Identification of subgroups of bladder cancer based on immune gene signature

Chen Yuan
Liuyang People Hospital
Xiao Ren
Liuyang People Hospital
Zhaoxia Nie
Liuyang People Hospital
Long Li
Liuyang People Hospital
Dunmang Peng
Liuyang People Hospital
Zhiguo Zhang (zhangzhiguo666@163.com)
Liuyang People Hospital

Research Article

Keywords: Immune related genes (IRGs), Bladder cancer(BC), Immune gene signature, Molecular subtypes

Posted Date: March 11th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1427681/v1

Abstract

Immunotherapy has become a new frontier in bladder cancer (BC) treatment. In this study, we utilized bioinformatic tools to identify an immune signature for BC. RNA-seq data of BC samples was downloaded from The Cancer Genome Atlas (TCGA) database, and GSE31684, GSE32894 and GSE77952 chip expression data were downloaded from the Gene Expression Omnibus (GEO) database. The immune related genes (IRGs) dataset was extracted from the ImmPort database. Unsupervised clustering method was used to determine genes with significant association with the prognosis of BC. Macrophages were enriched in C1/C2 subtypes whereas infiltration of B cells ware prominent in C3 subtype. The frequency of FGFR3 mutation was high in C3 subtypes whereas RB1 mutations were high in C1/C2 subtypes. The immune modules gSig1and gSig2 were associated with poor prognosis. gSig3 was associated with good prognosis. gSig1 was associated with activation and degranulation of neutrophils. On the other hand, gSig2 was associated with steroid hormone mediated signaling pathway whereas gSig3 was associated with secretion of extracellular matrix and muscle cell proliferation. gSig1 and gSig2 were positively correlated with expression of immune checkpoint genes (ICGs) in two different cohorts, whereas gSig3 scores was negatively correlated with expression of ICGs. Expression profile of immune-related genes could subclassify BC into three molecular subtypes with distinct histological characteristics, genetic and transcriptional changes. Our study have provided a novel insight into the immune-related state of BC and shed light on the prognostication of BC patient, all of them are of potential clinical implications.

1. Introduction

Bladder cancer is one of the most prevalent and aggressive cancers. Global statistics show that there are more than 430,000 cases of BC, with 165,000 deaths recorded yearly. Indeed the incidence of BC ranks 9th worldwide and 13th in overall tumor mortalities [1]. BC is categorized into two subtypes; non-muscle-invasive BC (stage Ta or T1; NMIBC) and muscle-invasive BC (stage T2-4; MIBC) [2]. Urothelial carcinoma is the predominant histological type of BC. According to pathological classification, muscle invasive bladder cancer (MIBC) accounts for 25% of all cases of bladder urothelial carcinomas. MIBC is a heterogeneous disease with multiple clinical outcomes [3]. At stage T2-T4, MIBC is characterized by high rates of recurrence and metastasis, high degree of malignancy, poor prognosis and high mortality [4]. Immunotherapy is a rapidly growing field, representing a paradigm shift in the treatment of MIBC compared to conventional chemotherapy and radiation treatment.

Tumor cells, various immune cells such as T\B lymphocytes, natural killer (NK) cells, macrophages dendritic cells and their secreted immune molecules constitute the immune microenvironment of tumors [5]. Immune escape is one of the top ten characteristics of tumors, a mechanism that aids in tumorigenesis. Therefore, targeting immune escape is emerging as a novel immunotherapeutic approach [6]. Mutations form in tumor cells that escape immune killing, and this reduces immunogenicity. Signaling pathways for immune checkpoints such as PD-1-PD-L1 and CTLA-4-CD80/CD86 are activated to induce formation of immunosuppressive tumor microenvironment [7]. Mechanisms that inhibit the

functions of immune cells such as CD4 + CD25 + FOXP3 + regulatory T cells (Treg) aid tumor cells to evade the immune microenvironment thus promoting tumorigenesis [7].

Numerous studies show that immune microenvironment affects the prognosis of tumors. Immune tumor responses are regulated by tumor-infiltration lymphocytes (TILs), which mainly include T cells and B cells and NK cells. Recent studies show that infiltration of TILs in breast, colorectal and ovarian cancer stroma is associated with better prognosis [8–10]. Compared to TILs, the number of neoantigens and tumor mutation burden influences immunotherapy response [11–13]. Although high expression of PD-1 \ PD-L1 is generally related to clinicopathological factors, its association with prognosis varies among studies [14]. Although many immunotherapeutic drugs have been shown to improve the prognosis of bladder cancer patients, these drugs are not always effective in all patients. Tumor heterogeneity triggers different therapeutic responses among patients. Therefore, stratification of treatments for patients with bladder cancer is necessary. For this, more molecular subtypes and immune markers are needed to identify patients who are likely to benefit from immunotherapy. Thus, in this study, gene expression profiles of 414 BC samples, including normal tissues adjacent to the tumor were obtained from the TCGA database. Non-negative Matrix Factorization (NMF) performed on gene expression profiles identified three immune-related molecular subtypes with distinct characteristics. There were significant differences in prognosis and molecular characteristics among the three immune subtypes. In addition, quantified immune molecules were further established. The gSig score showed the potential to be a powerful prognostic biomarker and predictor of immune checkpoint inhibitor responses.

2. Materials And Methods

2.1 Data collection and processing

The 433 RNA-seq samples data of BC were downloaded from the TCGA database. GSE31684, GSE32894 and GSE77952 chip expression data were downloaded from the GEO database which contained information on 93, 308 and 30 samples. In addition, gene expression profile data of bladder cancer in UROMOL was downloaded from EBI (https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-4321). For corresponding probe data, bioconductor package (hgu133plus2.db) was used to map microchip probe into human gene SYMBOL. IMvigor210 (n = 348 MIBC) was downloaded from GEO, ArrayExpress, and the Supplementary data in the study by Mariathasan et al. [15]. The immune related genes (IRGs) dataset was obtained from the ImmPort database (https://immport.niaid.nih.gov).

2.2 Identification of immune molecular subtypes

To investigate the relationship between expression profile of prognosis-related immune gene and phenotypes, we first performed univariate survival analysis to identify immune genes significantly associated with prognosis of BC. Nonnegative matrix factorization (NMF) was used in the discovery of genomics-based tumor molecular subtypes[16, 17]. The average contour width of the common member matrix was calculated using the R package NMF [18].

2.3 Characteristic gene score of IRGCluster

To profile genes of the three immune molecular subtypes, genes with significant association with prognosis were first identified and classified using the unsupervised clustering method. The expression patterns of each group were then analyzed. Genes expressed specifically in a subtype were identified as a subtype specific set of immune molecules. The importance of each immune gene was evaluated by dimension reduction analysis using random forest algorithm.

2.4 Relationship between IRGCluster and clinical features

To determine the relationship between IRGCluster and clinical phenotype, information on TNM, Stage and Age from the TCGA dataset for the samples was used to compare the distribution of clinical features in the IRGCluster. Specifically, we analyzed the distribution of Age, TNM and Stage in different IRGCluster. Chi-square test was used to assess the distribution of different clinical features in IRGCluster. $P < 0.05$ was selected as the threshold.

2.5 Relationship between IRGCluster and immune microenvironment

The tumor immune estimation resource (TIMER) was used to evaluate the relationship between IRGCluster and the immune microenvironment [19] to calculate scores for the six immune cells in the TCGA for BC samples. The difference in scores among the six immune cells in different IRGCluster was also analyzed. Nineteen immune checkpoint genes were obtained from TCGA database and the spearman correlation between each immune checkpoint gene and IRGCluster characteristic gene was calculated.

2.6 Relationship between IRGCluster and tumor genome variation

The copy number variation (CNV) data on BC was downloaded from TCGA database. CNV intervals were consolidated as follows: 1) Intervals with 50% regional overlap were considered the same. 2) The number of coverage probes with less than 5 intervals was removed. 3) The CNV range was mapped on to the corresponding gene using the GRh38 version of gencode.v22. 4) Multiple CNV regions in one gene region were consolidated, and the combined CNV values averaged.

The SSNV data was obtained from the mutect2 version in TCGA, which contain sequences for whole exomes. Data of neoantigens was derived from the references RESOURCE VOLUME 48, ISSUE 4, P812-830.E14, APRIL 17, 2018, The Immune Landscape for Cancer. Mutations in the silent and intron intervals were excluded from Tumor Mutation Burden (TMB) analyses, with a genome interval of 38.4 Mb, and the TMB for each sample was divided by 38.4 Mb for all mutations.

2.7 Gene ontology and pathway enrichment analysis

Functional enrichment was assessed using Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis. Functional enrichment analyses were performed using the R package clusterprofiler [20].

2.8 Statistical Analysis

Cox risk regression model and log grade were used for univariate survival analysis. The Kaplan-Meier method was used generate survival curves for the subgroups in each data set and the logrank test was utilized to compare groups, whereas the chi-square test was employed to test the significance of sample overlap between histological type and IRGCluster, as well as the distribution of clinical features across groups in IRGCluster. The Wilcox rank test was used to analyze continuous variables between two groups; whereas the Kruskal-Wallis rank test was used to compare more than two. Benjamini-Hochberg method was utilized for multiple test correction. R software v 3.5.1 was used for statistical analysis. $P < 0.05$ was considered statistically significant.

3. Results

3.1 Identification of BC molecular typing

Details of sample statistics are shown in Table 1. The expression profiles of IRGs was extracted from the TCGA dataset. In total, 646 IRGs were characterized after removing genes with no or low expression. Cox univariate regression analysis identified 81 genes that were significantly associated with prognosis. Three optimal molecular subtypes for IRGCluster were obtained using the NMF algorithm, with C1 (112), C2 (118) and C3 (165) (Fig. 1A). Unsupervised clustering was employed to cluster expression profiles of 81 genes. Three gene sets: gSig1 (23), gSig2 (37) and gSig3 (21) were obtained. gSig1 was highly expressed in C1 subtype but its expression was lowest in C3 subtype. gSig2 was highly expressed in C1/C2, but generally lowly expressed in C3. Expression of gSig3 was highest in C3 but lower in C1/C2 (Fig. 1A).

Table 1
Pre-processed dataset clinical information.

Characteristic		TCGA	GSE31684	GSE32894	GSE77952	UROMOL
Survival status	Alive	221		196		
	Dead	174		25		
Mean OS	Alive	1026		1291.72		
	Dead	564.4		596.3		
pT	T1	10	10	61	9	108
	T2/T3/T4	379	78	51	14	7
	Ta/CIS		5	109	7	343
	Un	6				
pN	N0	228				
	N1	44				
	N2	75				
	N3	7				
	Un	41				
pM	M0	189				
	M1	10				
	Un	196				
Stage	I	2				
	II	124				
	III	137				
	IV	130				
	Un	2				
Grade	High	374	87	44		179
	Low	18	6	175		272
	Un/PUNLMP	3		2		7
NewEventTissue	Bladder	11				
	Bone	22				
	Liver	12				

Characteristic		TCGA	GSE31684	GSE32894	GSE77952	UROMOL
	Lung	30				
	Lymph Node Only	22				
	Other,specify	37				
	Renal Pelvis	4				
	Un	257				
NewEventType	Distant metastatic	78				
	Locoregional	46				
	Metastatic	10				
	New primary tumor	7				
	Un	254				
Age	≤ 60	106				
	> 60	289				
Smoking	Lifelong Non-smoker	105				
	Current smoker	87				
	Current reformed smoker for > 15 years	108				
	Current reformed smoker for ≤ 15 years	71				
	Current reformed smoker, duration not specified	11				
	Un	13				

Overall, 81 genes were significantly differentially expressed across the three subtypes. In addition, FGFR3 mutations were significantly higher in C3 than in C1/C2 (Fig. 1A), while RB1 mutations were significantly higher in C1/C2 (Fig. 1A). Overall survival (OS) varied significantly among the subtypes (Fig. 1B). The expression of C3 correlated with the best prognosis whereas C1/C2 subtypes were associated with the worst prognosis. In the external independent verification set GSE32894, C1/C2/C3 had equally significant prognosis difference in OS, with C3 subtype having significantly better prognosis compared with C1/C2

(Fig. 1C). In bladder cancer, high-grade often associates with worse prognosis compared with low grade types.

When the high/low grade BC was compared in the IRGCluster, results showed that almost all of the low grade (17/18) were distributed in the best prognostic subtype C3, whereas high grade cancers contained the three subtypes, C1/C2/C3 (Fig. 1D). The expression of the three subtypes was compared with mRNA cluster typing reported by A Gordon Robertson et al. [21]. Our findings showed that 70.29% of samples in Cluster1 were Basal-squamous. Cluster 3 contained 92.59% of Luminal-papillary and 80.77% of Luminal type samples, whereas Cluster 2 contained 80.52% of Luminal-infiltrated and 68.42% of Neuronal subtype samples (Fig. 1E). Marker genes from a previous study [22] were also evaluated. Results showed that the expression of Basal marker gene was significantly high in C1 subgroup than the other two groups, whereas the expression of Immune/Infiltrate marker gene was significantly higher in C2 subgroup than in the other two groups. This may be one of the factors contributing to poor prognosis in C2 group. The expression of Luminal marker gene in C3 group was significantly higher than that in the other two groups. (Fig. 1F). These data indicated that high grade BCs are similar pathologically, but exhibit different immune characteristics. These differences are associated with different prognostic outcomes.

3.2 IRGCluster characteristic gene score and functional analysis

The scores for gSig1, gSig2 and gSig3 were calculated based on the average expression level of each set. The gSig scores were significantly different in the IRGCluster (Fig. 2A). These results were consistent with pathological features of the subtypes generated from the GSE32894 dataset (Fig. 2B). This further validated the uniformity between IRGCluster and gSig. Analysis of OS of gSig showed that gSig score was significantly associated with OS, with gSig1 and gSig2 being associated with poor prognosis, whereas gSig3 was associated with better prognosis (Fig. 2C). GO enrichment analysis revealed that gSig1 was mainly associated with degranulation of neutrophils and neutrophil activation (Fig. 2D). gSig2 was associated with steroid hormone mediated signaling pathway (Fig. 2E), whereas gSig3 was associated with secretion of extracellular matrix and muscle cell proliferation (Fig. 2F). Interestingly, no significant enrichment insight was found in the KEGG pathway. These results suggested that the three gene sets regulate different biological pathways, which may be one of the factors contributing to different clinical outcomes in patients with BC.

3.3 Clinical characteristics of IRGCluster and differential expression of characteristic immune gene sets

We first analyzed the distribution of T, N, M, Stage, Smoking and Age among the three molecular subtypes. Samples with advanced stage of tumors such as T, N, were more likely to fall in C2 subtype, and associate with worst prognosis (Table 2). However, no difference was observed in the distribution of characteristics such as smoking and metastasis between the three molecular subtypes. The worst prognosis category, C2, mainly comprised of old individuals between 70 to 100 years (Table 2). Based on

the high/low grade classification provided in the TCGA and GSE32894 and UROMOL data sets, we compared the gSig scores on different levels of samples and found that the gSig scores showed very high agreement on the three sets of data sets (Fig. 3A-C). Both gSig1 and gSig2 scores, which were associated with poor OS, were significantly higher in high grade than in low grade tumors. However, gSig3 score, which was associated with better OS, was significantly higher in low grade than in high grade tumors (Fig. 3A-C). The gSig score was consistently distributed between the muscle invasive and non-muscle invasive (Fig. 3D-F**).** These results imply that the gSig score could be a marker for BC.

Table 2
Comparison of clinical features with IRGCluster.

Characteristic		C1	C2	C3	p-value
pT	T1	2	3	5	0.00116
	T2	51	38	97	
	T3	44	58	49	
	T4	14	17	11	
	Un	1	2	3	
pN	N0	65	60	103	0.00751
	N1	17	15	12	
	N2	16	36	23	
	N3	2	1	4	
	Un	12	6	23	
pM	M0	55	36	98	0.19091
	M1	1	4	5	
	Un	56	78	62	
Stage	I	0	0	2	< 1e-5
	II	33	17	74	
	III	43	49	45	
	IV	36	52	42	
	Un	0	0	2	
Grade	High	111	117	146	0.00002
	Low	1	0	17	
	Un	0	1	2	
New Event Tissue	Bladder	4	2	5	0.12681
	Bone	4	5	13	
	Liver	5	3	4	
	Lung	7	11	12	
	Lymph Node Only	4	11	7	
*Chisq-test					

Characteristic		C1	C2	C3	p-value
	Other, specify	16	15	6	
	Renal Pelvis	1	1	2	
	Un	71	70	116	
New Event Type	Distant Metastasis	21	27	30	0.46538
	Locoregional	16	13	17	
	Metastatic	2	6	2	
	New Primary tumor	3	3	1	
	Un	70	69	115	
Age	≤ 60	29	23	54	0.0448
	> 60	83	95	111	
Smoking	Lifelong Non-smoker	28	27	50	0.29977
	Current smoker	28	20	39	
	Current reformed smoker for > 15 years	34	37	37	
	Current reformed smoker for ≤ 15 years	15	25	31	
	Current reformed smoker, duration not specified	2	5	4	
	Un	5	4	4	
*Chisq-test					

3.4 Immune landscape features for IRGCluster

The scores for six immune cells in the BC sample from the TCGA dataset were calculated using TIMER (tumor immune estimation resource) tool. Although there was significant difference between scores of the six immune cells in IRGCluster, C3 subtype displayed a different pattern from C1/C2. Except for B cell, the expression of other types of immune cells was significantly higher in C1/C2 than in C3 (Fig. 4A). The scores of the 22 immune cells calculated based on cibersort algorithm [23] also demonstrate that the score for memory B cells in C3 was significantly higher than that of C1/C2 (Fig. 4B). This expression pattern reflects the specific response pattern of the immune system in the IRGCluster. Analysis of immune expression signatures showed that there was a significant difference between C3 and C1/C2. For C3 that was found to be associated with positive prognosis, the scores of leukocyte fraction, proliferation, macrophage regulation, lymphocyte infiltration, IFN gamma response and TGF beta response were all significantly lower than those of C1/C2 (Fig. 4C).

3.5 The relationship between gSigs and immune checkpoints

There were 19 immune checkpoint genes (ICGs) in the TCGA dataset. gSig1 and gSig2 were positively correlated with expression of 18 ICGs, whereas gSig3 was negatively correlated with expression of ICGs (Fig. 5A). Similar findings were observed in independent GSE32894 dataset (Fig. 5B). The correlation analysis revealed that there was a strong positive association of gSig1 and gSig2 with 14 of 15 ICGs and 15 of 16 ICGs, respectively, but there was a strong negative association of gSig3 with 14 of 18 ICGs (Fig. 5C). Similar results were observed after an independent analysis on GSE32894 dataset (Fig. 5D). The relationship between the expression of the immune checkpoint gene IDO1 and the gSig score was further analyzed. In both TCGA and GSE32894 datasets, IDO1 was at a lower expression level in the best prognostic C3, but it was highly expressed in C1/C2 with poor prognosis (Fig. 5E-F).

3.6 Genome heterogeneity of IRGCluster

For the IRGCluster of BC, we analyzed gene mutations among C1/C2/C3. A total of 45 genes with significant mutations were obtained. The frequency of FGFR3 mutation was significantly high in C3 subtypes, whereas RB1 mutations were prominent in C1/C2 subtypes. The prevalence of TP53 and PIK3CA gene mutations was similar among the three molecular subtypes (Fig. 6A). Neoantigens and tumor mutation burden (TMB) between IRGclusters were similar among the three molecular types (C1/C2/C3) (Fig. 6B). There was no correlation between gSig score and TMB (Fig. 6C-E). These findings suggest that hot spot mutations of some important high frequency mutated genes may be associated with the prognosis of different IRGclusters compared to the overall tumor mutation burden.

4. Discussion

Several molecular classification schemes for muscle-infiltrating urothelial carcinomas have been reported, which provides opportunities for patients to stratify subtypes for customized treatment [21]. Urothelial carcinoma can be divided into two main molecular subtypes; basal and luminal type, each with unique prognosis and response to chemotherapy [24]. Immunotranscriptome analysis on silicone revealed that different expression patterns of immune genes in four muscle-infiltrating urothelial carcinoma clusters, previously profiled from data in the TCGA network [25]. In 2017, Robertson *et al*. classified muscular infiltrating urothelial carcinoma cohort in TCGA into five molecular subtypes, with the new classification scheme incorporating neural mRNA subtypes [21]. In this study, three clinically relevant molecular subtypes were established based on immune genes, and BC progression after initial treatment was predicted. The three molecular subtypes showed different genomic characteristics and immune microenvironments. These two factors are closely associated with the expression of immune checkpoint genes. Thus, they may potentially block the markers relevant to therapeutic benefits of immune checkpoint.

The emergence of immune checkpoint blockade therapy has opened a new frontier in molecular subtyping of muscle-infiltrating urothelial carcinoma [26]. This approach facilitates better understanding of tumor microenvironment (TME) cells and secreted factors together with their genomic evolution in cancer cells [26]. Bladder cancer is one of the few solid tumors in which a persistent response to immune checkpoint blockade therapy has been observed in some patients. In urinary tract metastatic carcinoma, a combination of adjuvant and neoadjuvant chemotherapy is underway for a large number of immune checkpoint blockade trials targeting PD-1-PD-L1 and CTLA-4-CD80/CD86 immune checkpoints [27]. Immune checkpoint inhibitors have been tested in clinical trials against various malignancies, including metastatic urothelial carcinoma, and they have been effective in modulating tumor progression with limited side effects [28, 29]. Recent trials on immune checkpoint blockade including IMvigor 210 and Check Mate 275 demonstrated that PD-L1 and PD-1 inhibitors were effective against platinum resistance muscle invasive urothelial [30]. In bladder cancer, atezolizumab and nivolumab are two immune checkpoint inhibitors approved as second-line treatments for the disease [31, 32].

In this study, we identified three molecular subtypes based on immune genes. These subtypes have different immune microenvironments. The genes specifically expressed by a subtype strongly correlate with the expression pattern of most immune checkpoint genes. Genomic heterogeneity analysis revealed that the frequency of RB1 and FGFR3 mutations was high in IRGCluster. These findings suggest that the three molecular subtypes may show different response patterns to immunotherapy. Accordingly, the characteristic genes of these three subtypes may serve as molecular markers of immunotherapy.

In this study, bioinformatics techniques were used to identify potential immunogenic markers involved in BC. However, further prospective analyses on immunotherapy should be performed to fully define the cutoff values to be used. Secondly, in view of the heterogeneity of different tumor regions, more clinical factors should be included in the prediction model to improve the reliability of the findings. The results of bioinformatics analysis are not sufficient, thus experimental studies are needed to validate the present findings.

5. Conclusion

Immune-related genes can be divided into three molecular subtypes with distinct histological types, genetic and transcriptional characteristics. Our study provides a conceptual framework for understanding BC microenvironment. Thus, this study provides a novel insight into the immune-related state of BC and shed light on the prognostication of BC patient, all of them are of potential clinical implications.

Declarations

Disclosure

TCGA and GEO belong to public databases. The patients involved in the database have obtained ethical approval. Users can download relevant data for free for research and publish relevant articles. Our study

is based on open source data, so there are no ethical issues and other conflicts of interest.

Conflicts of Interest

The authors have declared that no competing interest exists.

Funding

Not applicable.

Acknowledgments

Not applicable.

References

1. Antoni S, Ferlay J, Soerjomataram I, Znaor A, Jemal A, Bray F. Bladder Cancer Incidence and Mortality: A Global Overview and Recent Trends. Eur Urol. 2017; 71: 96–108.
2. Jordan B, Meeks JJ. T1 bladder cancer: current considerations for diagnosis and management. Nat Rev Urol. 2019; 16: 23–34.
3. Kamat AM, Hahn NM, Efstathiou JA, Lerner SP, Malmstrom PU, Choi W, et al. Bladder cancer. Lancet. 2016; 388: 2796–810.
4. Loehrer PJ, Sr., Einhorn LH, Elson PJ, Crawford ED, Kuebler P, Tannock I, et al. A randomized comparison of cisplatin alone or in combination with methotrexate, vinblastine, and doxorubicin in patients with metastatic urothelial carcinoma: a cooperative group study. J Clin Oncol. 1992; 10: 1066–73.
5. Tang D, Park S, Zhao H. NITUMID: Nonnegative Matrix Factorization-based Immune-TUmor MIcroenvironment Deconvolution. Bioinformatics. 2019.
6. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011; 144: 646–74.
7. Mittal D, Gubin MM, Schreiber RD, Smyth MJ. New insights into cancer immunoediting and its three component phases–elimination, equilibrium and escape. Curr Opin Immunol. 2014; 27: 16–25.
8. West NR, Milne K, Truong PT, Macpherson N, Nelson BH, Watson PH. Tumor-infiltrating lymphocytes predict response to anthracycline-based chemotherapy in estrogen receptor-negative breast cancer. Breast Cancer Res. 2011; 13: R126.
9. Pages F, Mlecnik B, Marliot F, Bindea G, Ou FS, Bifulco C, et al. International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. Lancet. 2018; 391: 2128–39.
10. Zhang L, Conejo-Garcia JR, Katsaros D, Gimotty PA, Massobrio M, Regnani G, et al. Intratumoral T cells, recurrence, and survival in epithelial ovarian cancer. N Engl J Med. 2003; 348: 203–13.

11. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, et al. TGFbeta attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. Nature. 2018; 554: 544–8.
12. Topalian SL, Drake CG, Pardoll DM. Immune checkpoint blockade: a common denominator approach to cancer therapy. Cancer Cell. 2015; 27: 450–61.
13. Schumacher TN, Schreiber RD. Neoantigens in cancer immunotherapy. Science. 2015; 348: 69–74.
14. Muenst S, Soysal SD, Gao F, Obermann EC, Oertli D, Gillanders WE. The presence of programmed death 1 (PD-1)-positive tumor-infiltrating lymphocytes is associated with poor prognosis in human breast cancer. Breast Cancer Res Treat. 2013; 139: 667–76.
15. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, et al. TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. Nature. 2018; 554: 544–8.
16. Mirzal A. Nonparametric Tikhonov Regularized NMF and Its Application in Cancer Clustering. IEEE/ACM Trans Comput Biol Bioinform. 2014; 11: 1208–17.
17. Yu N, Gao YL, Liu JX, Shang J, Zhu R, Dai LY. Co-differential Gene Selection and Clustering Based on Graph Regularized Multi-View NMF in Cancer Genomic Data. Genes (Basel). 2018; 9.
18. Ye C, Toyoda K, Ohtsuki T. Blind Source Separation on Non-contact Heartbeat Detection by Non-negative Matrix Factorization Algorithms. IEEE Trans Biomed Eng. 2019.
19. Li B, Severson E, Pignon JC, Zhao H, Li T, Novak J, et al. Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. Genome Biol. 2016; 17: 174.
20. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012; 16: 284–7.
21. Robertson AG, Kim J, Al-Ahmadie H, Bellmunt J, Guo G, Cherniack AD, et al. Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. Cell. 2017; 171: 540 – 56.e25.
22. Tan TZ, Rouanne M, Tan KT, Huang RY, Thiery JP. Molecular Subtypes of Urothelial Bladder Cancer: Results from a Meta-cohort Analysis of 2411 Tumors. Eur Urol. 2019; 75: 423–32.
23. Newman AM, Liu CL, Green MR. Robust enumeration of cell subsets from tissue expression profiles. 2015; 12: 453–7.
24. Dadhania V, Zhang M, Zhang L, Bondaruk J, Majewski T, Siefker-Radtke A, et al. Meta-Analysis of the Luminal and Basal Subtypes of Bladder Cancer and the Identification of Signature Immunohistochemical Markers for Clinical Use. EBioMedicine. 2016; 12: 105–17.
25. Ren R, Tyryshkin K, Graham CH, Koti M, Siemens DR. Comprehensive immune transcriptomic analysis in bladder cancer reveals subtype specific immune gene expression patterns of prognostic relevance. Oncotarget. 2017; 8: 70982–1001.
26. Vidotto T, Nersesian S, Graham C, Siemens DR, Koti M. DNA damage repair gene mutations and their association with tumor immune regulatory gene expression in muscle invasive bladder cancer subtypes. J Immunother Cancer. 2019; 7: 148.
27. Dallos MC, Drake CG. Blocking PD-1/PD-L1 in Genitourinary Malignancies: To Immunity and Beyond. Cancer J. 2018; 24: 20–30.

28. Ghate K, Amir E, Kuksis M, Hernandez-Barajas D, Rodriguez-Romo L, Booth CM, et al. PD-L1 expression and clinical outcomes in patients with advanced urothelial carcinoma treated with checkpoint inhibitors: A meta-analysis. Cancer Treat Rev. 2019; 76: 51–6.
29. Tripathi A, Plimack ER. Immunotherapy for Urothelial Carcinoma: Current Evidence and Future Directions. Curr Urol Rep. 2018; 19: 109.
30. Rosenberg JE, Hoffman-Censits J, Powles T, van der Heijden MS, Balar AV, Necchi A, et al. 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. 2016; 387: 1909–20.
31. Kim HS, Seo HK. Immune checkpoint inhibitors for urothelial carcinoma. Investig Clin Urol. 2018; 59: 285–96.
32. El Rassy E, Assi T, Bakouny Z, Pavlidis N, Kattan J. Beyond first-line systemic treatment for metastatic urothelial carcinoma of the bladder. Clin Transl Oncol. 2019; 21: 280–8.

Figures

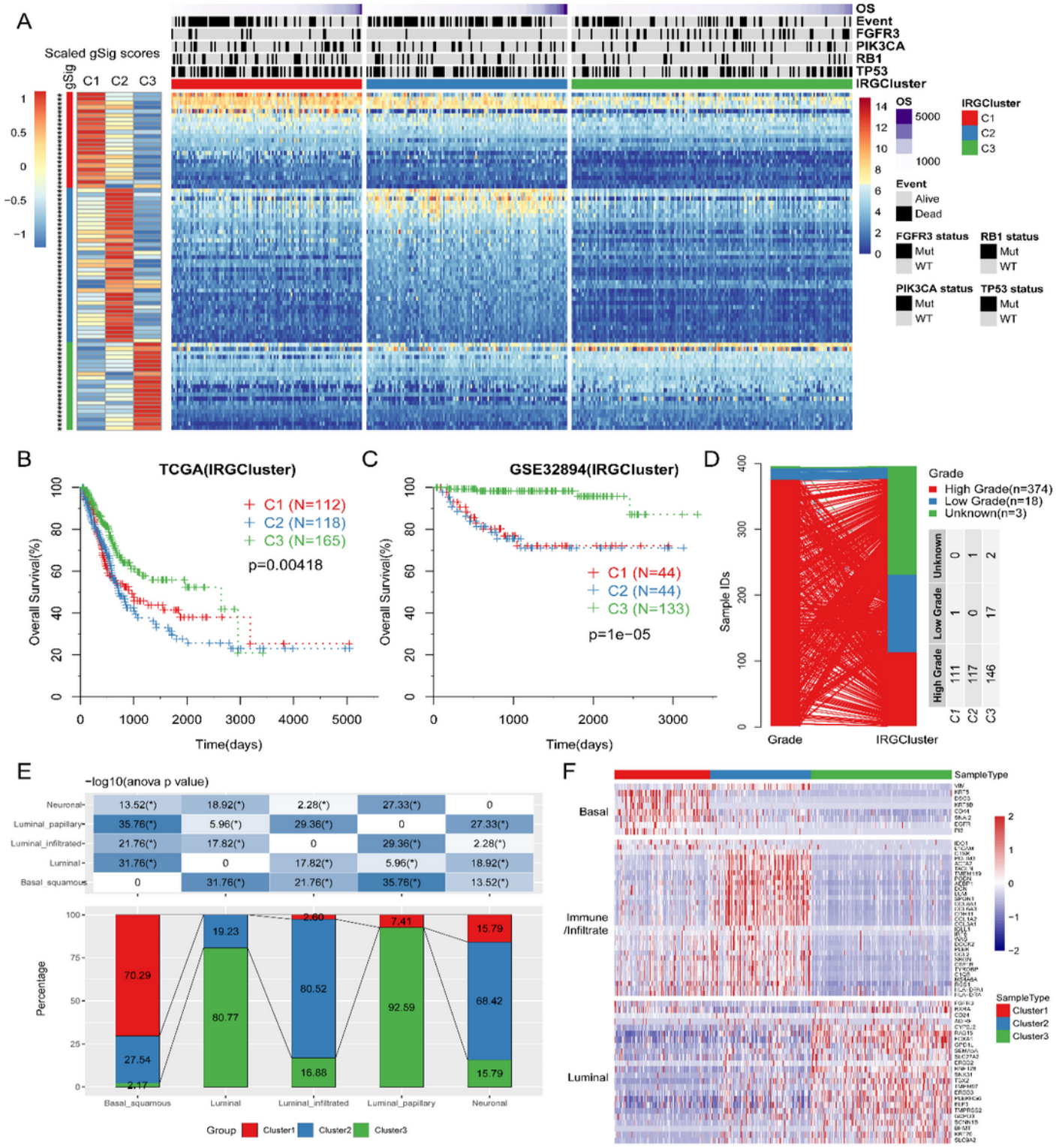

Figure 1

BC immunophenotyping based on IRGs. (A). Heatmap for BC immuno-types and expression of gene modules. (B). The overall survival KM curve of the immunological subtype from the TGGA data set. (C). The overall survival KM curve of the immunological subtype of the GSE32894data set. (D). Comparison between BC grade and IRGCluster. (E). Comparison between published subtypes and NMF clusters. (F). Distribution of marker gene in three subtypes.

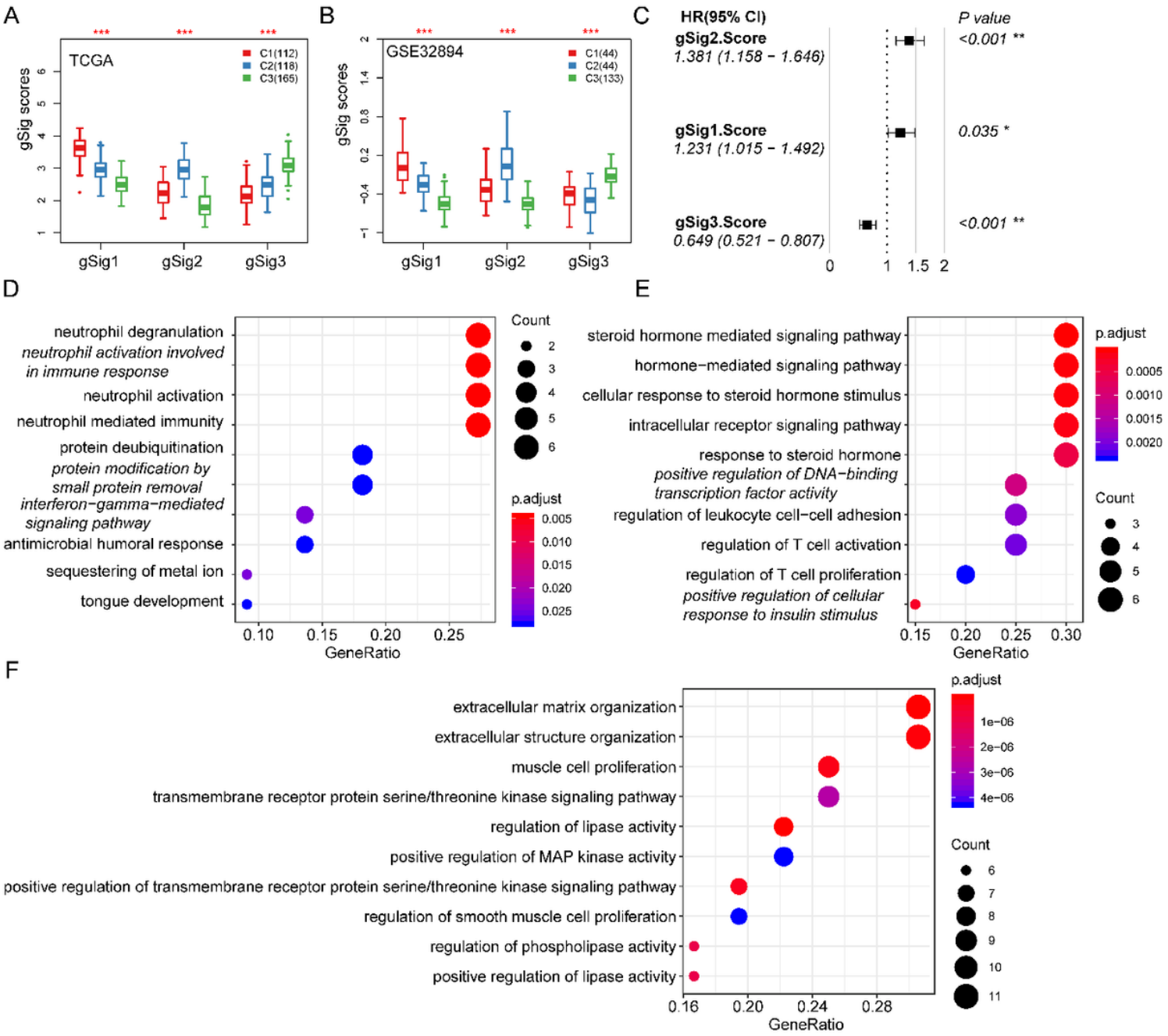

Figure 2

Expression characteristics and functional annotation of immune gene modules. (A). The distribution of gSigs scores in the IRGCluster. (B). The distribution of gSigs scores in the IRGCluster in GSE32894 data set. (C); The relationship between gSigs score and OS. (D)-(F). GO enrichment analysis for gSig1, gSig2 and gSig3. The Dot plot shows the GO enrichment results of the top 10, and the KEGG pathway has no significant enrichment results. The color in the figure indicates significance level, and the dot size indicates the number of genes.

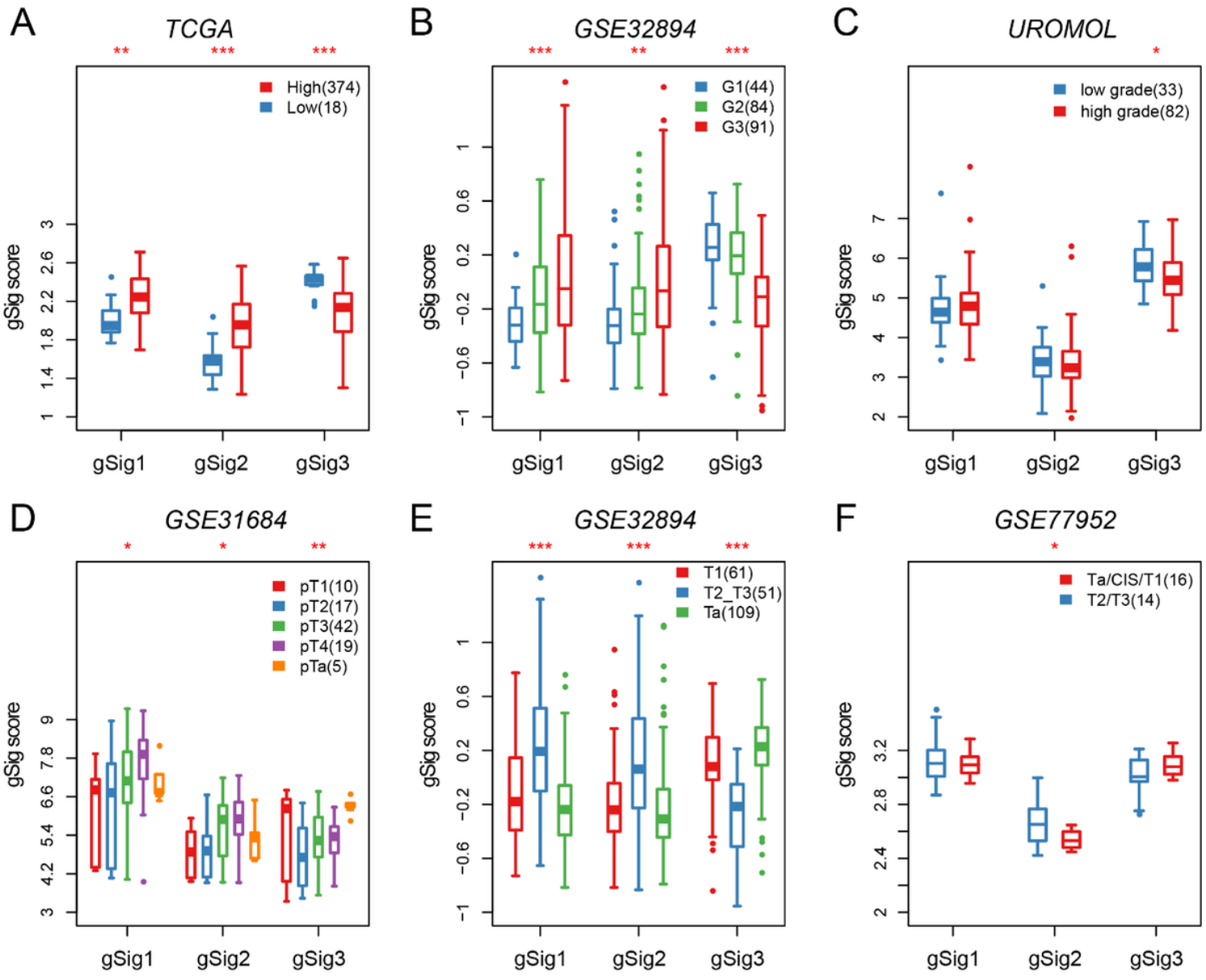

Figure 3

The relationship between gSig score and different clinical characteristics. (A)-(C). Distribution of gSig score in the TCGA data set for High/Low grade samples. (D)-(F). Distribution of gSig score on muscle invasive and non-muscle invasive.

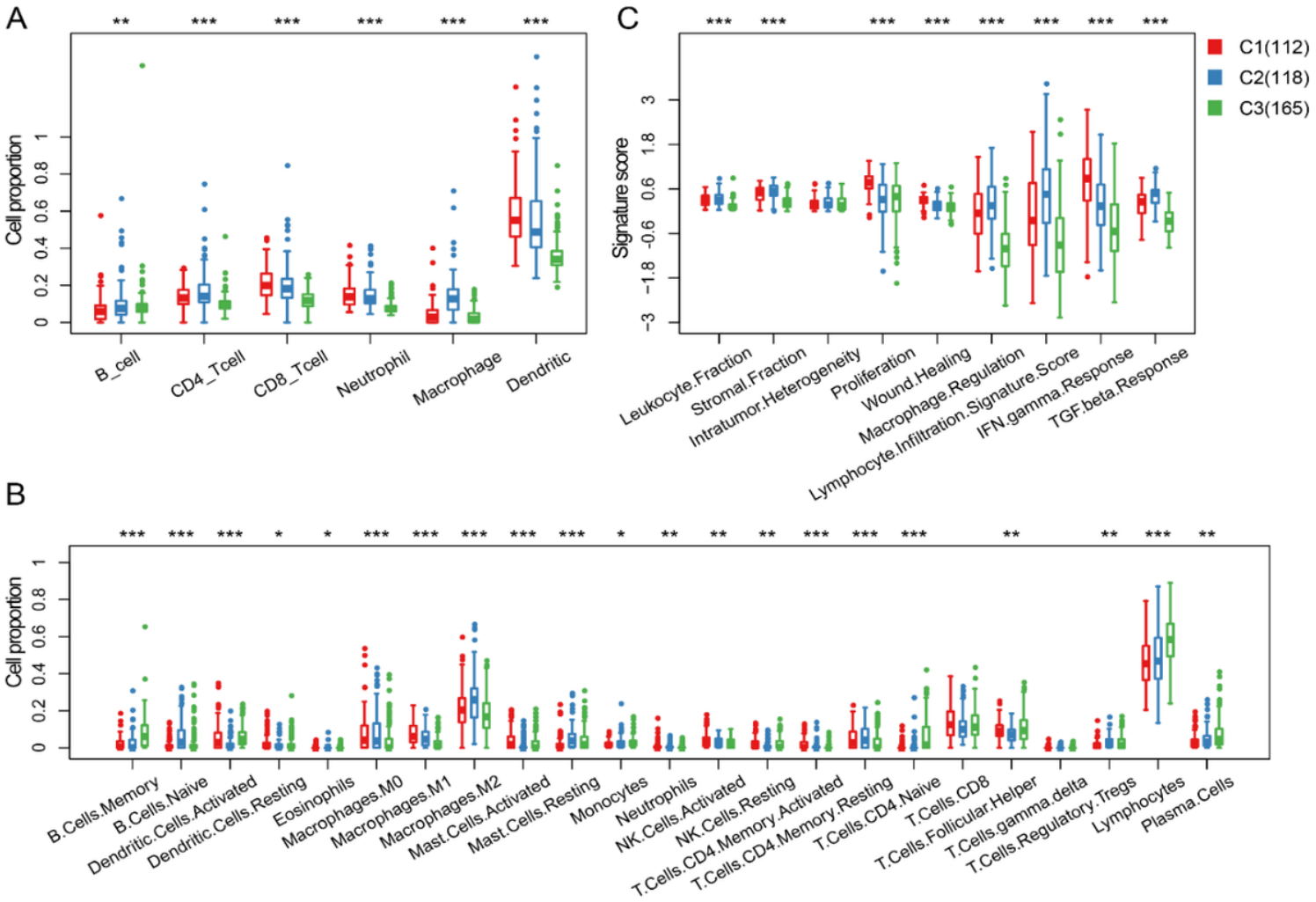

Figure 4

The immune landscape for BC IRGCluster. (A). Scores for 6 immune cells in the IRGCluster. (B). Scores of 22 immune cells in the IRGCluster. (C). Score of immune expression signature in IRGCluster.

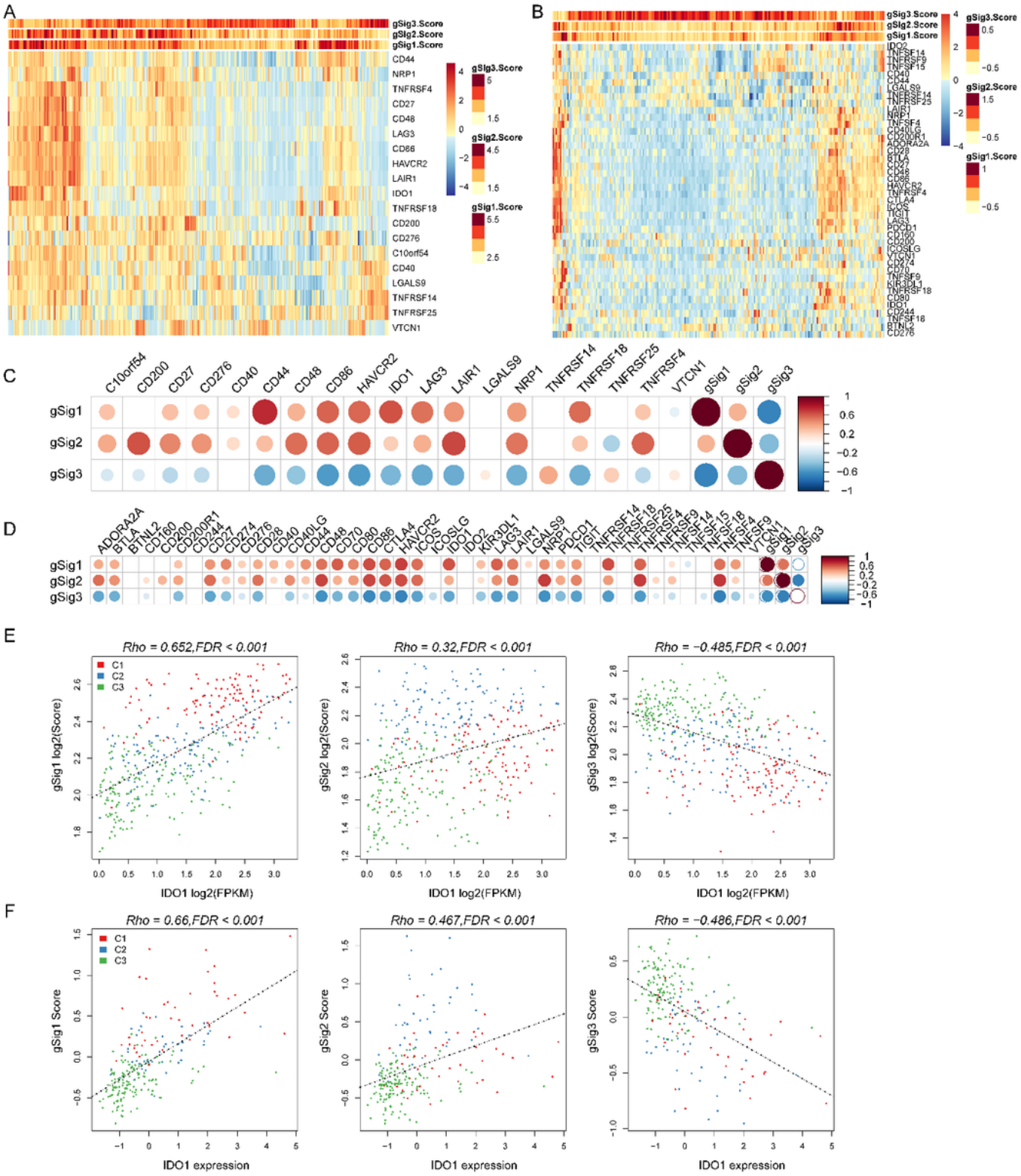

Figure 5

(A)-(B). Relationship between gSig scores and expression of ICGs in TCGA and GSE32894 data sets. (C)-(D). Correlation matrix between expression of ICGs in TCGA and GSE32894 data set and gSig score. (E). Scatter plots for IDO1 gene expression and gSig in TCGA data set. (F). Scatter plots for IDO1 gene expression and gSig in GSE32894 dataset.

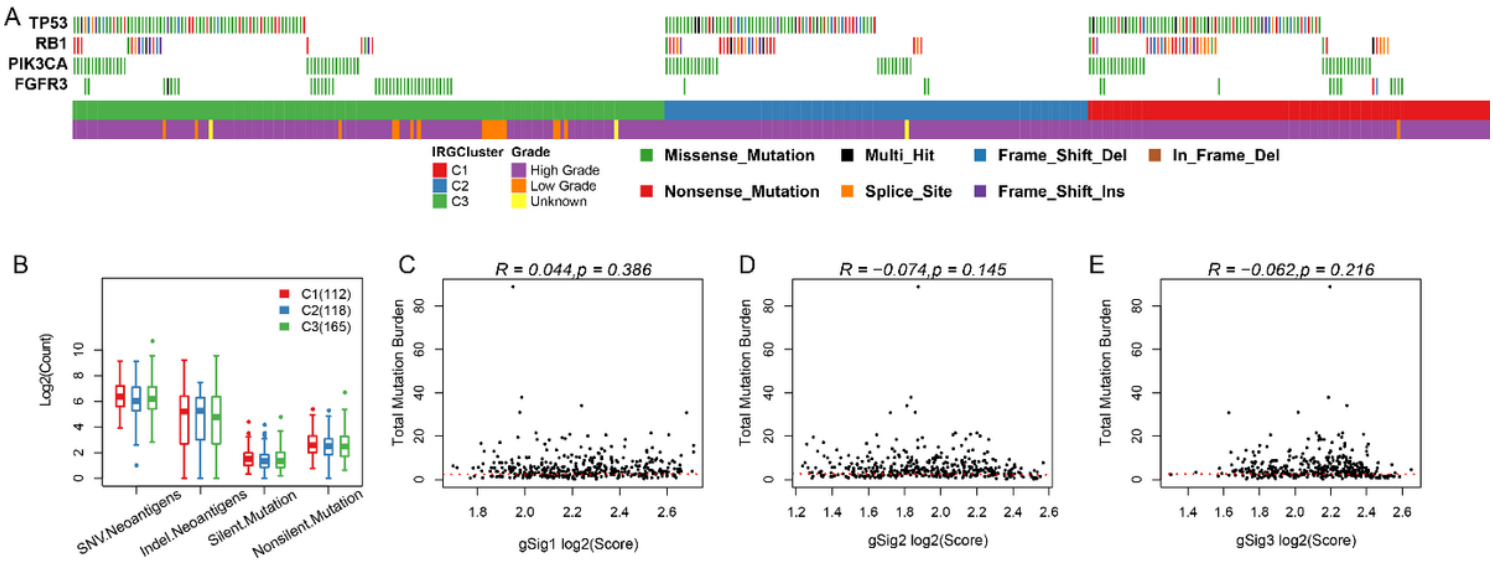

Figure 6

Genome heterogeneity analysis of IRGCluster. (A). Mutation of four high frequency mutant genes in IRGCluster. (B). The distribution of neoantigens and mutation burden on IRGCluster. (C). The relationship between gSig1 score and TMB. (D). The relationship between gSig2 score and TMB. (E). The relationship between gSig3 score and TMB.