


# Identification of testicular cancer immune infiltrates and novel immune cell subtypes

Zhiguo Zhu<sup>1,2</sup>, Xujun Xuan<sup>3</sup>, Xinkun Wang<sup>3</sup>, Miaomiao Wang<sup>4</sup>, Chunyang Meng<sup>5</sup> and Zhonghai Li<sup>1</sup> 

1 Department of Urology, Affiliated Hospital of Jining Medical University, Jining Medical University, China

2 Postdoctoral Mobile Station of Shandong University of Traditional Chinese Medicine, Jining, China

3 Department of Andrology, The Seventh Affiliated Hospital, Sun Yet-sen University, Shenzhen, China

4 Department of Medical, Affiliated Hospital of Jining Medical University, Jining Medical University, China

5 Medical Research Center, Affiliated Hospital of Jining Medical University, Jining Medical University, China

## Keywords

biomarker; immune infiltration; molecular subtyping; risk signature; STC1; testicular germ cell tumors

## Correspondence

Z. Li, Department of Urology, Affiliated Hospital of Jining Medical University, Jining Medical University, Guhuai Road 89#, Jining, Shandong 272029, China

E-mail: [li\\_urology@163.com](mailto:li_urology@163.com)

and

C. Meng, Medical Research Center, Affiliated Hospital of Jining Medical University, Jining Medical University, Guhuai Road 89#, Jining, Shandong 272029, China

E-mail: [mengchunyang1600@mail.jnmc.edu.cn](mailto:mengchunyang1600@mail.jnmc.edu.cn)

(Received 16 April 2023, revised 23 July 2023, accepted 3 August 2023)

doi:10.1002/2211-5463.13688

Testicular germ cell tumors (TGCT) are the most common type of testicular cancer, comprising 90–95% of cases and representing the most prevalent solid malignancy in young adult men. Immune infiltrates play important regulatory roles in tumors, but their role in TGCT remains unclear. Molecular subtyping is a promising way to provide precisely personalized treatment and avoid unnecessary toxicities. This study investigated immune infiltrates, key biomarkers, and immune subtyping of TGCT. In [GSE3218](#), 24 differentially expressed immune genes (immDEGs) were identified. A new risk signature consisting of six immDEGs was developed using these genes. Individuals in the high-risk group had poor overall survival (OS; hazard ratio of 4.61 and  $P$ -value  $< 0.001$ ). We validated the six-immDEGs risk signature in pure seminoma and mixed TGCT types. Two distinct immune patterns (Cluster 1 and Cluster 2) were identified using the `CONSENSUSCLUSTERPLUS`, and Cluster 1 possessed an unfavorable OS compared with Cluster 2 (hazard ratio, 2.56;  $P < 0.001$ ). Cluster 1 patients had significantly lower naive B cells, memory B cells, plasma cells, naive CD4 T cells, gamma delta T cells, and activated dendritic cells than Cluster 2 patients. Genes relating to the WNT signaling pathway, TGF- $\beta$  signaling pathway, antigen processing and presentation, and NK cell-mediated cytotoxicity were associated with TGCT. STC1 was elevated in TGCT tissues, and its high expression showed advanced clinicopathological characteristics and poor prognosis of TGCT. Our findings may contribute to an increased understanding of the onset and progression of TGCT.

Testicular cancer is the most prevalent solid malignancy among young adult men [1,2]. The incidence has been steadily increasing, especially in the developed

countries [3]. The etiology of testicular cancer is still unclear, and its pathological types are diverse, most of which (90–95%) are testicular germ cell tumors

## Abbreviations

BPs, biological processes; CCs, cellular components; DEGs, differentially expressed genes; FC, fold change; GEO, gene expression omnibus; GO, gene ontology; GSEA, gene set enrichment analysis; H&E, hematoxylin and eosin; IHC, immunohistochemical; immDEGs, differentially expressed immune genes; IRS, immunoreactivity score; KEGG, Kyoto Encyclopedia of Genes and Genome; KM, Kaplan–Meier; LASSO, least absolute shrinkage and selection operator; MFs, molecular functions; OS, overall survival; PD-L1, programmed death receptor ligand-1; PPI, protein–protein interaction; ROC, receiver operating characteristic; STC1, stanniocalcin-1; STRING, Search Tool for the Retrieval of Interacting Genes; TGCT, testicular germ cell tumors; TME, tumor microenvironment; UMAP, uniform manifold approximation and projection.

(TGCT). Most of these patients are cured with surgery alone or in combination with chemotherapy, with a 5-year survival rate of 98% for localized disease. However, primary resistance, disease progression after therapy, and adverse effects of treatment are still major clinical challenges [4,5]. Therefore, identifying the molecular mechanism and novel predictors of prognosis is important for diagnosis and personalized therapy.

The immune cells in the tumor microenvironment (TME) are critically involved in tumorigenesis [6,7], and effective immunotherapy has been achieved in multiple tumors [6]. Due to the existence of the blood-testis barrier, the testis is often regarded as a special site that is exempt from normal systemic immune surveillance. However, immune infiltrates are widespread in testicular tumors, particularly in TGCT. In 2002, Yakirevich *et al.* [8] found that the number of activated cytotoxic lymphocytes was increased in testicular seminomas. In 2015, Fankhauser *et al.* [9] highlighted that programmed death receptor ligand-1 (PD-L1) was frequently expressed in TGCT. Subsequent studies also confirmed PD-1 (programmed death 1)/PD-L1 as a potential therapeutic target of TGCT [10–13]. Although the immune infiltrate plays an important role in TGCT, it remains scarcely studied compared with other tumors. Especially, a systematic understanding of the immune milieu in TGCT is currently lacking.

The present research employed gene expression profiles to assess immune-related genes that exhibited a significant difference between individuals with TGCT and control samples. A risk signature comprising immune-related genes was successfully developed and is effective in predicting patient prognosis. Moreover, two distinct immune patterns with significant prognostic differences were identified. The identification of variations in immune infiltration patterns between TGCT and healthy tissues would aid in comprehending the onset and progression of TGCT and devising efficient treatment approaches.

## Materials and methods

### Data and human tissues collection

The datasets GSE3218 (GPL96) and GSE3218 (GPL97) were provided by the Gene Expression Omnibus (GEO) database [14]. To generate the integrated GEO dataset, Empirical Bayes methods were employed to adjust batch effects in microarray expression data [15]. The merged data contained 202 TCGT samples and 12 normal samples. TCGA dataset was downloaded from the assistant for clinical bioinformatics database platform (<https://www.aclbi.com/>).

The MSigDB databases were used to obtain 1811 immune-related genes [16].

A total of 24 TGCT (seminoma) tissues were retrieved from January 2016 to September 2022 in Affiliated Hospital of Jining Medical University. And normal testicular tissues ( $n = 8$ ) were obtained from patients with prostate cancer undergoing orchiectomy from January 2018 to January 2019. Two pathologists confirmed all tissue types using hematoxylin and eosin (H&E) staining. The study was conducted in accordance with the Declaration of Helsinki. All individuals granted their written informed consent prior to participation. The approval of this study was granted by the Ethics Committee of the Affiliated Hospital of Jining Medical University (number: 2023-04-C038).

### Identification of differentially expressed immune genes

Differentially expressed genes (DEGs) between tumors and control samples were identified utilizing the 'LIMMA' R package [17] (cutoff criteria, false discovery rate  $< 0.05$  and  $|\text{fold change}| > 2$ ). Volcano plots and heat maps were utilized to visualize the outcomes of DEGs. In order to identify the differentially expressed immune genes (immDEGs), the DEGs and immune-related genes were intersected.

### Construction of the immune-related genes risk signature

Univariate Cox regression analysis was conducted to identify immDEGs that were related to overall survival (OS). The R package 'GLMNET' [18] and least absolute shrinkage and selection operator (LASSO) were applied to obtain the characteristic genes of risk signature. The calculation formula of the model is as follows: risk score = sum (each gene's expression  $\times$  corresponding coefficient). To classify individuals into high- and low-risk groups, the median risk score was used as a cutoff value. The Kaplan–Meier (KM) method was utilized to examine differences in OS between both risk groups. Time-dependent receiver operating characteristic (ROC) curves were generated to assess the stability of the risk signature at 1-, 3-, and 5-year durations.

### Identification of immune molecular subtypes and extraction of DEGs between two distinct immune patterns

The R package 'CONSENSUSCLUSTERPLUS' [19] was employed to identify the immune molecular subtypes. The KM method was employed to analyze OS differences between immune molecular subtypes. The R package 'LIMMA' was utilized to extract DEGs between two distinct immune patterns (cutoff criteria, false discovery rate  $< 0.05$ , and  $|\text{fold change (FC)}| > 1.5$ ). The ROC curve was used to evaluate the ability of genes to distinguish immune subgroups.

## Construction of PPI network

The Search Tool for the Retrieval of Interacting Genes [20] (STRING) was utilized to construct the protein–protein interaction (PPI) network, which was then visualized using the CYTOSCAPE software (<https://cytoscape.org/>).

## Enrichment analysis

For gene ontology (GO) annotation analysis in terms of molecular functions (MFs), biological processes (BPs), and cellular components (CCs), as well as Kyoto Encyclopedia of Genes and Genome (KEGG) pathway enrichment analysis, the R packages ‘ORG.HS.EG.DB’ and ‘CLUSTERPROFILER’ [21] were employed. For statistical significance, a *P*-value of < 0.05 was set as the criterion. To explore the variations in BPs between distinct subgroups, gene set enrichment analysis (GSEA) [22] was carried out. To assess relevant pathways and underlying mechanisms, the ‘c2.cp.kegg.v7.4.symbols.gmt’ gene sets were downloaded from the Molecular Signatures Database. A *P*-value of < 0.05 was taken to be a statistically significant criterion for this analysis.

## Comparison of immune infiltrate cells

The analytical tool CIBERSORTx [23] can use gene expression data to perform linear support vector regression, thus estimating immune cell infiltration. We calculated 22 immune cell types in patients using CIBERSORTx and analyzed the correlation and difference between the proportion of different immune cells, the proportion of immune cells, and gene expression. A statistically significant *P*-value of < 0.05 was used to determine the significance of the results.

## Immunohistochemical assessment

The TGCT and normal testicular tissues were subjected to an immunohistochemical (IHC) assay using a 1 : 100 dilution of an antibody against STC1 (#20621-1-AP; Proteintech, Wuhan, Hubei, China). Paraffin wax was used for embedding the tissues. IHC was employed according to a standard method. A final immunoreactivity score (IRS) was obtained for each case by multiplying the positive cell percentage score (< 5%, 0; 5–25%, 1; 26–50%, 2; 51–75%, 3; > 75%, 4) and the intensity score (negative, 0; weak, 1; moderate, 2; strong, 3). Divided into four levels: 0–2: negative (–), 3–5: weak positive (+), 6–8: positive (++) , 9–12: strong positive (+++).

## Statistical analysis

All the statistical analyses were carried out utilizing RSTUDIO 4.1.3 (Posit, BOSTON, MA, USA). The independent Student’s *t*-test (normally distributed variables) and Mann–Whitney *U*-test (non-normally distributed variables) were conducted to compare two groups of continuous variables.

Pearson correlation analysis was conducted to estimate correlation coefficients between distinct variables. A *P*-value of < 0.05 suggested the significance level. Other statistical methods have been mentioned in the corresponding ‘Materials and methods’ section above.

## Results

### Dysregulation of immune-related genes in TGCT

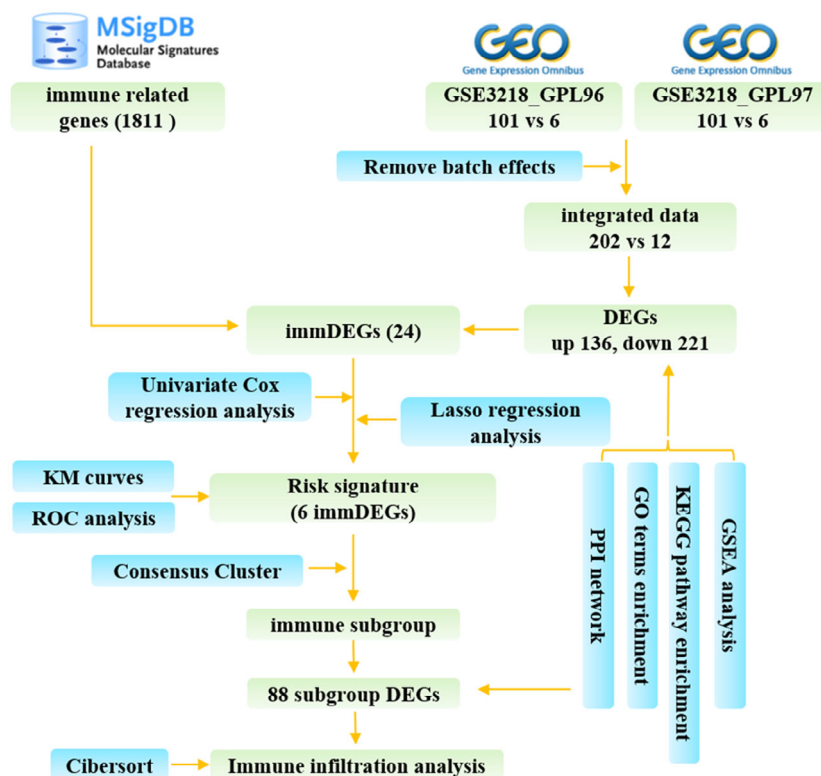
Figure 1 presents the flow chart for data collection and analysis. The integrated GEO dataset, including 202 TCGT samples and 12 normal samples, was obtained after removing batch effects (Fig. S1; Fig. 2A,B). The result of unsupervised clustering of the samples with all genes is shown in Fig. S2A. In total, 357 DEGs were screened out in TCGT samples compared with normal samples. Out of these, 136 were upregulated, whereas 221 were downregulated (Fig. 2C,D). The five most significantly upregulated genes were *ESRP1* (FC = 5.42), *SOCS3* (FC = 4.92), *PGKI* (FC = 4.77), *CTSB* (FC = 4.72), and *GBP1* (FC = 4.71). The five most significantly downregulated genes were *CABYR* (FC = 0.06), *KHDRBS3* (FC = 0.08), *PRND* (FC = 0.10), *TSGA10* (FC = 0.12), and *GSTM3* (FC = 0.13). By intersecting DEGs with immune-related genes, a total of 24 immDEGs were identified. Among these, 15 were upregulated, while nine were downregulated (Fig. 3A). The heat maps (Fig. 3B) and histograms (Fig. 3C) show their expression levels in samples.

### Construction of the immune-related genes risk signature

If the risk signature contains too many variables, it will increase the difficulty of applying the signature. Therefore, we first screened for genes that were associated with patients’ survival time. A total of eight immDEGs that were linked to the OS of patients were identified via univariate Cox regression (Fig. 3D). Except for *STC1*, other genes served as protective factors. But only the expression of *RORA* was decreased in TGCT tissues compared with its corresponding normal tissues. To prevent overfitting of the model, the LASSO algorithm was utilized to discover six characteristic genes out of eight immDEGs (Fig. S3).

$$\begin{aligned} \text{Risk Score} = & (-0.36) \times IGKC + (-0.20) \times IGLC1 + 0.45 \\ & \times STC1 + (-0.05) \times ITGB2 + (-0.02) \\ & \times CTSS + (-0.93) \times RORA. \end{aligned}$$

As per the median risk score calculated using the formula, TGCT samples were categorized into high- and



**Fig. 1.** Study flow diagram.

low-risk groups (Fig. 3E). The scatter plot depicted the survival rates of individuals based on their risk scores, and the heatmap demonstrated the differential expression profiles of six immDEGs between both risk groups. KM survival curves revealed that the individuals in the high-risk group possessed a poor OS (HR, 4.61; 95% CI, 2.43–8.78;  $P < 0.001$ ; Fig. 3F). The efficacy of the risk signature was validated using time-dependent ROC curves at 1-, 3-, and 5-year durations. (Fig. 3G).

Correlations between gene expression levels between six immDEGs were examined in all samples (Fig. 4A), TGCT samples (Fig. 4B), and normal samples (Fig. 4C). *STC1*, the only hazardous gene, had weak correlations with other genes in all samples. However, *STC1* had positive correlations with *RORA* and *CTSS* and a negative correlation with *IGKC* in subgroups. There was a clear correlation between *IGKC*, *IGLC1*, *ITGB2*, *CTSS*, and *RORA* in different groups. The collections between *CTSS* and *IGKC*, *CTSS* and *ITGB2*, *RORA*, and *ITGB2*, and *RORA* and *IGKC* showed considerable variation in TGCA and normal samples.

### Identification of immune molecular subtypes

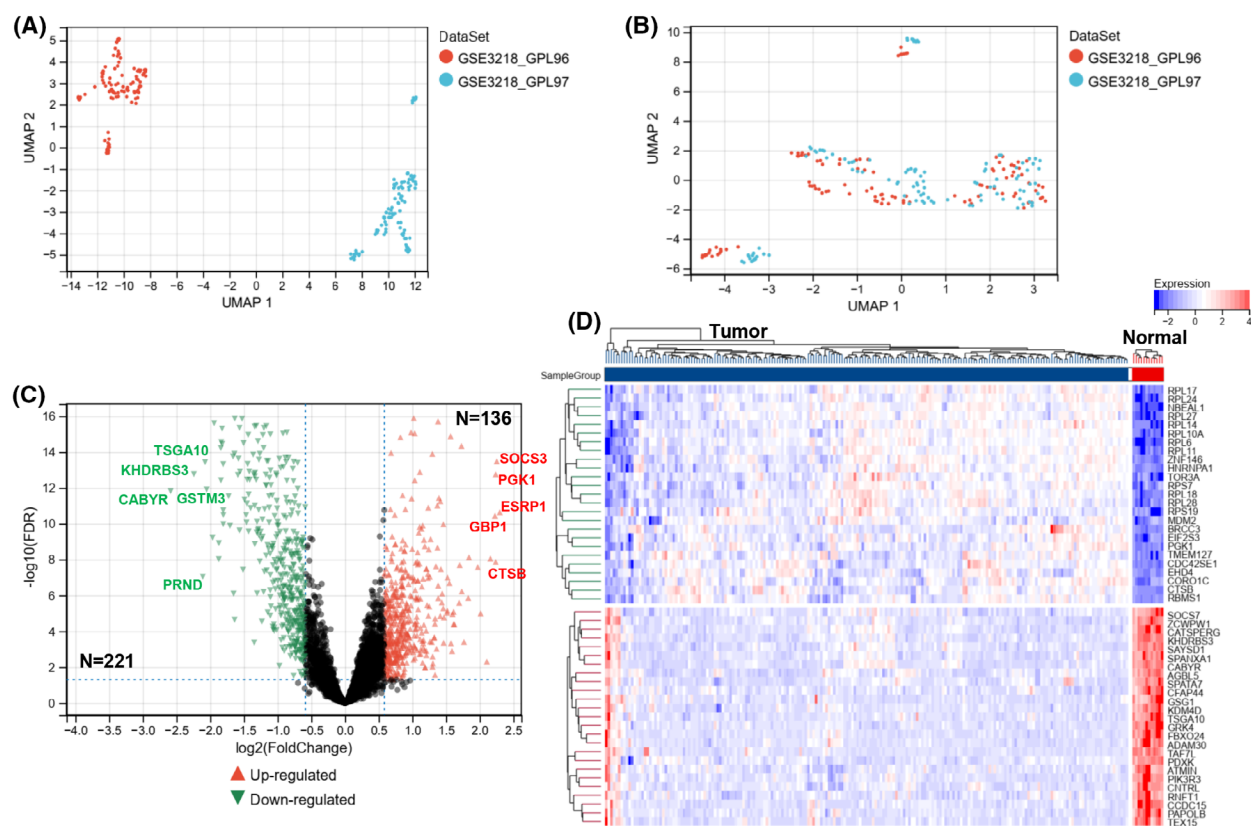
Based on six signature immDEGs, two immune molecular subtypes (Cluster 1 and Cluster 2) were identified

utilizing the R package ‘CONSENSUSCLUSTERPLUS’ (Fig. 5A–C). Cluster 1 comprised 108 samples, and Cluster 2 comprised 106 samples. When comparing Cluster 1 samples with Cluster 2 samples, 88 DEGs were determined, with 43 upregulated and 45 downregulated DEGs (Fig. 5D). Figure 5E shows the top 25 DEGs to show the considerable differences between the two clusters. Figure 5F shows 24 immDEGs expression levels between the two groupings. Expression levels of *ITGB2*, *ERAP2*, *IGKC*, *HLA-DPBI*, *CTSS*, *IGLC1*, and *RSAD2* were remarkably elevated in Cluster 2 than that in Cluster 1, whereas the expression levels of *STC1*, *INHBA*, *SORT1*, *PENK*, and *RORA* were considerably lowered in Cluster 2 as opposed to Cluster 1 (Fig. 5G). Moreover, KM survival curves revealed that Cluster 1 possessed a poor OS (HR, 2.56; 95% CI, 1.37–4.79;  $P < 0.001$ ; Fig. 5H). The ROC curves were used to assess the six signature immDEGs individually predicted values. The findings showed that all six signature immDEGs had good classification efficacy for two immune molecular subtypes (Fig. 6).

### PPI network and enrichment analysis

A PPI network of DEGs (tumor vs. normal) was constructed using CYTOSCAPE software, according to the





**Fig. 2.** Identification of DEGs. Uniform manifold approximation and projection (UMAP) between datasets before de-batching (A) and after de-batching (B). (C) The volcano plot was constructed based on fold change values and  $P$ -adjust, with red and blue dots representing upregulated and downregulated genes, respectively. (D) The heatmap depicting the top 50 genes identified through analysis of differential gene expression.

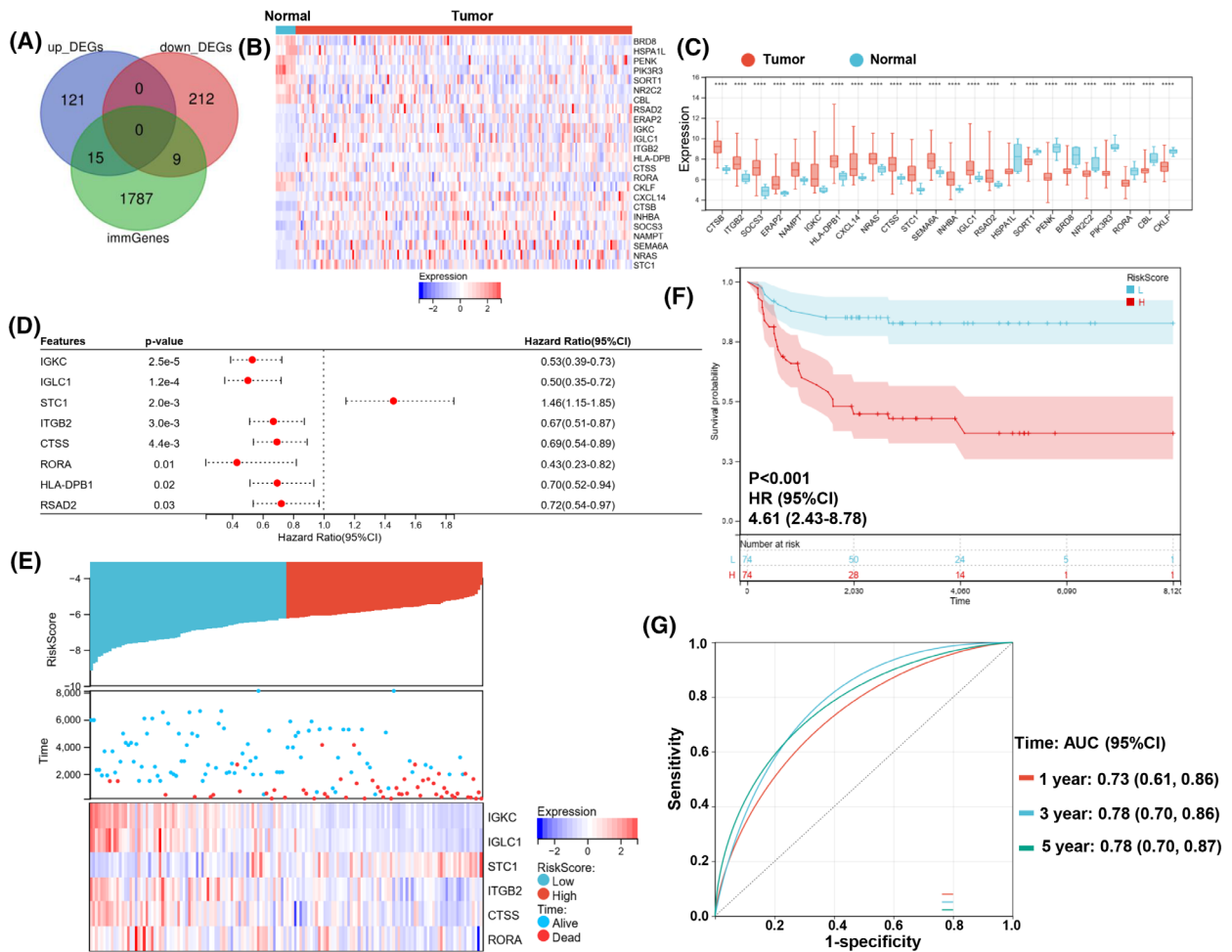
data from the STRING database, to explore their relationships. The PPI network of DEGs demonstrated that the distribution of the 24 immDEGs was relatively scattered (Fig. 7A). In the PPI network of immDEGs, *STC1* and *RORA* were linked to four immDEGs, *ITGB2* was related to nine immDEGs, and *CTSS* was linked to 10 immDEGs (Fig. 7B). Following the initial analysis, the role of DEGs between TGCT and normal samples in biological functions was investigated. Initially, the DEGs were found to have a significant association with fertility (Fig. 7C). Subsequently, KEGG pathway analysis revealed enrichment of these DEGs in immune, aging, cancer, and infection-related pathways (Fig. 7D).

A PPI network was constructed for the DEGs between the two immune patterns (Fig. 8A). Similarly, the distribution of 24 immDEGs was relatively scattered. The outcomes of the Go annotation analysis indicated that the DEGs were enriched in processes related to the immune system (Fig. 8B). Subsequently, KEGG pathway analysis revealed enrichment of these

DEGs in pathways related to the immune system and infection (Fig. 8C). Finally, a GSEA was performed on all genes between the two immune patterns (Fig. 8D,E). The results revealed that eight BPs, such as the WNT signaling pathway and TGF BETA signaling pathway, were activated in Cluster 1 compared with Cluster 2. Simultaneously, eight BPs, such as antigen processing and presentation, NK cell-mediated cytotoxicity, and primary immunodeficiency, were inhibited.

### Differences in immune characteristics

Immune cell infiltration was estimated using CIBERSORTx. Figure 9A displays the proportion of 22 immune cell types in each sample. The result of unsupervised clustering of the samples with immune cell types is shown in Fig. S2B. Notably, the levels of activated NK cells, resting dendritic cells, activated mast cells, and eosinophils were remarkably elevated in Cluster 1 compared with Cluster 2. In contrast, the



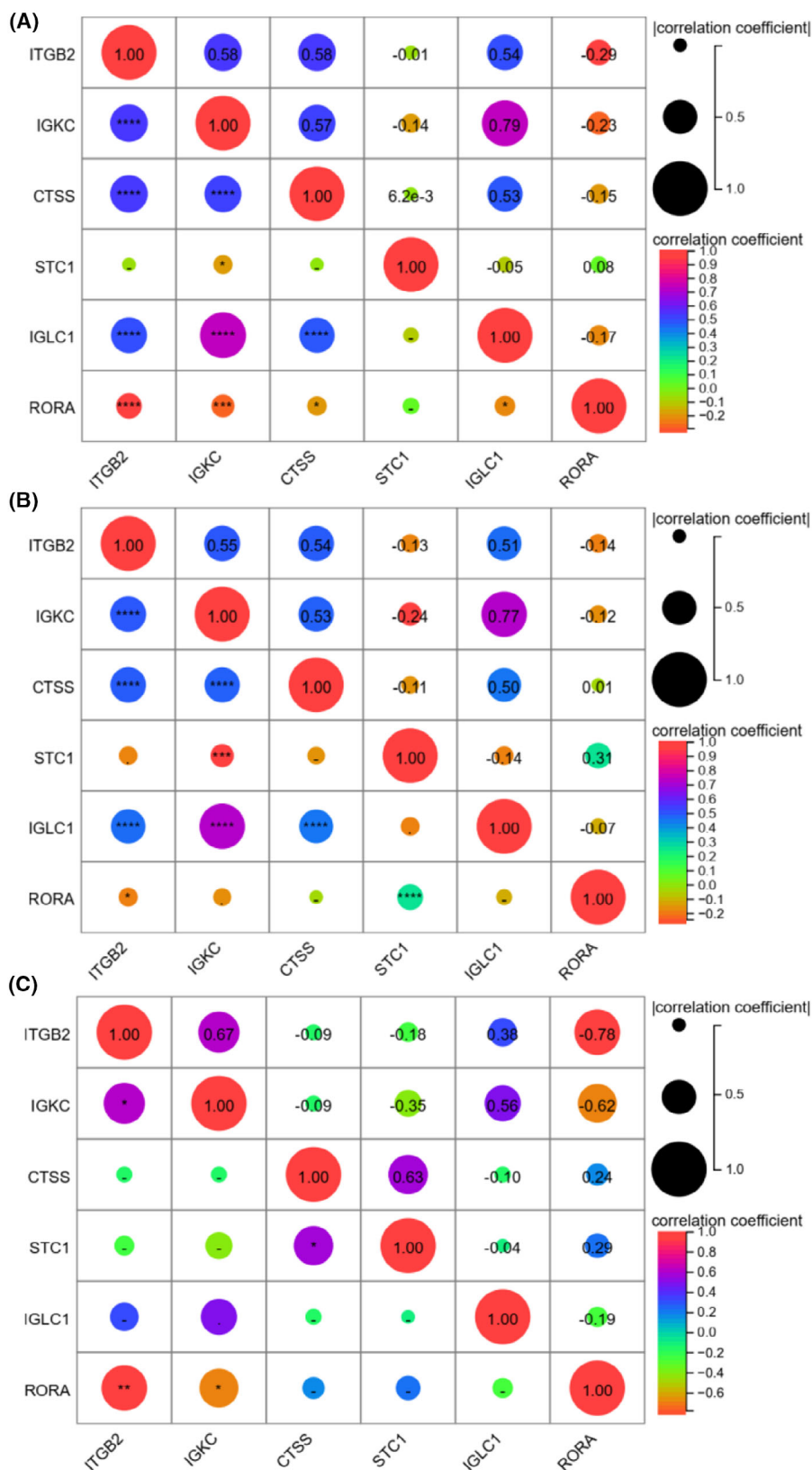
**Fig. 3.** Differentially expressed immune genes and construction of the immune-related genes risk signature. (A) Immune gene vs. DEG Venn diagram. (B) The heatmap of 24 immDEGs. (C) The expression histogram of 24 immDEGs in TGCT ( $n = 202$ ) and normal tissues ( $n = 12$ ). (D) Identification of eight immDEGs associated with OS of TGCT via univariate Cox regression. (E) The risk score, survival time, and survival status of TGCT. (F) Kaplan–Meier survival analysis of the two risk groups, TGCT samples were divided into high- and low-risk groups based on their median risk score. (G) Time-dependent ROC curves verified the predictive efficacy of the risk signature at 1, 3, and 5 years. Presentation of data as means  $\pm$  standard deviation.  $**P < 0.01$ ;  $***P < 0.0001$ ; rank-sum test. immGenes, immune genes.

levels of naïve B cells, memory B cells, plasma cells, native CD4<sup>+</sup> T cells, gamma delta T cells, and activated dendritic cells were considerably lowered in Cluster 1 than in Cluster 2 (Fig. 9B).

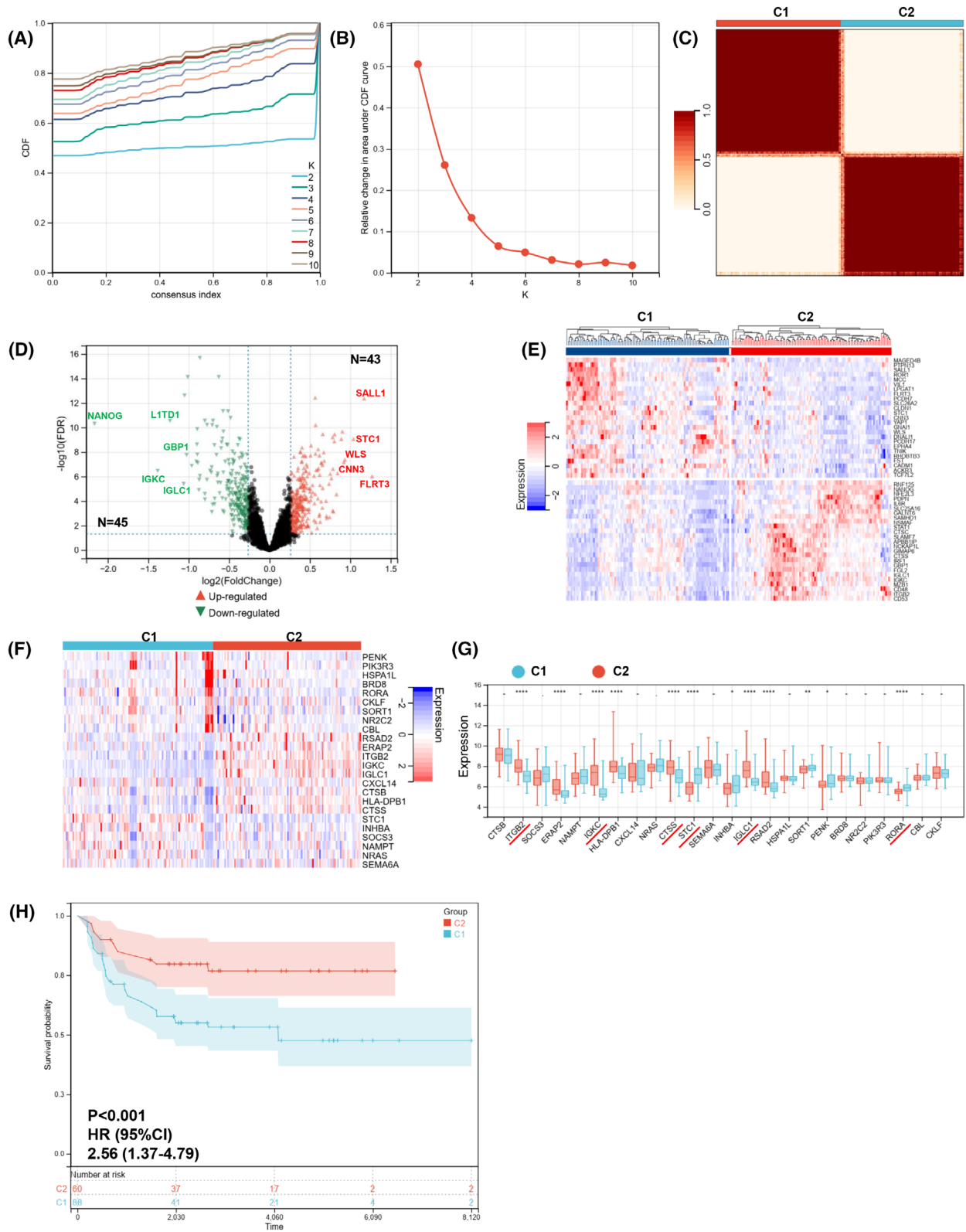
The correlation between the immune cell contents of samples in all individuals of Cluster 1 and Cluster 2 was calculated. In all individuals, a significant positive correlation was observed between naïve B cells and native CD4<sup>+</sup> T cells, whereas the remaining cell types primarily showed negative correlations (Fig. 9C). In Cluster 1, which had a poor prognosis, naïve B cells exhibited a positive correlation with native CD4<sup>+</sup> T cells. Moreover, M1 macrophages displayed a positive correlation with plasma cells,

follicular helper T cells, and activated mast cells. In contrast, M2 macrophages showed a negative correlation with plasma cells, follicular helper T cells, M1 macrophages, and resting dendritic cells. Additionally, eosinophils exhibited a negative correlation with active NK cells (Fig. 9D). In Cluster 2, the correlation between immune cell contents was like that observed in all patients (Fig. 9E).

Finally, the correlations between 24 immDEGs and immune cell types were calculated in all patients (Fig. 9F), individuals in Cluster 1 (Fig. 9G), and individuals in Cluster 2 (Fig. 9H). Follicular helper T cells, macrophages, and activated mast cells had a remarkable association with immDEGs.

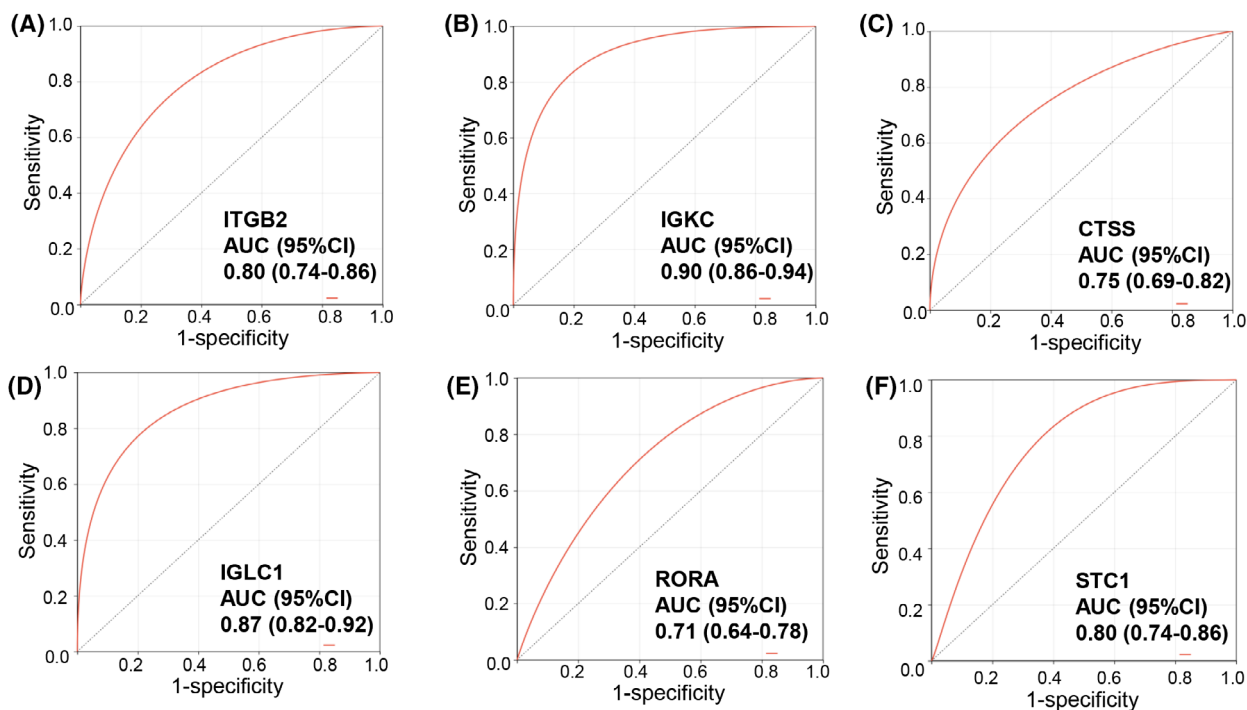


**Fig. 4.** Correlations between gene expression levels between six immDEGs in all samples (A), TGCT sample (B), and normal samples (C).





**Fig. 5.** Identification of immune molecular subtypes. (A–C) Two immune molecular subtypes were obtained using CONSENSUSCLUSTERPLUS. (D) DEGs between immune molecular subtypes visualized by volcano plot. (E) A heatmap depicting the top 50 genes identified through analysis of differential gene expression. (F) The heatmap of 24 immDEGs. (G) The expression histogram of 24 immDEGs in two immune molecular subtypes (C1,  $n = 108$ ; C2,  $n = 106$ ). (H) Kaplan–Meier survival analysis of the two immune molecular subtypes. Presentation of data as means  $\pm$  standard deviation. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\*\* $P < 0.0001$ ; rank-sum test.



**Fig. 6.** ROC curves of six characterized genes demonstrating their ability to distinguish between Cluster 1 and Cluster 2 independently. ROC curves of ITGB2 (A), IGKC (B), CTSS (C), IGLC1 (D), RORA (E), and STC1 (F).

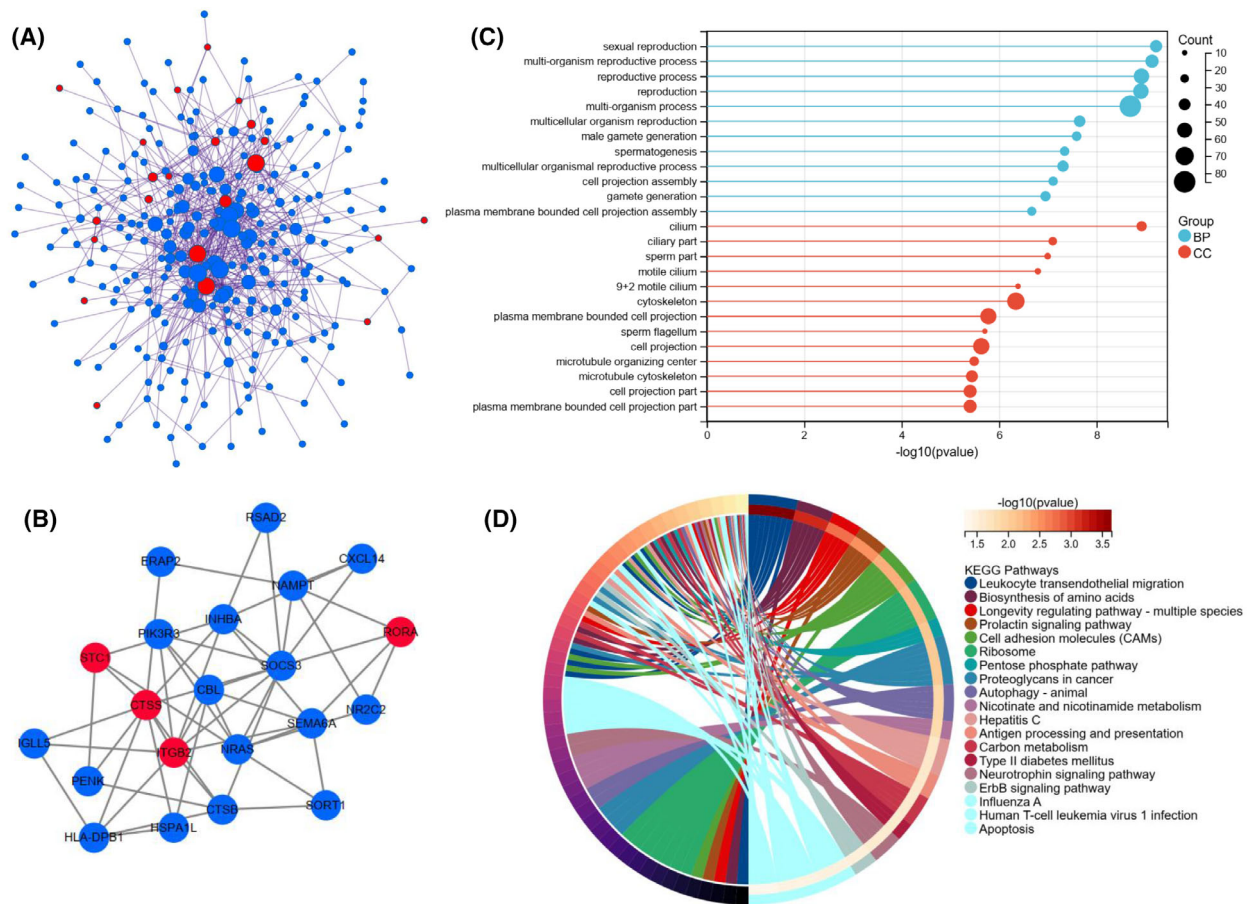
### Validation of the immDEGs risk signature in specific pathology types

TGCT is the most common type of testicular cancer. Seminoma is the most common histologic subtype of TGCT in young men. Mixed type of TGCT is the most common type of nonseminoma. Pure embryonal carcinoma, teratoma, and yolk sac tumor are rare. Therefore, we validated the validity of the six-immDEGs risk signature in pure seminoma or mixed type of TGCT. The 24 immDEGs expression levels were significantly different between normal samples and seminoma or mixed type of TGCT (Fig. 10A,B). All seminoma samples belonged to the Cluster 2 subtype. Cluster 2 subtype possessed a favorable OS compared with Cluster 1. This is consistent with a better prognosis in patients with seminoma. Due to the lack of survival data in seminoma samples, we are unable to assess the prognostic value of the risk signature in patients with

seminoma. For patients with mixed type of TGCT, KM survival curves revealed that the individuals in the high-risk group possessed a poor OS (HR, 2.82; 95% CI, 1.17–6.82;  $P = 0.02$ ; Fig. 10C). The 48 mixed type of TGCT samples belonged to Cluster 1, and 42 mixed type of TGCT samples belonged to Cluster 1. KM survival curves revealed that Cluster 1 possessed a poor OS (HR, 2.68; 95% CI, 1.06–6.75;  $P = 0.03$ ; Fig. 10D). The AUC of Time-dependent ROC at different time points were 0.79 (3 years) and 0.73 (5 years; Fig. 10E). This risk signature could serve as a predictor of survival for patients with mixed type of TGCT.

### Validation of STC1, RORA, and IGKC

Through multivariate Cox regression analysis of the 24 immDEGs, IGKC, STC1, and RORA were identified as being considerably linked to the patient OS (Fig. 11A). GSEA was conducted between different



**Fig. 7.** PPI network and functional analysis of DEGs between TGCT and normal samples. (A) DGEs PPI network: red nodes indicate immDEGs. (B) ImmDEGs PPI network: red nodes indicate characterized genes of risk signature. (C) GO functional enrichment analysis. (D) KEGG functional enrichment analysis.

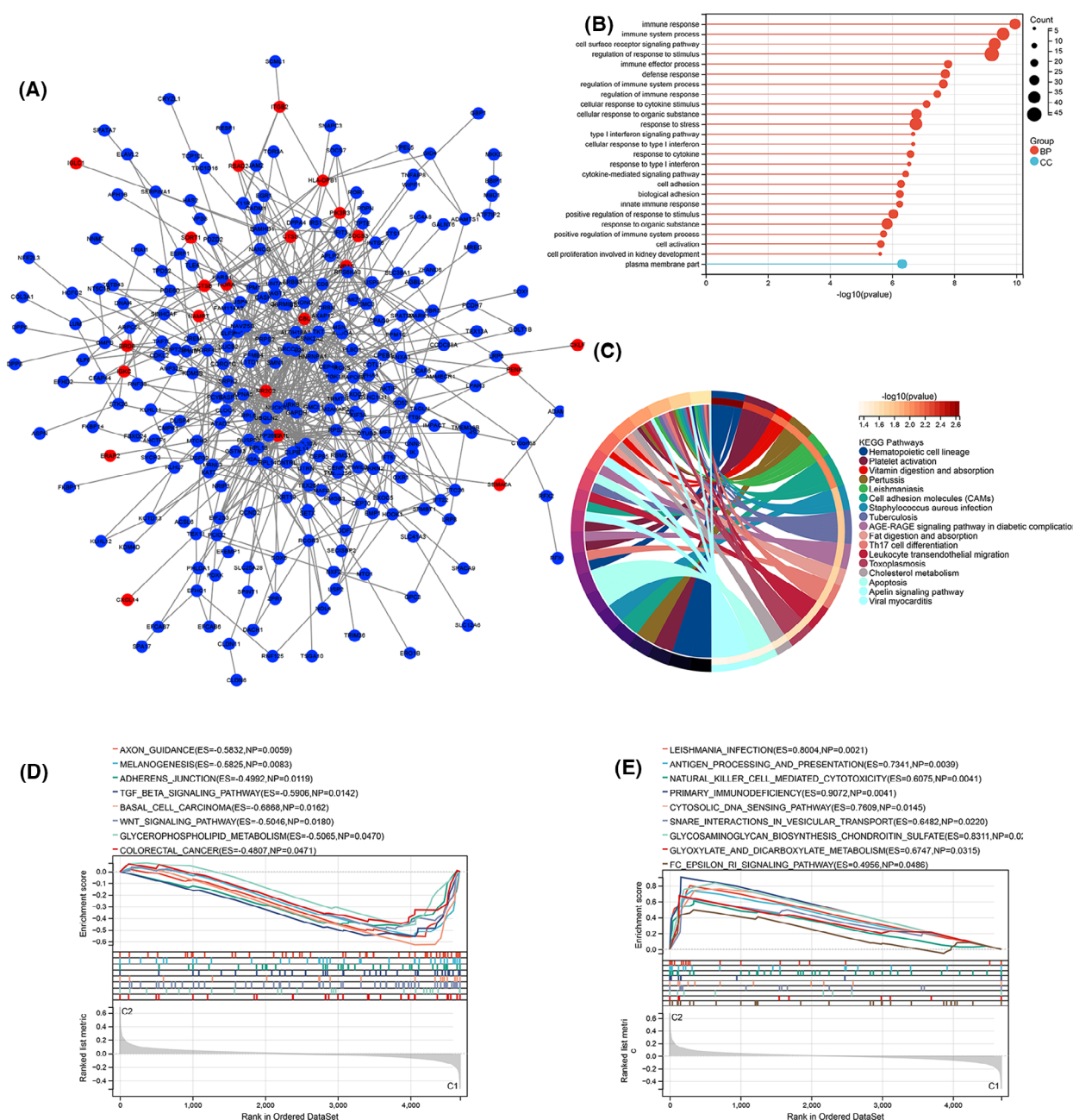
gene expression groups that were significantly different in terms of BPs. The high-*RORA* group showed inhibition of NOTCH signaling pathways and type-1 diabetes mellitus, while the VEGF signaling pathway and axon guidance were activated in this group (Fig. 11B). Numerous tumor-associated pathways were activated in the high-*STC1* group (Fig. 11C–F). Numerous tumor-associated pathways were also activated in the low-*IGKC* group, and only basal cell carcinoma, lysine degradation, and glycerophospholipid metabolism were activated in the high-*IGKC* group (Fig. 11G–I).

*STC1* was the only high-risk factor for OS (HR, 1.80; 95% CI, 1.32–2.45;  $P < 0.001$ ). *STC1* in TGCT was further characterized through bioinformatics and IHC experiments. First, the expression of *STC1* was evaluated in TGCT datasets (GSE3218, TCGA). As demonstrated in Fig. 12A,B, *STC1* expression levels were elevated in TGCT tissues compared with its corresponding normal tissues. Moreover, *STC1* was found to be upregulated in 32 out of 66 tumors in TCGA

(Fig. S4). Subsequently, the correlation between *STC1* expression and clinicopathological features was examined in TGCT patients from TCGA. *STC1* expression was found to be correlated with clinical T, metastasis, and tumor stage. (Fig. 12C–F). Finally, the expression of *STC1* was examined in clinical tissues by means of IHC. Results from IHC analysis of 24 seminoma tissues and eight normal tissues revealed a substantial increase in *STC1* expression levels in TGCT samples (Fig. 12G–I).

## Discussion

Although testicular cancer is a curable tumor, it is also associated with primary resistance, disease progression after therapy, and adverse effects of treatment. Identifying underlying molecular mechanisms and markers of poor prognosis is important for this subset of patients. The TME has been regarded as a critical factor in influencing tumorigenesis and progression

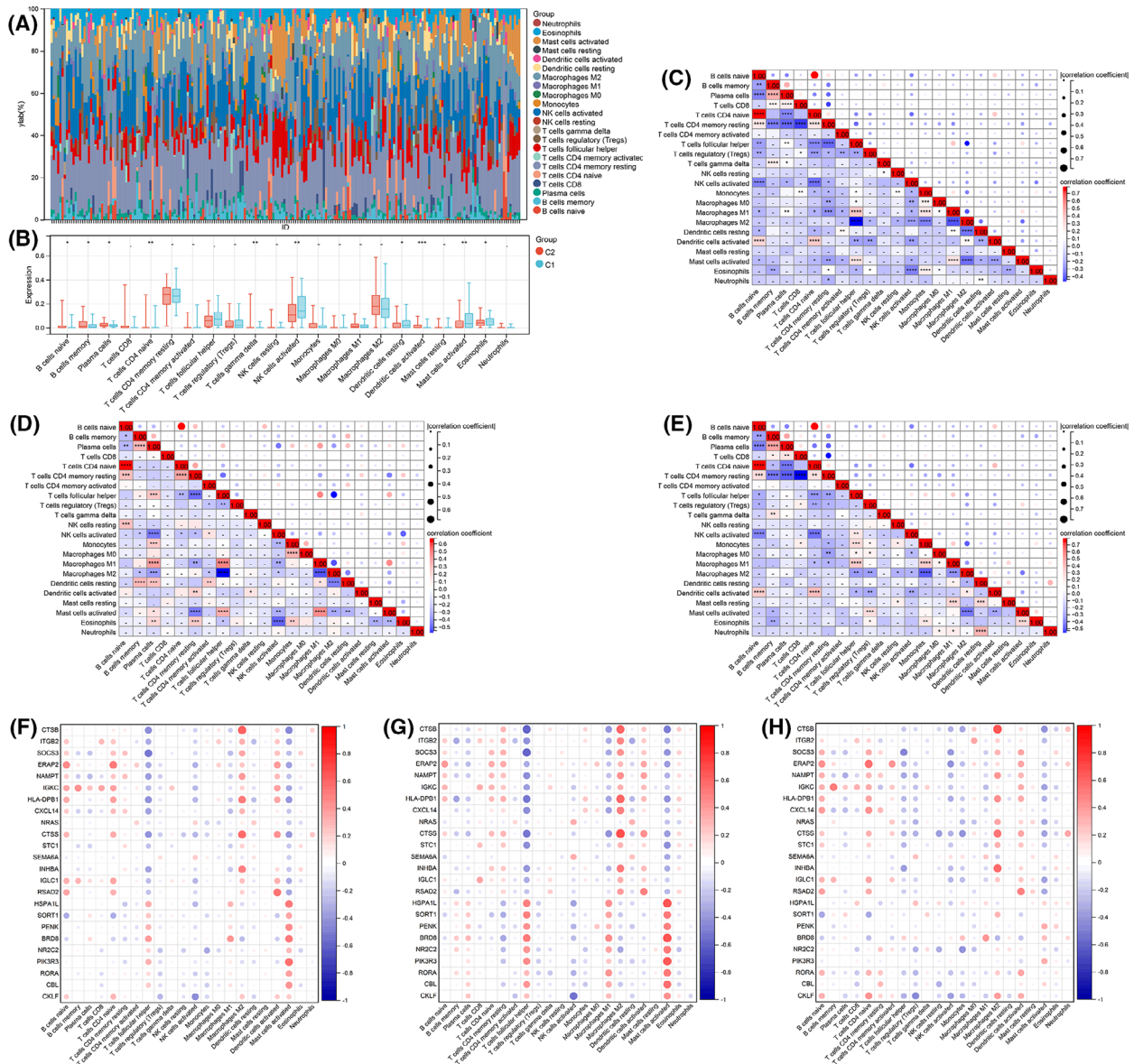


**Fig. 8.** PPI network and functional analysis of DEGs and GSEA analysis between two different immune patterns. (A) DGEs PPI network: red nodes indicate immDEGs. (B) GO functional enrichment analysis. (C) KEGG functional enrichment analysis. Activated (D) and inhibited (E) BPs in Cluster 1 compared with Cluster 2.

[24,25]. In TME, immune cells constitute a significant proportion of nontumor cells. But a systematic understanding of the immune milieu in testicular cancer is currently lacking. Hence, exploring patterns of immune infiltration in testicular cancer and finding novel targets have important clinical benefits for early diagnosis and personalized therapy [26–28].

This study obtained 202 TCGT samples and 12 normal samples to identify immDEGs. A risk signature consisting of six immDEGs was constructed for TGCT for the first time. This signature was used to identify two immune molecular subtypes, namely Cluster 1 and Cluster 2. Prognostic differences were found to be significant between these two immune subtypes.



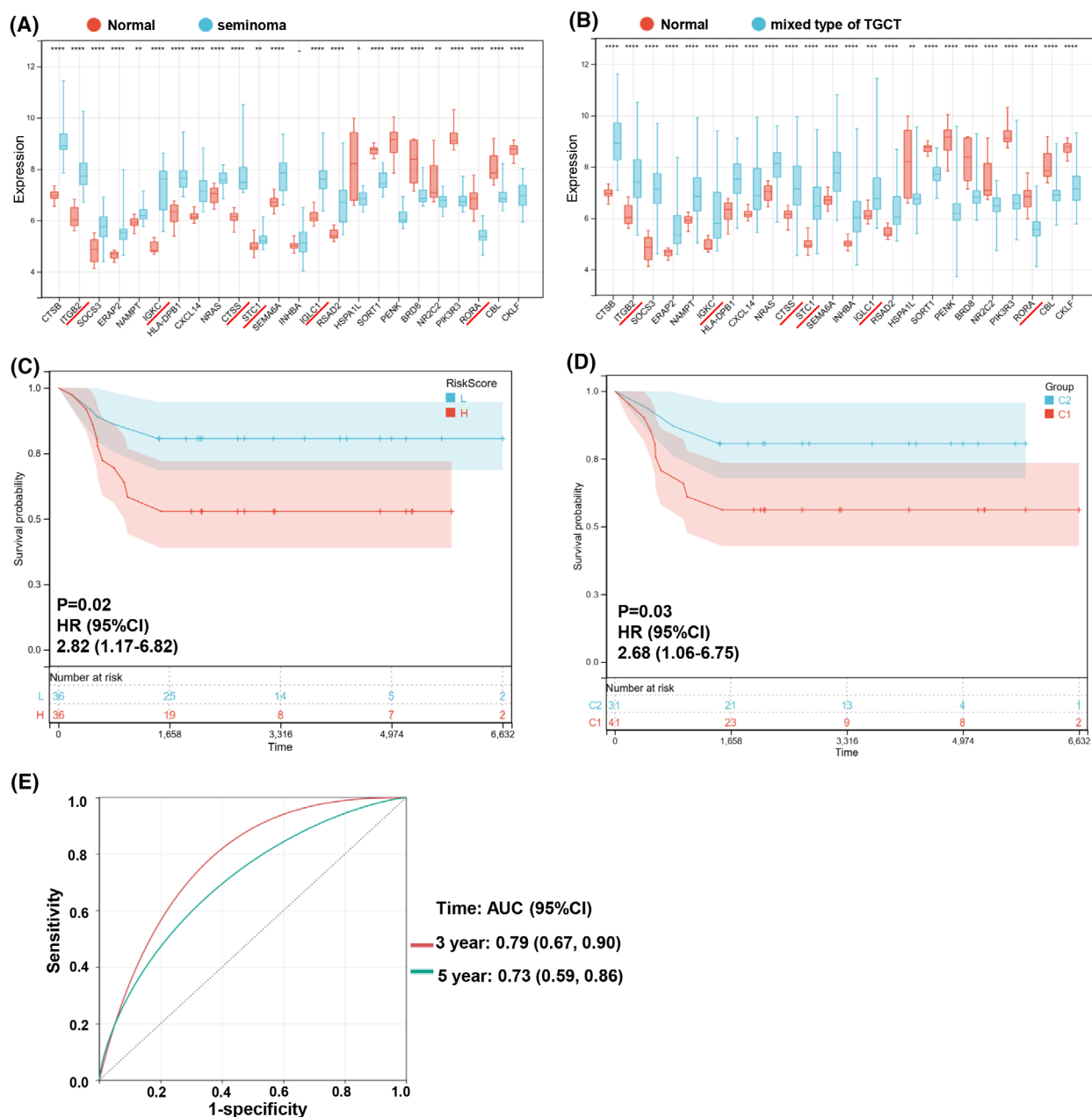


**Fig. 9.** Immune characteristics between two different immune patterns. (A) Immune cell content stacking plot for each sample. (B) Twenty-two immune cell types content histogram in two different immune patterns. Correlations between 22 immune cells in all patients (C), Clusters 1 patient (D), and Cluster 2 patients (E). Correlations between 24 immDEGs and immune cell types in all individuals (F), Cluster 1 patients (G), and Cluster 2 patients (H). Presentation of data as means  $\pm$  standard deviation. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ; rank-sum test.

A total of 357 DEGs (tumor vs. normal) and 88 DEGs (Cluster 1 vs. Cluster 2) were screened, respectively. The 24 immDEGs were identified in TCGT samples compared with normal samples. In the PPI network of two sets of DEGs, the 24 immDEGs were not closely linked and scattered throughout the network. This suggested that their functions might be relatively independent. In GO and KEGG analyses, DEGs between tumor and normal were mainly enriched in the

reproductive process. Studies have shown that TGCT can disrupt the hypothalamus-pituitary-gonadal axis and lead to obstruction of sperm production, which can be manifested as decreased sperm quality or azoospermia [29–31]. DEGs between the two immune patterns were enriched in various immune processes. Through GSEA, 16 BPs may affect the prognosis of TGCT. The pathways that have been extensively studied in different types of cancer, such as the WNT signaling pathway

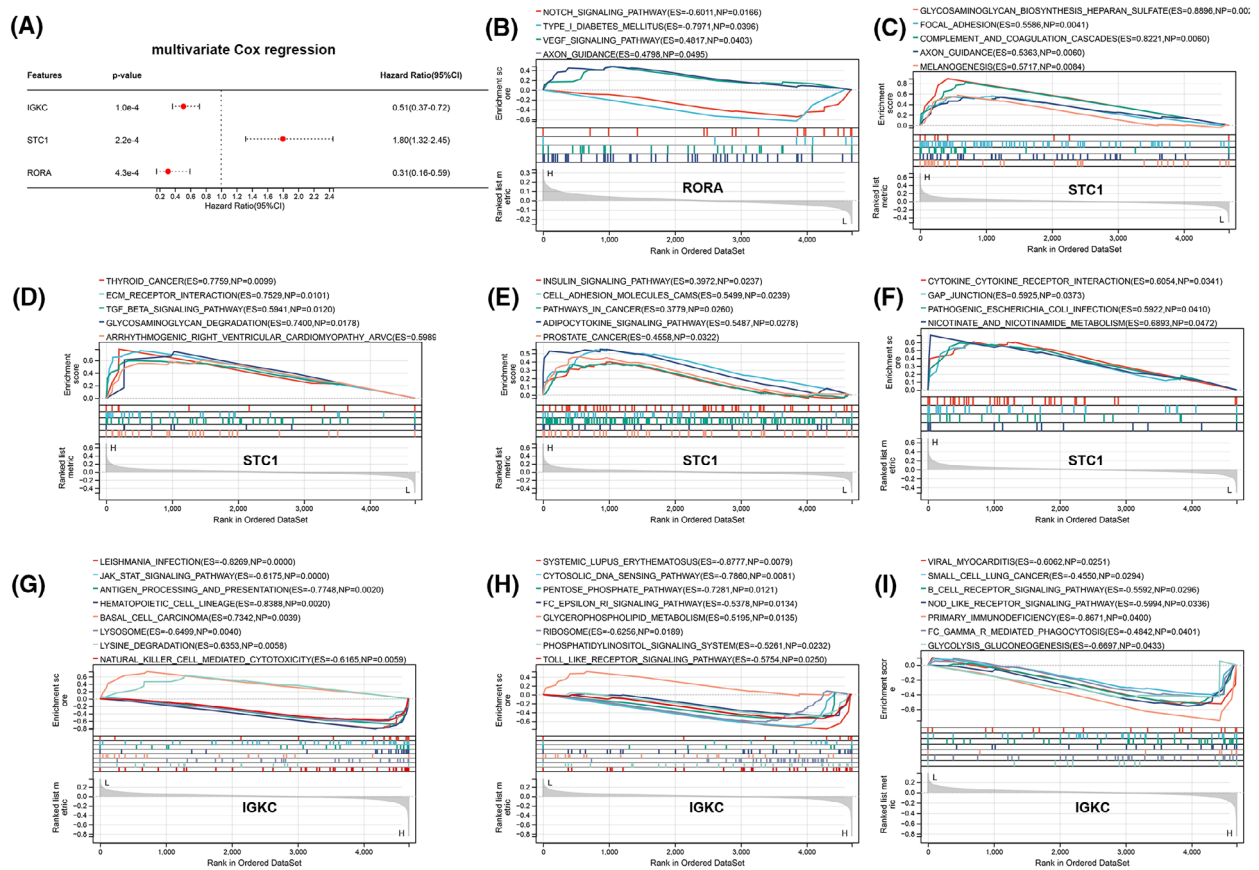




**Fig. 10.** (A) Expression histogram of 24 immDEGs in seminoma ( $n = 26$ ) and normal tissues ( $n = 12$ ). (B) The expression histogram of 24 immDEGs in mixed type of TGCT ( $n = 90$ ) and normal tissues ( $n = 12$ ). (C) Kaplan–Meier survival analysis of the two risk groups, mixed type of TGCT samples were divided into high- and low-risk groups based on their median risk score. (D) Kaplan–Meier survival analysis of the two immune molecular subtypes in mixed type of TGCT samples. (E) Time-dependent ROC curves verified the predictive efficacy of the risk signature at 3 and 5 years in mixed type of TGCT samples. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\*\* $P < 0.0001$ ; rank-sum test.

and TGF- $\beta$  signaling pathway, were also included. However, the studies of these two signaling pathways in testicular function and testicular cancer are limited. Young *et al.* [32] reported that the WNT pathway plays a role in normal spermatogenesis and that inhibiting canonical WNT signaling can attenuate the

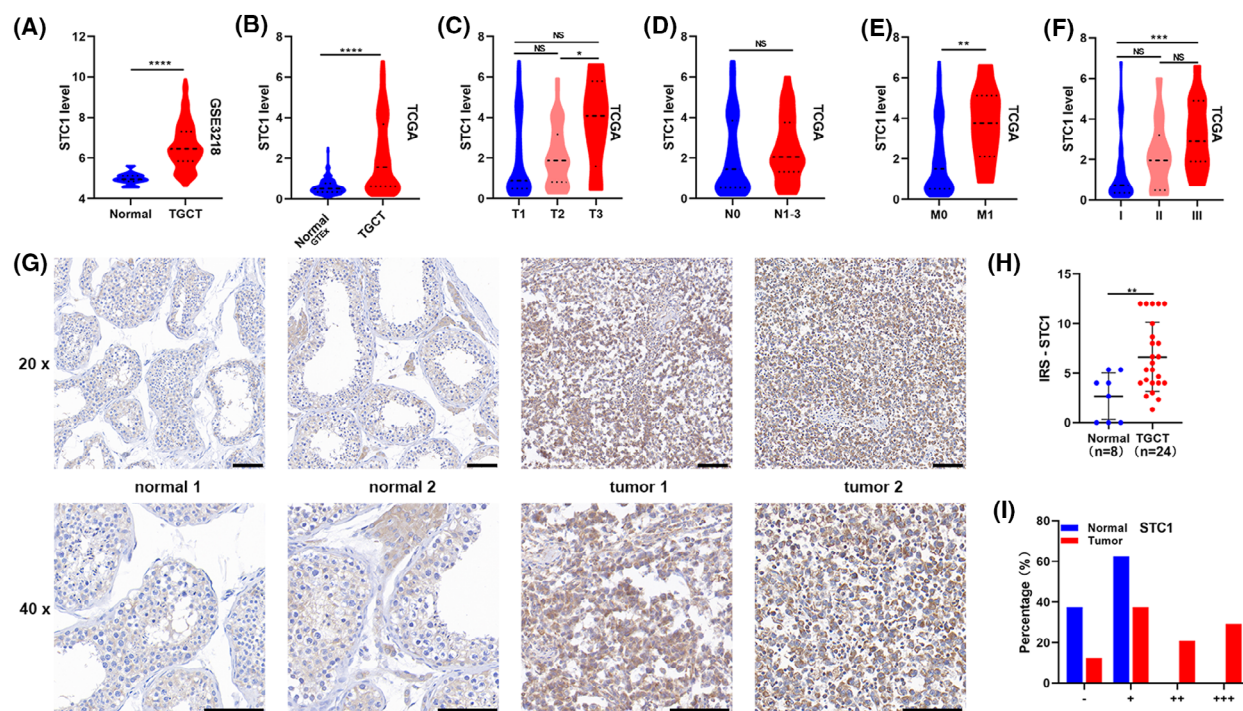
proliferation and migration ability of seminoma cells. *TCF7L1*, a WNT suppressor, can sensitively distinguish TGCT from nonseminomatous germ cell tumors [33]. The TGF- $\beta$  signaling pathway is essential for testis formation and can impact processes involved in testicular pathologies, including testicular cancer [34,35]. Immune



**Fig. 11.** GSEA analysis for RORA, STC1, and IGKC. (A) Identification of three immDEGs related to OS of TGCT via multivariate Cox regression. (B) GSEA analysis for RORA. (C–F) GSEA analysis for STC1. (G–I) GSEA analysis for IGKC.

pathways mainly include antigen processing and presentation, natural killer (NK) cell-mediated cytotoxicity, and primary immunodeficiency. Immunotherapy has revolutionized cancer treatment. The process of antigen presentation is the initial stage of the immune response and is critical for mounting an effective antitumor response [36]. It is well known that PDL1 expressed in tumor cell membranes binding to immune cell PD1 suppresses antitumor immunity [37]. The inhibitors of PD1/PDL1 have shown clinical efficacy in many tumors. TGCT has been the subject of various studies that indicate an elevation in PDL1 expression levels, and PD1/PDL1 is a potential therapeutic target of TGCT [9–13]. Cytotoxic T and NK cells are the main mediators of cytotoxicity against tumors [38]. By the directed release of lytic granules or by inducing apoptosis mediated by death receptors, NK cells can rapidly and directly kill tumor cells. Therefore, PD1/PDL1 inhibitors and NK cells have broad application prospects in TGCT, especially in patients who are resistant to conventional treatments or have relapsed.

A total of eight immDEGs (*IGKC*, *IGLC1*, *STC1*, *ITGB2*, *CTSS*, *RORA*, *HLA-DPB1*, and *RSAD2*) were found to be related to OS of patients of TGCT via univariate Cox regression. However, the expressions of *IGKC*, *IGLC1*, *ITGB2*, *CTSS*, *HLA-DPB1*, and *RSAD2* were elevated in tumor tissue and they served as protective factors. Such conflicting expression levels and prognostic value are not uncommon in bioinformatics analysis of tumors. I think this can be explained by the following reasons. First, unlike single-cell sequencing, in this study, we obtained the average transcriptome data of a population of cells, not just tumor cells. The changes in gene expression may be due to nontumor cells. Or the gene exerts biological effects through nontumor cell pathways. For example, Cao *et al.* [39] reported that high expression of *CXCL11* in colon cancer could improve the prognosis of patients by promoting antitumor immunity. Second, we measured the mRNA expression levels of genes only at one point in time. The level of gene transcription is regulated by many factors



**Fig. 12.** Elevated expression levels of STC1 positively correlated with advanced clinicopathological features and indicated poor prognosis. (A, B) The expression levels of STC1 increased in TGCT in *GSE3218* and TCGA. (C–F) STC1 expression levels in TGCT tissues with different tumor stages, lymph node metastasis status, metastasis status, and grade status. (G) STC1 expressions in 32 clinical samples were detected by IHC. Scale bars: 100  $\mu$ m. (H) Quantification of STC1 protein expression in TGCT tissues and normal tissues. (I) The proportion of clinical tissues with negative (–), weak (+), moderate (++), and strong (+++) STC1 staining intensity. Presentation of data as means  $\pm$  standard deviation. <sup>NS</sup> $P > 0.05$ ; \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ; \*\*\*\* $P < 0.0001$ ; rank-sum test. T, tumor stage; N, lymph node metastasis; M, metastasis.

and changes rapidly. Third, mRNA is subject to post-transcriptional regulation. The mRNA expression levels are not equivalent to protein expression levels. Proteins are the main bodies that perform biological functions. The protein expression levels of genes need to be further verified.

Except for *HLA-DPB1*, the remaining seven genes have not been studied in testicular cancer. According to Gotovac *et al.* [40], there was an observed increase in the frequency of the *HLA-DPB1*\*1701 allele in individuals diagnosed with TGCT, suggesting that the HLA region might have a function in the onset and progression of this tumor. Then, three genes (*IGKC*, *STC1*, and *RORA*) related to the OS of patients were identified via multivariate Cox regression by 24 immDEGs. In breast carcinoma, non-small-cell lung carcinoma, and colorectal carcinoma, *IGKC* has been linked to a favorable prognosis and may be utilized as a compatible prognostic indicator in human solid tumors [41,42]. However, *IGKC* was significantly higher in ovarian cancer and clear cell renal cell carcinoma than in normal tissues [43,44]. *RORA*, one of the

circadian genes, inhibited tumorigenesis and progression in various tumors, including breast cancer [44], prostate cancer [45,46], lung carcinoma [45,47], endometrial cancer [48], and gastric cancer [49]. Only *STC1* behaves as an oncogene in our results. Stanniocalcin-1 (*STC1*) promoted different types of cancer progression [50–52]. According to Lin *et al.* [53], *STC1* is involved in immune evasion and resistance to immunotherapy and is negatively correlated with patient survival in various types of cancer. The expression level of *STC1* in clinical tissue samples was examined and confirmed to be elevated in TGCT. GSEA revealed that numerous tumor-associated pathways were activated in the high-*STC1* group. The mechanism by which *STC1* promotes the initiation and progression of TGCT is still unknown and requires further study.

A novel immunophenotyping of TGCT was presented. Based on six signature immDEGs, two immune molecular subtypes (Cluster 1 and Cluster 2) were identified, and Cluster 1 possessed an unfavorable OS compared with Cluster 2. Traditional histological classification has certain limitations, especially for high-

grade tumors, chemotherapy-resistant tumors, and recurrent tumors. In patients with the same histological classification, there may be large differences in the efficacy of the same treatment. The fast advancement of molecular technologies, such as sequencing, has led to an increasing application of molecular subtyping in various types of tumors [54–56]. Molecular subtyping is a promising way to provide precise personalized treatment and avoid unnecessary toxicities. There are 10 immune cell types that are differently expressed in the two molecular subtypes. Cluster 1 exhibited a significant reduction in T and B cells compared with Cluster 2. Combined with the pathway analysis, the worse prognosis of Cluster 1 might be the result of decreased ability of antigen presentation and the failure of activation of T cells and B cells leads to inhibiting antitumor immunity.

Seminoma is the most common histologic subtype of TGCT in young men. Mixed type of TGCT is the most common type of nonseminoma. Pure embryonal carcinoma, teratoma, and yolk sac tumor are rare. Medvedev *et al.* and Savelyeva *et al.* [57,58] identified two distinct seminoma subtypes in 64 pure seminoma samples from TCGA. They found that seminoma subtype 2 shows signs of differentiation into nonseminoma TGCT and may have higher resistance to platinum-based chemotherapy, higher immune score, and overexpression of 21 genes related to senescence-associated secretory phenotype. We validated the validity of the six-immDEGs risk signature in pure seminoma or mixed type of TGCT. All seminoma samples belonged to the Cluster 2 subtype. Cluster 2 subtype possessed a favorable OS compared with Cluster 1. This is consistent with a better prognosis in patients with seminoma. This six-immDEGs risk signature could serve as a reliable predictor of survival for patients with mixed type of TGCT.

Although we got some meaningful results, there are some inherent limitations in our study. First, the study was based on single-center data. This six-immDEGs risk signature needs to be validated in more clinical cohorts. Second, we used a dataset that included seminoma, mixed type of TGCT, embryonal carcinoma, teratoma, and yolk sac tumor. Although we validated the six-immDEG risk marker in pure seminoma and mixed type of TGCT, a different risk model might be obtained if one of the TGCT types were analyzed alone. Third, the genes identified in this study need to be experimentally verified to explore their mechanisms in TGCT.

To summarize, bioinformatics analyses were conducted to compare immune infiltration in the TGCT and normal samples. A six-immDEGs risk signature was constructed, and two immune molecular subtypes

with significant prognostic differences were identified. WNT signaling pathway, TGF- $\beta$  signaling pathway, antigen processing and presentation, and NK cell-mediated cytotoxicity were linked to TGCT. The expression level of STC1 was found to be elevated in TCGA, GSE3218, and clinical tissues, and its high expression level was linked to advanced clinicopathological features and poor prognosis. These findings contribute to a better understanding of TGCT.

## Acknowledgements

This work was supported by Guangdong Basic and Applied Basic Research Foundation (2020A1515110946), PhD Research Foundation of Affiliated Hospital of Jining Medical University (2022-BS-001), Research Fund for Academician LinHe New Medicine (JYHL2022FMS09), Jining Key Research and Development Foundation (2021YXNS055, 2022YXN006), and Shandong Province Medical and Health Science and Technology Development Plan Research Project (202204050782). We thank Bullet Edits for language editing service.

## Conflict of interest

The authors declare no conflict of interest.

## Data accessibility

The data that support the findings of this study are openly available in GEO and TCGA databases.

## Author contributions

ZZ contributed to writing—original draft, funding acquisition, and resources. XX contributed to conceptualization and writing—review and editing. XW and MW contributed to methodology, software, and formal analysis. CM contributed to resources and writing—review and editing. ZL contributed to conceptualization, software, funding acquisition, and writing—review and editing.

## References

- Cheng L, Albers P, Berney DM, Feldman DR, Daugaard G, Gilligan T and Looijenga LHJ (2018) Testicular cancer. *Nat Rev Dis Primers* **4**, 29.
- Siegel RL, Miller KD, Fuchs HE and Jemal A (2022) Cancer statistics, 2022. *CA Cancer J Clin* **72**, 7–33.
- Smith ZL, Werntz RP and Eggener SE (2018) Testicular cancer: epidemiology, diagnosis, and management. *Med Clin North Am* **102**, 251–264.



- 4 Kalavaska K, Conteduca V, De Giorgi U and Mego M (2018) Molecular mechanisms of resistance in testicular germ cell tumors – clinical implications. *Curr Cancer Drug Targets* **18**, 967–978.
- 5 O’Shaughnessy MJ, Feldman DR, Carver BS and Sheinfeld J (2015) Late relapse of testicular germ cell tumors. *Urol Clin North Am* **42**, 359–368.
- 6 Galon J and Bruni D (2019) Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. *Nat Rev Drug Discov* **18**, 197–218.
- 7 Lei X, Lei Y, Li JK, Du WX, Li RG, Yang J, Li J, Li F and Tan HB (2020) Immune cells within the tumor microenvironment: biological functions and roles in cancer immunotherapy. *Cancer Lett* **470**, 126–133.
- 8 Yakirevich E, Lefel O, Sova Y, Stein A, Cohen O, Izhak OB and Resnick MB (2002) Activated status of tumour-infiltrating lymphocytes and apoptosis in testicular seminoma. *J Pathol* **196**, 67–75.
- 9 Fankhauser CD, Curioni-Fontecedro A, Allmann V, Beyer J, Tischler V, Sulser T, Moch H and Bode PK (2015) Frequent PD-L1 expression in testicular germ cell tumors. *Br J Cancer* **113**, 411–413.
- 10 Cierna Z, Mego M, Miskovska V, Machalekova K, Chovanec M, Svetlovska D, Hainova K, Rejlekova K, Macak D, Spanik S *et al.* (2016) Prognostic value of programmed-death-1 receptor (PD-1) and its ligand 1 (PD-L1) in testicular germ cell tumors. *Ann Oncol* **27**, 300–305.
- 11 Shah S, Ward JE, Bao R, Hall CR, Brockstein BE and Luke JJ (2016) Clinical response of a patient to anti-PD-1 immunotherapy and the immune landscape of testicular germ cell tumors. *Cancer Immunol Res* **4**, 903–909.
- 12 Siska PJ, Johnpulle RAN, Zhou A, Bordeaux J, Kim JY, Dabbas B, Dakappagari N, Rathmell JC, Rathmell WK, Morgans AK *et al.* (2017) Deep exploration of the immune infiltrate and outcome prediction in testicular cancer by quantitative multiplexed immunohistochemistry and gene expression profiling. *Onco Targets Ther* **6**, e1305535.
- 13 Lobo J, Rodrigues Á, Guimarães R, Cantante M, Lopes P, Maurício J, Oliveira J, Jerónimo C and Henrique R (2019) Detailed characterization of immune cell infiltrate and expression of immune checkpoint molecules PD-L1/CTLA-4 and MMR proteins in testicular germ cell tumors disclose novel disease biomarkers. *Cancer* **11**, 1535.
- 14 Korkola JE, Houldsworth J, Chadalavada RS, Olshen AB, Dobrzynski D, Reuter VE, Bosl GJ and Chaganti RS (2006) Down-regulation of stem cell genes, including those in a 200-kb gene cluster at 12p13.31, is associated with in vivo differentiation of human male germ cell tumors. *Cancer Res* **66**, 820–827.
- 15 Johnson WE, Li C and Rabinovic A (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118–127.
- 16 Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P and Mesirov JP (2011) Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740.
- 17 Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W and Smyth GK (2015) limma powers differential expression analyses for RNA-seq and microarray studies. *Nucleic Acids Res* **43**, e47.
- 18 Engebretsen S and Bohlin J (2019) Statistical predictions with glmnet. *Clin Epigenetics* **11**, 123.
- 19 Wilkerson MD and Hayes DN (2010) ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* **26**, 1572–1573.
- 20 Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P *et al.* (2019) STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* **47**, D607–D613.
- 21 Yu G, Wang LG, Han Y and He QY (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287.
- 22 Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES *et al.* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* **102**, 15545–15550.
- 23 Rusk N (2019) Expanded CIBERSORTx. *Nat Methods* **16**, 577.
- 24 Quail DF and Joyce JA (2013) Microenvironmental regulation of tumor progression and metastasis. *Nat Med* **19**, 1423–1437.
- 25 Bader JE, Voss K and Rathmell JC (2020) Targeting metabolism to improve the tumor microenvironment for cancer immunotherapy. *Mol Cell* **78**, 1019–1033.
- 26 Christofides A, Strauss L, Yeo A, Cao C, Charest A and Boussiotis VA (2022) The complex role of tumor-infiltrating macrophages. *Nat Immunol* **23**, 1148–1156.
- 27 Khalaf K, Hana D, Chou JT, Singh C, Mackiewicz A and Kaczmarek M (2021) Aspects of the tumor microenvironment involved in immune resistance and drug resistance. *Front Immunol* **12**, 656364.
- 28 Wang G, Wang D, Sun M, Liu X and Yang Q (2020) Identification of prognostic and immune-related gene signatures in the tumor microenvironment of endometrial cancer. *Int Immunopharmacol* **88**, 106931.
- 29 Morrish DW, Venner PM, Siy O, Barron G, Bhardwaj D and Outhet D (1990) Mechanisms of endocrine

- dysfunction in patients with testicular cancer. *J Natl Cancer Inst* **82**, 412–418.
- 30 Petersen PM, Skakkebaek NE, Vistisen K, Rørth M and Giwercman A (1999) Semen quality and reproductive hormones before orchiectomy in men with testicular cancer. *J Clin Oncol* **17**, 941–947.
  - 31 Dias TR, Agarwal A, Pushparaj PN, Ahmad G and Sharma R (2020) Reduced semen quality in patients with testicular cancer seminoma is associated with alterations in the expression of sperm proteins. *Asian J Androl* **22**, 88–93.
  - 32 Young JC, Kerr G, Micati D, Nielsen JE, Rajpert-De Meyts E, Abud HE and Loveland KL (2020) WNT signalling in the normal human adult testis and in male germ cell neoplasms. *Hum Reprod* **35**, 1991–2003.
  - 33 Bu L, Yang Q, McMahan L, Xiao GQ and Li F (2019) Wnt suppressor and stem cell regulator TCF7L1 is a sensitive immunohistochemical marker to differentiate testicular seminoma from non-seminomatous germ cell tumor. *Exp Mol Pathol* **110**, 104293.
  - 34 Young JC, Wakitani S and Loveland KL (2015) TGF- $\beta$  superfamily signaling in testis formation and early male germline development. *Semin Cell Dev Biol* **45**, 94–103.
  - 35 Dias VL, Rajpert-De Meyts E, McLachlan R and Loveland KL (2009) Analysis of activin/TGF $\beta$ -signaling modulators within the normal and dysfunctional adult human testis reveals evidence of altered signaling capacity in a subset of seminomas. *Reproduction* **138**, 801–811.
  - 36 Pishesha N, Harmand TJ and Ploegh HL (2022) A guide to antigen processing and presentation. *Nat Rev Immunol* **22**, 751–764.
  - 37 Kornepati AVR, Vadlamudi RK and Curiel TJ (2022) Programmed death ligand 1 signals in cancer cells. *Nat Rev Cancer* **22**, 174–189.
  - 38 Prager I and Watzl C (2019) Mechanisms of natural killer cell-mediated cellular cytotoxicity. *J Leukoc Biol* **105**, 1319–1329.
  - 39 Cao Y, Jiao N, Sun T, Ma Y, Zhang X, Chen H, Hong J and Zhang Y (2021) CXCL11 correlates with antitumor immunity and an improved prognosis in colon cancer. *Front Cell Dev Biol* **9**, 646252.
  - 40 Gotovac K, Grubic Z, Kastelan Z, Stingl K, Kulis T, Krhen I, Hudolin T, Kastelan M and Zunec R (2011) Pilot study of the association between the HLA region and testicular carcinoma among Croatian patients. *Urol Int* **87**, 288–292.
  - 41 Schmidt M, Hellwig B, Hammad S, Othman A, Lohr M, Chen Z, Boehm D, Gebhard S, Petry I, Lebrecht A *et al.* (2012) A comprehensive analysis of human gene expression profiles identifies stromal immunoglobulin  $\kappa$  C as a compatible prognostic marker in human solid tumors. *Clin Cancer Res* **18**, 2695–2703.
  - 42 Schmidt M, Edlund K, Hengstler JG, Heimes AS, Almstedt K, Lebrecht A, Krajnak S, Battista MJ, Brenner W, Hasenburger A *et al.* (2021) Prognostic impact of immunoglobulin kappa C (IGKC) in early breast cancer. *Cancer* **13**, 3626.
  - 43 Lundgren S, Berntsson J, Nodin B, Micke P and Jirstrom K (2016) Prognostic impact of tumour-associated B cells and plasma cells in epithelial ovarian cancer. *J Ovarian Res* **9**, 21.
  - 44 Zhou J, Yang Z, Wu X, Zhang J, Zhai W and Chen Y (2021) Identification of genes that correlate clear cell renal cell carcinoma and obesity and exhibit potential prognostic value. *Transl Androl Urol* **10**, 680–691.
  - 45 Mocellin S, Tropea S, Benna C and Rossi CR (2018) Circadian pathway genetic variation and cancer risk: evidence from genome-wide association studies. *BMC Med* **16**, 20.
  - 46 Li Y, He J, Yu L, Yang Q, Du J, Chen Y and Tang W (2022) Hsa-miR-1290 is associated with stemness and invasiveness in prostate cancer cell lines by targeting RORA. *Andrologia* **54**, e14396.
  - 47 Xian H, Li Y, Zou B, Chen Y, Yin H, Li X and Pan Y (2022) Identification of TIMELESS and RORA as key clock molecules of non-small cell lung cancer and the comprehensive analysis. *BMC Cancer* **22**, 107.
  - 48 Sun X, Dongol S, Qiu C, Xu Y, Sun C, Zhang Z, Yang X, Zhang Q and Kong B (2018) miR-652 promotes tumor proliferation and metastasis by targeting RORA in endometrial cancer. *Mol Cancer Res* **16**, 1927–1939.
  - 49 Ma X, Chen H, Li L, Yang F, Wu C and Tao K (2021) CircGSK3B promotes RORA expression and suppresses gastric cancer progression through the prevention of EZH2 trans-inhibition. *J Exp Clin Cancer Res* **40**, 330.
  - 50 Sun J, Wei X, You J, Yue W, Ouyang J, Ling Z and Hou J (2021) STC1 is a novel biomarker associated with immune characteristics and prognosis of bladder cancer. *Int J Gen Med* **14**, 5505–5516.
  - 51 Peña C, Céspedes MV, Lindh MB, Kiflemariam S, Mezheyeuski A, Edqvist PH, Hägglöf C, Birgisson H, Bojmar L, Jirstrom K *et al.* (2013) STC1 expression by cancer-associated fibroblasts drives metastasis of colorectal cancer. *Cancer Res* **73**, 1287–1297.
  - 52 Sakata J, Sasayama T, Tanaka K, Nagashima H, Nakada M, Tanaka H, Hashimoto N, Kagawa N, Kinoshita M, Nakamizo S *et al.* (2019) MicroRNA regulating stanniocalcin-1 is a metastasis and dissemination promoting factor in glioblastoma. *J Neurooncol* **142**, 241–251.
  - 53 Lin H, Kryczek I, Li S, Green MD, Ali A, Hamasha R, Wei S, Vatan L, Szeliga W, Grove S *et al.* (2021) Stanniocalcin 1 is a phagocytosis checkpoint driving tumor immune resistance. *Cancer Cell* **39**, 480–493.e6.
  - 54 Cosgrove N, Varešlija D, Keelan S, Elangovan A, Atkinson JM, Cocchiglia S, Bane FT, Singh V, Furney S, Hu C *et al.* (2022) Mapping molecular subtype specific alterations in breast cancer brain metastases

identifies clinically relevant vulnerabilities. *Nat Commun* **13**, 514.

- 55 Seiler R, Ashab HAD, Erho N, van Rhijn BWG, Winters B, Douglas J, Van Kessel KE, Fransen van de Putte EE, Sommerlad M, Wang NQ *et al.* (2017) Impact of molecular subtypes in muscle-invasive bladder cancer on predicting response and survival after neoadjuvant chemotherapy. *Eur Urol* **72**, 544–554.
- 56 Jang HJ, Lee HS, Ramos D, Park IK, Kang CH, Burt BM and Kim YT (2020) Transcriptome-based molecular subtyping of non-small cell lung cancer may predict response to immune checkpoint inhibitors. *J Thorac Cardiovasc Surg* **159**, 1598–1610.e3.
- 57 Medvedev KE, Savelyeva AV, Chen KS, Bagrodia A, Jia L and Grishin NV (2022) Integrated molecular analysis reveals 2 distinct subtypes of pure seminoma of the testis. *Cancer Inform* **21**, 11769351221132634.
- 58 Savelyeva AV and Medvedev KE (2023) Seminoma subtypes differ in the organization and functional state of the immune microenvironment. *3 Biotech* **13**, 110.

## Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Fig. S1.** Sample distribution before de-batching (A) and after de-batching (B).

**Fig. S2.** Unsupervised clustering of the samples with all genes (A) or 22 immune cell types (B).

**Fig. S3.** Screening out six characteristic genes out of eight immDEGs using the LASSO algorithm. (A) LASSO coefficient path diagram. (B) LASSO regression analysis cross validation curve.

**Fig. S4.** Pan-cancer analysis of STC1. -  $P > 0.05$ ; \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ; rank-sum test. Tumor types (sample number): ACC (tumor = 79, normal = 0, GTE<sub>x</sub> = 258); BLCA (tumor = 406, normal = 19, GTE<sub>x</sub> = 21); BRCA (tumor = 1101, normal = 113, GTE<sub>x</sub> = 459); CEST (tumor = 306, normal = 3, GTE<sub>x</sub> = 19); CHOL (tumor = 35, normal = 9); COAD (tumor = 455, normal = 41, GTE<sub>x</sub> = 779); DLBC (tumor = 48, normal = 0, GTE<sub>x</sub> = 929); ESCA (tumor = 163, normal = 11, GTE<sub>x</sub> = 1445); GBM (tumor = 153, normal = 5, GTE<sub>x</sub> = 2642); HNSC (tumor = 504, normal = 44); KICH (tumor = 65, normal = 25, GTE<sub>x</sub> = 89); KIRC (tumor = 532, normal = 72, GTE<sub>x</sub> = 89); KIRP (tumor = 290, normal = 32, GTE<sub>x</sub> = 89); LAML (tumor = 150, normal = 0); LGG (tumor = 513, normal = 0, GTE<sub>x</sub> = 2642); LIHC (tumor = 371, normal = 50, GTE<sub>x</sub> = 226); LUAD (tumor = 516, normal = 59, GTE<sub>x</sub> = 578); LUSC (tumor = 501, normal = 49, GTE<sub>x</sub> = 623); MESO (tumor = 87, normal = 0); OV (tumor = 376, normal = 0, GTE<sub>x</sub> = 180); PAAD (tumor = 179, normal = 4, GTE<sub>x</sub> = 328); PCPG (tumor = 181, normal = 3); PRAD (tumor = 498, normal = 52, GTE<sub>x</sub> = 245); READ (tumor = 165, normal = 10, GTE<sub>x</sub> = 779); SARC (tumor = 260, normal = 2); SKCM (tumor = 471, normal = 1, GTE<sub>x</sub> = 1809); STAD (tumor = 375, normal = 32, GTE<sub>x</sub> = 359); TGCT (tumor = 134, normal = 0, GTE<sub>x</sub> = 391); THCA (tumor = 512, normal = 59, GTE<sub>x</sub> = 653); THYM (tumor = 120, normal = 2); UCEC (tumor = 545, normal = 35, GTE<sub>x</sub> = 142); UCS (tumor = 57, normal = 0, GTE<sub>x</sub> = 142); UVM (tumor = 80, normal = 0).