Molecular Therapy
Oncolytics
Original Article

Tumor immune microenvironment-based classifications of bladder cancer for enhancing the response rate of immunotherapy

Jialin Meng,[1,6] Xiaofan Lu,[2,6] Yujie Zhou,[3,6] Meng Zhang,[1,4] Qintao Ge,[1] Jun Zhou,[1] Zongyao Hao,[1] Shenglin Gao,[5] Fangrong Yan,[2] and Chaozhao Liang[1]

[1]Department of Urology, The First Affiliated Hospital of Anhui Medical University, Institute of Urology, Anhui Medical University, Anhui Province Key Laboratory of Genitourinary Diseases, Anhui Medical University, Hefei 230022, P.R. China; [2]State Key Laboratory of Natural Medicines, Research Center of Biostatistics and Computational Pharmacy, China Pharmaceutical University, Nanjing 211198, P.R. China; [3]Division of Gastroenterology and Hepatology, Key Laboratory of Gastroenterology and Hepatology, Ministry of Health, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai Institute of Digestive Disease, Shanghai 200127, P.R. China; [4]Urology Institute of Shenzhen University, The Third Affiliated Hospital of Shenzhen University, Shenzhen University, Shenzhen 518000, P.R. China; [5]Department of Urology, The Affiliated Changzhou No. 2 People's Hospital of Nanjing Medical University, Changzhou, Jiangsu, P.R. China

Immunotherapy is a potential way to save the lives of patients with bladder cancer, but it only benefits approximately 20% of them. A total of 4,028 bladder cancer patients were collected for this study. Unsupervised non-negative matrix factorization and the nearest template prediction algorithms were employed for the classification. We identified the immune and non-immune classes from The Cancer Genome Atlas Bladder Urothelial Carcinoma (TCGA-BLCA) training cohort. The 150 most differentially expressed genes between these two classes were extracted, and the classification reappeared in 20 validation cohorts. For the activated and exhausted subgroups, a stromal activation signature was assessed by the NTP algorithm. Patients in the immune class showed highly enriched signatures of immunocytes, while the exhausted subgroup also exhibited activated transforming growth factor (TGF)-β1, and cancer-associated extracellular matrix signatures. Patients in the immune-activated subgroup showed a lower genetic alteration and better overall survival. Anti-PD-1/PD-L1 immunotherapy was more beneficial for the immune-activated subgroup, while immune checkpoint blockade therapy plus a TGF-β inhibitor or an EP300 inhibitor might achieve greater efficacy for patients in the immune-exhausted subgroup. Novel immune molecular classifier was identified for the innovative immunotherapy of patients with bladder cancer.

INTRODUCTION

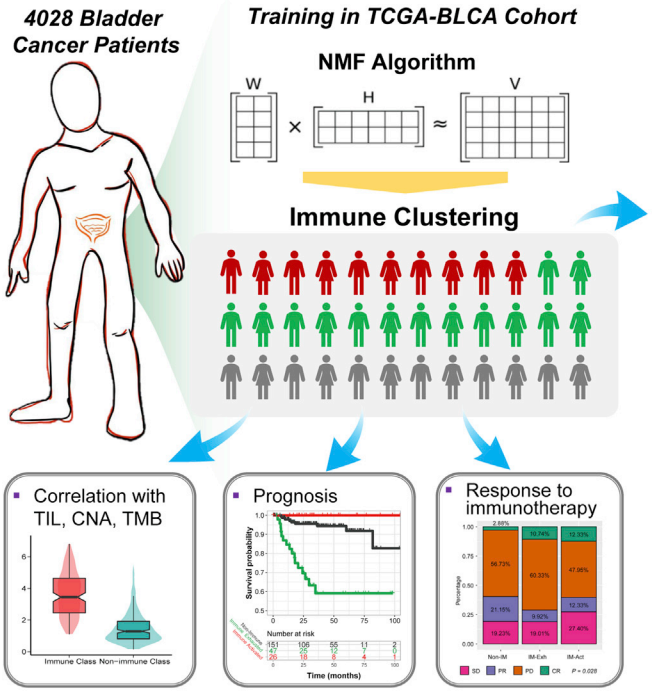
Bladder cancer is the 10th most frequent tumor globally and exhibits a high rate of recurrence.[1] The major challenge of clinical care of bladder cancer is the short-term recurrence of non-muscle-invasive bladder cancer (NMIBC), as well as the shortened overall survival of muscle-invasive bladder cancer (MIBC) patients, especially those with distant metastases, the 5-year survival rate of whom less than 10%.[2,3] In the tumor mass, the normal cells, blood vessels, and cytokines that surround and support the vitality of tumor cells compose the tumor microenvironment (TME). Crosstalk exists between the tumor and the TME. Tumor cells can alter the TME, and the TME can promote the growth and spread of tumors.

From the molecular side, bladder cancer is composed of a multitude of heterogenetic characteristics, including gene mutations, gene copy number alterations, and neoantigens, as well as the infiltration of immunocytes. Several teams have established molecular classifications among bladder cancer. Mo et al.[4] generated an 18-gene tumor signature in MIBC patients that can reflect the urothelial differentiation and predict clinical outcomes; basal and differentiated groups are the two groups with the highest and lowest risk scores, respectively. Damrauer et al.[5] developed BASE47, a 47 gene-based classifier, for the separate of luminal-like or basal-like subtypes of MIBC tumors. Robertson et al.[6] further generated a consensus hierarchical clustering of luminal-papillary, luminal-infiltrated, luminal, basal/squamous, and neuronal subtypes. However, most of the molecular classifiers only focused on the clinical outcomes, not the tumor immune microenvironment. Therefore, our goals were to provide comprehensive insight into the immune response of bladder cancer patients with

Received 31 December 2020; accepted 1 February 2021;
https://doi.org/10.1016/j.omto.2021.02.001.

[6]These authors contributed equally

Correspondence: Shenglin Gao, Department of Urology, The Affiliated Changzhou No. 2 People's Hospital of Nanjing Medical University, Xinglong Road 29, Changzhou, Jiangsu 213003, P.R. China.
E-mail: gsl_cmu@163.com
Correspondence: Fangrong Yan, State Key Laboratory of Natural Medicines, Research Center of Biostatistics and Computational Pharmacy, China Pharmaceutical University, 24th, Tongjia Alley, Nanjing 211198, P.R. China.
E-mail: f.r.yan@163.com
Correspondence: Chaozhao Liang, Department of Urology, The First Affiliated Hospital of Anhui Medical University, Institute of Urology and Anhui Province Key Laboratory of Genitourinary Diseases, Anhui Medical University, 218th Jixi Road, Hefei 230022, P.R. China.
E-mail: liang_chaozhao@ahmu.edu.cn



Validation in 20 External Cohorts

Dataset	N
E-MTAB-4321	476
IMvigor210	348
GSE32894	308
GSE83586	307
GSE87304	305
GSE128702	256
GSE13507	164
GSE129871	158
GSE120736	145
GSE39016	141
GSE128701	136
GSE124035	133
GSE86411	132
GSE48276	116
GSE128192	112
GSE31684	93
GSE134292	80
GSE93527	79
E-MTAB-1803	70
GSE69795	61

Figure 1. Flow chart of the analyses performed in this study

NMF, non-negative matrix factorization; TCGA-BLCA, The Cancer Genome Atlas Bladder Urothelial Carcinoma; TILs, tumor-infiltrating lymphocytes; CNA, copy number alteration; TMB, tumor mutation burden.

diverse inner molecular features and generate the classifier to screen the patients suited for immunotherapy.

The non-negative matrix factorization (NMF) algorithm is a multiplicative updates algorithm; it can decompose a non-negative matrix V into two non-negative matrices, W and H.[7] Similar to principal component analysis (PCA) or independent component analysis (ICA), the NMF algorithm can also use a limited number of components to reflect the original observed data, which might contain huge volumes.[8] NMF has been applied to reveal biomarkers, classify tumor subtypes, and predict the prognosis of tumors in recent works.[9–12]

We enrolled 4,028 patients with bladder cancer from independent cohorts. The NMF and nearest template prediction (NTP) algorithms were applied to distinguish patients with different immunophenotypes in The Cancer Genome Atlas Bladder Urothelial Carcinoma (TCGA-BLCA) training cohort and reappeared in the validation cohorts. The novel definition of these immunophenotypes could provide illumination for the immunotherapy of patients with bladder cancer.

RESULTS

Identification of the immune module and derivation of the immune class of bladder cancer

4,028 bladder cancer patients were involved, along with the matched overall survival data, clinicopathological information, and gene expression profiles (Figure 1). We performed virtual microdissection using the NMF algorithm in the TCGA-BLCA training cohort. To obtain the robust immune module, we preset the respective module numbers as five to 10. When the total module number was nine, the first module strongly enriched patients with high immune enrichment scores, which were defined as the immune module (Figure 2A). The top 150 weighted genes in the immune module were defined as exemplar genes that reflected the characteristics of the immune module (Table S1). According to the ontological analysis of biological processes, these genes are associated with the activation of immunocytes, T helper 1 (Th1)/Th2 cell differentiation, T cell receptor signaling, and B cell receptor signaling (all $p < 0.05$; Table S2). Subsequently, we redefined the 408 bladder patients into immune-enriched or non-immune-enriched groups via the consensus clustering analysis of the 150 exemplar genes (Figure 2B). Furthermore, multidimensional scaling (MDS) random forest (RF) was further applied to define a more precise classification for the immune and non-immune classes (Figure 2C). In Figure 2D, the distributions of the 408 bladder cancer patients among the NMF modules, immune module weight, exemplar gene clustering, final immune classes, and immune enrichment score are shown.

Several immune-associated signatures (Table S3) were collected to help confirm the classification of the immune or non-immune classes, and the enrichment score of each signature for each patient was determined by single-sample gene set enrichment analysis (ssGSEA). We observed the increased enrichment of immunocytes in the immune class as compare with the non-immune class, including T cells (as reflected by the signatures of 13 T cell signature, T cells, CD8+ T cells, and T. NK. Metagene), B cells (as reflected by the signatures of B cell clusters and B.P. metagene), macrophages, tertiary lymphoid structure (TLS), cytolytic activity score (CYT), and interferon (IFN) signatures (all $p < 0.05$; Figure 3A). We also analyzed the activated KEGG signaling pathways by GSEA, revealing that immune cell pathways (including T cell-, B cell-, natural killer cell-, and leukocyte-associated pathways), immune response pathways (including chemokine signaling pathways, antigen processing presentation, cell adhesion molecules, and complement coagulation cascades), and proinflammatory pathways (including FC-Epsilon-RI-, NOD-like receptor-, and FC gamma R-mediated phagocytosis pathways) were all activated in the immune class (Figure S1). From the results of Figures 2, 3A (top panel), and S1 and Tables S1, S2, and S3, we microdissected the immune and non-immune classes in the TCGA-BLCA cohort, and activated immune-associated signatures and signaling pathways were observed in the immune class.

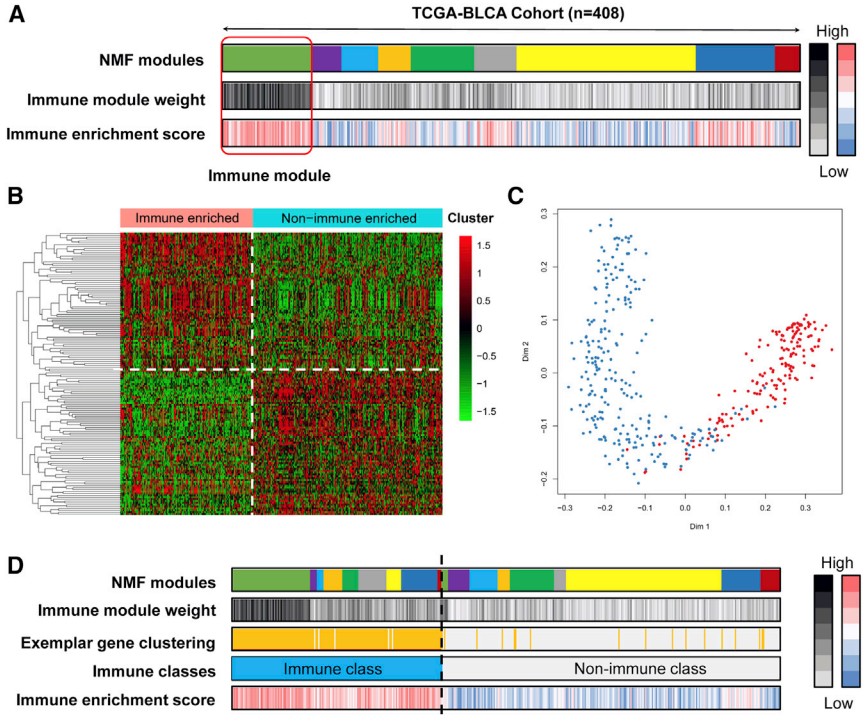


Figure 2. Recognition of the immune classes by the non-negative matrix factorization (NMF) algorithm

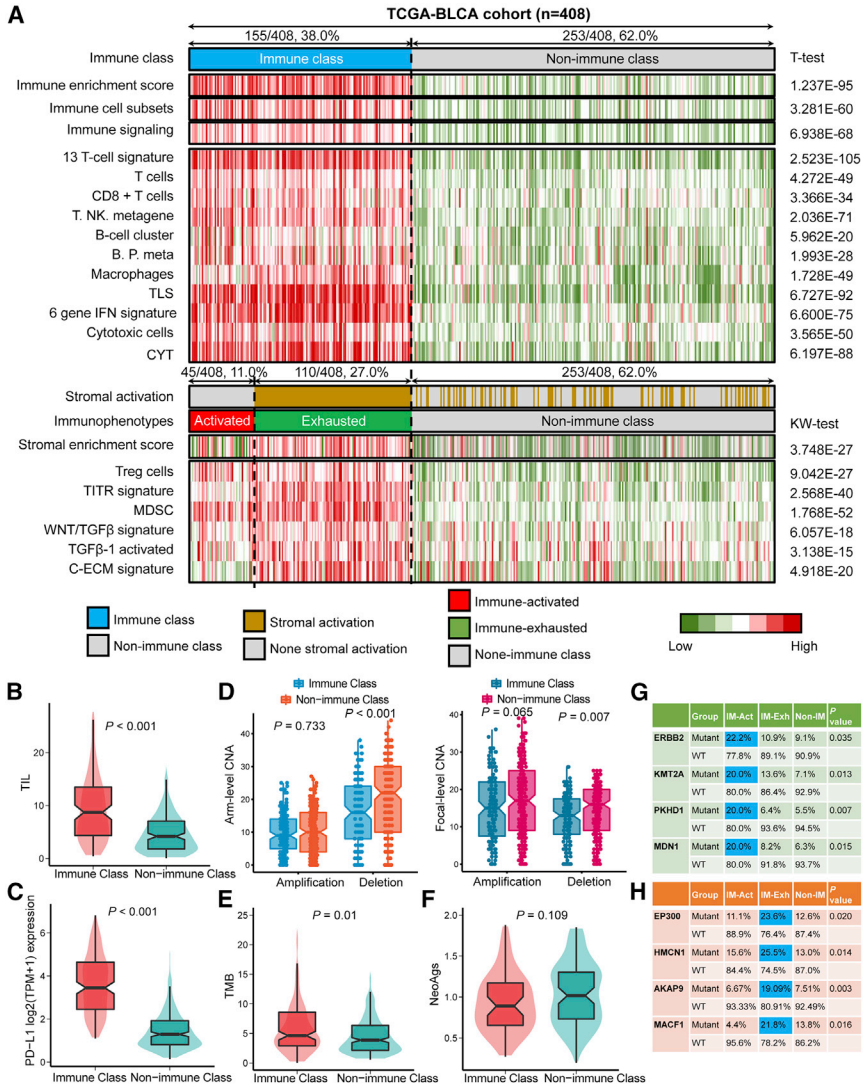
(A) Nine modules were generated from the NMF algorithm, and the module gathered patients with high immune enrichment score were recognized as the immune module. (B) Heatmap showing the top 150 exemplar genes expression among immune-enriched and non-immune-enriched clusters, divided by consensus clustering. (C) The multidimensional scaling random forest further modified the clusters to immune and non-immune classes. (D) The distributions of patients in different NMF modules, immune module weight, exemplar gene clustering, final immune classes, and immune enrichment score.

Tumor immune microenvironment immunophenotypes are distinguished by the activation of stromal cells

Fibroblasts, mesenchymal stromal cells (MSCs), and the extracellular matrix (ECM) are the key components of the tumor stroma and support and connect tumor cells.[13] Especially during the late stages of tumors, genetic and epigenetic alterations of tumor cells are driven by activated stromal components.[14] MSCs act as inherent regulators of tumors and can secrete inhibiting soluble factors and alter cell surface markers to suppress the immune microenvironment. MSCs can regulate the expression of PD-L1 and impact the proliferation and induction of T regulatory cells (Tregs).[15,16] MSCs suppress immune processes by decreasing the expression of proinflammatory factors, including IFN-γ, tumor necrosis factor (TNF)-α, and interleukin (IL)-1β, or by promoting the expression of type 2 factors IL-10 and IL13.[17–19] For this reason, the previously defined stromal activated signature was used for the further separation of patients with high immunocyte infiltration to immune-activated and immune-exhausted immunophenotypes, which could reflect the immune response status. A total of 11.0% (45/408) of bladder cancer patients in the TCGA-BLCA cohort were recognized as the immune-activated subgroup, while the remaining 110 patients (27.0%, 110/408) belonged to the immune-exhausted subgroup, with the activated stromal phenotype (Table 2; Figure 3A). Cancer-associated ECM (C-EC) regulated by the fibroblasts can recruit immunosuppressive cells, transforming growth factor (TGF)-β is an accepted immunosuppressor in the immune microenvironment, and Tregs and MDSCs can reflect the immune-exhausted status in the TME.[20–23] These signatures were evaluated by ssGSEA, and we revealed that the tumor-infiltrating Tregs (TITRs), WNT/TGF-β, TGF-β1-activated, and C-ECM signatures were higher in the immune-exhausted subgroup than in the immune-activated subgroup (all $p < 0.05$; Figure 3A; Figure S2). TIM-3 and LAG3 are reported to be associated with immune exhaustion status,[24,25] and we also found similar results in the immune-exhausted subgroups; increased TIM-3 ($p = 0.008$) and LAG3 ($p = 0.218$) were observed in the immune-exhausted subgroup (Figure S2). Based on the results from Figure 3A (bottom panel) and Figure S2, we separated the immune class into the immune-activated and immune-exhausted subgroups. The stromal enrichment score, TITR, myeloid-derived suppressor cell (MDSC), and WNT/TGF-β signatures increased in the immune-exhausted subgroup, as validated by the immune-exhausted markers TIM-3 and LAG3.

Heterogeneity of genetic phenotypes among the immune classes

To confirm the infiltration of immunocytes among the immune and non-immune classes, as distinguished by the mRNA expression profiles of the exemplar genes, we compared the tumor-infiltrating lymphocyte (TIL) abundance of 408 bladder cancer patients, which was anteriorly estimated by hematoxylin and eosin (H&E) staining,[26] and observed that the TIL abundance was higher in the immune class than in the non-immune class ($p < 0.001$; Figure 3B), consistent with the definition of these two groups. Furthermore, we also observed a higher PD-L1 expression in the immune class than in the non-immune class ($p < 0.001$; Figure 3C). Gene copy number alteration (CNA), tumor mutant burden (TMB), and neoantigens were reported to exhibit crosstalk with tumor immune activation. Patients in the non-immune class showed increased levels of deletion at both the arm and focal levels ($p_{Arm\text{-}del} < 0.001$, $p_{Focal\text{-}del} = 0.007$) but not CNA amplification ($p_{Arm\text{-}Amp} = 0.733$, and $p_{Focal\text{-}Amp} = 0.065$) (Figure 3D), which reflected the positive association of immune infiltration and gene CNA deletion. With the online tool TIMER, we twice confirmed that the association between immune infiltration and gene CNA deletion; deep deletion and arm-level deletion of PD-1, PD-L1 and CTLA-4; and the three major immune checkpoints



	Group	IM-Act	IM-Exh	Non-IM	P value
ERBB2	Mutant	22.2%	10.9%	9.1%	0.035
	WT	77.8%	89.1%	90.9%	
KMT2A	Mutant	20.0%	13.6%	7.1%	0.013
	WT	80.0%	86.4%	92.9%	
PKHD1	Mutant	20.0%	6.4%	5.5%	0.007
	WT	80.0%	93.6%	94.5%	
MDN1	Mutant	20.0%	8.2%	6.3%	0.015
	WT	80.0%	91.8%	93.7%	

	Group	IM-Act	IM-Exh	Non-IM	P value
EP300	Mutant	11.1%	23.6%	12.6%	0.020
	WT	88.9%	76.4%	87.4%	
HMCN1	Mutant	15.6%	25.5%	13.0%	0.014
	WT	84.4%	74.5%	87.0%	
AKAP9	Mutant	6.67%	19.09%	7.51%	0.003
	WT	93.33%	80.91%	92.49%	
MACF1	Mutant	4.4%	21.8%	13.8%	0.016
	WT	95.6%	78.2%	86.2%	

Figure 3. The diverse immune characteristics and heterogeneity of genetic phenotypes of non-immune class, immune-activated subgroup, and immune-exhausted subgroup

(A) Division and characterization of three immunophenotypes. CYT, cytolytic activity score; TITR, tumor-infiltrating Tregs; MDSC, myeloid-derived suppressor cell; TLS, tertiary lymphoid structure; C-ECM, cancer-associated extracellular matrix. (B) Difference of tumor-infiltrating lymphocyte abundance. (C) Difference in the PD-L1 mRNA expression level. (D) Difference in gene copy number alterations, including amplification and deletion, among arm levels and focal levels. (E) Difference the tumor mutation burden. (F) Difference in tumor neoantigens. (G) Specific mutant genes in the immune-activated subgroup. (H) Specific mutant genes in the immune-exhausted subgroup. WT, wild-type; IM-Act, immune-activated subgroup; IM-Exh, immune-exhausted subgroup.

were linked with decreased immunocyte infiltration, especially for CD4+ T cells, neutrophils, and dendritic cells (Figure S3).

The TMB in the immune class was higher than that in the non-immune class ($p = 0.01$; Figure 3E), while the neoantigen level exhibited no difference ($p = 0.109$; Figure 3F). We further compared the specific gene mutations in the immune subgroups (Figure S4A). Mutations of *TP53* (53.5% versus 43.1%, $p = 0.051$), TTN (52.9% versus 39.5%, $p = 0.011$), *PIK3CA* (28.0% versus 17.0%, $p = 0.007$), and RB1 (26.0% versus 13.0%, $p < 0.001$) appeared more frequently in the immune class than in the non-immune class (Figure S4B). *ERBB2* ($p = 0.035$), *KMT2A* ($p = 0.013$), *PKHD1* ($p = 0.007$), and *MDN1* ($p = 0.015$) were the specific mutations noted in the immune-activated subgroup (Figure 3G), and patients with *EP300* ($p = 0.020$), *HMCN1* ($p = 0.014$), *AKAP9* ($p = 0.003$), and *MACF1* ($p = 0.016$) mutations were more highly enriched in the immune-exhausted subgroups (Figure 3H). The mutation of *EP300* lead to the increased expression of EP300 (Figure S4C). Based on the results from Figures 3B–3H, S3, and S4, we can conclude that the immune class exhibits lower copy number deletion, higher TIL abundance, higher TMB, and higher PD-L1 level. The specific mutant genes in the immunophenotypes are diverse.

Reappearance of the three immunophenotypes in external cohorts

External cohorts with mRNA expression profiles were collected to recapitulate the three immunophenotypes defined by the NMF algorithm with respect to the microdissected and activated stroma signature (Figure 1; Tables 1 and 2). The top 150 increased differentially expressed genes (DEGs) between the immune and non-immune classes (Table S4) were chosen as the seed genes to regenerate the immune subclasses in the external cohorts using the GenePattern module NMFConsensus, and then immune-activated and immune-exhausted subgroups were further separated via the NTP method.

In the GSE32894 cohort, 60.7% (187/308) of patients belonged to the non-immune class, with the lower enrichment of immune-associated signatures; as for the remaining 121 patients, compared with the signatures of stromal enrichment, 42 patients divided into the immune-activated subgroup and 79 belonged to the immune-exhausted subgroup. High enrichment scores of TITR, MDSC, WNT/TGFβ, TGFβ-1-activated, and C-ECM signatures were observed in the immune-exhausted subgroup (all $p < 0.01$; Table 3; Figure S5).

In the other cohorts, we also replicated the three immunophenotypes, and the results are displayed in Table 3 and Figures S5 and

Table 1. Summary of the detailed information of the enrolled bladder cancer cohorts

Dataset	Data array	Patients	Reference
TCGA-BLCA	RNA sequencing	408	https://xenabrowser.net/datapages/?cohort=GDC%20TCGA%20Prostate%20Cancer%20(PRAD)&removeHub=https%3A%2F%2Fxena.treehouse.gi.ucsc.edu%3A443
E-MTAB-4321	RNA sequencing	476	https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-4321/
IMvigor210	Illumina HiSeq 2500	348	http://research-pub.gene.com/IMvigor210CoreBiologies/
GSE32894	Illumina HumanHT-12 V3.0 expression beadchip	308	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE32894
GSE83586	Affymetrix Human Gene 1.0 ST Array	307	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE83586
GSE87304	Affymetrix Human Exon 1.0 ST Array	305	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE87304
GSE128702	Affymetrix Human Exon 1.0 ST Array	256	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE128702
GSE13507	Illumina human-6 v2.0 expression beadchip	164	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE13507
GSE129871	Illumina HiSeq 2000 (Homo sapiens)	158	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE129871
GSE120736	Illumina HumanHT-12 V4.0 expression beadchip	145	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE120736
GSE39016	Affymetrix Human Exon 1.0 ST Array	141	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE39016
GSE128701	Affymetrix Human Exon 1.0 ST Array	136	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE128701
GSE124035	Affymetrix Human Exon 1.0 ST Array	133	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE124305
GSE86411	Illumina HumanHT-12 WG-DASL V4.0 R2 expression beadchip	132	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE86411
GSE48276	Illumina HumanHT-12 WG-DASL V4.0 R2 expression beadchip	116	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE48276
GSE128192	Illumina HumanHT-12 WG-DASL V4.0 R2 expression beadchip	112	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE128192
GSE31684	Affymetrix Human Genome U133 Plus 2.0 Array	93	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE31684
GSE134292	Illumina HiSeq 4000 (Homo sapiens)	80	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE134292
GSE93527	Affymetrix Human Transcriptome Array 2.0	79	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE93527
E-MTAB-1803	Affymetrix GeneChip Human Genome U133 Plus 2.0	70	https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-1803/
GSE69795	Illumina HumanHT-12 WG-DASL V4.0 R2 expression beadchip	61	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE69795

S6. In these cohorts, the distribution of immune-activated subgroups ranged from 11.3% to 30.9%, while the proportion of immune-exhausted subgroups ranged from 17.1% to 40.8%. We also observed increased scores for the immune enrichment signature and immune signaling signature in the 18 validation cohorts, as well as the other immunocyte signatures. As expected, the immune-exhausted subgroup showed increased enrichment scores for Tregs, TITR, MDSC, WNT/TGFβ, and C-ECM signatures. With the combined results from Tables 1, 2, 3, and S4 and Figures S5 and S6, our results suggest that the NMF and NTP algorithms could stably and precisely divide bladder patients into immune-activated, immune-exhausted, and non-immune phenotypes. The specific immune characteristics can reappear in all the enrolled bladder cancer cohorts.

Table 2. Summary of the clinicopathological parameters of the TCGA-BLCA, GSE32894, and E-MTAB-1803 cohorts

	TCGA-BLCA (n = 408)	GSE32894 (n = 308)	E-MTAB-1803 (n = 70)
Age			
≤70	230	143	42
>70	178	165	28
Sex			
Male	301	228	59
Female	107	80	11
Stage[a]			
Ta	–	116	–
T1	11	97	–
T2	191	85	24
T3	157	7	28
T4	43	1	18
Grade[b]			
G1/low	21	48	–
G2	–	103	4
G3/high	384	154	66
Smoking[c]			
No	109	–	–
Yes	286	–	–
Status			
Alive	229	199	28
Dead	177	25	42

[a]Six samples lacked T stage data in the TCGA database, and two samples lacked data in GEO: GSE32894.
[b]Three samples lacked grade data in the TCGA database, and three samples lacked date in GEO: GSE32894.
[c]Two samples lacked alive status data in the TCGA database, and 84 samples lacked data in GEO: GSE32894.

Favorable response to anti-PD-L1 therapy was observed in the immune-activated subgroup

To evaluate the response to immunotherapy of the bladder cancer patients with newly defined immunophenotypes, we collected the gene expression profile and clinical outcomes of 348 patients from the IMvigor210 cohort, a large phase II trial investigating the clinical response of PD-L1 blockade with atezolizumab. The observed endpoint is the response to anti-PD-L1 treatment, including complete response (CR), partial response (PR), stable disease (SD), and progressive disease (PD). The objective response rate (ORR) included patients with CR and PR, while the disease control rate (DCR) included patients with CR, PR, and SD. Patients were first separated into immune (237/348, 65.2%) and non-immune (121/348, 34.8%) classes by the 150 DEGs generated from the TCGA-BLCA training cohort. Then, 85 patients (85/348, 24.4%) with a negative result of stromal activation prediction in the immune class were recognized as the immune-activated subgroup, while the remaining 142 patients (142/348, 40.8%) belonged to the immune-exhausted subgroup (Figure 4A). It is obvious that patients in the immune class showed an abundance of immune signatures, including predefined immune enrichment score, immune cell subset score, immune signaling score, as well as the signatures of several immunocytes (all $p < 0.05$). For the separation of the immune-activated and immune-exhausted subgroups, MDSC, TITR, Tregs, WNT/TGFβ, TGFβ-1, and C-ECM signatures showed a higher enrichment score in the immune-exhausted subgroup, consistent with the characteristics in the TCGA-BLCA training cohort. We observed that more than half of the patients in the immune-activated subgroup benefited from anti-PD-L1 therapy; the ORR was 24.66%, and the DCR was 52.06%. However, most patients in the immune-exhausted subgroup showed PD results after anti-PD-L1 treatment ($p = 0.028$; Figure 4B). Moreover, if we only focused on the treatment results of CR and PD, we found that patients in the immune-activated subgroup showed a higher response rate than other subgroups (20.5% versus 15.1% versus 4.8%; Figure 4C). In general, the results in Figure 4 show that the newly defined classifier can distinguish the appropriate patients to receive anti-PD-L1 therapy, of whom in the immune-activated subgroup.

Immune-activated subgroup shows a favorable prognosis and benefits more from anti-PD-1 therapy

To verify the excellent effectiveness of immune therapy in the immune-activated subgroup, we used subclass mapping (SubMap) analysis to compare the gene expression distribution of the responders in the IMvigor210 cohort and the MD Anderson melanoma cohort, with the immune-activated subgroups of TCGA-BLCA, GSE32894, and E-MTAB-1803 cohorts. We found that patients belonging to immune-activated subgroups contained a more similar gene distribution to the responders than non-responders of anti-PD-L1 treatment in the IMvigor210 cohort but not those patients in the non-immune-activated subgroup (all Bonferroni corrected $p < 0.05$) (Figure 5A). Moreover, we also observed consistent results in the comparison with the MD Anderson melanoma cohort, and patients in the immune-activated subgroup can profit more from anti-PD-1 treatment (all Bonferroni-corrected $p < 0.05$) (Figure 5B). In addition, we revealed that patients belonging to the immune-activated subgroup illustrated a better prognosis, while patients separated into the immune-exhausted subgroup suffered from a poor prognosis (Figure 5C). Taken together with the results from Figure 5, we concluded that patients in the immune-activated subgroup can profit more from anti-PD-1 or anti-PD-L1 therapy and had the longest average overall survival time.

Comparison of the three immunophenotypes with the proposed molecular subtypes and reappearance in pan cancer

The newly defined immunophenotypes are the supplement to the proposed molecular subtypes, which can help to precisely depict the molecular characters of bladder cancer patients. Six immune molecular features were previously defined by Thorsson et al.,[27] reflecting wound healing (C1), IFN-γ-dominant (C2), inflammatory (C3), lymphocyte-depleted (C4), immunologically quiet (C5), and TGF-β-dominant (C6) specific features. We found that most patients in the immune-activated subgroup (75.00%) belonged to the IFN-γ group, which is associated with a strong CD8 signal and a high

Table 3. The distribution of three newly defined immunophenotypes in all enrolled cohorts

Dataset	No. of patients	Immunophenotype distribution, n (%)		
		Immune activated	Immune exhausted	Non-immune
TCGA-BLCA	408	45 (11.03)	110 (26.96)	253 (62.01)
E-MTAB-4321	476	74 (15.55)	111 (23.32)	291 (61.13)
IMvigor210	348	85 (24.43)	142 (40.8)	121 (34.77)
GSE32894	308	42 (13.64)	79 (25.65)	187 (60.71)
GSE83586	307	63 (20.52)	93 (30.29)	151 (49.19)
GSE87304	305	59 (19.34)	85 (27.87)	161 (52.79)
GSE128702	256	72 (28.13)	88 (34.38)	96 (37.50)
GSE13507	164	23 (14.02)	36 (21.95)	105 (64.02)
GSE129871	158	26 (16.46)	27 (17.09)	105 (66.46)
GSE120736	145	21 (14.48)	35 (24.14)	89 (61.38)
GSE39016	141	16 (11.35)	31 (21.99)	94 (66.67)
GSE128701	136	42 (30.88)	34 (25.00)	60 (44.12)
GSE124035	133	32 (24.06)	54 (40.6)	47 (35.34)
GSE86411	132	22 (16.67)	36 (27.27)	74 (56.06)
GSE48276	116	24 (20.69)	29 (25.00)	63 (54.31)
GSE128192	112	26 (23.21)	36 (32.14)	50 (44.64)
GSE31684	93	14 (15.05)	34 (36.56)	45 (48.39)
GSE134292	80	13 (16.25)	16 (20.00)	51 (63.75)
GSE93527	79	13 (16.46)	15 (18.99)	51 (64.56)
E-MTAB-1803	70	13 (18.57)	19 (27.14)	38 (54.29)
GSE69795	61	9 (14.75)	19 (31.15)	33 (54.10)

proliferation rate, and that approximately 68.22% of patients in the immune-exhausted subgroup also belonged to the IFN-γ group (Figure S7A). Another classical molecular classifier was defined by Kamoun et al.;[28] luminal papillary, luminal nonspecified, luminal unstable, stroma-rich, basal/squamous, and neuroendocrine-like subtypes cover almost every aspect of bladder cancer. In the current study, we revealed that most immune-exhausted patients and half of the immune-activated patients belonged to the basal/squamous classes (Figure S7B), with *TP53* being the most frequently mutated gene, consistent with previous results. Unquestionably, the Ba/sq subclass was associated with poor prognosis, and patients in the immune-exhausted subgroup also exhibited poor prognosis. Along with the results shown in Figure S7, we can conclude that the newly defined bladder cancer immunophenotypes have a high consistency with the proposed molecular subtypes.

We also tried to reproduce the three immunophenotypes in other cancer types of the TCGA dataset. With our recently developed R package MOVICS, we first selected the specific DEGs for each immunophenotype and then distinguished the non-immune, immune-activated, and immune-exhausted subgroups in pan cancer cohorts. Finally, we successfully distinguished the three immunophenotypes in genitourinary system tumors, including papillary renal cell carcinoma (KIRP), prostate cancer (PRAD), and testicular germ cell tumor (TGCT) (Figure S8). Patients in the immune-exhausted subgroup showed a poor prognosis (Figure S9), and the prognosis results of PRAD are consistent with our recent work.[11] In addition, we also found that the classification system is suitable for adrenocortical carcinoma (ACC), brain lower-grade glioma (LGG), mesothelioma (MESO), stomach adenocarcinoma (STAD), and uveal melanoma (UVM) (Figures S8 and S9).

Dimensionality reduction of the 150 DEGs for the distinguishment of immune and non-immune classes

The definition of the three immunophenotypes was a two-step process. First, we separated the patients into immune and non-immune classes by NMF consensus via the 150 DEGs from the TCGA training cohort. Second, we partitioned patients in the immune class into immune-activated and immune-exhausted subtypes by the NTP method based on the expression of 47 stromal-associated genes defined by Moffitt et al.[29] The smaller the panel is, the easier it is to implement clinically. Therefore, we used the random forest algorithm to generate a predictive model of immune and non-immune classes via a smaller number of genes. A total of 21 genes were selected (Table S5), the functions of which were enriched in several immune-associated biological pathways (Figure S10A). The 21 genes can 100% predict the distribution of patients in two immune classes in the TCGA-BLCA training cohort, with the predicted accuracy of 0.85 in the E-MTAB-1803 validation cohort, 0.97 in the GSE32894 validation cohort, and 0.93 in the IMvigor210 validation cohort (Figures S10B and S10C).

DISCUSSION

Bladder cancer is a heavy health burden worldwide, especially in Europe and North America.[30] There are approximately 550,000 new cases and 200,000 specific deaths caused by bladder cancer each year. The incidence rate of bladder cancer varies around the world, with the highest rate in southern Europe and the lowest rate in Middle Africa.[30] In the United States, bladder cancer is the 6th most common type of cancer, with approximately 81,000 new cases and 18,000 new deaths reported in 2020.[31] Most patients are diagnosed with early-stage bladder cancer, also known as NMIBC, in which processes do not involve the muscle layer and can be removed easily through transurethral resection (TUR) or with intravesical therapy with bacillus Calmette-Guérin (BCG) or by another chemotherapeutic treatment.[32–34] Recurrence is extremely common in NMIBC; approximately 70% of patients will suffer from the health burden of bladder cancer again within 10 years, and one-third of patients will progress to an advanced stage, also called MIBC.[3] The standard care for MIBC is radical cystectomy with or without neoadjuvant chemotherapy or chemoradiation. Even after treatment, almost 50% of MIBC patients will experience recurrence and death from the metastatic stage within 3 years.[35]

Several studies investigated the diagnostic markers, prognostic signatures, or therapeutic targets for malignancy of tumors based on the TME, as well as in bladder cancer. BCG is the earliest immune therapy

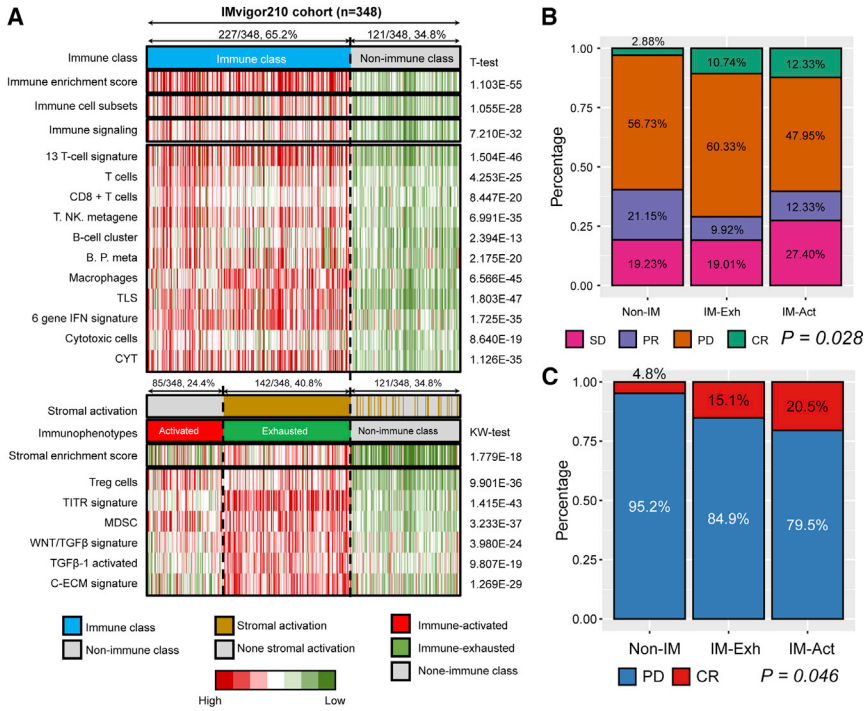


Figure 4. The immunophenotypes in the IMvigor210 cohort reflected the different responses to anti-PD-L1 immunotherapy

(A) Reappearing three immunophenotypes in IMvigor210 cohort. (B) The distribution of the best confirmed overall response to anti-PD-L1 treatment in the three immunophenotypes. (C) The distribution of complete response and progressive disease in the three immunophenotypes. Non-IM, non-immune class; IM-Act, immune-activated subgroup; IM-Exh, immune-exhausted subgroup.

approved for bladder cancer treatment, which can stimulate an immunologic reaction and induce proinflammatory cytokines and direct cell-to-cell cytotoxicity.[36] BCG is still the standard therapy for NMIBC, which reflects that bladder cancer patients could benefit from immunotherapy. The blockade of immune checkpoints is also applied in the treatment of bladder cancer. Two PD-1 inhibitors (pembrolizumab and nivolumab) and three PD-L1 inhibitors (atezolizumab, durvalumab, and avelumab) have been approved by the US Food and Drug Administration (FDA) for the treatment of bladder cancer (http://www.accessdata.fda.gov/scripts/cder/daf/). In the IMvigor210 clinical study, atezolizumab was used to block PD-L1: the ORR for the IMvigor210 cohort 1 was only 23%, and it was only 15% for cohort 2.[37,38] The ORR values for nivolumab and durvalumab are similar, ranging from 17% to 24.4%.[39–42] Therefore, a full understanding of the immunophenotypes of bladder cancer is essential and could serve as guidance in choosing patients to receive the appropriate immunotherapy.

For the enrolled 4,028 bladder cancer patients, we used NMF and NTP algorithms to investigate a robust classification of three immunophenotypes. First, we identified the immune-activated subgroup, immune-exhausted subgroup, and non-immune class in the TCGA-BLCA training cohort. The top weighted 150 exemplar genes selected from the immune module represent the immune features of bladder cancer patients and further divide the entire cohort into immune and non-immune classes. The other 150 DEGs among the immune and non-immune classes were extracted as the input profile for the reappearance of the classification in validation cohorts. Stromal activation status was evaluated via the NTP algorithm to further separate the immune class into immune-activated and immune-exhausted subgroups. The features of these three immunophenotypes were illuminated by several verified signatures of immunocytes or immune signaling pathways. Patients in the immune classes showed highly enriched signatures of T cells, B cells, IFN, and CYT (L.Q.M. Chow et al., 2016, J. Clin. Oncol., abstract),[43,44] while the exhausted subgroup also showed increased signatures of TITR, WNT/TGF-β, TGF-β1-activated, and C-ECM signatures;[45–47] this was not the case for the immune-activated subgroup. B Based on our results, we revealed that only approximately 11% to 30.9% of bladder cancer patients belong to the immune-activated subgroup, and can benifit from immunotherapy.

Clinical outcome is an important factor on which we focused with respect to the newly defined immunophenotypes. With the clinical information of the TCGA-BLCA, GSE32894, and E-MTAB-1803 cohorts, we found that patients belonging to the immune-activated subgroup exhibited the most favorable overall survival, while patients belonging to the immune-exhausted subgroup showed the worst clinical outcome of a shortened overall survival time. Immune exhaustion, which is mostly focused on the exhaustion of T cells, is reflected by altered inflammatory and tissue microenvironments, lymphocytes, and inhibitory signals from cytokines.[48] These alterations in the TME could lead to the escape from immune recognition by blocking immune checkpoints, which is related to unfavorable overall survival for patients.[49] We predicted the potential response to immunotherapy of the bladder cancer patients by comparing their mRNA expression profiles with those of melanoma samples receiving anti-CTLA-4 or anti-PD-1 checkpoint therapy. As expected, patients in the immune-activated subgroup could benefit from anti-PD-1 therapy, but this was not the case for patients in the non-immune-activated subgroup, which included the combined immune-exhausted and non-immune classes.

To further understand the molecular diversity among the three immunophenotypes, we compared the CNA, TMB, and gene mutations. Recent studies have reported the association of CNA with increased

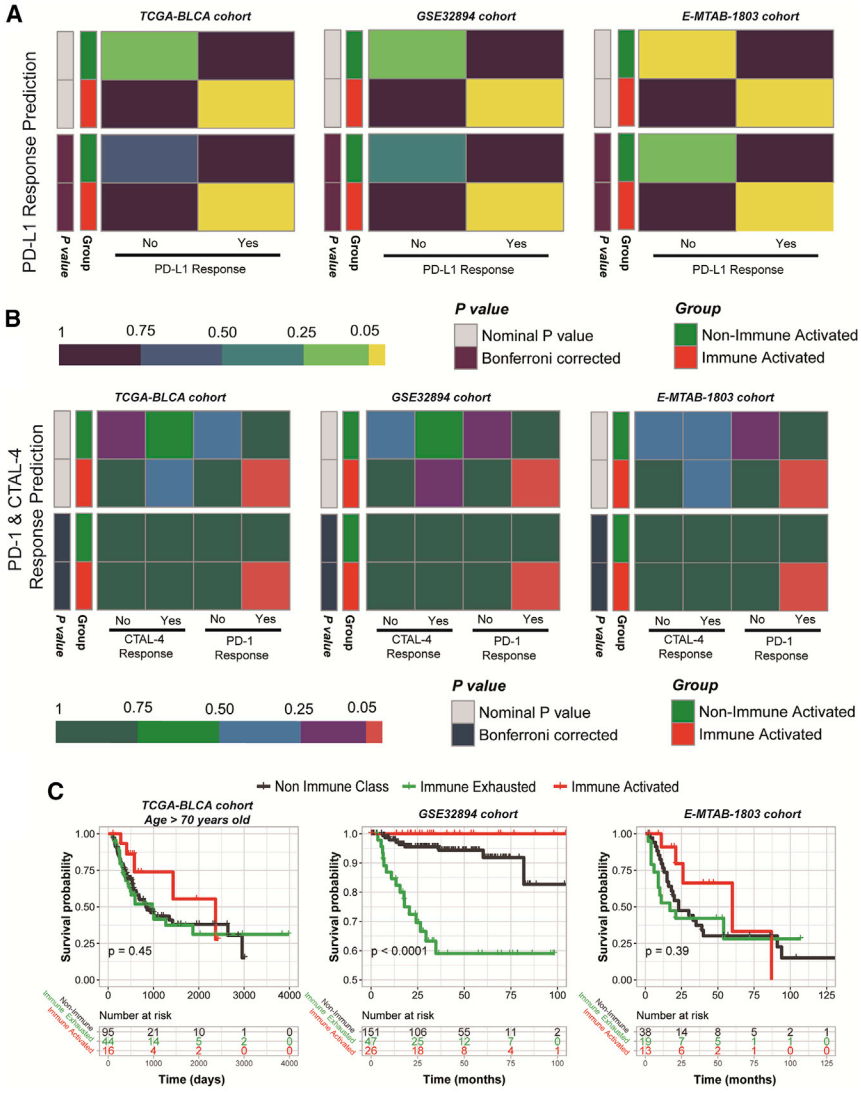


Figure 5. Predicting the response to immune checkpoint blockade therapy and revealed diverse overall survival outcomes in the three immunophenotypes

(A) Prediction results of the response to anti-PD-L1 therapy. (B) Prediction results of the response to anti-CTAL-4 and anti-PD-1 therapy. (C) Different overall survival outcomes in the three immunophenotypes.

immune infiltration and the outcome of immune checkpoint blockade therapy.[50,51] Patients in the immune class showed a lower CNA burden with gene deletion at the arm and focal levels. The association was verified twice by the deletion copy numbers of PD-1 and PD-L1, and CTLA-4 was positively associated with the decreased infiltration level of immunocytes. Tripathi et al.[52] found that antigen presentation through the major histocompatibility complex (MHC) class I pathway was suppressed in tumors with high chromosomal instability, also known as high CNA, which plays a pivotal role in immune evasion. In addition, Lu et al.[53] revealed that patients treated with immune checkpoint blockade therapy could receive a durable clinical benefit and achieve better survival if they exhibited a lower CNA burden. Gene mutation is another key component upon which we focused with respect to the three immunophenotypes. We extracted the specific mutant genes for each subgroup. The proportions of mutant *TP53, TTN, PIC3CA*, and *RB1* were higher in the immune class than that in the non-immune class. Nusrat et al. (2019, J. Clin. Oncol., abstract) reported that colorectal cancer patients with mutant *PIK3CA* had higher median densities of CD3+ and CD8+ cells, as well as a high rate of clinical benefit from immunotherapy (50% versus 8.6%). Furthermore, we observed a high rate of *ERBB2* mutations in the immune-activated subgroup. ERBB2 amplification or overexpression is a biomarker of anti-ERBB2-targeted therapy in breast cancer, while the activity of trastuzumab and infiltration of immune cells can both be regulated by the activated function of ERBB2 via the inducing of CCL2 and PD-1 ligands.[54] The V659E mutation of the *ERBB2* gene has also been reported to be associated with altered sensitivity to afatinib and lapatinib treatment *in vitro*.[55,56] The highest mutation proportion of *EP300* was observed in the immune-exhausted subgroup, and the mutation leads to the increased expression of EP300. The function of effector T cell responses to tumor antigens can be dampened by the Tregs and cause the immunosuppressive status of the TME; this is associated with poor prognosis.[57] Tregs mediated immune suppressive function was reported to be impaired by the inhibition of EP300 as well;[58] therefore, it is considerable to combine the immune checkpoint blockade therapy plus an EP300 inhibitor for the treatment of immune-exhausted patients. The limitation of the current study is the lack of the prospective validation of the newly defined immunophenotypes. In the future, we will use the immunophenotype classification system to choose appropriate bladder cancer patients to receive the immunotherapy.

We defined and validated a novel classifier among 4,028 bladder cancer patients to separate them into immune-activated, immune-exhausted, and non-immune subgroups. Patients in the immune-activated subgroup could benefit more from a single treatment of anti-PD-1 immunotherapy; for patients in the immune-exhausted subgroup, immune checkpoint blockade therapy plus a TGF-β inhibitor or an EP300 inhibitor might have greater efficacy. In summary, our novel classifier provides a better understanding of enhanced immunotherapy to treat bladder cancer patients.

MATERIALS AND METHODS

Bladder cancer patient cohorts

For the TCGA-BLCA cohort, we obtained the level three gene expression profiles of 408 patients from the TCGA Data Portal (https://www.cancer.gov/tcga/), and only genes expressed in at least 50% of the samples were retained for the subsequent analyses. For further external validation, we enrolled the Gene Expression Omnibus (GEO): GSE32894, GSE83586, GSE87304, GSE128702, GSE13507, GSE129871, GSE120736, GSE39016, GSE128701, GSE124035, GSE86411, GSE48276, GSE31684, GSE134292, GSE93527, and GSE69795 cohorts; the gene expression profiles and clinical information were collected from the GEO (https://www.ncbi.nlm.nih.gov/geo/). For the E-MTAB-4321 and E-MTAB-1803 cohorts, the gene expression profiles and paired clinical features were downloaded from ArrayExpress (https://www.ebi.ac.uk/arrayexpress/). Detailed information of these datasets is displayed in Tables 2 and 3.

Bioinformatic analyses

The mRNA expression profiles of the TCGA-BLCA training cohort were microdissected using the NMF algorithm.[59] The identification of an immune class was achieved as reported by Sia et al.[60] The immune module was selected, with the gathering of patients with the highest immune enrichment score, which was estimated by ssGSEA.[61] Then, the top 150 exemplar genes with the highest weight in the immune module were extracted as the key genes to dichotomize the immune and non-immune classes, which were further modified by the MDS random forest method.[60] In addition, immune-activated and immune-exhausted subgroups were recognized by the stromal activation signature by the NTP method.[29] To depict the characteristics of these three immunophenotypes, several published immune gene sets were manually collected to reflect the immune status, and the scores of each gene set for each patient were also calculated via ssGSEA (Table S3). The different genetic types among the immune and non-immune classes were evaluated, including TIL abundance, PD-L1 expression, CNAs, TMB, neoantigens, and mutant genes. To recreate the immunophenotypes in validation cohorts, the expression profiles of the top 150 DEGs were used to dichotomize the immune classes via the NMF consensus method, and the immune class was subsequently divided into the activated and exhausted subgroups. To predict the immunotherapy response of bladder cancer patients, we collected the expression profile of the IMvigor210 cohort, along with the response results of anti-PD-L1 treatment from IMvigor210CoreBiologies.[62] The MD Anderson melanoma cohort that received anti-CTLA-4 or anti-PD-1 therapy was also considered for the prediction of immunotherapy response.[63] SubMap[64] analysis was used to compare the similarity of gene expression between the newly defined bladder cancer immunophenotypes and the responders of anti-CTLA-4, anti-PD-1, and anti-PD-L1 therapy in the IMvigor210 cohort or MD Anderson melanoma cohort. We developed a predictive model to estimate immune classes based on the random forest algorithm using R package varSelRF. To this end, we obtained the DEGs using R package limma between immune and non-immune classes in the TCGA training cohort to select informative genes, which served as input to random forest model (foldchange > 2 or < 0.5 with false discovery rate [FDR] < 0.05). A backward elimination procedure was applied to search for optimal markers where an out-of-bag (OOB) error was used as the minimization criterion. In this procedure, variable elimination was conducted by setting the dropping fraction of each iteration at 0.2, and the gene combination was identified to develop the final predictive model when random forest reached the smallest OOB error rate. The predictive performance was assessed in TCGA, EMATB-1803, GSE32894, and IMvigor210 cohorts using receiver operating characteristic (ROC) curves. The details of the enrolled cohorts, as well as the specific method used for each step, are provided in the Supplemental methods.

Statistical analysis

For normally distributed continuous data, t test was used for the comparisons, while non-normally distributed data were compared via Wilcoxon rank-sum test. The comparisons between factors with more than two groups were performed by Kruskal-Wallis test. The difference in survival between different groups was analyzed by Kaplan-Meier plots and log-rank test. Chi-square test was used to illustrate the correlations between newly defined immunophenotypes and proposed molecular subtypes. Two-side p value less than 0.05 was considered as the statistically significant difference. All analyses were performed by GenePattern[65] and R version 4.0.2 (http://www.r-project.org).

Availability of data and materials

All data used in this work can be acquired from the GDC portal (https://portal.gdc.cancer.gov/), GEO (https://www.ncbi.nlm.nih.gov/geo/), and ArrayExpress (https://www.ebi.ac.uk/arrayexpress/).

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.omto.2021.02.001.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (grant numbers 81802827, 81630019, and 31701162); the Scientific Research Foundation of the Institute for Translational Medicine of Anhui Province (grant number 2017ZHYX02); the Natural Science Foundation of Guangdong Province, China (grant number 2017A030313800); the Key Project of Provincial Natural Science Research Project of Anhui Colleges (grant number KJ2019A0278); the Supporting Project for Distinguished Young Scholar of Anhui Colleges (grant number gxyqZD2019018); and the 2017 Anhui Province special program for guiding local science and technology development by the central government (grant number 2017070802D148).

AUTHORS CONTRIBUTIONS

Conception and design, J.M., S.G., F.Y., and C.L.; collection and assembly of data, Y.Z., X.L., M.Z., and J.Z.; data analysis and interpretation, J.M., Y.Z., X.L., and Z.H.; manuscript writing, J.M., X.L., Y.Z., and S.G.; final approval of manuscript, all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Sanchez, A., Wszolek, M.F., Niemierko, A., Clayman, R.H., Drumm, M., Rodríguez, D., Feldman, A.S., Dahl, D.M., Heney, N.M., Shipley, W.U., et al. (2018). Incidence, Clinicopathological Risk Factors, Management and Outcomes of Nonmuscle Invasive Recurrence after Complete Response to Trimodality Therapy for Muscle Invasive Bladder Cancer. J. Urol. *199*, 407–415.
2. Hautmann, R.E., Gschwend, J.E., de Petriconi, R.C., Kron, M., and Volkmer, B.G. (2006). Cystectomy for transitional cell carcinoma of the bladder: results of a surgery only series in the neobladder era. J. Urol. *176*, 486–492, discussion 491–492.
3. Chamie, K., Litwin, M.S., Bassett, J.C., Daskivich, T.J., Lai, J., Hanley, J.M., Konety, B.R., and Saigal, C.S.; Urologic Diseases in America Project (2013). Recurrence of high-risk bladder cancer: a population-based analysis. Cancer *119*, 3219–3227.
4. Mo, Q., Nikolos, F., Chen, F., Tramel, Z., Lee, Y.C., Hayashi, K., Xiao, J., Shen, J., and Chan, K.S. (2018). Prognostic Power of a Tumor Differentiation Gene Signature for Bladder Urothelial Carcinomas. J. Natl. Cancer Inst. *110*, 448–459.
5. Damrauer, J.S., Hoadley, K.A., Chism, D.D., Fan, C., Tiganelli, C.J., Wobker, S.E., Yeh, J.J., Milowsky, M.I., Iyer, G., Parker, J.S., and Kim, W.Y. (2014). Intrinsic subtypes of high-grade bladder cancer reflect the hallmarks of breast cancer biology. Proc. Natl. Acad. Sci. USA *111*, 3110–3115.
6. Robertson, A.G., Kim, J., Al-Ahmadie, H., Bellmunt, J., Guo, G., Cherniack, A.D., Hinoue, T., Laird, P.W., Hoadley, K.A., Akbani, R., et al.; TCGA Research Network (2017). Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. Cell *171*, 540–556.e25.
7. Devarajan, K. (2008). Nonnegative matrix factorization: an analytical and interpretive tool in computational biology. PLoS Comput. Biol. *4*, e1000029.
8. Gaujoux, R., and Seoighe, C. (2010). A flexible R package for nonnegative matrix factorization. BMC Bioinformatics *11*, 367.
9. Zeng, Z., Vo, A.H., Mao, C., Clare, S.E., Khan, S.A., and Luo, Y. (2019). Cancer classification and pathway discovery using non-negative matrix factorization. J. Biomed. Inform. *96*, 103247.
10. Esposito, F., Boccarelli, A., and Del Buono, N. (2020). An NMF-Based Methodology for Selecting Biomarkers in the Landscape of Genes of Heterogeneous Cancer-Associated Fibroblast Populations. Bioinform. Biol. Insights *14*, 1177932220906827.
11. Meng, J., Zhou, Y., Lu, X., Bian, Z., Chen, Y., Zhou, J., Zhang, L., Hao, Z., Zhang, M., and Liang, C. (2020). Immune response drives outcomes in prostate cancer: implications for immunotherapy. Mol. Oncol. Published online December 18, 2020. https://doi.org/10.1002/1878-0261.12887.
12. Zhou, Y.J., Zhu, G.Q., Lu, X.F., Zheng, K.I., Wang, Q.W., Chen, J.N., Zhang, Q.W., Yan, F.R., and Li, X.B. (2020). Identification and validation of tumour microenvironment-based immune molecular subgroups for gastric cancer: immunotherapeutic implications. Cancer Immunol. Immunother. *69*, 1057–1069.
13. Valkenburg, K.C., de Groot, A.E., and Pienta, K.J. (2018). Targeting the tumour stroma to improve cancer therapy. Nat. Rev. Clin. Oncol. *15*, 366–381.
14. Hanahan, D., and Coussens, L.M. (2012). Accessories to the crime: functions of cells recruited to the tumor microenvironment. Cancer Cell *21*, 309–322.
15. Sivanathan, K.N., Gronthos, S., Rojas-Canales, D., Thierry, B., and Coates, P.T. (2014). Interferon-gamma modification of mesenchymal stem cells: implications of autologous and allogeneic mesenchymal stem cell therapy in allotransplantation. Stem Cell Rev. Rep. *10*, 351–375.
16. van Megen, K.M., van 't Wout, E.T., Lages Motta, J., Dekker, B., Nikolic, T., and Roep, B.O. (2019). Activated Mesenchymal Stromal Cells Process and Present Antigens Regulating Adaptive Immunity. Front. Immunol. *10*, 694.
17. Soboslay, P.T., Lüder, C.G., Riesch, S., Geiger, S.M., Banla, M., Batchassi, E., Stadler, A., and Schulz-Key, H. (1999). Regulatory effects of Th1-type (IFN-gamma, IL-12) and Th2-type cytokines (IL-10, IL-13) on parasite-specific cellular responsiveness in Onchocerca volvulus-infected humans and exposed endemic controls. Immunology *97*, 219–225.
18. Aggarwal, S., and Pittenger, M.F. (2005). Human mesenchymal stem cells modulate allogeneic immune cell responses. Blood *105*, 1815–1822.
19. Selleri, S., Dieng, M.M., Nicoletti, S., Louis, I., Beausejour, C., Le Deist, F., and Haddad, E. (2013). Cord-blood-derived mesenchymal stromal cells downmodulate CD4+ T-cell activation by inducing IL-10-producing Th1 cells. Stem Cells Dev. *22*, 1063–1075.
20. Batlle, E., and Massagué, J. (2019). Transforming Growth Factor-β Signaling in Immunity and Cancer. Immunity *50*, 924–940.
21. Furukawa, A., Wisel, S.A., and Tang, Q. (2016). Impact of Immune-Modulatory Drugs on Regulatory T Cell. Transplantation *100*, 2288–2300.
22. Groth, C., Hu, X., Weber, R., Fleming, V., Altevogt, P., Utikal, J., and Umansky, V. (2019). Immunosuppression mediated by myeloid-derived suppressor cells (MDSCs) during tumour progression. Br. J. Cancer *120*, 16–25.
23. Berraondo, P., Sanmamed, M.F., Ochoa, M.C., Etxeberria, I., Aznar, M.A., Pérez-Gracia, J.L., Rodríguez-Ruiz, M.E., Ponz-Sarvise, M., Castañón, E., and Melero, I. (2019). Cytokines in clinical cancer immunotherapy. Br. J. Cancer *120*, 6–15.
24. Dong, Y., Li, X., Zhang, L., Zhu, Q., Chen, C., Bao, J., and Chen, Y. (2019). $CD4^+$ T cell exhaustion revealed by high PD-1 and LAG-3 expression and the loss of helper T cell function in chronic hepatitis B. BMC Immunol. *20*, 27.
25. Liu, J.F., Wu, L., Yang, L.L., Deng, W.W., Mao, L., Wu, H., Zhang, W.F., and Sun, Z.J. (2018). Blockade of TIM3 relieves immunosuppression through reducing regulatory T cells in head and neck cancer. J. Exp. Clin. Cancer Res. *37*, 44.
26. Saltz, J., Gupta, R., Hou, L., Kurc, T., Singh, P., Nguyen, V., Samaras, D., Shroyer, K.R., Zhao, T., Batiste, R., et al.; Cancer Genome Atlas Research Network (2018). Spatial Organization and Molecular Correlation of Tumor-Infiltrating Lymphocytes Using Deep Learning on Pathology Images. Cell Rep. *23*, 181–193.e7.
27. 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., et al.; Cancer Genome Atlas Research Network (2018). The Immune Landscape of Cancer. Immunity *48*, 812–830.e14.
28. Kamoun, A., de Reyniès, A., Allory, Y., Sjödahl, G., Robertson, A.G., Seiler, R., Hoadley, K.A., Groeneveld, C.S., Al-Ahmadie, H., Choi, W., et al.; Bladder Cancer Molecular Taxonomy Group (2020). A Consensus Molecular Classification of Muscle-invasive Bladder Cancer. Eur. Urol. *77*, 420–433.
29. Moffitt, R.A., Marayati, R., Flate, E.L., Volmar, K.E., Loeza, S.G., Hoadley, K.A., Rashid, N.U., Williams, L.A., Eaton, S.C., Chung, A.H., et al. (2015). Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. Nat. Genet. *47*, 1168–1178.
30. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J. Clin. *68*, 394–424.
31. Siegel, R.L., Miller, K.D., and Jemal, A. (2020). Cancer statistics, 2020. CA Cancer J. Clin. *70*, 7–30.
32. Kamat, A.M., Hahn, N.M., Efstathiou, J.A., Lerner, S.P., Malmström, P.U., Choi, W., Guo, C.C., Lotan, Y., and Kassouf, W. (2016). Bladder cancer. Lancet *388*, 2796–2810.
33. Gandhi, N.M., Morales, A., and Lamm, D.L. (2013). Bacillus Calmette-Guérin immunotherapy for genitourinary cancer. BJU Int. *112*, 288–297.
34. Bellmunt, J., Orsola, A., Leow, J.J., Wiegel, T., De Santis, M., and Horwich, A.; ESMO Guidelines Working Group (2014). Bladder cancer: ESMO Practice Guidelines for diagnosis, treatment and follow-up. Ann. Oncol. *25* (*Suppl 3*), iii40–iii48.
35. Pectasides, D., Pectasides, M., and Nikolaou, M. (2005). Adjuvant and neoadjuvant chemotherapy in muscle invasive bladder cancer: literature review. Eur. Urol. *48*, 60–67, discussion 67–68.
36. Akaza, H. (1995). BCG treatment of existing Ta, T1 tumours or carcinoma in situ of the bladder. Eur. Urol. *27* (*Suppl 1*), 9–12.
37. Balar, A.V., Galsky, M.D., Rosenberg, J.E., Powles, T., Petrylak, D.P., Bellmunt, J., Loriot, Y., Necchi, A., Hoffman-Censits, J., Perez-Gracia, J.L., et al.; IMvigor210 Study Group (2017). Atezolizumab as first-line treatment in cisplatin-ineligible patients with locally advanced and metastatic urothelial carcinoma: a single-arm, multicentre, phase 2 trial. Lancet *389*, 67–76.
38. Rosenberg, J.E., Hoffman-Censits, J., Powles, T., van der Heijden, M.S., Balar, A.V., Necchi, A., Dawson, N., O'Donnell, P.H., Balmanoukian, A., Loriot, Y., et al. (2016). Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. Lancet *387*, 1909–1920.

39. Sharma, P., Callahan, M.K., Bono, P., Kim, J., Spiliopoulou, P., Calvo, E., Pillai, R.N., Ott, P.A., de Braud, F., Morse, M., et al. (2016). Nivolumab monotherapy in recurrent metastatic urothelial carcinoma (CheckMate 032): a multicentre, open-label, two-stage, multi-arm, phase 1/2 trial. Lancet Oncol. *17*, 1590–1598.

40. Sharma, P., Retz, M., Siefker-Radtke, A., Baron, A., Necchi, A., Bedke, J., Plimack, E.R., Vaena, D., Grimm, M.O., Bracarda, S., et al. (2017). Nivolumab in metastatic urothelial carcinoma after platinum therapy (CheckMate 275): a multicentre, single-arm, phase 2 trial. Lancet Oncol. *18*, 312–322.

41. Powles, T., O'Donnell, P.H., Massard, C., Arkenau, H.T., Friedlander, T.W., Hoimes, C.J., Lee, J.L., Ong, M., Sridhar, S.S., Vogelzang, N.J., et al. (2017). Efficacy and Safety of Durvalumab in Locally Advanced or Metastatic Urothelial Carcinoma: Updated Results From a Phase 1/2 Open-label Study. JAMA Oncol. *3*, e172411.

42. Patel, M.R., Ellerton, J., Infante, J.R., Agrawal, M., Gordon, M., Aljumaily, R., Britten, C.D., Dirix, L., Lee, K.W., Taylor, M., et al. (2018). Avelumab in metastatic urothelial carcinoma after platinum failure (JAVELIN Solid Tumor): pooled results from two expansion cohorts of an open-label, phase 1 trial. Lancet Oncol. *19*, 51–64.

43. Iglesia, M.D., Vincent, B.G., Parker, J.S., Hoadley, K.A., Carey, L.A., Perou, C.M., and Serody, J.S. (2014). Prognostic B-cell signatures using mRNA-seq in patients with subtype-specific breast and ovarian cancer. Clin. Cancer Res. *20*, 3818–3829.

44. Bindea, G., Mlecnik, B., Tosolini, M., Kirilovsky, A., Waldner, M., Obenauf, A.C., Angell, H., Fredriksen, T., Lafontaine, L., Berger, A., et al. (2013). Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity *39*, 782–795.

45. Magnuson, A.M., Kiner, E., Ergun, A., Park, J.S., Asinovski, N., Ortiz-Lopez, A., Kilcoyne, A., Paoluzzi-Tomada, E., Weissleder, R., Mathis, D., and Benoist, C. (2018). Identification and validation of a tumor-infiltrating Treg transcriptional signature conserved across species and tumor types. Proc. Natl. Acad. Sci. USA *115*, E10672–E10681.

46. Lachenmayer, A., Alsinet, C., Savic, R., Cabellos, L., Toffanin, S., Hoshida, Y., Villanueva, A., Minguez, B., Newell, P., Tsai, H.W., et al. (2012). Wnt-pathway activation in two molecular classes of hepatocellular carcinoma and experimental modulation by sorafenib. Clin. Cancer Res. *18*, 4997–5007.

47. Chakravarthy, A., Khan, L., Bensler, N.P., Bose, P., and De Carvalho, D.D. (2018). TGF-β-associated extracellular matrix genes link cancer-associated fibroblasts to immune evasion and immunotherapy failure. Nat. Commun. *9*, 4692.

48. Wherry, E.J., and Kurachi, M. (2015). Molecular and cellular insights into T cell exhaustion. Nat. Rev. Immunol. *15*, 486–499.

49. Sharma, P., Hu-Lieskovan, S., Wargo, J.A., and Ribas, A. (2017). Primary, Adaptive, and Acquired Resistance to Cancer Immunotherapy. Cell *168*, 707–723.

50. Chan, T.A., Wolchok, J.D., and Snyder, A. (2015). Genetic Basis for Clinical Response to CTLA-4 Blockade in Melanoma. N. Engl. J. Med. *373*, 1984.

51. Davoli, T., Uno, H., Wooten, E.C., and Elledge, S.J. (2017). Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. Science *355*, eaaf8399.

52. Tripathi, R., Modur, V., Senovilla, L., Kroemer, G., and Komurov, K. (2019). Suppression of tumor antigen presentation during aneuploid tumor evolution contributes to immune evasion. OncoImmunology *8*, 1657374.

53. Lu, Z., Chen, H., Li, S., Gong, J., Li, J., Zou, J., Wu, L., Yu, J., Han, W., Sun, H., et al. (2020). Tumor copy-number alterations predict response to immune-checkpoint-blockade in gastrointestinal cancer. J. Immunother. Cancer *8*, e000374.

54. Triulzi, T., Forte, L., Regondi, V., Di Modica, M., Ghirelli, C., Carcangiu, M.L., Sfondrini, L., Balsari, A., and Tagliabue, E. (2018). HER2 signaling regulates the tumor immune microenvironment and trastuzumab efficacy. OncoImmunology *8*, e1512942.

55. Serra, V., Vivancos, A., Puente, X.S., Felip, E., Silberschmidt, D., Caratù, G., Parra, J.L., De Mattos-Arruda, L., Grueso, J., Hernández-Losa, J., et al. (2013). Clinical response to a lapatinib-based therapy for a Li-Fraumeni syndrome patient with a novel HER2V659E mutation. Cancer Discov. *3*, 1238–1244.

56. Yamamoto, H., Toyooka, S., Ninomiya, T., Matsumoto, S., Kanai, M., Tomida, S., Kiura, K., Muto, M., Suzawa, K., Desmeules, P., et al. (2018). Therapeutic Potential of Afatinib for Cancers with *ERBB2* (*HER2*) Transmembrane Domain Mutations G660D and V659E. Oncologist *23*, 150–154.

57. Curiel, T.J., Coukos, G., Zou, L., Alvarez, X., Cheng, P., Mottram, P., Evdemon-Hogan, M., Conejo-Garcia, J.R., Zhang, L., Burow, M., et al. (2004). Specific recruitment of regulatory T cells in ovarian carcinoma fosters immune privilege and predicts reduced survival. Nat. Med. *10*, 942–949.

58. Liu, Y., Wang, L., Predina, J., Han, R., Beier, U.H., Wang, L.C., Kapoor, V., Bhatti, T.R., Akimova, T., Singhal, S., et al. (2013). Inhibition of p300 impairs Foxp3^{+} T regulatory cell function and promotes antitumor immunity. Nat. Med. *19*, 1173–1177.

59. Lee, D.D., and Seung, H.S. (1999). Learning the parts of objects by non-negative matrix factorization. Nature *401*, 788–791.

60. Sia, D., Jiao, Y., Martinez-Quetglas, I., Kuchuk, O., Villacorta-Martin, C., Castro de Moura, M., Putra, J., Camprecios, G., Bassaganyas, L., Akers, N., et al. (2017). Identification of an Immune-specific Class of Hepatocellular Carcinoma, Based on Molecular Features. Gastroenterology *153*, 812–826.

61. Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-Garcia, W., Treviño, V., Shen, H., Laird, P.W., Levine, D.A., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. Nat. Commun. *4*, 2612.

62. Mariathasan, S., Turley, S.J., Nickles, D., Castiglioni, A., Yuen, K., Wang, Y., Kadel, E.E., III, Koeppen, H., Astarita, J.L., Cubas, R., et al. (2018). TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. Nature *554*, 544–548.

63. Roh, W., Chen, P.L., Reuben, A., Spencer, C.N., Prieto, P.A., Miller, J.P., Gopalakrishnan, V., Wang, F., Cooper, Z.A., Reddy, S.M., et al. (2017). Integrated molecular analysis of tumor biopsies on sequential CTLA-4 and PD-1 blockade reveals markers of response and resistance. Sci. Transl. Med. *9*, eaah3560.

64. Hoshida, Y., Brunet, J.P., Tamayo, P., Golub, T.R., and Mesirov, J.P. (2007). Subclass mapping: identifying common subtypes in independent disease data sets. PLoS ONE *2*, e1195.

65. Reich, M., Liefeld, T., Gould, J., Lerner, J., Tamayo, P., and Mesirov, J.P. (2006). GenePattern 2.0. Nat. Genet. *38*, 500–501.

OMTO, Volume *20*

Supplemental Information

Tumor immune microenvironment-based classifications of bladder cancer for enhancing the response rate of immunotherapy

Jialin Meng, Xiaofan Lu, Yujie Zhou, Meng Zhang, Qintao Ge, Jun Zhou, Zongyao Hao, Shenglin Gao, Fangrong Yan, and Chaozhao Liang

Supplementary methods

Pattern discovery of immune expression and unsupervised analysis

In TCGA training cohort, tumour, stromal, and immune cell transcriptome profiling data were virtually microdissected employing unsupervised NMF method as previously described [1] via GenePattern [2]. The NMF algorithm, which is suitable for decomposing biological data, can factorize the gene expression matrix V (*n* genes × *m* samples) into two matrixes: gene factor matrix *W* of (*n* genes × *k* factors) and sample factor matrix *H* of (*m* samples × *k* factors) [3]. We chose $k = 9$ as the number of factors or expression patterns, given it could produce a high cophenetic coefficient [1] as well as effectively decompose the dataset in our TCGA training cohort. The identification of an immune class, as reported similarly by Sia *et al.* [4], involved the following steps. Firstly, identification of immune-related NMF factors was achieved through single-sample set enrichment analysis (GenePattern module "ssGSEA") of immune enrichment score (IES) gene signature [5]. To obtain the robust immune module, we pre-set the numbers of module as five to 10, respectively. When the total modules is nine, the first module strongly enriched the patients with a highly IES while the average IES of other factors are low, therefore, this module was then named as the "immune module" .

The top 150 weighted genes (**Table S1**) in the immune module were defined as the exemplar genes which could inflect the characteristics of the immune module, these genes were ranked according to the descending order by difference between factor loading value in first column of matrix *W* (immune factor weight) and the largest factor loading in other columns of *W*. Secondly, the top 150 exemplar genes were selected to classify into two

preliminary subgroups, immune and non-immune for the TCGA training cohort. This procedure was accomplished by supervised clustering via GenePattern module "NMFConsensus". Finally, the immune and non-immune classes were adjusted by the multidimensional scaling (MDS) random forest method, which could visualize the level of similarity of individual cases of a dataset[4]. The immune class was furthermore divided into immune-exhausted, and immune-activated subtypes by the nearest template prediction (GenePattern module 'NTP') of the activated stroma [6].

Correlation of Immune class with copy number alterations, tumour-infiltrating lymphocytes

The tumour-infiltrating lymphocytes (TIL) abundance estimated by H&E stained whole-slide images of TCGA samples were obtained from a previous study [7]. Copy number alterations (CNA) data were generated by GISTIC2.0 from GDAC Firehose (https://gdac.broadinstitute.org). We compared the differences in amplification or deletion events of both focal and arm level between Immune and non-Immune classes. The neoantigen number was accessed from a previous study by Rooney *et al.* [8]. The mutation data were retrieved from TCGA (https://tcga-data.nci.nih.gov); we calculated the number of nonsynonymous mutations per million bases to evaluate tumour mutation burden (TMB). Whats more, the mutation landscape Oncoprint was drawn by R package "maftools"[9]. The different distribution of gene mutations among immunosubtypes were evaluated by the Chi-square test.

Molecular characterization of Immune class

Hand-curated gene signatures representing various immune cell types or host anti-tumour immunity (**Table S3**) from literature and databases were used to characterize immune class in TCGA cohort. Immune-exhausted and activation subtypes was identified by using ssGSEA (GenePattern module "ssGSEA") and nearest template prediction (GenePattern module "NTP") of stroma activation[2]. The signature of stroma activation was derived from Figure 2 of Moffitt et al.'s work [6]. Overexpression or downregulation of genes in immune vs. non-immune classes was performed by "limma" package with R, genes with a false discovery rate (FDR) < 0.05 and a $\log_2$ fold change (FC) ≥ 1 were considered differentially expressed between two groups. Subsequently, gene set enrichment analysis (GSEA, http://www.broadinstitute.org/gsea/index.jsp) was performed to determine gene sets and pathways enriched in Immune vs. non-Immune classes.

Validation of immune molecular subtypes in independent external datasets

We identified top 150 upregulated genes between immune and non-immune classes (**Table S4**). Then NMF-based consensus clustering based on the immune classifier was applied to identify the three immunophenotypes in three independent external datasets (**Table 3**) using GenePattern module "NMFConsensus" with the 150 DEGs. Immune-related gene signature ssGSEA scores were calculated to feature molecular characteristics and validate the existence of abovementioned immune molecular subtypes in each dataset, and then, the immune class divided to activated and exhausted subgroups by the and nearest template prediction (NTP) module.

References:

1. Brunet, J.P., Tamayo, P., Golub, T.R., and Mesirov, J.P. (2004). Metagenes and molecular pattern discovery using matrix factorization. Proc Natl Acad Sci U S A *101,* 4164-4169.
2. Reich, M., Liefeld, T., Gould, J., Lerner, J., Tamayo, P., and Mesirov, J.P. (2006). GenePattern 2.0. Nat Genet *38,* 500-501.
3. Lee, D.D., and Seung, H.S. (1999). Learning the parts of objects by non-negative matrix factorization. Nature *401,* 788-791.
4. Sia, D., Jiao, Y., Martinez-Quetglas, I., Kuchuk, O., Villacorta-Martin, C., Castro de Moura, M., Putra, J., Camprecios, G., Bassaganyas, L., Akers, N., et al. (2017). Identification of an Immune-specific Class of Hepatocellular Carcinoma, Based on Molecular Features. Gastroenterology *153,* 812-826.
5. Yoshihara, K., Shahmoradgoli, M., Martinez, E., Vegesna, R., Kim, H., Torres-Garcia, W., Trevino, V., Shen, H., Laird, P.W., Levine, D.A., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun *4,* 2612.
6. Moffitt, R.A., Marayati, R., Flate, E.L., Volmar, K.E., Loeza, S.G., Hoadley, K.A., Rashid, N.U., Williams, L.A., Eaton, S.C., Chung, A.H., et al. (2015). Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. Nat Genet *47,* 1168-1178.
7. Saltz, J., Gupta, R., Hou, L., Kurc, T., Singh, P., Nguyen, V., Samaras, D., Shroyer, K.R., Zhao, T., Batiste, R., et al. (2018). Spatial Organization and Molecular Correlation of Tumor-Infiltrating Lymphocytes Using Deep Learning on Pathology Images. Cell Rep *23,* 181-193 e187.

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

9. Mayakonda, A., Lin, D.C., Assenov, Y., Plass, C., and Koeffler, H.P. (2018). Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res. *28*, 1747-1756.

Supplementary Tables

Table S1. Top 150 weighted genes.

Table S2. Top 150 exemplar genes of the immune module enriched in the immune associated ontology biological process and KEGG pathways.

ID	Description	P adjust
Ontology Biological Process		
GO:0042110	T cell activation	5.48E-49
GO:0019882	antigen processing and presentation	8.12E-10
GO:0042113	B cell activation	1.37E-09
GO:0038110	interleukin-2-mediated signaling pathway	4.30E-03
GO:2000316	regulation of T-helper 17 type immune response	1.91E-01
GO:0070098	chemokine-mediated signaling pathway	3.33E-19
KEGG Pathway		
hsa04658	Th1 and Th2 cell differentiation	1.80E-10
hsa04060	Cytokine-cytokine receptor interaction	1.56E-21
hsa04650	Natural killer cell mediated cytotoxicity	1.03E-06
hsa04660	T cell receptor signaling pathway	3.65E-07
hsa04662	B cell receptor signaling pathway	3.48E-04
hsa05235	PD-L1 expression and PD-1 checkpoint pathway in cancer	5.89E-04

Table S3. Immune associated gene signatures used in this study.

Signature Name	Reference
Immune enrichment score	Yoshihara *et al.* Nat Commun. 2013;4:2612
Stromal enrichment score	Yoshihara *et al.* Nat Commun. 2013;4:2612
Immune signalling molecules	Cancer Genome Atlas Network. Cell. 2015;161:1681-96
13 T-cell signature	Spranger *et al.* Proc Natl Acad Sci U S A. 2016;113(48):E7759-E7768.
T cells	Bindea *et al.* Immunity. 2013;39:782-95
CD8 T cells	Bindea *et al.* Immunity. 2013;39:782-95
Treg cells	Angelova *et al.* Genome Biol. 2015;16:64
TITR signature	Magnuson *et al.* PANS. 2018;115(45):E10672-e81
MDSC	Angelova *et al.* Genome Biol. 2015;16:64
T.NK. metagene	Alistar *et al.* Genome Med. 2014;6:80
B-cell cluster	Iglesia *et al.* Clin Cancer Res. 2014;20(14):3818–3829.
B.P. metagene	Alistar *et al.* Genome Med. 2014;6:80
Macrophages	Bindea *et al.* Immunity. 2013;39:782-95
TLS	Finkin *et al.* Nat Immunol. 2015;16:1235-44
6-gene IFN signature	Chow *et al.* J Clin Oncol. 34, (suppl; abstr 6010) 2016
CYT	Iglesia *et al.* Clin Cancer Res. 2014;20(14):3818–3829.
WNT/TGF-β signature	Lachenmayer *et al.* Clin Cancer Res. 2012;18:4997-5007
TGF-β1 activated	Ingenuity Pathway Analysis
C-ECM signature	Chakravarthy *et al.* Nature Communications. 2018;9(1)
Six immune subtypes of Pan-Cancer Atlas	Thorsson *et al.* Immunity. 2018;5(5):489-500
PAM50 pan-cancer	Zhao *et al.* Clin Cancer Res. 2019;25(8):2450-2457

Abbreviations: TITR, tumour-infiltrating Tregs; MDSC, myeloid-derived suppressor cell; IFN: interferon; TLS, tertiary lymphoid structure; CYT, cytolytic activity score; C-ECM, cancer-associated extracellular matrix.

Table S4. Top 150 difference genes in immune class compare with non-immune class in TCGA training cohort.

Table S5. List of the 21 genes selected form random forest algorithm.

Gene name

BCL2A1	GZMB
C1QA	GZMH
C1QB	HAVCR2
C1QC	HK3
CCL3	LAG3
CD2	LILRB2
CD8A	NKG7
CTLA4	PRF1
CXCL9	SLAMF8
FCER1G	TNFSF13B
GBP5	

Supplementary Figures

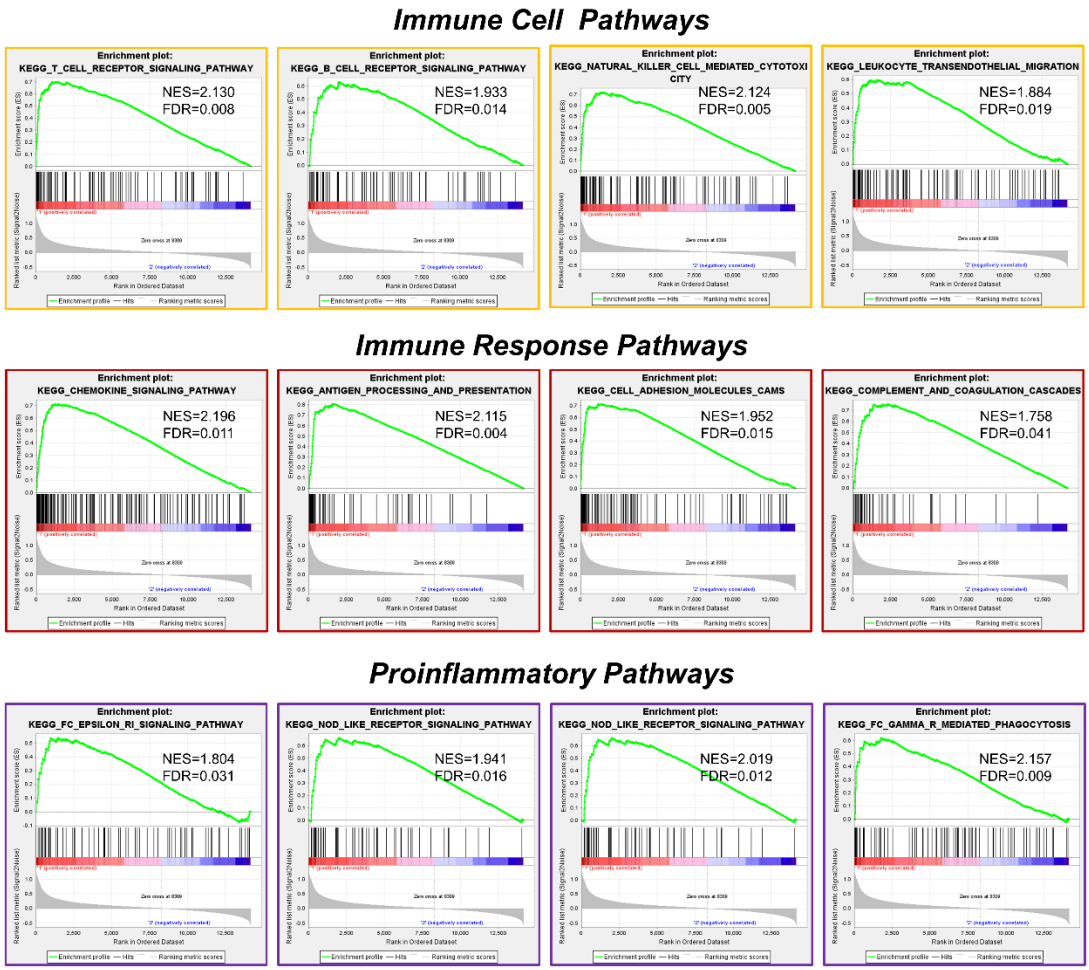


Figure S1

Figure S1. GSEA results showing the activated signaling pathways in the immune class. NES, normalized enrichment score; FDR, false discovery rate; FDR less than 0.05 indicates statistical significance.

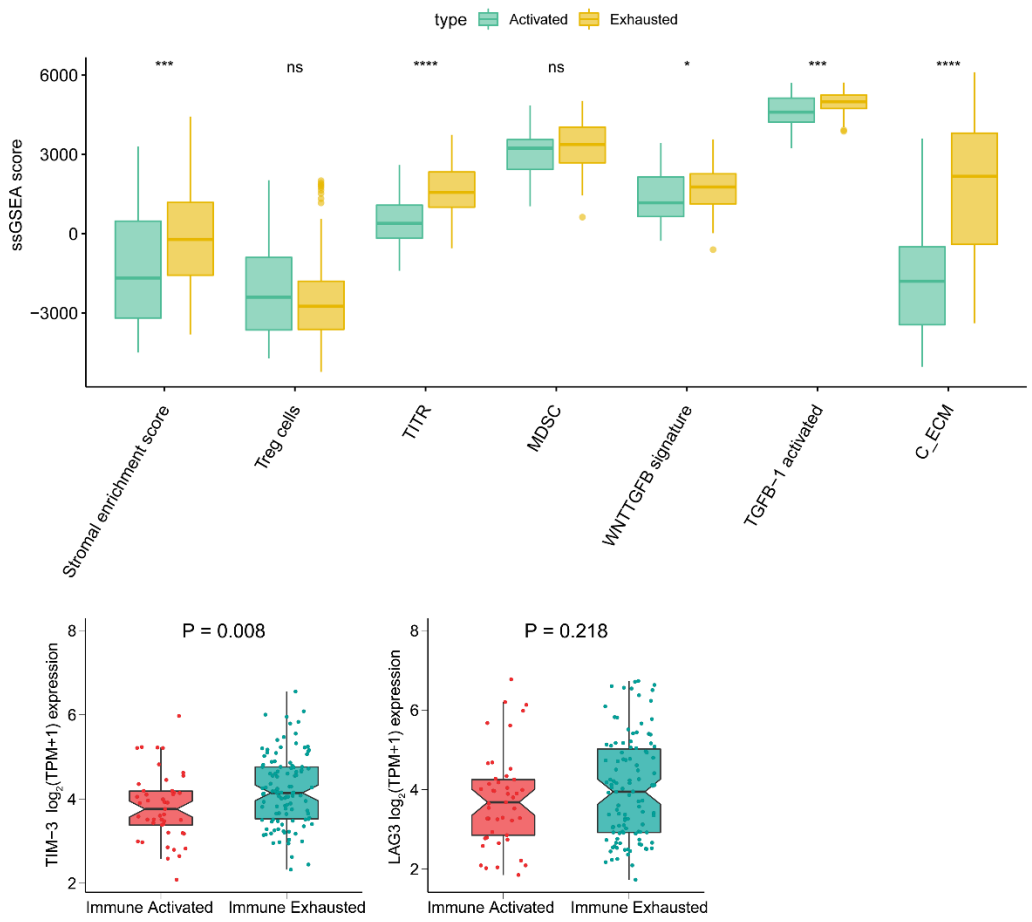


Figure S2

Figure S2. Stromal representative signatures and markers between immune-activated and immune-exhausted subgroups. ****, P<0.0001; ***, P<0.001; **, P<0.01; *, P<0.05; ns, no significance.

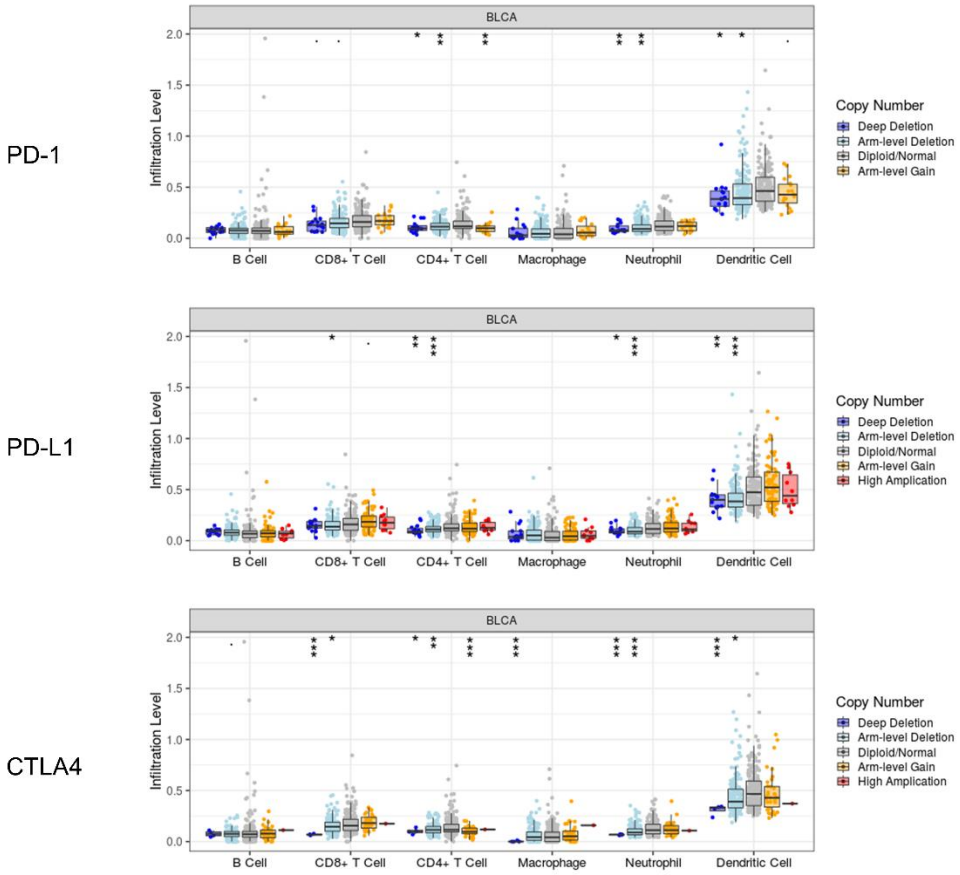


Figure S3

Figure S3. The association between copy number alteration of immune checkpoints and immunocyte infiltration. ***, P<0.001; **, P<0.01; *, P<0.05.

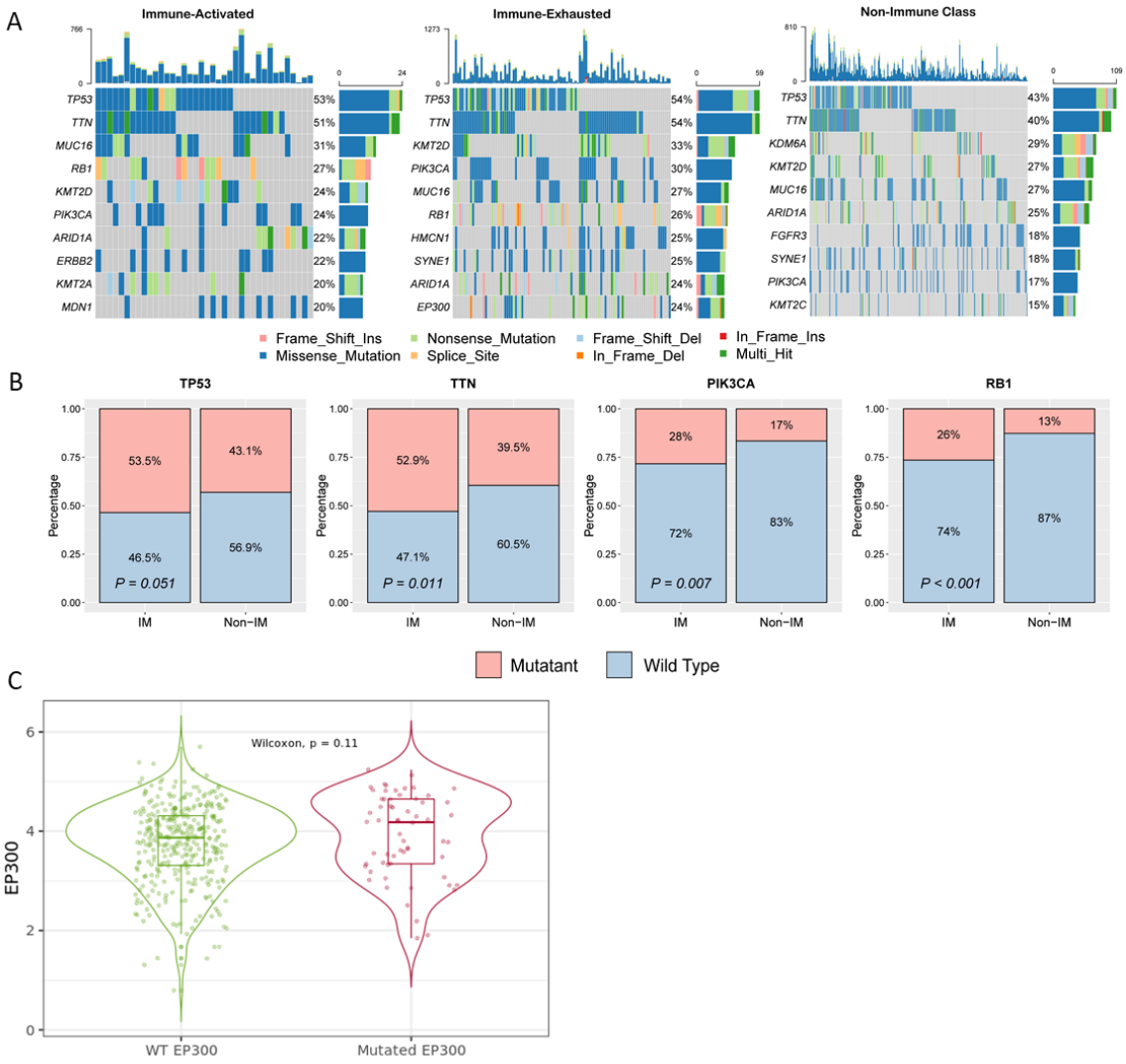


Figure S4

Figure S4. The top mutant genes of non-immune class, immune-activated subgroup, and immune-exhausted subgroup.

(A) Different distribution of mutant genes in three immunophenotypes. (B) TP53, TTN, PIK3CA and RB1 are the specific mutant genes in immune class compared with non-immune class. (C) The expression of EP300 in EP300wild type and mutated patients.

Figure S5. Reappearing the diverse immune characteristics among three immunophenotypes in E-MTAB-4321, GSE32894, GSE83586, GSE87304, GSE128702, GSE13507, GSE129871, GSE120736, GSE39016 cohorts.

CYT, cytolytic activity score; TITR, tumor-infiltrating Tregs; MDSC, myeloid-derived suppressor cell; TLS, tertiary lymphoid structure; C-ECM, cancer-associated extracellular matrix.

Figure S6. Reappearing the diverse immune characteristics among three immunophenotypes in GSE128701, GSE124035, GSE86411, GSE48276, GSE128192, GSE31684, GSE134292, GSE93257, E-MTAB-1803, GSE69795.

CYT, cytolytic activity score; TITR, tumor-infiltrating Tregs; MDSC, myeloid-derived suppressor cell; TLS, tertiary lymphoid structure; C-ECM, cancer-associated extracellular matrix.

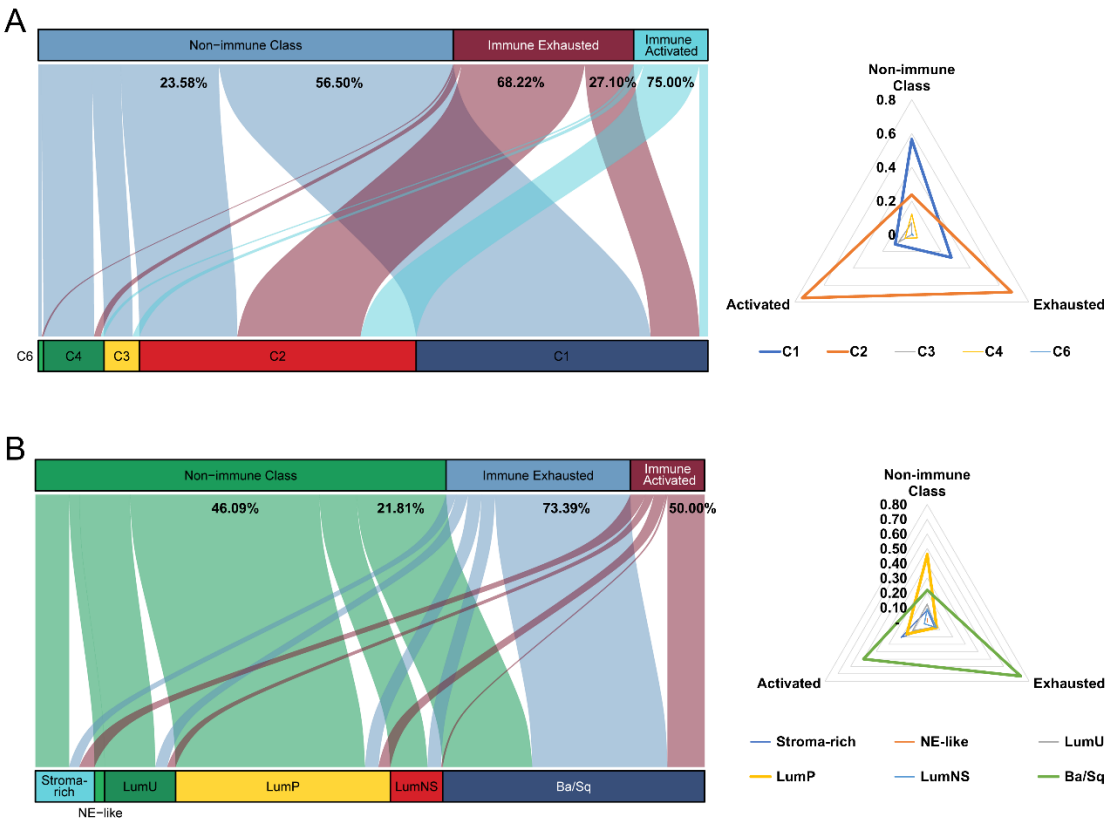


Figure S7

Figure S7. Correlate the three immunophenotypes with proposed molecular subtypes.

(A) Association with Thorsson et al. generated pan-cancer six immune molecular features;

(B) Association with Kamoun et al. identified the consensus set of six molecular classes.

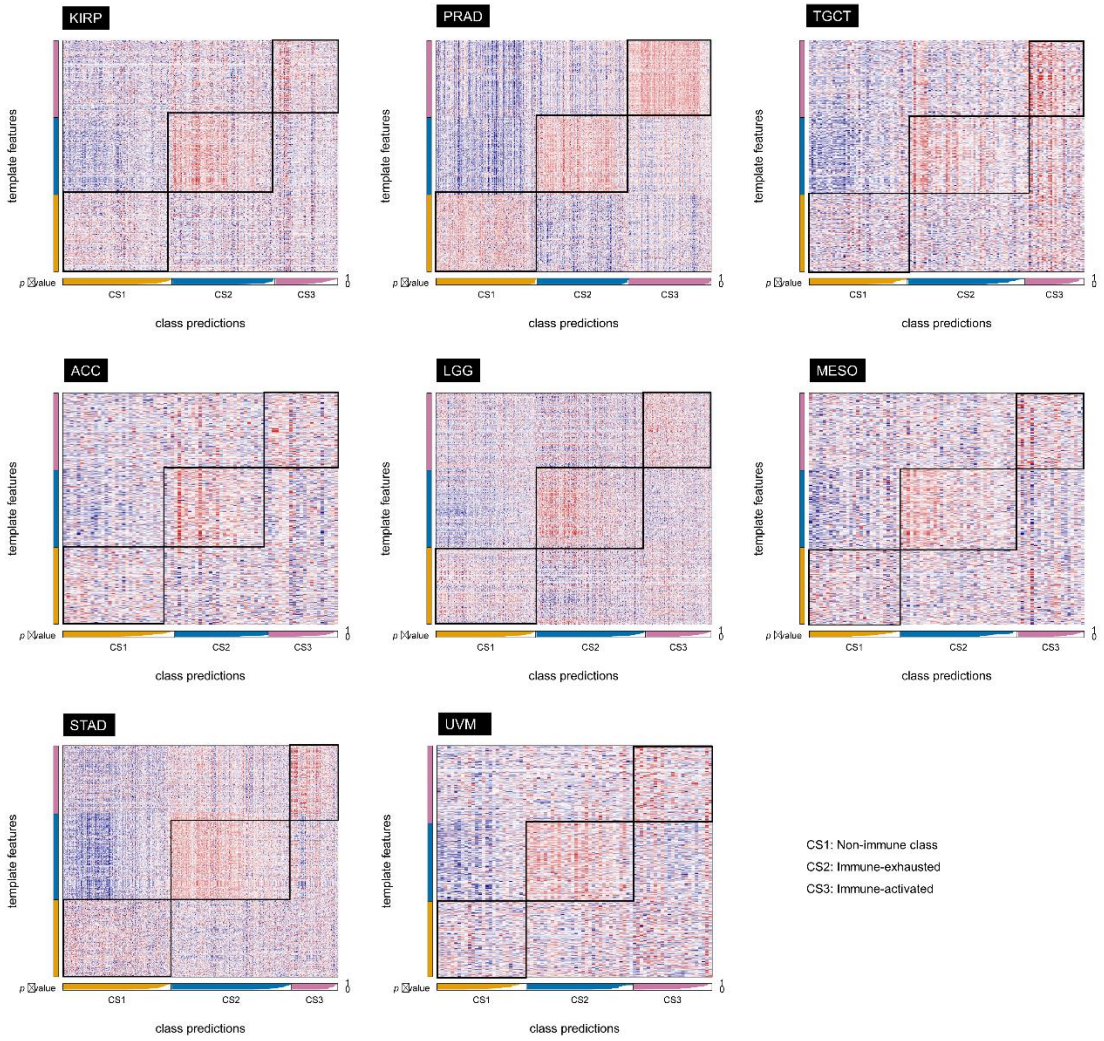


Figure S8

Figure S8. Verify the three immunophenotypes in pan-cancer.

KIRP, papillary renal cell carcinoma; PRAD: prostate cancer; TGCT: testicular germ cell tumor; ACC, adrenocortical carcinoma; LGG, brain lower grade glioma; MESO, mesothelioma; STAD, stomach adenocarcinoma; UVM, uveal melanoma.

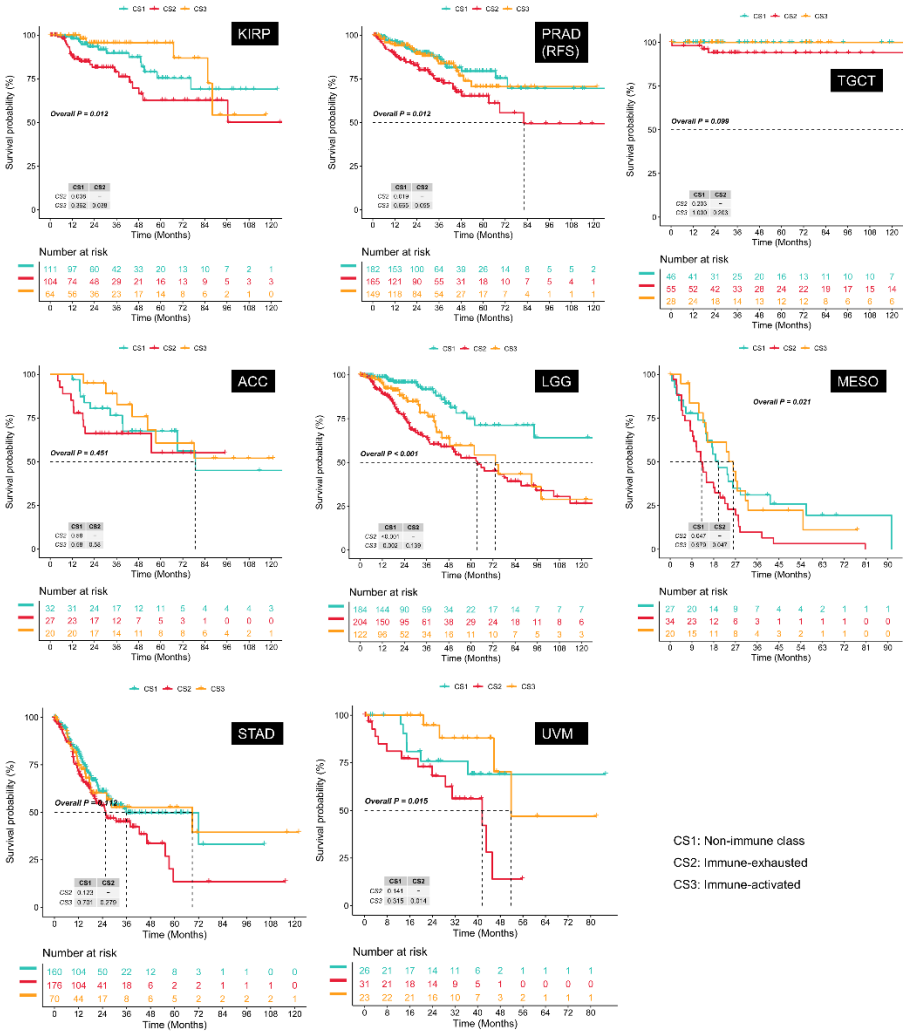


Figure S9

Figure S9. Different overall survival outcome of the three immunophenotypes in pan-cancer.

KIRP, papillary renal cell carcinoma; PRAD: prostate cancer; TGCT: testicular germ cell tumor; ACC, adrenocortical carcinoma; LGG, brain lower grade glioma; MESO, mesothelioma; STAD, stomach adenocarcinoma; UVM, uveal melanoma.

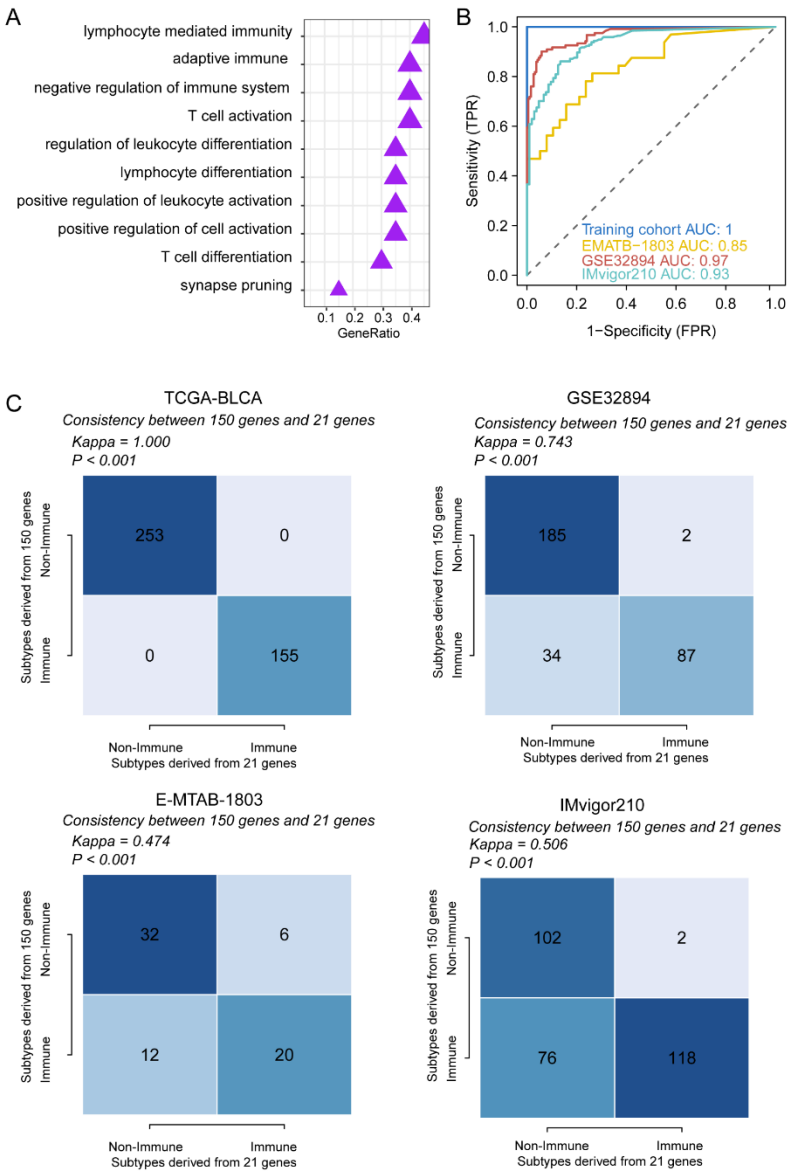


Figure S10. Dimensionality reduction of the 150 DEGs for the distinguishment of immune and non-immune classes

(A) Biological pathway enrichment of 21 genes; (B) ROC curve showing the distinguish value of the 21 genes in different cohort; (C) The consistency between 150 genes and 21 genes defined immune and non-immune classes.