Identification of tumor antigens and immune subtypes of bladder cancer for mRNA vaccine development

Yongwen Luo (luoywen@whu.edu.cn)
Zhongnan Hospital of Wuhan University
Rui Cao
Beijing Friendship Hospital
Zhenghao Chen
Beijing Friendship Hospital
Bo Ma
Beijing Shijitan Hospital
Gang Wang
Zhongnan Hospital of Wuhan University
Ye Tian
Beijing Friendship Hospital

Research Article

Keywords: mRNA vaccine, immunotype, bladder cancer, Tumor immune microenvironment, Immune landscape

Posted Date: March 11th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1433880/v1

Abstract

Background: Bladder cancer (BLCA) is a common malignancy from urinary tract. Although the diagnosis and treatment of bladder cancer has made great progress in the past few decades, the effects of existing treatment methods are still limited. Therefore, it is still necessary to develop new methods to assist in the disease management and treatment. Tumor antigens are tumor-specific surface molecules and are generally considered to be the main components of a typical cancer vaccine, which could initiate and active immune cells to recognize and eliminate cancer cells. In the context of the COVID-19 pandemic, mRNA vaccines have re-entered people's vision.

Methods: The genomic and clinical data of 411 BLCA and 19 normal tissues were acquired from The Cancer Genome Atlas (TCGA) and GSE13507 cohorts. Differential expression genes and mutation analysis were performed to screen out potential antigens, Kaplan-Meier curves were carried out to investigate the correlation between the level of potential antigens and OS of patients. Immuno-phenotyping of 411 tumor samples was based on the single-sample gene sets enrichment analysis (ssGSEA). The tumor immune microenvironment characteristics was explored in each immune subtype. Weighted gene co-expression network analysis (WGCNA) was used to clusterimmune-related genes and screen the hub genes, and pathway enrichment analyses were performed on the hub modules related to immune subtypes in the WGCNA.

Results: Through genetic and transcriptional analysis on TCGA and GSE13507 datasets, we have identified 6 genes as potential candidate genes for BLCA specific tumor antigens. We also identified 3 immune subtypes of BLCA, which displayed distinct clinical, molecular and immune-related characteristics. In addition, we have constructed immune landscape to identify the immune cell components of each BLCA patient, which could predict clinical outcome of the patients, and assist in the development of personalized mRNA vaccines.

Conclusions: our findings indicated that 6 genes such as PTPN6 may be potential tumor antigens, and provide a reliable reference for the further development and management of cancer vaccines.

Background

Bladder cancer (BLCA), a common malignancy of the urinary tract, is the ninth most common cancer (1). About 70–80% of first diagnosis BLCA patients are non–muscle-invasive bladder cancer (NMIBC), of which 50–70% have a high rate of recurrence, and 10–30% are inclined to develop into muscle-invasive bladder cancer (MIBC) (2). Although transurethral resection (TUR) is the preferred regimen for patients with NMIBC, the five-year overall survival (OS) is only 60% for stage pT2, 35% for pT3, and 25% for pT4 BLCA(3, 4) Therefore, it is necessary to explore other new strategies to improve clinical outcome of BLCA.

Tumor antigens are cancer specific surface molecules, which are often overexpressed and easily recognized by the immune system(5). They are usually regarded as the main component of typical cancer vaccine, assisting immune cells in recognizing and eliminating cancer cells(6). Cancer vaccines

can be divided into peptide, tumor cell, dendritic cell, DNA and RNA types according to their antigenic forms(7).

Recently mRNA-cancer vaccine has become increasingly attractive to scientists and oncologists and is a hotspot in cancer immunotherapy(8, 9). Actually, mRNA-based therapy was not common before the 2000s due to the instability of mRNA and related excessive inflammation responses. Although mRNA-based vaccine has been proposed as a promising approach of combatting tumors two decades ago, it is coming back our vision under the background of coronavirus disease-2019 (COVID-19) pandemic(10, 11). Nowadays, many technological breakthroughs, including incorporation of modified nucleosides, purification of IVT mRNA, optimization of coding sequences, and development of efficacious delivery material have been changing the situation by enabling mRNA optimal form to carry tumor antigens(12, 13). Compared with conventional approaches, mRNA vaccine has its major advantages as follows(9, 14): (1) Safety profile: mRNA vaccine is non-infectious, which means it is not made with pathogen particles or inactivated pathogen. And mRNA does not integrate itself into the host genome or excluded irrelevantly. Once the protein is made, mRNA is degraded by cellular RNases with a short and regulatable half-life in vivo. (2) High efficacy: In the context of clinical trial, the mRNA vaccine is well tolerated by healthy individuals, with few side effects after eliciting a reliable immune response. Plus, mRNA sequences can be easily designed to encode any pathological antigen, which is conducive to individualized therapies. (3) Easily manufactured: mRNA vaccines can be produced more rapidly under standardized processes improving responsiveness to emerging outbreaks.

At present, the immunotherapy of BLCA is mainly based on the Bacillus Calmette-Gueri (BCG) vaccine as well as immune checkpoint inhibitors. BCG immunotherapy reduces the incidence of cancer recurrence and progression. (15) Typically, patients receive "6 + 3" intravesical BCG instillations per week, with a 6-week break between induction and maintenance instillations. Maintenance therapy improves recurrence-free survival. (16) However there are still tones of side effects ranging from common, mild and transient symptoms, such as dysuria and flu-like symptoms, to more severe and rarely occurring life-threatening complications. (17). AS for ICIs, Typical immune-related AEs reported were pneumonitis, rash, and elevated levels of liver enzymes, requiring specific management by the clinicians(18). Though ICIs have shown manageable safety profiles as single agents(19), they were far less tolerable in combination with chemotherapy. In fact, as reported in previous studies, the rates of all-grades of AE were much higher than ICIs alone (up to 96% for AEs, and 81% for severe AEs), leading to treatment withdrawal in up to one out of three patients.(20)

Therefore, there is still a long way to go to solve the above problems perfectly. Considering that BLCA still lacks an effective mRNA vaccine, it is still a challenge to isolate effective antigens for anti-BLCA mRNA vaccines from thousands of mutated candidate genes, it is also a worthwhile direction given the advantages of mRNA vaccines. Furthermore, due to the heterogeneity of tumors and their complex immune microenvironment (TIME), traditional methods are insufficient to screen candidate genes suitable for mRNA vaccines from the perspective of immune regulation. Stratification based on immune

gene expression profiles may be suitable for identifying patients receiving mRNA vaccination from immune heterogeneous populations

In this study, we identified six new potential tumor antigens for anti-BLCA mRNA vaccine development. Immuno-typing for identifying suitable BLCA patients for vaccination was also investigated. Six tumor antigens were correlated with superior prognosis in BLCA. Then the BLCA patients were stratified into three immune subtypes, which were associated with differential cellular, molecular and clinical features in different cohorts. Our findings might be a valuable and reliable reference for further developing and administering cancer vaccines.

Material And Method

1. Data extraction

The normalized mRNA-seq data and corresponding clinical information of 411 BLCA and 19 normal tissues were acquired from The Cancer Genome Atlas (TCGA). Then, the mRNA data were merged and normalized as one cohort by the R package “limma”. TCGA-BLCA somatic mutation data were downloaded from the Genomic Data Commons (https://portal.gdc.cancer.gov/) using the package TCGAbiolinks in R. The called somatic variants processed with the MuTect2 algorithm were utilized as the raw mutation count. Then we selected the mutation data of the corresponding sample in the BLCA expression profile from the mutation data, and use the oncoplot function in the "maftools" package to display the mutation.

2. Differential expression analysis to identify overexpressed genes

We use the R package "limma" to identify differentially expressed genes in cancer and normal samples (threshold: adj.P.Val < 0.05 & |log(FC) |>= 1). Next, R package "EnhancedVolcano" was used to visualize differentially expressed genes, and "ComplexHeatmap" was used to display the expression of overexpressed genes in normal and cancer samples. Finally, the position of the differential gene in the genome was displayed with GEPIA (http://gepia.cancer-pku.cn/) to display.

3. Identification over-expressed and mutated genes as potential antigen candidate genes and functional analysis

The intersection of overexpressed genes and mutant genes is selected as potential antigen candidate genes (764). Then R package "clusterProfiler" was used to perform enrichment analysis of the intersection genes on GO/KEGG/HALLMARK/REACTOME type pathways (pvalueCutoff < 0.05).

4. Survival analysis of potential antigen candidate genes to identify prognostic-related potential antigens

We used R packages "survival" and "survminer" to perform univariate Cox regression analysis of potential antigen candidate genes, and the cut-off was $P < 0.05$. The survival curve of the most significant potential antigen in OS is displayed with K-M curves.

5. Correlation analysis of prognostic-related potential antigens and immune cell infiltration

The CIBERSORT algorithm could calculate the proportion of 22 kinds of immune cells through the expression of some genes. We use Spearman correlation analysis to demonstrate the correlation between the expression of prognostic-related potential antigens and immune cell infiltration through the "ggplot2" package.

6. Immune subtype recognition

In order to identify immune subtypes, we clustered the cancer samples in BLCA based on the expression of immune gene sets. The clustering were verified by principal component analysis, and the expression levels of immune genes in each subtype were displayed. In addition, the survival curves of each subtype are shown.

7. Clinical features between subtypes

We use the "ggplot2" R package to visualize the distribution of subtypes in each clinical feature, and perform analysis of variance between each feature

8. CNV analysis between subtypes

We first showed the distribution of tumor mutation burden (TMB) among the subtypes, and then extracted the display the mutation distribution of the corresponding samples from the mutation data by subtype. And the oncoplot function in the "maftools" package was used to show the mutations distribution separately. Finally, the samples with the absolute value of copy number G-score greater than 0.4 are selected, and the scores in each subtype are displayed. The small tool "CNV Distribution Map" was used to display the distribution of copy numbers in chromosomes.

9. Analysis of immunological activity between subtypes

We obtained the immune activity score from previous study (21), and used the "ggplot2" and "ggpubr" R packages to compare the differences between the different subtypes of immune cells

10. Analysis of different indicators in subtypes

The homologous recombination defect (HRD), neoantigen load and chromosomal instability, and difference in dryness index of TCGA-BLCA dataset was obtained from Malta TM and Knijnenburg TA (22, 23), and then showed the distribution of each index in the subtype And the difference. Then we obtained immune checkpoints and immunogenic cell death regulatory factors from previous study (6), and showed the expression of each factor among subtypes, and the difference between each factor among subtypes.

11. Analysis of immune cells between subtypes

We used the R package "estimate" to calculate the immune score, stroma score, and tumor purity of cancer samples. Then R package "ggplot2" was used to to display the distribution of the above scores in different subtypes. And the R package "ComplexHeatmap" was used to display the heat map of immune infiltration scores and above. In addition, we also compared the subtypes identified in this paper with the existing pan-cancer subtypes of TCGA pan-cancer cohort, showing the proportion of the existing pan-cancer subtypes.

12. Construction of immune map

The dimensionality reduction function of Monocle package of Gaussian distribution is used to perform dimensionality reduction analysis based on graph learning to visualize the distribution of immune subtypes among individual patients. Finally, different colors are used to show the immune landscape of the cell trajectories of immune subtypes. Then the correlation between cell infiltration and principal components was calculated. Finally, samples in extreme locations are selected to show their survival.

13. WGCNA

Use the "WGCNA" R package to construct a WGCNA co-expression network for immune gene expression, showing the number of genes in the generated modules and the ME scores of samples in different subtypes in each module. Then, univariate cox regression analysis was used to identify the prognostic module, and GO and KEGG enrichment analysis were performed on the genes in the identified prognostic-related modules. For the prognostic module, we calculated the MM score of the gene (the correlation between the gene and the module) and selected genes with MM greater than 0.5 as the hub gene of the prognostic module. Subsequently, a multivariate cox analysis was performed on the identified hub genes, and the obtained coef value was used as a coefficient to construct a risk score. Then, the distribution of

the score, the survival of the high and low risk groups were displayed, and the AUC curve and the expression heat map of the hub gene were displayed.

14. Differential gene recognition between subtypes

We used ANOVA and t.test to identify differentially expressed genes in cluster1 and cluster2, cluster2 and cluster3, and cluster1 and cluster3 samples (threshold: anova p < 0.05 & group p < 0.05), and selected 4,505 intersecting genes with three differences. Next, "ComplexHeatmap" was used to show the expression of differentially expressed genes in cluster1, cluster2, and cluster3 samples and in each phenotype.

15. LASSO regression to construct a risk score

First, we selected the inter-subtype differential genes (4505), potential antigen candidate genes (764) and the prognostic module hub gene (518) of the intersection genes (10), and defined them as key tumor antigen markers. We used univariate cox regression to screen key tumor antigen markers, selected genes with P less than 0.05 (3) for multivariate cox regression, and used the obtained coef value to construct a risk model. Finally, it shows the distribution of immune subtype risk scores and the distribution of high and low risk scores in each subtype.

16. Statistical Analysis

Unpaired Student’s t-test was used to compare the two groups with distributed variables. One-way analysis and Kruskal-Wallis tests of variance were adopted as parametric and nonparametric methods, respectively, for comparing multiple groups. Contingent variables were analyzed with the chi-square test or Fisher’s exact test. Pearson’s test or Spearman’s test was conducted to analyze the correlation between gene expression and the abundance of immune cells or gene expression. All statistical analyses were performed on GraphPad Prism 7.0 or R software (Version 3.6.0, https://www. r-project.org/). A two-tailed P value < 0.05 was considered statistically significant.

Results

Identification of potential antigens of BLCA

The workflow was presented in Figure S1. 900 differentially overexpressed genes were screened out within a total of 1810 aberrantly expressed genes to identify potential antigens of BLCA (p < 0.05 & |log(FC)| >= 1) (Figure S2 A-B, SupplementaryTable1 (all differentially expressed genes) and SupplementaryTable2 (overexpressed genes)). It can be clearly found that the difference in the expression of over-expressed genes between cancer and normal samples. The chromosomes distribution

of these genes was showed in Fig. 1A. We identified 15751 genes that mutated in TCGA-BLCA cohort and the top 30 mutated genes were showed in Fig. 1B. The altered genome fraction and mutation counts in individual samples were demonstrated in Fig. 1C-D. Of note, AHNAK2 and ARID1A were also the most frequently mutated genes with both altered genome fraction and mutation counts (Fig. 1E-F). In addition to AHNAK2 and ARID1A in top 10 candidates with altered genome fractions, ANK2, SAXL2, CHD2, CREBBP, CSMD3, DCC, DMD as well as DNAH5 all have the high mutation frequency (Fig. 1F). High mutation counts were also observed in FRGFR3, KDM6A, RYR2, TP53, TTN, KMT2D, SPTANI, ZFP36L1(Fig. 1E). Overall 764 genes, which were highly expressed and mutated in TCGA-BLCA cohort, were identified as potential candidate antigens (Fig. 2A, Figure S2f).

Then Gene Ontology (GO) analysis and KEGG HALLMARK and REACTOME signaling pathway enrichment of these 764 genes were respectively shown in Fig. 2B and Figure S2C-E. The results indicated that these genes might involve in tumor-specific immune response and could be the potential candidates for mRNA vaccine development.

Identification of tumor antigens associated with BLCA prognosis and antigen presenting cells

To identify potential prognostic-related antigens, we performed univariate cox regression analysis on candidate genes based on overall survival (OS) in TCGA-BLCA cohort ($P < 0.05$), and finally found 102 prognostic-related potential antigens (Fig. 2A). We selected the 6 potential antigens, PTPN6, OAS1, PLA2G2F, ETV7, CHMP4C, and SPA4G, which were most associated with OS (Fig. 2C-D). We further used CIBERSORT algorithm to analyze the immune infiltration in TCGA-BLCA cohort. Then the spearman correlation between the expression of 6 potential antigens and B Cell, Macrophages, and Dendritic scores was detected. The final correlation score is distributed between 0.0011 (CHMP4C in B_cells) and 0.32 (CHMP4C in Macrophages), indicating that the correlation between potential antigens and immune infiltration scores is not evenly distributed, which also indicated the differences in immune levels of different genes. (Figure S3).

Immune subtypes identification

Tumor immune microenvironment might impact the efficacy of immunotherapy. And immune subtypes might be helpful to identify patients who can response to mRNA vaccination. We analyzed the transcriptional profiles of patients in TCGA-BLCA cohort and found 2323 immune-related genes. According to consensus clustering algorithm, by considering its cumulative distribution function and incremental area map, we chosed $k = 3$ for stable clustering of immune-related genes, and established 3 immune subtypes, named Cluster 1, Cluster 2, and Cluster 3, respectively(Figure 3A-C,3E). Principal component analysis showed the three subtypes were distinguished very well (Fig. 1D) with distinct OS. Cluster 3 associated with better prognosis, while Cluster 1 had the poorest survival probability (Fig. 3F).

Subtype distribution across different tumor stages and grades indicated that patients diagnosed as differential stage were irregularly clustered And we could found that with the TNM staging progress, the proportion of Cluster 1 increases significantly (Figure S4). Taken together, the immunotyping could be used to predict prognosis of BLCA patients and its accuracy was superior to traditional grading and staging. We also validate the feasibility of this classification in the testing dataset GSE13507 (P = 0.014) (Fig. 3G).

The association of immune subtypes with mutational status

Previous studies demonstrated that tumor mutational burden (TMB) and mutation used to quantify the number of tumor antigens is closely correlated to immunotherapeutic efficacy, including mRNA vaccine(24). Therefore the TMB and mutations frequency of each immune subtypes were also analyzed(Fig. 4). Our results showed that there was no difference between groups regarding TMB, while there were significantly difference of mutation number among these three groups, and Cluster 1 was obviously higher than the other groups. The waterfall diagrams of different subtypes were showed in Fig. 4A-B. We also visualized the G-score scores of copy number variation in different subtypes (Fig. 4F) and found significant differences between each subtype. Furthermore we also obtained the immune activity scores of the samples from TIP (http://biocc.hrbmu.edu.cn/TIP/), and showed the distribution of Cluster 1, Cluster 2, and Cluster 3 scores in immune cells, and showed statistical significance (Fig. 5A). It can be found that the immune subtypes we have identified are significantly related to immune activity. And then we analyzed differences in homologous recombination defects (HRD), neoantigen load, chromosomal instability, and dryness index between subtypes (Figure S5). These findings suggest that the immune subtype can predict TMB mutation number in BLCA patients, and that patients of Cluster 1 may respond positively to the mRNA vaccine.

Immune subtypes with immune checkpoints (ICPs) and immune infiltrating cells

Previous studies indicated that both ICPs (such as PD-L) (10) and immunogenic cell death (ICD) modulators (such as HMGB1)(25) play critical roles in modulating the host anti-tumor immunity, which could influence the efficacy of mRNA vaccine. Therefore, the differential expression of ICPs and ICD modulators was assessed in these immune subtypes (Fig. 5B-C, 5F). Our results showed that, 43 (91.5%) of ICPs in TCGA cohort were differentially expressed among the immune subtypes. And the expression levels of almost all genes in Cluster 1 and Cluster 3, such as CD274, CTLA4, PDCD1, IDO1 and TIGIT, were lower than those in Cluster 2. Interestingly, some specific ICPs in Cluster 3 seemed to be lower than those in Cluster 1 such as BLTA, CD200, CD200R1, CD27, CD274, CD276, CD28, CD48, CD77, CD80, CD86, CTLA4, ICOS, IDO1, LAIR1, PDCD1, PDCD1LG2, TIGIT, TNFRSF18, TNFRSF4, TNFRSF9, TNFRSF14. This result indicated that Cluster 3 might also be a potentially effective population for mRNA vaccines despite lower levels of TMB. For immune cell deaths (ICDs), The expression level of 17 ICDs of all 25 kinds of

ICDs were significantly different in three immune subtypes. The level of CALR, CXCL10, FPR1, HGF, IFNAR2, MET, P2RX7, TLR4, PANX1 and TLR3 were highest in Cluster 2 than Cluster 1 and Cluster 3, EIF2AK2, EIF2AK4, ANXA1 were overexpressed in Cluster 1. While EIF2A, IF2AK1were both overexpressed in Cluster 1 and Cluster 3. Moreover, we showed that Cluster 2 had markedly higher immune score, higher stromal score and lower tumor purity (Figure S6) than Cluster 1 and Cluster 3. The results showed that the subtype scan reflect the expression levels of ICPs and ICD regulators, and can be used as a potential therapeutic biomarker for mRNA vaccines. Also, considering the previous results of immune cell infiltration, Cluster 2 is an immune “hot” and immunosuppressive phenotype, while Cluster 1 and Cluster 3 were more likely to be immune “cold” phenotypes.

In a previous study, Thorsson et al. identified six immune categories (C1-C6) based on the immunogenomic analysis of more than 1000 tumor samples among 33 cancer types (26). These categories were significantly associated with prognosis, genetic, and immune-modulatory alterations in tumors. Thus, the distribution of six categories was also investigated in our study. A distinct distribution over Cluster 1, Cluster 2 and Cluster 3 was observed, and the individual immune categories substantially varied in their proportion in the three immune subtypes (Fig. 5E). Our result indicated that proportion of C2 in Cluster 2 was obviously higher than that in Cluster 1 and Cluster 3, while the proportion of C4 in cluster3 was significantly more than that of C4 in Cluster1 and Cluster2 and there was barely no C4 in Cluster 2. This result further validated the “hot” phenotype of Cluster 2 and the “cold” phenotype of Cluster 1 and Cluster 3. Therefore, mRNA vaccine administration in Cluster 1 and Cluster 3 might stimulate the immune response, namely turning “cold” tumor to “hot”.

Immune landscape of BLCA

Then we used the immune transcriptional profile of patients to construct immune landscape (Fig. 6A). As shown in Fig. 6B, Macrophages.M1 displayed the most obviously correlation among different subtypes. We found the distribution of samples in Cluster 1 and Cluster 2 was relatively discrete in immune landscape (Fig. 6A) and they can be further divided into two subgroups according to the location of immune cell populations (Fig. 6C). We then found the relative amount of immune cells were significantly differences among multiple subgroups (Fig. 6F-G). In addition, patients in group 5 had a higher survival probability when comparing with the samples in the extreme locations within the immune landscape (Fig. 6D-E). The immune landscape based on immune subtypes can accurately identify the immune components of BLCA patients and predict their prognosis, providing favorable conditions for personalized treatment options for mRNA vaccines.

Identification of immune gene co-expression modules of BLCA

Immune gene co-expression modules were identified by clustering the samples by WGCNA.Using the default parameters of the "WGCNA" R package with a soft threshold of 4 for scale-free network. The representation matrix was then converted to an adjacency and then to a topological matrix. We finally got 5 modules, of which the gray modules are not grouped together with other modules (Figure S7 and Fig. 6A). The survival analysis was also identified (Fig. 6B). It is found that the Turquoise module had significant prognostic efficacy, with differently expressed pathways such as cytokine pathway, natural killer meditated cytotoxicity pathway as well as antigen progressing and presentation pathway (Fig. 6C-D) .

Identification of differential expression genes (DEGs) between immune subtypes

To explore our immunotyping more deeply, we intersected the differentially expressed genes between distinct three clusters ($p < 0.05$) (Figure S7 D and SupplementaryTable3), and showed it in the heat map (Figure S7E).

Then, by intersecting the differential expression genes between subtypes (4505), potential antigen candidate genes (764) and prognostic module hub genes (518), we have finally defined 10 genes as key tumor antigen markers (Fig. 7A). We performed a univariate cox regression analysis on all key tumor antigen markers and screened 3 genes with $P < 0.05$ (Fig. 6B-C). Through LASSO regression analysis, we built a risk model (Fig. 2D). We found the expression of MMP9 and INHBA in the low-risk group was significantly lower than that in the high-risk group, while ERBB3 was significantly higher in low-risk group, which also confirmed their risk and provides a direction for clinical research (Fig. 2E). In addition, we also found that MMP9 is also a potential prognostic-related antigen, which is consistent with our previous study. It can be found that the risk scores are significantly different between subtypes, and the proportion of high-risk samples in Cluster 2 is relatively high, while the proportion of low-risk samples in Cluster 3 is relatively high (Fig. 2F). Taken together, we believed that ERBB3, MMP9 and INHBA were potential BLCA antigens biomarkers for mRNA vaccine development.

Discussion

Nowadays, immunotherapy plays an important role in treatment of BLCA. Immune checkpoint inhibitors (ICIs) were approved by FDA- as a second-line systemic therapy for patients with metastatic and locally advanced bladder cancer, and as a first-line therapy for cisplatin-ineligible patients. The enthusiasm for checkpoint inhibition has consequently carried over into treatment for NMIBC, where there is a significant unmet need for efficacious second-line treatment options for patients who have failed BCG(27). In parallel to the developments of systemic immunotherapy and intravesical agents, there have been several noteworthy advancements in vaccine-based immunotherapy. Most tumor antigens exhibit immunogenicity that is too weak to elicit an effective antitumor immune response. The immune response to tumor antigens leads to the reduction or loss of tumor surface antigens, so that tumor cells are not

recognized by the immune system and escape immune attack (antigen modulation) (28). Tumor antigens may be coated with substances, such as salivary mucopolysaccharides, that are not recognized by the host's lymphocytes and thus cannot induce tumor cell killing effect. Therefore, when endogenous tumor antigens fail to elicit an effective immune response, it is necessary to reactivate and modulate the immune response through exogenous vaccines(29) .

Therapeutic cancer vaccines eliminate tumors by activation of acquired cellular immunity against specific tumor-associated antigens and come in various forms(30). Recent years, a cancer vaccine named PANVAC-VF that utilizes recombinant vaccinia and fowlpox viruses engineered to contain genes for human carcinoembryonic antigen and mucin 1 (both known to be overexpressed in urothelial carcinoma) (31) as well as an agent, which is called HS-410 (Vesigenurtacel-L), designed by Heat Biologics to serve as an intradermal vaccine to enhance the effects of BCG(32). However, the development of mRNA vaccines for bladder cancer is still in a blank state, and there is still a lot of research value.

In this study, we focused on finding neoantigens unique to BLCA that contribute to creating the personalized mRNA-based vaccine. The differentially expressed and mutational profile of BLCA was constructed, and 6 targetable antigens (PTPN6, OAS1, PLA2G2F, ETV7, CHMP4C and SPA4G) were further confirmed. Their overexpressions were associated with poor OS as well as connection with APCs and B cell infiltration. The findings revealed that these candidates have an outsize impact on the progression and prognosis of BLCA. Although these antigens need more indepth clinical evaluation, their potential for BLCA-related mRNA development has been consolidated by previous studies. For instance, PTPN6 was proved to be overexpressed in BLCA tissues compared with normal bladder tissues and significantly correlated with grade, stage, T, and N, while survival analysis showed that low expression of PTPN6 was significantly related to the poor overall survival (OS) in BLCA patients(33). OAS1 had been shown in many previous studies to be a class of classical immune-related genes(IRGs) and has been used as an adjunct to predict bladder cancer prognosis(34–36). ETV7 was identified as the key differentially expressed transcription factors DE-TFs due to its association with the autophagy activation pathway and various immune cells in bladder cancer(37, 38) as well as its functional role in breast cancer(39), gastric cancer(40), ovarian cancer (41) et al. Although the study of PLA2G2F in bladder cancer is still relatively lacking, previous studies have demonstrated that PLA2G2F, as a functional monocyte expressed in the suprabasal epidermis, secretes phospholipase A2, can regulate skin homeostasis and hyperplasic disorders(42), and Role in human colorectal adenocarcinomas(43). Similarly, CHMP4C has not been well studied in bladder cancer, but has demonstrated its role in other tumors(44–47). Of course, we must admit that further research is essential, and the function of these identified vaccine antigens must be verified before actual clinical development and application.

To screen suitable populations for optimal mRNA vaccination, we subdivided BLCA patients into three immune subtypes based on immune gene expression profiles. We observed distinct molecular, cellular and clinical features in the three immune subtypes Cluster 1–3. Patients in Cluster 1 had the poorest prognosis among all the three subtypes of BLCA. In addition, malignant phenotypes such as higher

histological grade as well as worse TNM stage were confirmed in patients with subtype Cluster 1 compared with the other subtypes, this suggests the immunotyping can be used for predicting the prognoses of PAAD patients.

In addition to prognostic prediction, immunotyping is also indicative of the therapeutic response to mRNA vaccine. High tumor mutation burden has been usually proposed as a predictive biomarker for response to immune checkpoint inhibitors (ICIs), largely due to the potential for tumor mutations to generate immunogenic neoantigens (48, 49), and in our study, subtype Cluster 1 has the highest levels of TMB indicating Cluster 1 was the suitable candidates for mRNA vaccine against BLCA. And this conclusion is also consistent with the conclusion of ICPs and ICD.

mRNA vaccine could be more effective in patients with upregulation of ICD modulators, while patients with higher ICP expressions were not suitable for mRNA vaccine (50). In our study, Both Cluster 1 and Cluster 3 had lower levels of ICPs, while Cluster 2 had higher expression of ICPs, and higher ICPs usually suggests an immunosuppressive tumor microenvironment, which may inhibit the mRNA vaccine from eliciting an effective immune response. Therefore, considering the performance of ICPs and ICDs, we believe that Cluster1 is a more suitable mRNA vaccine candidate.

Biomarkers of immune subtypes are the hub of linkage mechanism research, population screening, and typing specificity. Our WGCNA revealed key module closely associated with each immune subtype and were of great significance to explore the potential biological mechanism of subtypes. KEGG and GO analysis showed that the Turquoise modules had apparent differences in biology and involved pathways, which further suggested that the classification based on this study was of a high degree of discrimination.

Conclusion

In summary, our study identified PTPN6, OAS1, PLA2G2F, ETV7, CHMP4C and SPA4G as potential effective neoantigens for BLCA mRNA vaccine development, and patients with Cluster 1 and Cluster 3 tumor might benefit more from mRNA vaccination. Also ERBB3, MMP9 and INHBA were potential BLCA antigens for mRNA vaccine development. Our study paved a way for future mRNA vaccine development and define the suitable population for vaccination.

Declarations

Ethics approval and consent to participate

This type of observational study does not require approval.

Consent for publication

All authors wrote and approved the final manuscript

Availability of data and materials

The datasets supporting the conclusions of this article are included within the article and supplementary files.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by grants from Zhongnan Hospital of Wuhan University, Excellent Doctor Fund Project (ZNYB2020005) to Yongwen Luo and National Natural Science Foundation of China (Grant No. 81802550) to Rui Cao

Authors’contributions

YWL conceived the project. YWL, RC, and ZHC analyzed the data. BM ,GW and YT contributed towards the interpretation of the data.

Acknowledgements

We acknowledge the Cancer Genome Atlas Research Network for the clinicopathological and genetic alteration data.

Authors' information

1 Department of Urology, Beijing Friendship Hospital, Capital Medical University, Beijing, 100050, China.

2. Department of Stomatology, Beijing Shijitan Hospital, Capital Medical University, Beijing, 100038, China

3. Department of Biological Repositories, Zhongnan Hospital of Wuhan University, Wuhan, 430071, China

4. Department of Urology, Zhongnan Hospital of Wuhan University, Wuhan, 430071, China

References

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2015;136(5):E359-86.
2. Jacobs BL, Lee CT, Montie JE. Bladder cancer in 2010: how far have we come? CA Cancer J Clin. 2010;60(4):244-72.
3. Morgan TM, Clark PE. Bladder cancer. Curr Opin Oncol. 2010;22(3):242-9.
4. Shariat SF, Karakiewicz PI, Palapattu GS, Lotan Y, Rogers CG, Amiel GE, et al. Outcomes of radical cystectomy for transitional cell carcinoma of the bladder: a contemporary series from the Bladder Cancer Research Consortium. J Urol. 2006;176(6 Pt 1):2414-22; discussion 22.
5. Wuttke M, Papewalis C, Jacobs B, Schott M. Identifying tumor antigens in endocrine malignancies. Trends Endocrinol Metab. 2009;20(3):122-9.
6. Huang X, Zhang G, Tang T, Liang T. Identification of tumor antigens and immune subtypes of pancreatic adenocarcinoma for mRNA vaccine development. Mol Cancer. 2021;20(1):44.
7. Bouzid R, Peppelenbosch M, Buschow SI. Opportunities for Conventional and in Situ Cancer Vaccine Strategies and Combination with Immunotherapy for Gastrointestinal Cancers, A Review. Cancers (Basel). 2020;12(5).
8. Sullenger BA, Nair S. From the RNA world to the clinic. Science. 2016;352(6292):1417-20.
9. Pardi N, Hogan MJ, Porter FW, Weissman D. mRNA vaccines - a new era in vaccinology. Nat Rev Drug Discov. 2018;17(4):261-79.
10. Wang Y, Zhang L, Xu Z, Miao L, Huang L. mRNA Vaccine with Antigen-Specific Checkpoint Blockade Induces an Enhanced Immune Response against Established Melanoma. Mol Ther. 2018;26(2):420-34.
11. Xu H, Zheng X, Zhang S, Yi X, Zhang T, Wei Q, et al. Tumor antigens and immune subtypes guided mRNA vaccine development for kidney renal clear cell carcinoma. Mol Cancer. 2021;20(1):159.
12. Pardi N, Hogan MJ, Weissman D. Recent advances in mRNA vaccine technology. Curr Opin Immunol. 2020;65:14-20.
13. Gu YZ, Zhao X, Song XR. Ex vivo pulsed dendritic cell vaccination against cancer. Acta Pharmacol Sin. 2020;41(7):959-69.
14. Emens LA. Roadmap to a better therapeutic tumor vaccine. Int Rev Immunol. 2006;25(5-6):415-43.
15. Lamm DL. Efficacy and safety of bacille Calmette-Guerin immunotherapy in superficial bladder cancer. Clin Infect Dis. 2000;31 Suppl 3:S86-90.
16. Zhu S, Tang Y, Li K, Shang Z, Jiang N, Nian X, et al. Optimal schedule of bacillus calmette-guerin for non-muscle-invasive bladder cancer: a meta-analysis of comparative studies. BMC Cancer. 2013;13:332.
17. Larsen ES, Joensen UN, Poulsen AM, Goletti D, Johansen IS. Bacillus Calmette-Guerin immunotherapy for bladder cancer: a review of immunological aspects, clinical effects and BCG infections. APMIS. 2020;128(2):92-103.

18. Conroy M, Naidoo J. Immune-related adverse events and the balancing act of immunotherapy. Nat Commun. 2022;13(1):392.
19. Brahmer JR, Lacchetti C, Schneider BJ, Atkins MB, Brassil KJ, Caterino JM, et al. Management of Immune-Related Adverse Events in Patients Treated With Immune Checkpoint Inhibitor Therapy: American Society of Clinical Oncology Clinical Practice Guideline. J Clin Oncol. 2018;36(17):1714-68.
20. Maiorano BA, De Giorgi U, Ciardiello D, Schinzari G, Cisternino A, Tortora G, et al. Immune-Checkpoint Inhibitors in Advanced Bladder Cancer: Seize the Day. Biomedicines. 2022;10(2).
21. Xu L, Deng C, Pang B, Zhang X, Liu W, Liao G, et al. TIP: A Web Server for Resolving Tumor Immunophenotype Profiling. Cancer Res. 2018;78(23):6575-80.
22. Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, et al. Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. Cell. 2018;173(2):338-54 e15.
23. Knijnenburg TA, Wang L, Zimmermann MT, Chambwe N, Gao GF, Cherniack AD, et al. Genomic and Molecular Landscape of DNA Damage Repair Deficiency across The Cancer Genome Atlas. Cell Rep. 2018;23(1):239-54 e6.
24. Sha D, Jin Z, Budczies J, Kluck K, Stenzinger A, Sinicrope FA. Tumor Mutational Burden as a Predictive Biomarker in Solid Tumors. Cancer Discov. 2020;10(12):1808-25.
25. Gonzalez FE, Gleisner A, Falcon-Beas F, Osorio F, Lopez MN, Salazar-Onfray F. Tumor cell lysates as immunogenic sources for cancer vaccine design. Hum Vaccin Immunother. 2014;10(11):3261-9.
26. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, et al. The Immune Landscape of Cancer. Immunity. 2018;48(4):812-30 e14.
27. Cancer Genome Atlas Research N. Comprehensive molecular characterization of urothelial bladder carcinoma. Nature. 2014;507(7492):315-22.
28. Girard-Pierce KR, Stowell SR, Smith NH, Arthur CM, Sullivan HC, Hendrickson JE, et al. A novel role for C3 in antibody-induced red blood cell clearance and antigen modulation. Blood. 2013;122(10):1793-801.
29. Zhou Q, Yan X, Zhu H, Xin Z, Zhao J, Shen W, et al. Identification of three tumor antigens and immune subtypes for mRNA vaccine development in diffuse glioma. Theranostics. 2021;11(20):9775-90.
30. Siefker-Radtke AO, Apolo AB, Bivalacqua TJ, Spiess PE, Black PC. Immunotherapy with Checkpoint Blockade in the Treatment of Urothelial Carcinoma. J Urol. 2018;199(5):1129-42.
31. Petrulio CA, Kaufman HL. Development of the PANVAC-VF vaccine for pancreatic cancer. Expert Rev Vaccines. 2006;5(1):9-19.
32. Cisarovsky C, Latifyan S, Perrinjaquet C, Berthold D, Orcurto A. [Immunotherapy for bladder cancer in 2021]. Rev Med Suisse. 2021;17(739):974-7.
33. Shen C, Liu J, Wang J, Yang X, Niu H, Wang Y. The Analysis of PTPN6 for Bladder Cancer: An Exploratory Study Based on TCGA. Dis Markers. 2020;2020:4312629.

34. Jin K, Qiu S, Jin D, Zhou X, Zheng X, Li J, et al. Development of prognostic signature based on immune-related genes in muscle-invasive bladder cancer: bioinformatics analysis of TCGA database. Aging (Albany NY). 2021;13(2):1859-71.
35. Chen F, Wang Q, Zhou Y. The construction and validation of an RNA binding protein-related prognostic model for bladder cancer. BMC Cancer. 2021;21(1):244.
36. Qiu H, Hu X, He C, Yu B, Li Y, Li J. Identification and Validation of an Individualized Prognostic Signature of Bladder Cancer Based on Seven Immune Related Genes. Front Genet. 2020;11:12.
37. Li H, Zhang Y, Zheng S. Comprehensive Analysis Identified ETV7 as a Potential Prognostic Biomarker in Bladder Cancer. Biomed Res Int. 2021;2021:8530186.
38. Wang Y, Yan K, Lin J, Liu Y, Wang J, Li X, et al. CD8+ T Cell Co-Expressed Genes Correlate With Clinical Phenotype and Microenvironments of Urothelial Cancer. Front Oncol. 2020;10:553399.
39. Alessandrini F, Pezze L, Menendez D, Resnick MA, Ciribilli Y. ETV7-Mediated DNAJC15 Repression Leads to Doxorubicin Resistance in Breast Cancer Cells. Neoplasia. 2018;20(8):857-70.
40. Maeda O, Ando T, Ohmiya N, Ishiguro K, Watanabe O, Miyahara R, et al. Alteration of gene expression and DNA methylation in drug-resistant gastric cancer. Oncol Rep. 2014;31(4):1883-90.
41. Wu Y, Xia L, Zhao P, Deng Y, Guo Q, Zhu J, et al. Immune profiling reveals prognostic genes in high-grade serous ovarian cancer. Aging (Albany NY). 2020;12(12):11398-415.
42. Yamamoto K, Miki Y, Sato M, Taketomi Y, Nishito Y, Taya C, et al. The role of group IIF-secreted phospholipase A2 in epidermal homeostasis and hyperplasia. J Exp Med. 2015;212(11):1901-19.
43. Mounier CM, Wendum D, Greenspan E, Flejou JF, Rosenberg DW, Lambeau G. Distinct expression pattern of the full set of secreted phospholipases A2 in human colorectal adenocarcinomas: sPLA2-III as a biomarker candidate. Br J Cancer. 2008;98(3):587-95.
44. Tauro BJ, Greening DW, Mathias RA, Ji H, Mathivanan S, Scott AM, et al. Comparison of ultracentrifugation, density gradient separation, and immunoaffinity capture methods for isolating human colon cancer cell line LIM1863-derived exosomes. Methods. 2012;56(2):293-304.
45. Sadler JBA, Wenzel DM, Williams LK, Guindo-Martinez M, Alam SL, Mercader JM, et al. A cancer-associated polymorphism in ESCRT-III disrupts the abscission checkpoint and promotes genome instability. Proc Natl Acad Sci U S A. 2018;115(38):E8900-E8.
46. Liu B, Guo S, Li GH, Liu Y, Liu XZ, Yue JB, et al. CHMP4C regulates lung squamous carcinogenesis and progression through cell cycle pathway. J Thorac Dis. 2021;13(8):4762-74.
47. Lin SL, Wang M, Cao QQ, Li Q. Chromatin modified protein 4C (CHMP4C) facilitates the malignant development of cervical cancer cells. FEBS Open Bio. 2020;10(7):1295-303.
48. Gubin MM, Artyomov MN, Mardis ER, Schreiber RD. Tumor neoantigens: building a framework for personalized cancer immunotherapy. J Clin Invest. 2015;125(9):3413-21.
49. Chan TA, Yarchoan M, Jaffee E, Swanton C, Quezada SA, Stenzinger A, et al. Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. Ann Oncol. 2019;30(1):44-56.

50. Huang X, Tang T, Zhang G, Liang T. Identification of tumor antigens and immune subtypes of cholangiocarcinoma for mRNA vaccine development. Mol Cancer. 2021;20(1):50.

Figures

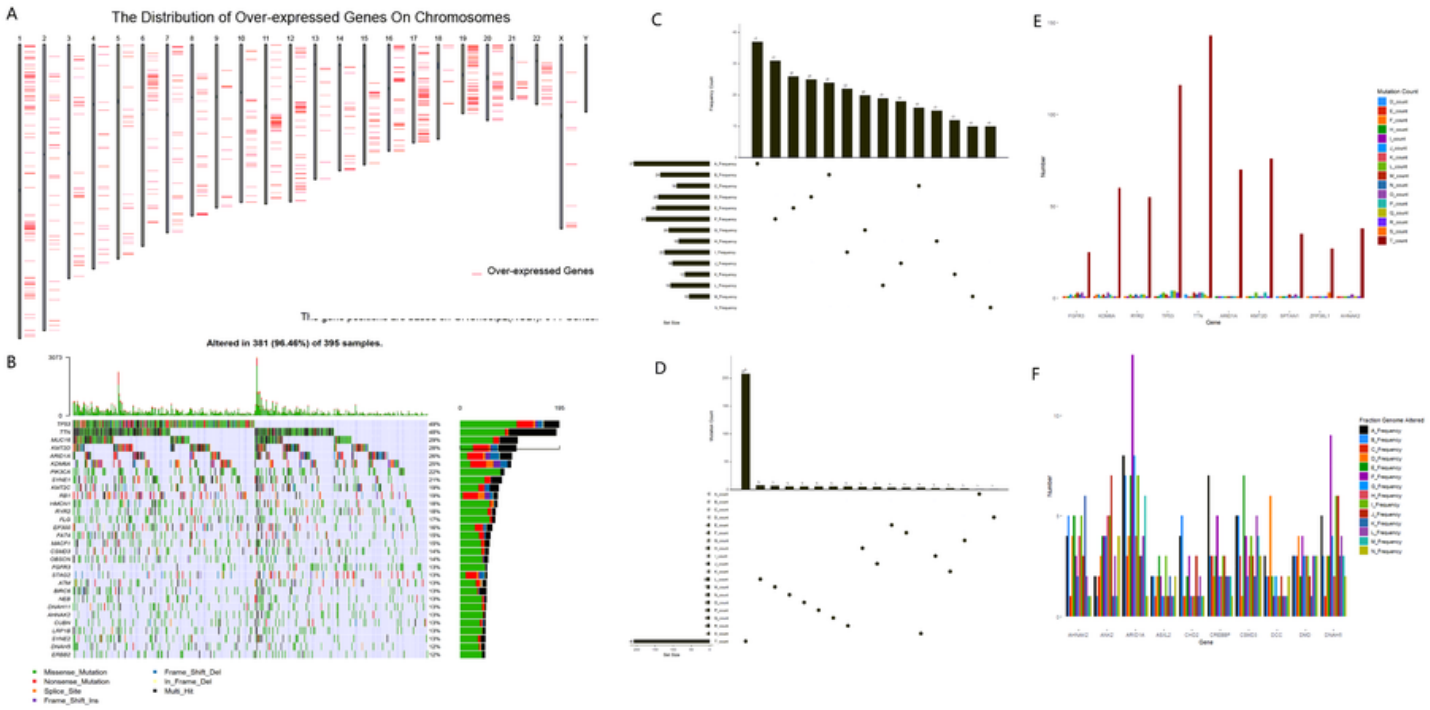

Figure 1

Identification of potential antigens of BLCA. A) Chromosomal distribution of up- and down-regulated genes in BLCA; B) Waterfall diagram of the top 30 mutant genes; C) Distribution of mutation frequency; D) Distribution of mutation number; E) Distribution of mutation number of the top 10 genes; F) Distribution of mutation frequency of the top 10 genes.

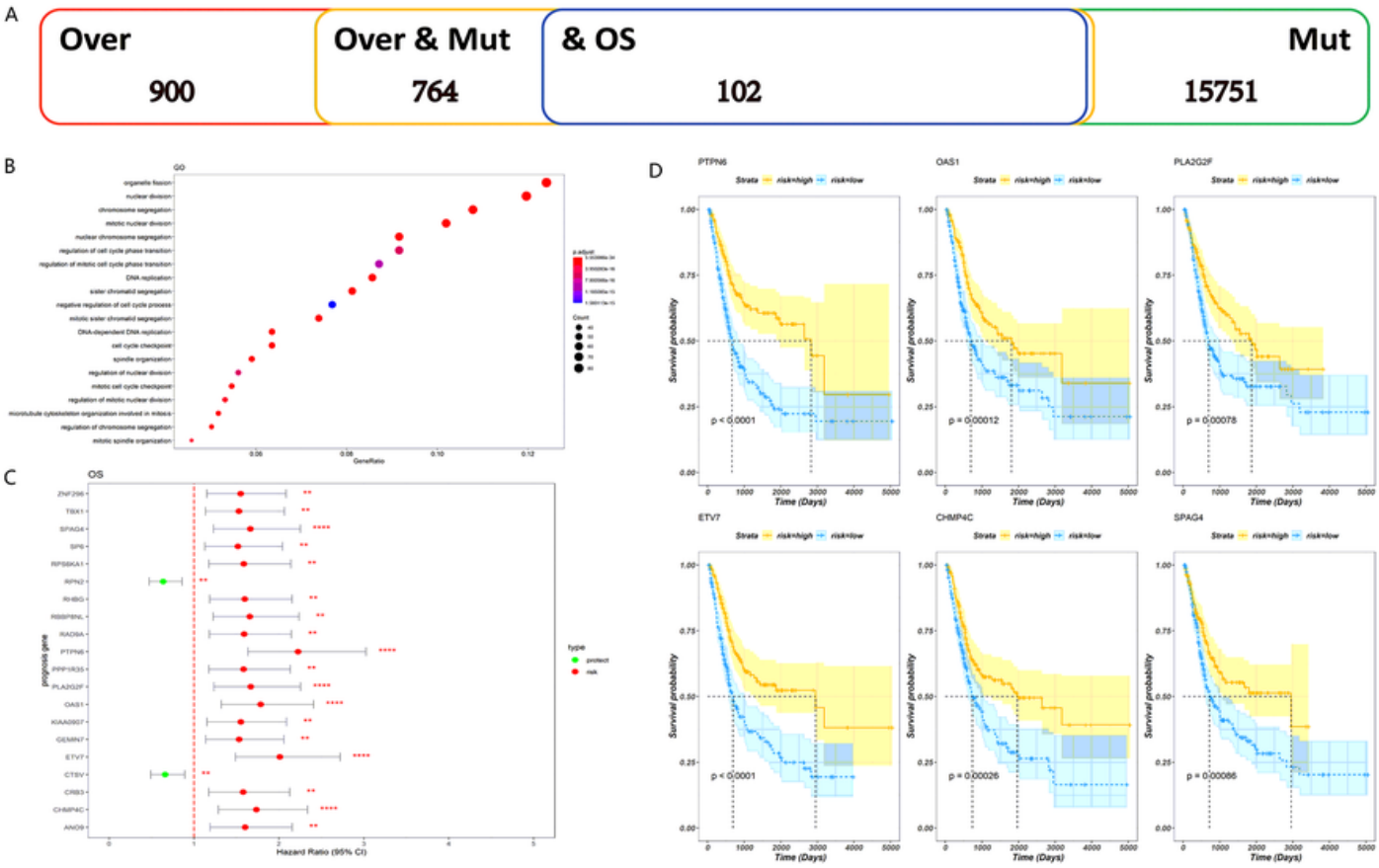

Figure 2

Identification of tumor antigens associated with BLCA prognosis and antigen presenting cells. A) Overlapped genes identified through intersection; B) GO enrichment analysis of 764 genes after intersection of overexpressed and mutated genes; C) Univariate Cox regression analysis of the top 20 potential antigens for OS; D) Kaplan-Meier curves of the association of PTPN6, OAS1, PLA2G2F, ETV7, CHMP4C and SPA4G with OS ;

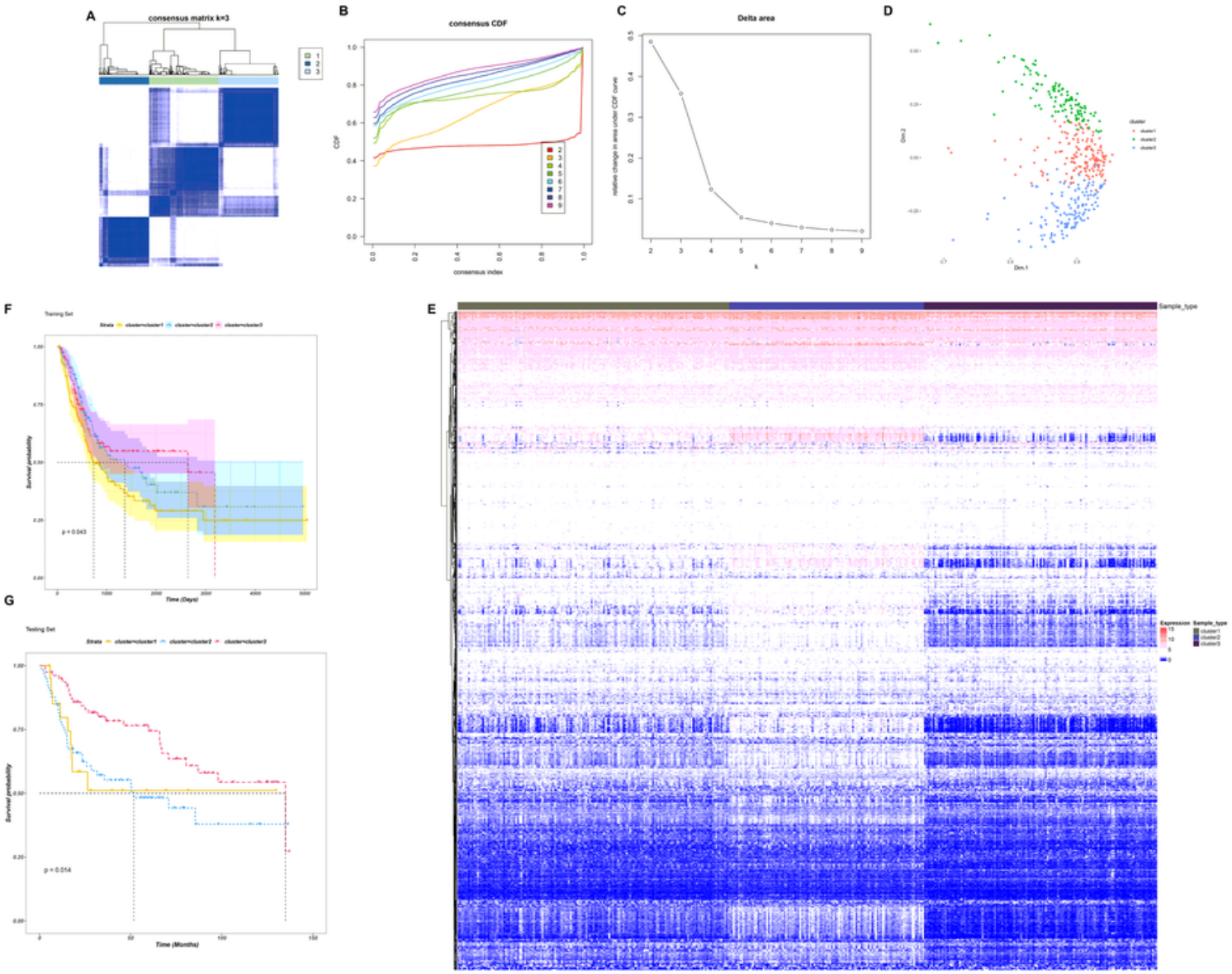

Figure 3

Immune subtypes identification. A) Sample clustering heatmap; B) Cumulative distribution function curve; C) delta area of immune-related genes in TCGA cohort ; D) principal component analysis of three immune subtypes; E) Immune related gene heatmap of immune subtypes; F) Kaplan-Meier curves of the association of immune subtypes with overall survival of TCGA cohort; G Kaplan-Meier curves of the association of immune subtypes with overall survival of GSE13507 cohort

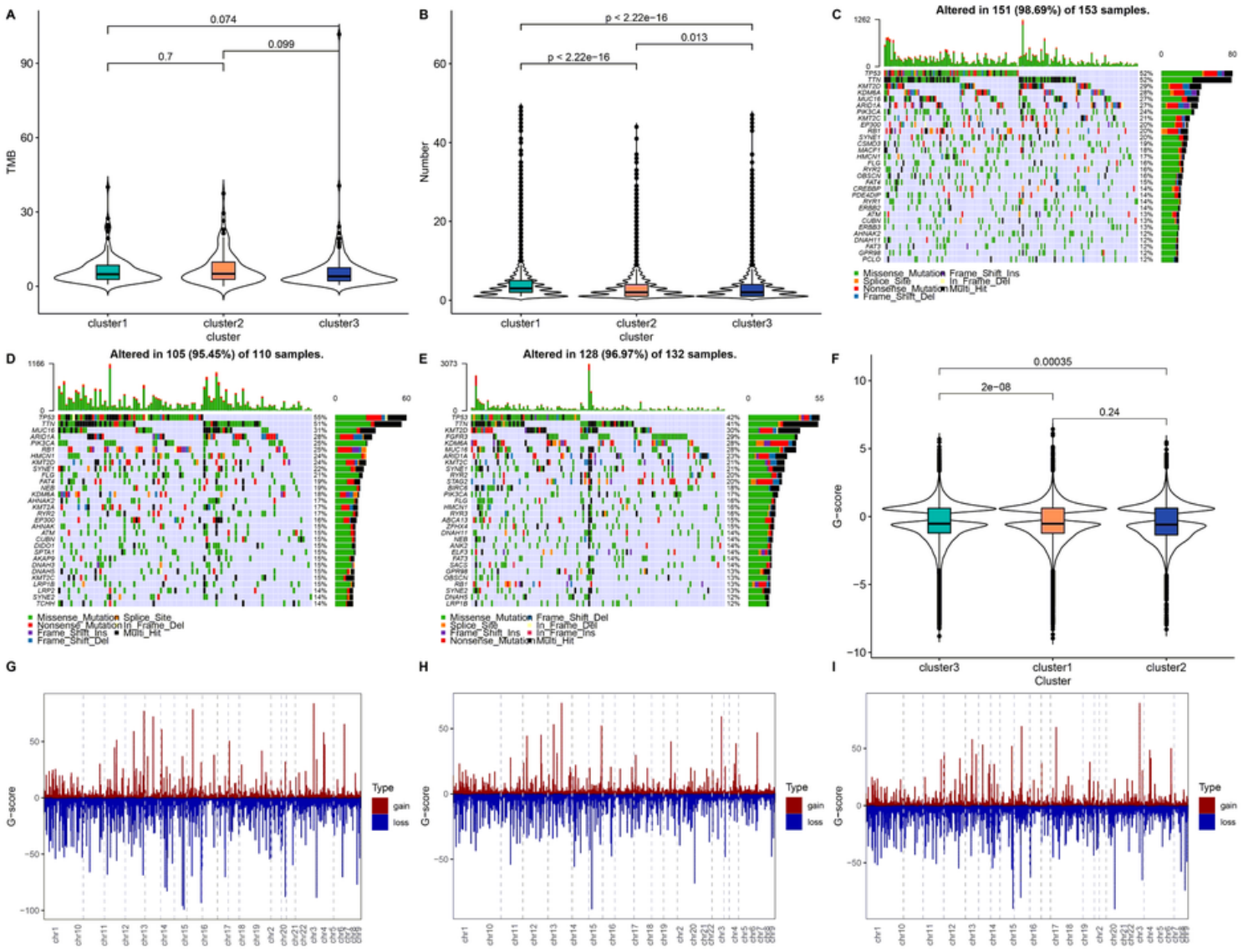

Figure 4

The association of immune subtypes with mutational status. A) Distribution of TMB scores in subtypes; B) The number of genes in subtypes respectively; C) Waterfall diagram in cluster1; D) Waterfall diagram in cluster2; E) Waterfall diagram in cluster3; F) G-score distribution in subtypes G) Distribution of copy numbers in cluster1 on chromosomes; H) distribution of copy numbers in cluster2 on chromosomes; I) distribution of copy numbers in cluster3 on chromosomes

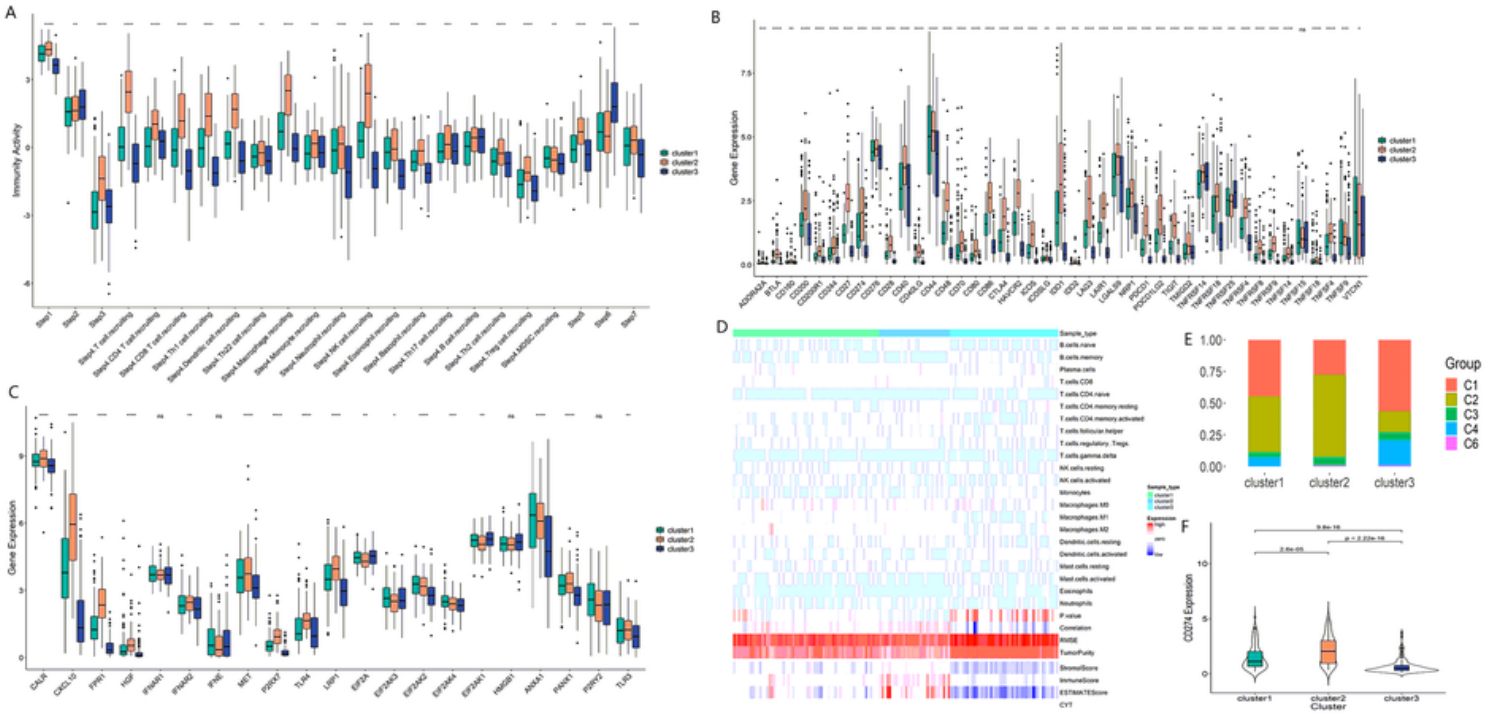

Figure 5

The association of immune subtypes with mutational status. A) Distribution of TMB scores in subtypes; B) The number of genes in subtypes respectively; C) Waterfall diagram in cluster1; D) Waterfall diagram in cluster2; E) Waterfall diagram in cluster3; F) G-score distribution in subtypes G) Distribution of copy numbers in cluster1 on chromosomes; H) distribution of copy numbers in cluster2 on chromosomes; I) distribution of copy numbers in cluster3 on chromosomes

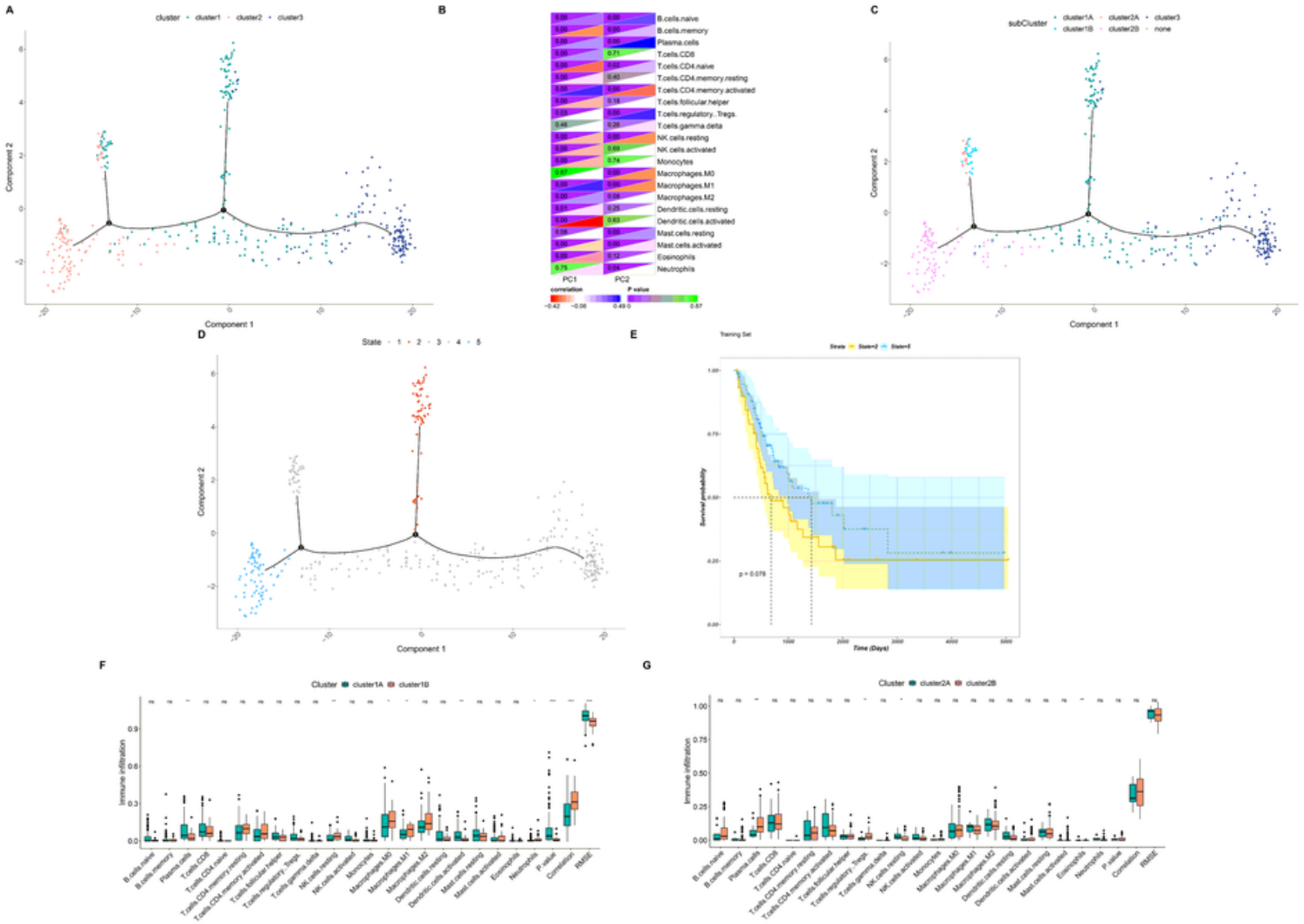

Figure 6

Immune landscape of BLCA. A) Immune landscape in BLCA, Each dot represents one patient, and the immune subtype is color coded. The horizontal axis represents the first principal component, and the vertical axis represents the second principal component; B) The heatmap of the correlation of immune cells between the two principal components; C) The immune landscape of subgroup of BLCA immune subtype; D) Immune profiles of samples from 2 extreme locations; E) Prognosis of 2 extreme locations; F) and G) differentially enriched fractions of immune cells in the aforementioned subsets

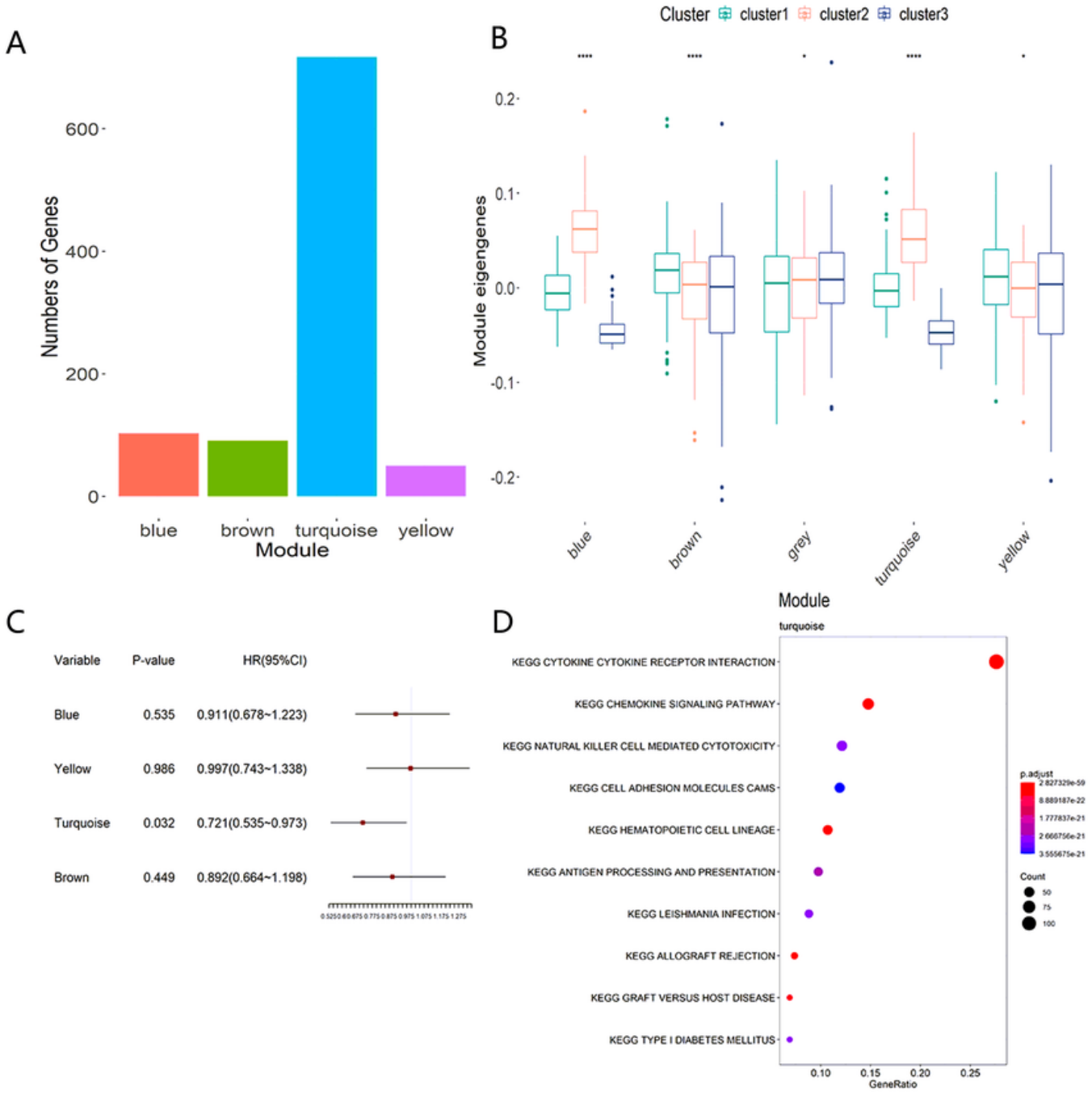

Figure 7

WGCNA of BLCA; A) Gene numbers of each module in WGCNA module; B) Differential distribution of feature vectors of each module in BLCA subtypes; C) Forest maps of single factor survival analysis of 4 modules of BLCA; D) Dot plot showing top 10 KEGG terms in the turquoise module.

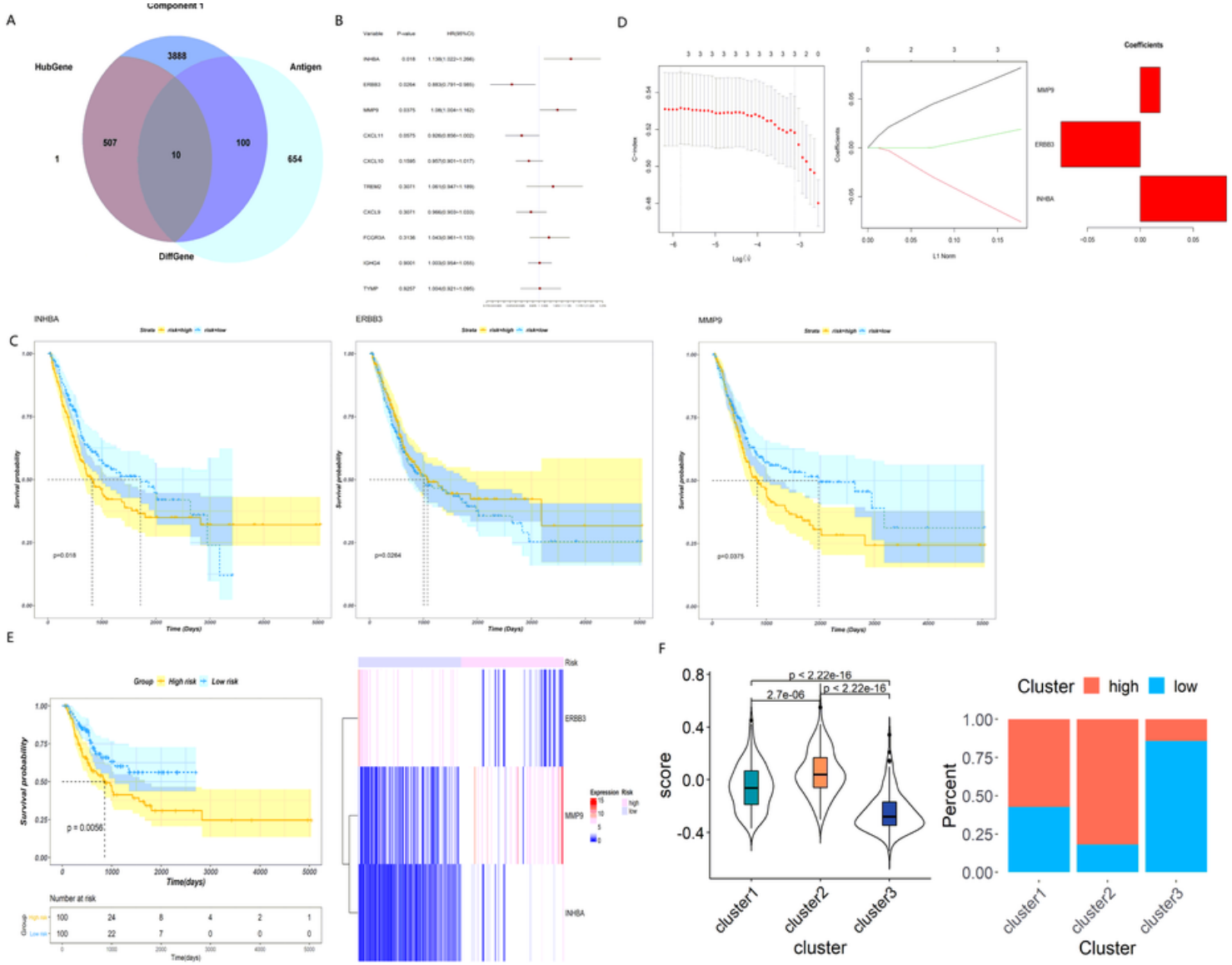

Figure 8

Identification of potential antigen biomarkers for mRNA vaccine development; A) Key tumor antigen markers obtained from the intersection of differential genes between subtypes (4505) and potential antigen candidate genes (764) and module hub genes (518). B-C) Univariate Cox regression analysis of tumor antigen marker; D) Risk model was constructed based on key tumor antigen marker; E) Survival score related survival curve and gene expression heat map; F) Sample distribution of high and low score among subtypes.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarymaterial.rar