PML, PNOC, PPARG, PPBP, PPBPP2, PPIA, PPP3CA, PPP3CB, PPP3CC, PPP3R2, PPP4C, PRDX1, PRDX2, PRF1, PRKCA, PRKCB, PRKCG, PRKCQ, PROC, PROK2, PSMB8, PSMC1, PSMC2, PSMC3, PSMC4, PSMD1, PSMD10, PSMD11, PSMD13, PSMD14, PSMD2, PSMD3, PSMD4, PSMD6, PSMD7, PSMD8, PSME1, PSME2, PSME3, PSPN, PTAFR, PTGDR, PTGDR2, PTGDS, PTGER2, PTGER3, PTGER4, PTGFR, PTH1R, PTHLH, PTK2B, PTN, PTPN11, PTPN6, PTPRC, PTX3, QRFP, RABEP1, RABEP2, RAC1, RAC3, RAET1E, RAET1L, RAF1, RARA, RARB, RASGRP3, RBP1, RBP2, RBP4, RBP5, REG1A, RELB, RETN, RFX5, RFXANK, RFXAP, RHOA, RLN2, RLN3, RNASE2, RNASE7, RNASEL, ROBO1, ROBO2, ROBO3, RORA, RORC, RSAD2, RXFP1, RXFP2, RXRA, RXRB, RXRG, S100A1, S100A10, S100A11, S100A13, S100A14, S100A16, S100A2, S100A3, S100A6, S100A7, S100A9, S100B, S100P, S100Z, S1PR1, S1PR2, SAA1, SAA2, SBDS, SCGB3A1, SDC1, SDC2, SDC3, SDC4, SECTM1, SEMA3B, SEMA3D, SEMA3E, SEMA3F, SEMA3G, SEMA4B, SEMA4C, SEMA4F, SEMA4G, SEMA5A, SEMA6A, SEMA6C, SEMA6D, SEMA7A, SERPIND1, SFTPA1, SFTPA2, SFTPD, SH2D1A, SH2D1B, SH3BP2, SHC1, SHC2, SHC3, SKIV2L, SLC11A1, SLC40A1, SLIT2, SLPI, SOCS1, SOD1, SORT1, SOS1, SOS2, SP1, SPINK5, SPP1, SST, SSTR1, SSTR5, STAB2, STAT1, STAT3, SYTL1, TACR1, TANK, TAP1, TAP2, TAPBP, TBK1, TCF7L2, TEC, TEK, TFR2, TFRC, TGFA, TGFB2, TGFBR2, TGFBR3, THRA, THRB, TIE1, TINAGL1, TK2, TKFC, TLR1, TLR3, TLR4, TLR8, TMPRSS6, TMSB10, TMSB15A, TMSB15B, TMSB4X, TMSB4XP8, TNC, TNFAIP3, TNFRSF10A, TNFRSF10B, TNFRSF10C, TNFRSF10D, TNFRSF11A, TNFRSF12A, TNFRSF13B, TNFRSF13C, TNFRSF14, TNFRSF17, TNFRSF18, TNFRSF19, TNFRSF1B, TNFRSF21, TNFRSF4, TNFRSF6B, TNFRSF8, TNFRSF9, TNFSF11, TNFSF12, TNFSF13, TNFSF14, TNFSF15, TNFSF18, TNFSF4, TNFSF8, TNFSF9, TPM2, TRAF3, TRAV19, TRAV2, TRAV21, TRAV22, TRAV27, TRAV29DV5, TRAV3, TRAV4, TRAV5, TRBV19, TRBV27, TRBV28, TRBV30, TRBV9, TRDC, TRDV3, TRGC1, TRGC2, TRGV2, TRGV3, TRGV4, TRGV5, TRGV8, TRGV9, TRH, TRHR, TRIM22, TRIM27, TRPC4AP, TSHB, TSLP, TUBB3, TXK, TXLNA, TYK2, TYMP, TYROBP, UBR1, UCN, UCN2, UCN3, ULBP2, ULBP3, UMODL1, UNC93B1, UTS2B, VAV2, VAV3, VCAM1, VDR, VEGFA, VEGFB, VEGFC, VEGFD, VGF, VIM, VIP, VIPR1, WFDC2, WFIKKN1, WNT5A, XCL1, XCL2, XCR1, ZAP70, ZC3HAV1, ZC3HAV1L, ZYX

https://dx.doi.org/10.21037/tlcr-23-800

Table S5 Gene expression levels with prognostic significance in the univariate Cox regression analysis

Gene	P value	HR	95% CI_L	95% CI_U
CCR9	0.017	3.28E-05	6.65E-09	0.162
KLRC4	0.023	1.66E-04	8.95E-08	0.309
KIR2DL3	0.050	0.004	1.69E-05	0.996
KLRC1	0.050	0.007	5.36E-05	0.997
KLRD1	0.004	0.009	3.46E-04	0.218
NCR1	0.016	0.009	2.06E-04	0.413
IL27	0.033	0.017	4.26E-04	0.718
TXK	0.023	0.130	0.022	0.758
PTGDR	0.040	0.137	0.021	0.915
CD244	0.025	0.160	0.032	0.793
IL18RAP	0.044	0.196	0.040	0.960
DUOX2	0.013	0.241	0.078	0.745
IL18R1	0.029	0.251	0.073	0.866
LCN10	0.049	0.266	0.071	0.995
CXCR6	0.007	0.266	0.102	0.698
KLRK1	0.018	0.290	0.105	0.805
GPR17	0.044	0.310	0.099	0.970
PPP3CC	0.045	0.318	0.104	0.974
HDGFL3	0.047	0.347	0.123	0.984
PLCG2	0.009	0.366	0.172	0.778
TRAV3	0.045	0.379	0.147	0.979
IL34	0.049	0.396	0.158	0.995
SEMA6A	0.042	0.415	0.178	0.969
GIPR	0.032	0.421	0.191	0.928
ICAM2	0.024	0.438	0.214	0.896
PTK2B	0.003	0.442	0.257	0.758
JAK2	0.035	0.447	0.211	0.944
BMP6	0.012	0.461	0.253	0.841
TRIM22	0.007	0.475	0.276	0.819
S1PR1	0.029	0.492	0.260	0.930
EPOR	0.047	0.494	0.246	0.991
VIPR1	0.042	0.526	0.283	0.977
ADRB2	0.030	0.624	0.408	0.955
PI3	0.030	0.637	0.424	0.957
INHBB	0.026	0.679	0.482	0.956
DUOX1	0.027	0.708	0.522	0.961
CD79A	0.028	0.724	0.542	0.966
IGHG3	0.041	0.805	0.655	0.991
PLAU	0.031	1.310	1.020	1.670
SPP1	0.005	1.347	1.090	1.660
TUBB3	0.050	1.381	1.000	1.910
HTR3A	0.046	1.453	1.010	2.100
CRABP2	0.001	1.532	1.190	1.970
PLXNB3	0.040	1.542	1.020	2.330
F2RL1	0.027	1.585	1.050	2.380
PPIA	0.035	1.609	1.030	2.500
IL13RA2	0.015	1.621	1.100	2.390
EGFR	0.028	1.643	1.060	2.560
PLAUR	0.017	1.663	1.090	2.530
PGLYRP4	0.043	1.715	1.020	2.890
ULBP2	0.008	1.747	1.160	2.640
LGR4	0.006	1.880	1.190	2.960
IL31RA	0.003	2.062	1.290	3.310
IL11	0.002	2.163	1.330	3.510
PSMD4	0.041	2.207	1.030	4.720
GDF5	0.016	2.486	1.180	5.230
NGF	0.038	3.060	1.060	8.800
CCR8	0.034	3.826	1.100	13.300
HFE	0.044	3.925	1.040	14.800
EPGN	0.038	4.016	1.080	14.900
IL1A	0.008	4.366	1.470	13.000
CRHR1	0.022	100.150	1.940	5.17E+03
HTR3B	0.044	1.03E+03	1.190	8.97E+05
PPBPP2	0.002	2.32E+04	39.500	1.36E+07
MBL2	0.010	1.42E+05	17.500	1.16E+09

HR, hazard ratio; CI_L, lower bounds of the 95% confidence interval; CI_U, upper bounds of the 95% confidence interval.

https://dx.doi.org/10.21037/tlcr-23-800

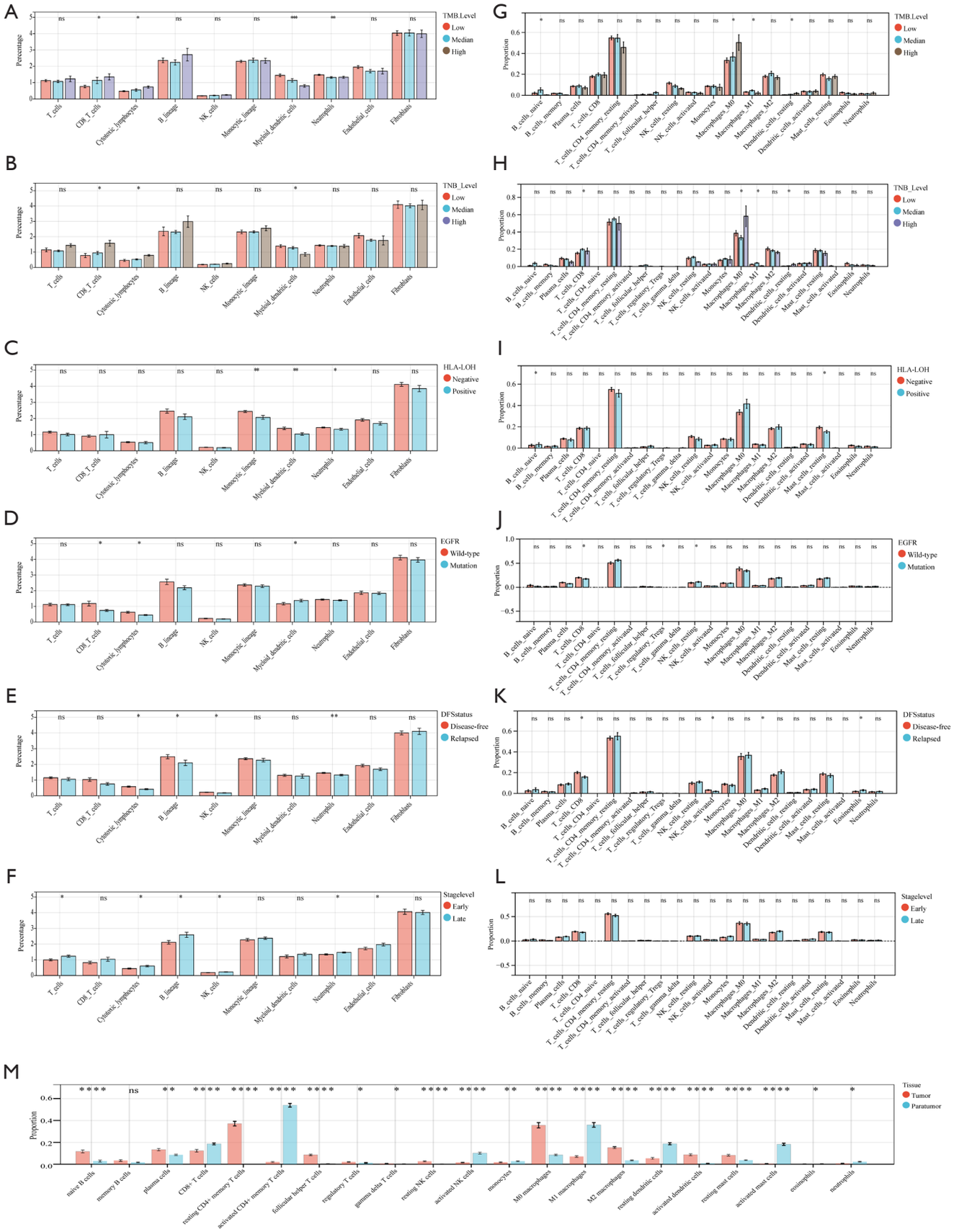

Figure S3 Immune infiltration differences between tumor nest and adjacent tissues. Intratumoral immune infiltration differences between different tumor mutational burden (TMB) levels (A), tumor neoantigen burden (TNB) levels (B), human leukocyte antigen loss of heterozygosity (HLA-LOH) status (C), EGFR mutation status (D), disease-free survival (DFS) status (E), and cTNM stage level (F), as evaluated by the MCP-counter algorithm. Intrastromal immune infiltration differences between different TMB levels (G), TNB levels (H), HLA-LOH status (I), EGFR mutation status (J), DFS status (K), and cTNM stage level (L), as evaluated by the CIBERSORT algorithm. Comparison of immune infiltration difference between tumor nest and paratumor tissue (M). Comparison of continuous data by Kruskal-Wallis test. *, P<0.05; **, P<0.01; ****, P<0.0001; ns, non-significant.

https://dx.doi.org/10.21037/tlcr-23-800

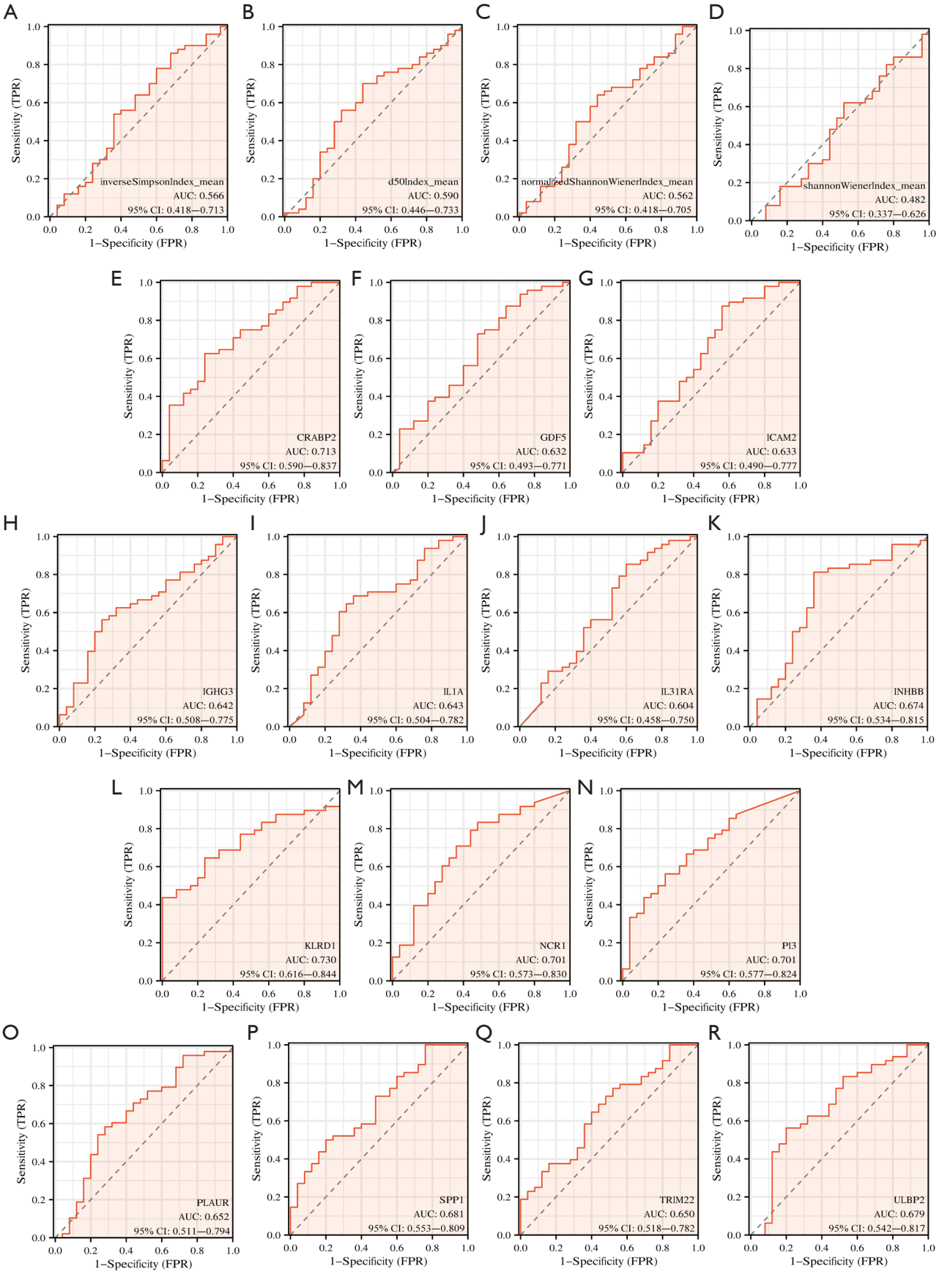

Figure S4 The predictive accuracy of single omics/biomarker in disease-free survival. T cell receptor repertoire diversity features (A-D), and transcriptomic characteristics (E-R) showed limited performance in predicting prognosis.

https://dx.doi.org/10.21037/tlcr-23-800

Table S6 Performance of the machine learning algorithms in the training cohort

	Accuracy	Precision	Recall	F1-score	AUC
Clinical + RNA					
SVM	0.833	0.813	0.650	0.722	0.818
DTC	0.933	0.900	0.900	0.900	0.978
ETC	1.000	1.000	1.000	1.000	1.000
GBC	0.950	1.000	0.850	0.919	0.998
GPC	0.975	0.975	0.975	0.975	0.999
KNN	0.767	0.800	0.400	0.533	0.815
RF	0.983	1.000	0.950	0.974	0.975
Clinical +DNA+RNA+TCR					
SVM	0.733	0.625	0.500	0.556	0.735
DTC	0.917	0.857	0.900	0.878	0.973
ETC	1.000	1.000	1.000	1.000	1.000
GBC	0.667	0	0	0	0.986
GPC	0.813	0.931	0.675	0.783	0.904
KNN	0.800	0.833	0.500	0.625	0.869
RF	0.933	1.000	0.800	0.889	1.000
Clinical + DNA					
SVM	0.767	0.750	0.450	0.563	0.829
DTC	0.983	1.000	0.950	0.974	0.999
ETC	0.900	1.000	0.700	0.824	0.988
GBC	1.000	1.000	1.000	1.000	1.000
GPC	0.975	0.975	0.975	0.975	0.997
KNN	0.783	0.733	0.550	0.629	0.803
RF	0.833	0.917	0.550	0.687	0.936
Clinical + TCR					
SVM	0.717	0.600	0.450	0.514	0.685
DTC	0.917	0.941	0.800	0.865	0.968
ETC	1.000	1.000	1.000	1.000	1.000
GBC	1.000	1.000	1.000	1.000	1.000
GPC	0.875	0.917	0.825	0.868	0.890
KNN	0.767	0.800	0.400	0.533	0.804
RF	0.817	0.846	0.550	0.667	0.926
DNA+RNA					
SVM	0.783	0.733	0.550	0.629	0.860
DTC	0.933	0.833	1.000	0.909	0.972
ETC	0.917	1.000	0.750	0.857	0.992
GBC	0.667	0	0	0	0.973
GPC	1.000	1.000	1.000	1.000	1.000
KNN	0.733	0.667	0.400	0.500	0.819
RF	0.983	1.000	0.950	0.974	1.000
DNA+RNA+TCR					
SVM	0.650	0.455	0.250	0.323	0.695
DTC	0.967	0.950	0.950	0.950	0.997
ETC	0.983	1.000	0.950	0.974	1.000
GBC	1.000	1.000	1.000	1.000	1.000
GPC	0.788	0.829	0.725	0.773	0.869
KNN	0.800	0.900	0.450	0.600	0.850
RF	0.967	1.000	0.900	0.947	1.000

SVM, support vector machine; DTC, decision tree classifier; ETC, extra tree classifier; GBC, gradient boosting classifier; GPC, Gaussian process classifier; KNN, K-nearest neighbors; RF, random forest.

https://dx.doi.org/10.21037/tlcr-23-800

Table S7 Performance of the machine learning algorithms in the testing cohort

	Accuracy	Precision	Recall	F1-score	AUC
Clinical + RNA					
DTC	0.688	0.600	0.500	0.545	0.692
ETC	0.875	1.000	0.667	0.800	0.933
GBC	0.688	1.000	0.167	0.286	0.800
GPC	0.688	1.000	0.167	0.286	0.733
KNN	0.750	1.000	0.333	0.500	0.867
RF	0.813	1.000	0.500	0.667	0.833
SVM	0.813	1.000	0.500	0.667	0.900
Clinical +DNA+RNA+TCR					
DTC	0.625	0.500	0.333	0.400	0.667
ETC	0.688	0.667	0.333	0.444	0.833
GBC	0.625	0	0	0	0.717
GPC	0.750	0.625	0.833	0.714	0.817
KNN	0.625	0	0	0	0.933
RF	0.625	0	0	0	0.817
SVM	0.563	0.400	0.333	0.364	0.617
Clinical + DNA					
DTC	0.625	0.500	0.333	0.400	0.567
ETC	0.625	0.500	0.167	0.250	0.717
GBC	0.625	0.500	0.333	0.400	0.433
GPC	0.563	0	0	0	0.517
KNN	0.875	0.833	0.833	0.833	0.767
RF	0.625	0	0	0	0.733
SVM	0.688	1.000	0.167	0.286	0.750
Clinical + TCR					
DTC	0.563	0.400	0.333	0.364	0.508
ETC	0.688	1.000	0.167	0.286	0.650
GBC	0.625	0	0	0	0.650
GPC	0.813	0.714	0.833	0.769	0.850
KNN	0.750	1.000	0.333	0.500	0.833
RF	0.688	1.000	0.167	0.286	0.600
SVM	0.625	0.500	0.167	0.250	0.450
DNA+RNA					
DTC	0.750	0.750	0.500	0.600	0.683
ETC	0.625	0.500	0.167	0.250	0.700
GBC	0.625	0	0	0	0.833
GPC	0.625	0.500	0.333	0.400	0.817
KNN	0.688	0.667	0.333	0.444	0.817
RF	0.688	1	0.167	0.286	0.767
SVM	0.625	0.500	0.167	0.250	0.667
DNA+RNA+TCR					
DTC	0.625	0.500	0.167	0.250	0.608
ETC	0.625	0.500	0.167	0.250	0.683
GBC	0.688	1.000	0.167	0.286	0.633
GPC	0.875	0.833	0.833	0.833	0.900
KNN	0.625	0	0	0	0.758
RF	0.625	0	0	0	0.683
SVM	0.563	0	0	0	0.483

DTC, decision tree classifier; ETC, extra tree classifier; GBC, gradient boosting classifier; GPC, Gaussian process classifier; KNN, K-nearest neighbors; RF, random forest; SVM, support vector machine.

https://dx.doi.org/10.21037/tlcr-23-800