Heliyon 10 (2024) e24320

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Subgrouping testicular germ cell tumors based on immunotherapy and chemotherapy associated lncRNAs

Jian Cao [a,1], Zhizhong Liu [a,1], Junbin Yuan [b,1], Yanwei Luo [c], Jinrong Wang [d], Jianye Liu [d], Hao Bo [e,f,**], Jie Guo [g,h,i,*]

[a] Hunan Cancer Hospital, Department of Urology, The Affiliated Cancer Hospital of Xiangya School of Medicine of Central South University, Changsha, 410013, Hunan, China
[b] Department of Urology, Xiangya Hospital, Central South University, Changsha, 410013, Hunan, China
[c] Department of Blood Transfusion, the Third Xiangya Hospital of Central South University, Changsha, 410013, Hunan, China
[d] Department of Urology, The Third Xiangya Hospital of Central South University, No.138, Tongzipo Road, Changsha, 410013, Hunan, China
[e] Clinical Research Center for Reproduction and Genetics in Hunan Province, Reproductive and Genetic Hospital of CITIC-Xiangya, Changsha, 410078, Hunan, China
[f] NHC Key Laboratory of Human Stem Cell and Reproductive Engineering, Institute of Reproductive and Stem Cell Engineering, Central South University, Changsha, 410078, Hunan, China
[g] National Institution of Drug Clinical Trial, Xiangya Hospital, Central South University, Changsha, Hunan, China
[h] China National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha, Hunan, China
[i] International Science and Technology Innovation Cooperation Base for Early Clinical Trials of Biological Agents in Hunan Province, Changsha, Hunan, China

ARTICLE INFO

Keywords:
Testicular germ cell tumors
Classification
lncRNA
Immunotherapy
Chemotherapy
Precise treatment

ABSTRACT

Testicular germ cell tumors (TGCT) are the most common reproductive system malignancies in men aged 15–44 years, accounting for 95 % of all testicular tumors. Our previous studies have been shown that long non-coding RNAs (lncRNAs), such as LINC00313, TTTY14 and RFPL3S, were associated with development of TGCT. Subgrouping TGCT according to differential expressed lncRNAs and immunological characteristics is helpful to comprehensively describe the characteristics of TGCT and implement precise treatment. In this study, the TGCT transcriptome data in The Cancer Genome Atlas Program (TCGA) database was used to perform consensus clustering analysis to construct a prognostic model for TGCT. TGCT was divided into 3 subtypes C1, C2, and C3 based on the differentially expressed lncRNAs. C1 subtype was sensitive to chemotherapy drugs, while the C2 subtype was not sensitive to chemotherapy drugs, and C3 subtype may benefit from immunotherapy. We defined the C1 subtype as epidermal progression subtype, the C2 subtype as mesenchymal progression subtype, and the C3 subtype as T cell activation subtype. Subgrouping based on differentially expressed genes (DEGs) and immunological characteristics is helpful for the precise treatment of TGCT.

* Corresponding author. National Institution of Drug Clinical Trial, Xiangya Hospital, Central South University, Xiangya Road 87, Changsha, 410008, Hunan, China.
** Corresponding author. Clinical Research Center for Reproduction and Genetics in Hunan Province, Reproductive and Genetic Hospital of CITIC-Xiangya, Xiangya Road 84, Changsha, 410078, Hunan, China.
E-mail addresses: 1172881652@qq.com (H. Bo), guojiexy@csu.edu.cn (J. Guo).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.heliyon.2024.e24320
Received 8 May 2023; Received in revised form 1 December 2023; Accepted 7 January 2024
Available online 8 January 2024

1. Introduction

Testicular germ cell tumors (TGCT) are the most common reproductive system malignancies in men aged 15–44 years [1]. Since the 20th century, the global incidence of TGCT has gradually increased, especially in some western countries, with an annual growth rate of 1 %–2 % [2]. According to the epidemiology, clinical manifestations, phenotypic characteristics, chromosomal composition, and genomic imprinting of TGCTs, World Health Organization divides TGCTs into three groups, a total of 5 subtypes: infant/prepubertal TGCTs (including teratoma and yolk sac tumor), postpubertal TGCTs and TGCTs in young men (including seminoma and non-seminoma, which includes embryonal carcinoma, yolk sac tumor, choriocarcinoma, and teratoma) and spermatocyte or spermatogonia cell tumor in older men [3]. At present, the pathogenesis of testicular tumors is not clear. Epidemiological analysis has found that various factors increase the chance of TGCT, including congenital factors such as cryptorchidism or undescended testes, Klinefelter syndrome, testicular feminization syndrome, acquired factors such as injury, infection, nutritional factors, and excessive use of exogenous estrogen by the mother during pregnancy [4,5].

In general, seminoma is quite sensitive to radiotherapy and cisplatin-based chemotherapy, but non-seminoma tumors are only sensitive to chemotherapy [6]. Although surgery, radiotherapy and chemotherapy can effectively treat more than 95 % of TGCT patients, it leads to ineffective treatment adverse reactions such as fertility and hypogonadism cause severe physical and mental distress to patients [7]. Therefore, new treatment regimens should try to avoid these adverse events. With the development of high-throughput genome sequencing technology and the advancement of bioinformatics, cancer-related genes have been further studied, and immune-related messenger ribonucleic acids (mRNAs) and long non-coding RNAs (lncRNAs) may play an important role in various stages of tumorigenesis, development, and transformation [8]. Immune cells are involved in the biological process of tumorigenesis, development, and prognosis [9]. Tumor cells can be specifically recognized and eliminated by specific immune cells, and tumor cells can escape the immune surveillance, thereby further leading to the occurrence and development of tumors [10]. Using genome sequencing technology and bioinformatics analysis methods to reclassify TGCT according to differential gene expression and immunological characteristics is helpful to comprehensively describe the characteristics of TGCT and implement precise treatment [11,12].

In this study, the TGCT transcriptome data in the TCGA database was combined with our team previously published data to perform consensus clustering analysis combined with immune-related genes to reclassify TGCT. We identified three subtypes of TGCT based on differentially expressed lncRNAs: C1, epidermal cell progressive type; C2, mesenchymal cell progressive type; C3, T cell activation type. The C3 subtype has the lowest degree of malignancy and the greatest benefit from immunotherapy, the C1 subtype is the most sensitive to chemotherapy drugs, and the C2 subtype is not sensitive to both immunotherapy and chemotherapy and has the worst prognosis. Thus, we construct a prognostic model for TGCT to predict the responsiveness to immunotherapy and chemotherapy and would be helpful for clinical decision-making of TGCT.

2. Materials and methods

2.1. Data source

The RNA-seq data, mutation data and clinical information of TGCT were obtained from The Cancer Genome Atlas (TCGA) dataset (https://portal.gdc.com). Our previously published data on lncRNA sequencing in TGCT samples was used [13]. The differential lncRNAs were analyzed by the Hiplot online tool, and a Venn diagram was drawn (https://hiplot-academic.com/) [14].

2.2. Gene expression and cluster analysis

Somatic mutations information in patients were downloaded and visualized using the maftools package in R software (Metascape: http://metascape.org/gp/index.html#/main/step1) [15,16]. R software package ConsensusClusterPlus (v1.54.0) was used to conduct consensus clustering analysis, the maximum number of clusters is 6, and 80 % of the total samples are extracted 100 times [17]. The cluster heatmaps were analyzed by the R package pheatmap (v1.0.12). Differential expression of mRNA was investigated using Limma package of R software (version: 3.40.2) [18]. Adjusted P-values for multiple testing using Benjamini-Hochberg were analyzed in TCGA to correct for false-positive results. "Adjusted $P < 0.05$ and log2 (fold change) > 1 or log2 (fold change) $< -1''$ was defined as a threshold mRNA differential expression screen. To further confirm the potential functions of potential targets, the data were analyzed by functional enrichment. The ClusterProfiler package in R was used to analyze the Gene Ontology (GO) function of underlying mRNAs [19]. Gene expression heatmaps retain variance at 0.1 for the above genes, if the number of input target genes is more than 1000, the top 25 % genes are extracted and displayed after sorting according to the variance value from high to low.

2.3. Analysis of clinicopathological features

The Sankey diagram is analyzed and drawn by the online tool of the Clinical Bioinformatics (https://www.aclbi.com/static/index.html#/). The clinicopathological characteristics of patients with different subtypes were analyzed statistically by R software v4.0.3 and ggplot2 (v3.3.2) [20]. The data in TPM format from the dataset was normalized to log2(TPM+1), and the samples with RNAseq data and clinical information were retained. The R software survival package (v4.0.3, R Foundation for Statistical Computing, 2020) was used for subsequent analysis. The log rank was used to test the Kaplan-Meier survival analysis to compare the survival difference between the above two or more groups, and the timeROC analysis was performed to judge the accuracy of the prediction model.

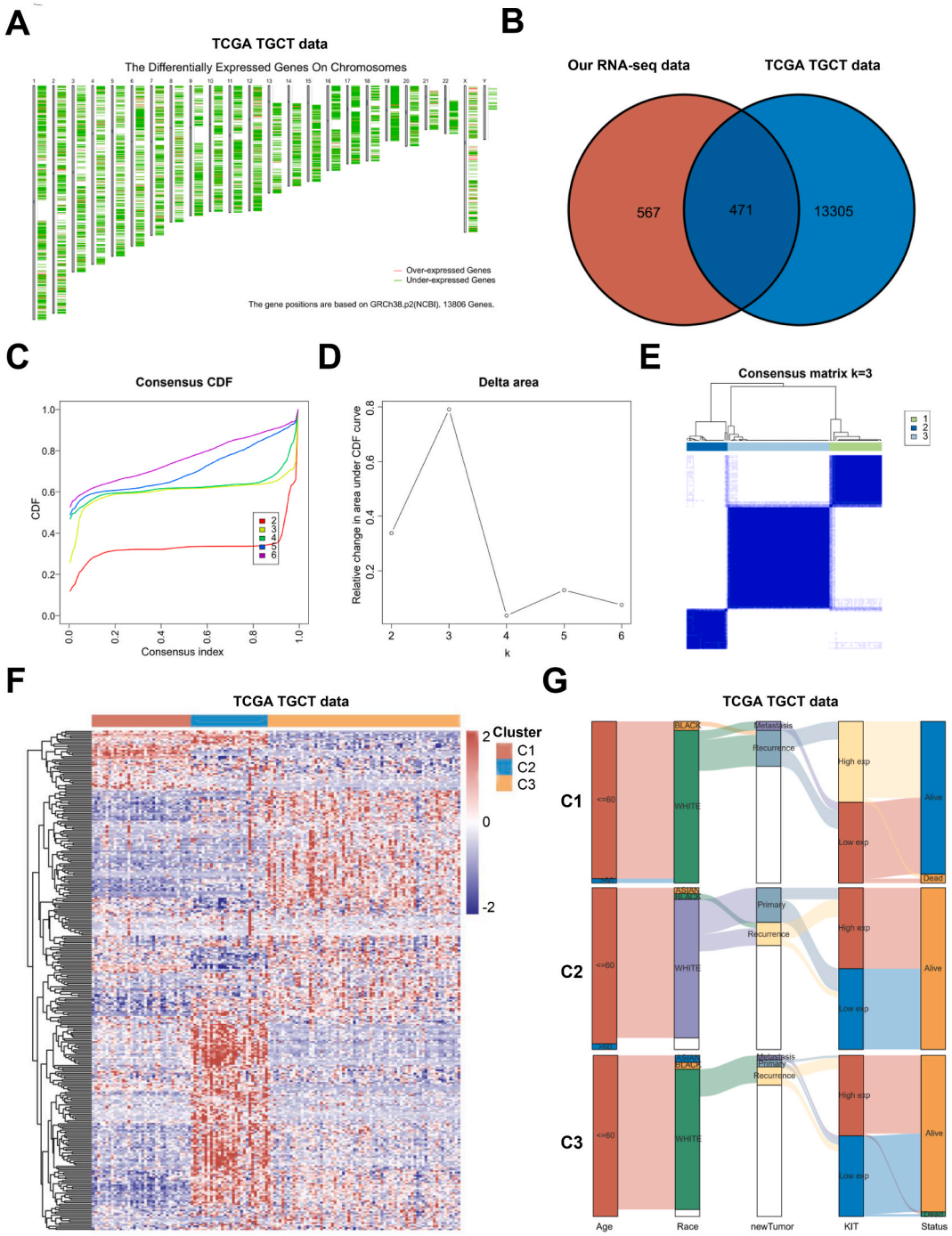

Fig. 1. TGCT subtypes based on differentially expressed lncRNAs. (A) The GEPIA database was used to screen the differentially expressed genes (DEGs) based on the TCGA TGCT cohort data, and the chromosomal locations of DEGs were displayed (red represents up-regulation, green represents down-regulation). (B) The DEGs of the TCGA TGCT cohort data were intersected with our previously published dataset of TGCT lncRNA sequencing to screen 471 common differentially expressed lncRNAs. (C–E) Consensus clustering analysis from the TCGA TGCT cohort data based on co-differentially expressed lncRNAs. C, consensus cumulative distribution graph CDF, using a histogram of 100 bins to calculate the cumulative distribution, the CDF graph shows that the best K value is 3; D, delta area plot, for each K, calculate K and K-1 in comparison, for the relative change of the area under the CDF curve, the point with insignificant increase is selected as the optimal K value; E, consensus matrix heat map, the value represents the possibility of two belong to the same cluster, the value ranges from 0 to 1, Colors range from white to dark blue. (F) Gene clustering heatmap based on common differential lncRNAs. (G) Sankey Diagram showing baseline data for the three subtypes of TGCT patients.

Multivariate cox regression analysis was used to construct a prognostic model. For Kaplan-Meier curves, p-value, and hazard ratios (HR) with 95 % confidence intervals (CI) were obtained by logrank test and univariate Cox regression. $p < 0.05$ was considered statistically significant. According to the largest publicly available pharmacogenomics database Genomics of Cancer Drug Sensitivity (GDSC) (https://www.cancerrxgene.org/), drug treatment response for each sample was predicted based on the sample transcriptome [21].

2.4. Immune infiltration and immune score

An R package immunedeconv that integrates six state-of-the-art algorithms, including TIMER, xCell, MCP-counter, CIBERSORT, EPIC and quanTIseq were used for immune score assessment [22,23]. The OCLR algorithm was used to calculate mRNA stemness index (mRNAsi) [24]. The mRNA expression-based signature contains a gene expression profile containing 11,774 genes, we used the Spearman correlation (RNA expression data), and then used a linear transformation mapping mRNAsi to the [0,1] range.

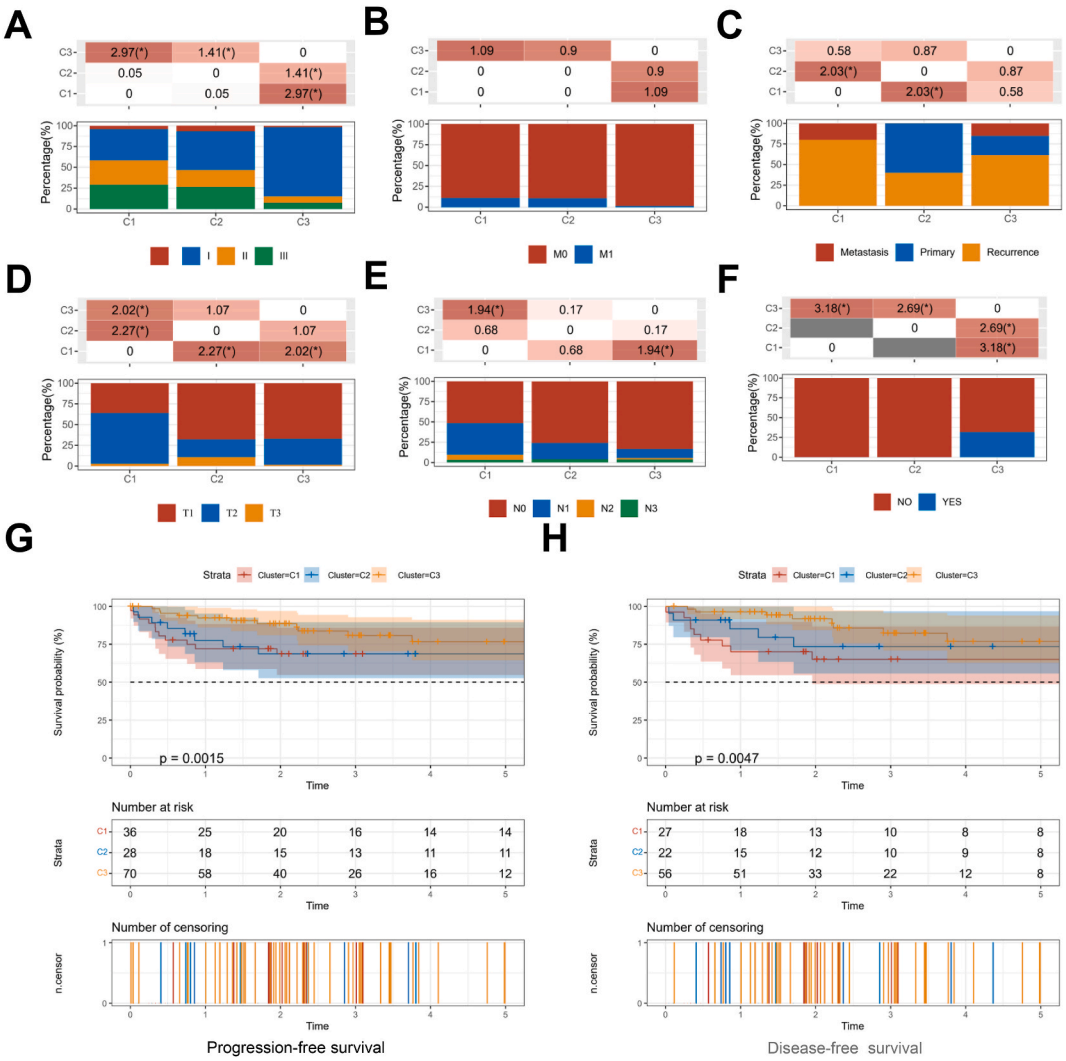

Fig. 2. Clinicopathological characteristics of patients with the three subtypes. (A–F). Percentage of patients with different clinicopathological characteristics of three TCGT subtypes C1, C2, C3 (Lower), including clinical grade (A), M grade (B), metastatic recurrence (C), T grade (D), lymph node metastasis (E), radiotherapy (F); the values in the table represent the logP value of the comparison between the two groups, and * means the difference is statistically significant (Upper). (G–H) Progression-free survival (G) and Disease-free survival (H) of three TCGT subtypes C1, C2, and C3. C3 subtype has better Progression-free survival and Disease-free survival, while Progression-free survival and Disease-free survival of C1 and C2 subtype were poor.

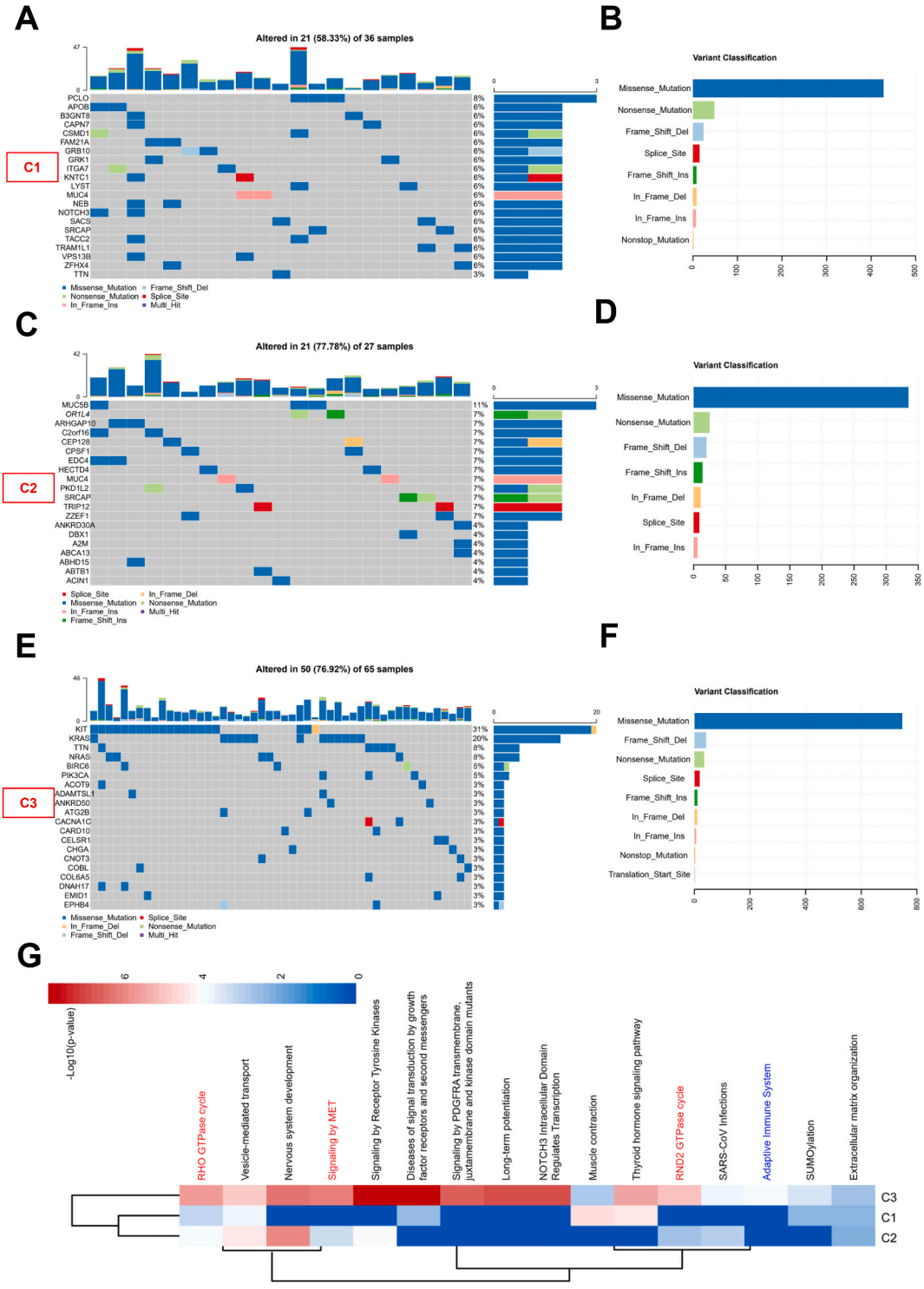

Fig. 3. The gene mutation map of the TGCT subtypes. (A–B) TOP20 mutated genes of subtype C1, and distribution of mutation types. (C–D) TOP20 mutated genes of C2 subtype, and distribution of mutation types. (E–F) TOP20 mutated genes of C3 subtype, and distribution of mutation types. Oncoplot showing the somatic landscape of TGCT. Genes are ordered by mutation frequency and samples are ordered by disease histology. Sidebar plots show log10-transformed Q-values estimated by MutSigCV. Different colors with specific annotations at the bottom indicate various mutation types. Right: Cohort summary plot showing variant distribution according to variant classification, type and SNV category. Bottom indicates mutational load for each sample (variant classification type). (G) KEGG enrichment analysis for all mutated genes in the three subtype samples.

3. Results

3.1. TGCT was classified into three subtypes based on differentially expressed lncRNAs

TGCTs vary widely in immune cell infiltration and treatment response, suggesting a high degree of heterogeneity, therefore precise classification of TGCTs is important to TGCT treatment and outcome prediction. We here utilized the TCGA TGCT cohort data as well as our previous sequencing results to screen for differentially expressed lncRNAs. These differentially expressed lncRNAs were distributed on different chromosomes (Fig. 1A). Combining the TGCT cohort data and our previous sequencing results, we screened 471 common differentially expressed lncRNAs (Fig. 1B). We used a consensus clustering algorithm to perform a consensus clustering analysis on the TGCT cohort according to these common differentially expressed lncRNAs (Fig. 1 C, D, E). TGCT was divided into 3 subtypes C1, C2, and C3 (Fig. 1F). As shown in the Sankey Diagram, the C1 group has a higher recurrence rate, the C2 group has a higher primary rate, and the C3 group has a lower recurrence rate; there is no significant difference in the age and ethnic distribution of the 3 subtypes (Fig. 1G).

3.2. Clinicopathological features of the TGCT subtypes

We further analyzed the clinicopathological characteristics of the TGCT subtypes, C1, C2, and C3, including clinical stage, M grade, metastatic recurrence, T grade, and lymph node metastasis. The results showed that the C3 subtype had a higher proportion of Stage I,

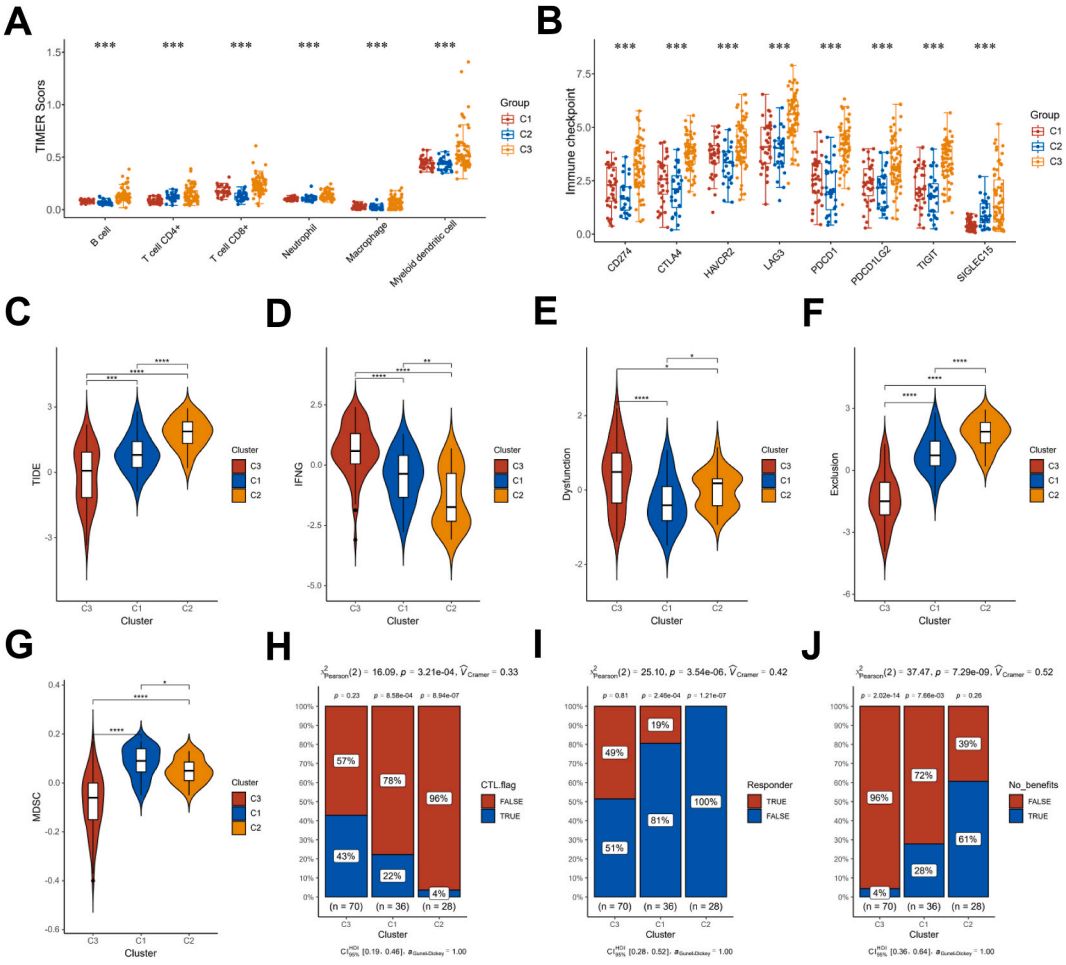

Fig. 4. Correlation of the TGCT subtypes with immunotherapy. (A) Using the TIMER algorithm to analyze the degree of immune cell infiltration in the subtypes based on the TCGA TGCT cohort data. (B) Differences in immune checkpoint molecule expression among three subtype samples were analyzed based on TCGA TGCT cohort data. (C) Differences between the subtypes of tumor immune dysfunction and rejection (TIDE) scores. (D) Differences among the subtypes of cytotoxic T cell effector molecule IFNG. (E) Differences among the subtypes of immune dysregulation scores. (F) Differences among the subtypes of immune escape scores. (G) Differences among the three subtypes of myeloid-derived immunosuppressive cell scores. (H) Differences between the presence and absence of the subtypes of cytotoxic T cell signatures. (I) Differences among the subtypes of immune checkpoint inhibitor treatment responses. (J) Differences between the benefits of immune checkpoint therapy among the subtypes.

while the C1 and C2 subtypes had a higher proportion of Stage III (Fig. 2A); the C3 subtype had a significantly lower percentage of M1 than the C1 and C2 subtypes (Fig. 2B); C1 subtype had the highest recurrence proportion, C2 subtype had the highest primary proportion, while C3 subtype was distributed in primary, metastasis and recurrence (Fig. 2C); C1 subtype had a higher T2 grade in patients, C2 subtype had the highest proportion of T2 and T3 grades, while the C3 subtype had a higher proportion of T1 grade (Fig. 2D); in terms of lymph node metastasis, the C1 subtype had highest N1, N2, and N3 patients, and the C2 subtype had more patients with N1 and N2, while 80 % of the C3 subtype belonged to N0 (Fig. 2E). In addition, in terms of radiotherapy response, both C1 and C2 subtypes were non-responsive to radiotherapy, whereas C3 subtypes responded to radiotherapy in 70 % of cases (Fig. 2F). We next analyzed the survival prognosis of the three subtypes. We found that the C3 subtype patients had better progression-free survival and disease-free survival, while the C1 and C2 subtypes had poorer progression-free survival and disease-free survival (Fig. 2G and H). These results suggest that the C3 subtype has a better prognosis, while the C1 and C2 subtypes have a high TNM grade, a high proportion of metastatic recurrence, and no response to radiotherapy, so their prognosis was poorer.

3.3. Gene mutation map of the TGCT subtypes

We next analyzed the gene mutation profiles of the TGCT subtypes. Gene mutations were found in 58 %, 77.8 % and 76.9 % of samples of C1, C2 and C3 subtypes, respectively. The highest type of gene mutation was missense mutation. The gene with the highest mutation frequency in C1 subtype was PCLO, C2 subtype was MUC5B, and C3 subtype was KIT and KRAS (Fig. 3A–F). Through gene

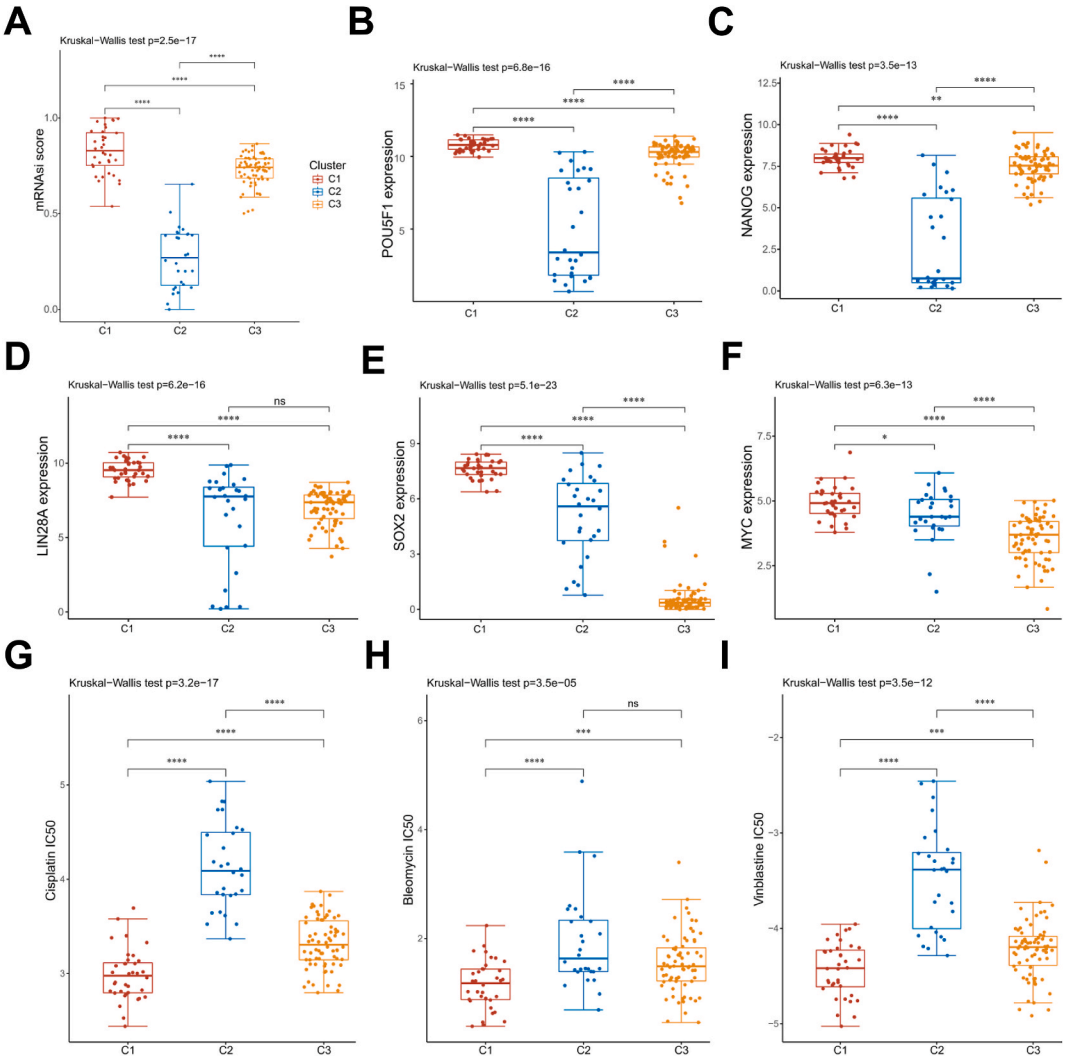

Fig. 5. TGCT subtypes are associated with tumor stemness and drug sensitivity. (A) Differences between the stemness index (mRNAsi) scores among the subtypes, the stemness index was calculated based on the transcriptome of the sample and used to assess the degree of stemness of the sample). (B–F) Differential expression of stem cell-related genes POU5F1/NANOG/LIN28A/SOX2/MYC among the subtypes. (G–I) Differences between the three chemotherapeutic drugs cisplatin, bleomycin, vinblastine, IC50 between the three subtypes.

enrichment of these mutant genes, it was found that the mutant genes in C3 subtype were enriched and activated in receptor tyrosine kinases, disease of signal transduction by growth factor receptors and second messengers, and NOTCH3 intracellular domain regulates transcription pathway-related genes, while these signaling in C1 and C2 subtypes were inactivated (Fig. 3G).

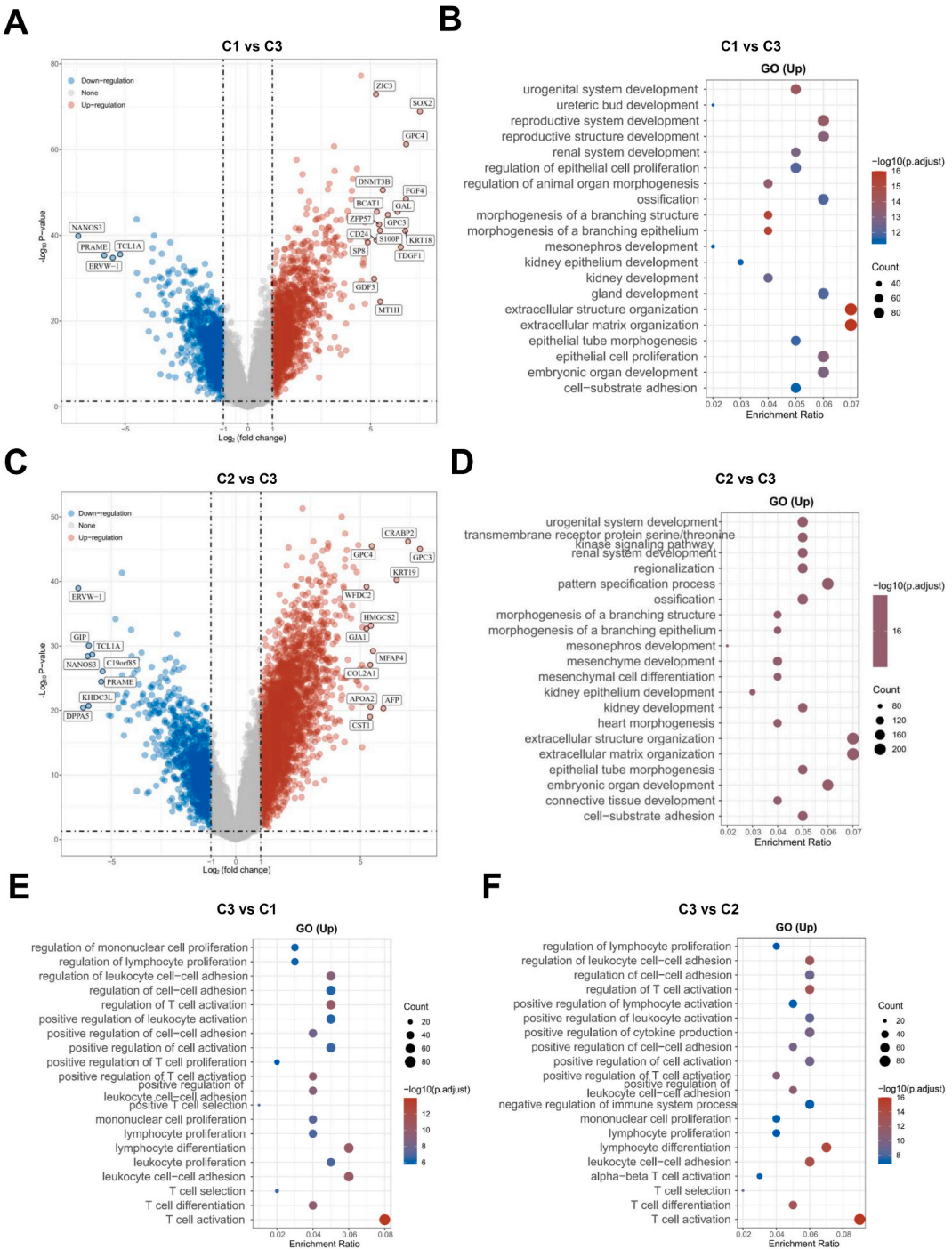

Fig. 6. Functional enrichment analysis of three subtype samples. (A–F) Differential gene volcano plots and GO enrichment analysis (bubble plots) of up-regulated genes for the subtype samples.

3.4. Correlation of TGCT subtypes with immunotherapy

The TIMER algorithm was used to analyze the degree of immune cell infiltration in the three TGCT subtypes based on the TCGA TGCT cohort data. The infiltration of B cells, $CD4^+$ T cells, $CD8^+$ T cells, neutrophils, macrophages, and dendritic cells of the C3 subtype was significantly higher than that of the C1 and C2 subtypes (Fig. 4A). The expression of immune checkpoint molecules in the C3 subtype, including CD274, CTLA4, HAVCR2, LAG3, PDCD1, PDCD1LG2, TIGIT, SIGLEC15, was significantly higher than that in the C1 and C2 subtypes (Fig. 4B), whereas tumor immune dysfunction and rejection scores (TIDE) were significantly lower than C1 and C2 subtypes (Fig. 4C). The expression of the cytotoxic T cell effector molecule IFNG was significantly higher in the C3 subtype than in the C1 and C2 subtypes (Fig. 4D). The C3 subtype was suggested to have lower immune dysregulation (Fig. 4E), lower probability of immune escape (Fig. 4F), and lower probability of immune suppression (Fig. 4G). Analysis of the differences between the 3 TGCT subtypes and response to immune checkpoint inhibitor treatment revealed that 43 % of the C3 subtypes had a cytotoxic T cell signature, which was significantly higher than 22 % for C1 and 4 % for C2 (Fig. 4H). 49 % of C3 subtypes responded to immune checkpoint inhibitor therapy, C1 was 19 %, and C2 was almost non-responsive (Fig. 4I). In addition, C3 subtypes also benefited the most from immune checkpoint therapy (96 %), the C1 subtype had a benefit of 72 % and the C2 subtype only 39 % (Fig. 4J).

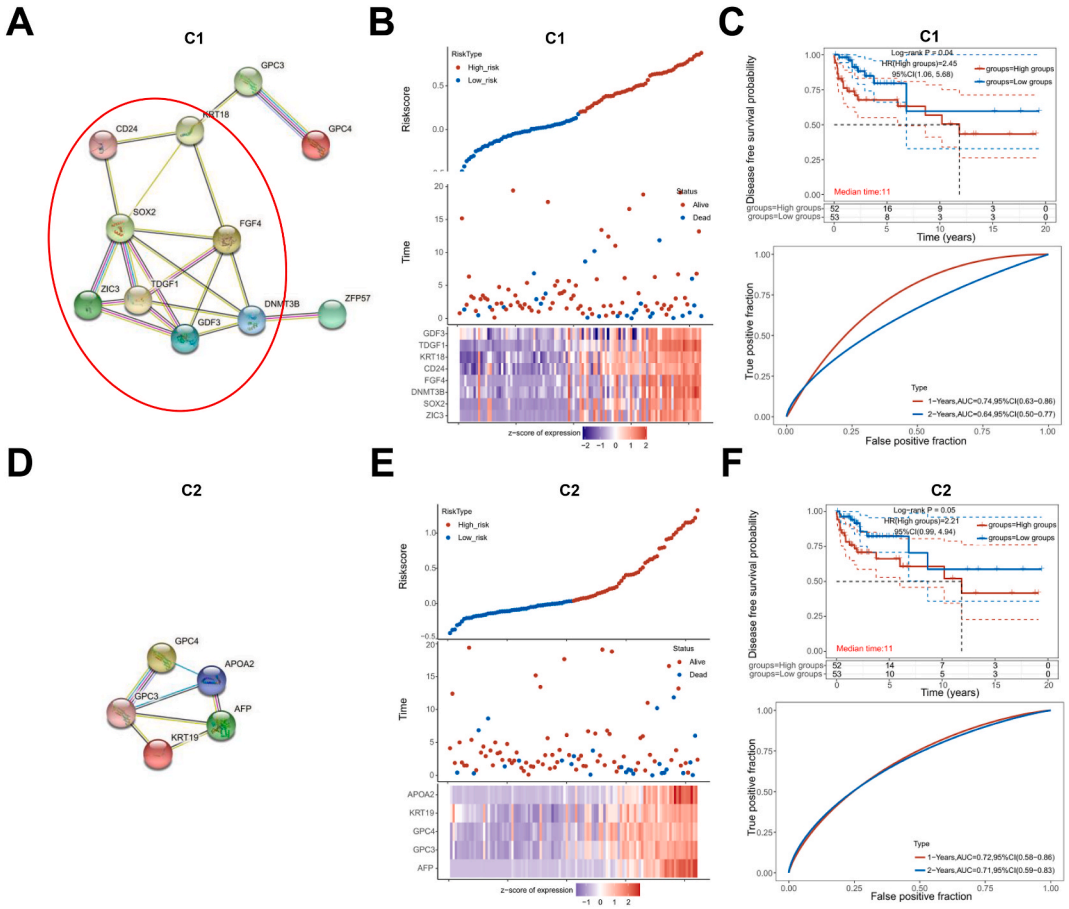

Fig. 7. Prognostic model of the C1 and C2 subtypes based on the core genes. (A) protein-protein interaction network of C1 subtype core genes. (B) Riskscore scatter plot from low to high, red indicates high risk group, blue indicates low risk group (upper); the scatter plot distribution of survival time and survival status corresponding to Riskscore of different samples (Median); the expression heat of C1 subtype signature gene Figure (Lower). (C) The distribution of the KM survival curve of the risk model of the C1 subtype in the data set, HR (High risk) represents the risk coefficient of the high-risk group relative to the low-risk group sample; 95 % CL represents the HR confidence interval (upper). The ROC curve and AUC of the C1 subtype risk model at different times. The higher the AUC value, the stronger the predictive ability of the model (Lower). (D) protein-protein interaction network of C2 subtype core genes. (E) COX regression prognostic model based on genes in the core network of C2 subtype. Riskscore scatter plot from low to high, red indicates high risk group, blue indicates low risk group (upper); the scatter plot distribution of survival time and survival status corresponding to Riskscore of different samples (Median); the expression heat of C2 subtype signature gene Figure (Lower). (F) The KM survival curve distribution of the risk model of C2 subtype in the data set, HR (High risk) represents the risk coefficient of the high-risk group relative to the low-risk group sample; 95 % CL represents the HR confidence interval (upper). The ROC curve and AUC of the C2 subtype risk model at different times. The higher the AUC value, the stronger the predictive ability of the model (Lower).

3.5. The association of TGCT subtypes with tumor stemness and drug sensitivity

Cancer stem cell stemness can reflect sensitivity to drugs. We used a logistic regression machine learning algorithm (OCLR, One Class Linear Regression) to analyze the mRNA stemness index (mRNAsi) among the three subtypes. The results showed that the mRNAsi of the C2 subgroup was the lowest, the mRNAsi of the C1 subgroup was the highest, and the mRNAsi of the C3 subgroup was in the middle (Fig. 5A). By analyzing the differences in the expression of stem cell-related genes POU5F1/NANOG/LIN28A/SOX2/MYC among the three subtypes, we found that the expression of POU5F1/NANOG/LIN28A/SOX2/MYC was significantly increased in the C1 subtype compared with the C2 subtype. In contrast, the expression of POU5F1/NANOG was significantly increased and SOX2/MYC was significantly decreased in the C3 subtype (Fig. 5B–F). We next analyzed the differences between the three subtypes against the chemotherapeutic drugs cisplatin, bleomycin, vinblastine. The results showed that the C2 subgroup had the highest IC50 for these three chemotherapy drugs, the C1 subgroup had the lowest IC50, and the C3 subgroup had the middle IC50 (Fig. 5G–I), suggesting that the C1 subgroup was sensitive to chemotherapy drugs, while the C2 subgroup was not sensitive to chemotherapy drugs.

3.6. The prognostic model of TGCT subtypes

The above results suggest that the C3 subtype has the lowest degree of malignancy, followed by the C1 subtype, and the C2 subtype was the highest. Compared with C3 subtype, C1 subtype significantly higher expression of ZIC3, SOX2, GPC4 and other stemness-related genes, C2 subtype significantly higher expression of CST1, KRT19, CARBP2 and other pro-invasion, metastasis and immune escape-related genes (Fig. 6A–C). Compared with the C3 subtype, the C1 subtype was significantly enriched for extracellular structure organization, extracellular matrix organization, embryonic organ development, and multiple epidermal cell proliferation and differentiation-related pathways (Fig. 6B), while the C2 subtype was significantly enriched for extracellular structure organization, extracellular matrix organization, embryonic organ development, and multiple mesenchymal cell proliferation and differentiation-related signaling pathways (Fig. 6D); compared with C1 and C2 subtypes, the C3 subtype was significantly enriched in immune cell-related pathways, especially T cell-related pathways including T cell activation, T cell differentiation, leukocyte cell-cell adhesion and other pathways (Fig. 6E and F). Therefore, we defined the C1 subtype as epidermal progression subtype, the C2 subtype as mesenchymal progression subtype, and the C3 subtype as T cell activation subtype. These definitions tightly link the TGCT subtypes to cellular origin and the tumor immune microenvironment.

We next used the STRING online tool to construct a protein-protein interaction network of genes that were significantly upregulated in C1 and C2 subtypes (Fig. 7A–D). The core genes of the C1 subtype include APOA2, KRT19, GPC4, GPC3, and AFP, and the higher the expression of these genes, the higher the Riskscore score (Fig. 7B). KM survival curve analysis found that patients with high expression of C1 subtype core genes had significantly reduced disease-free survival, and these core gene signatures could be used to distinguish high-risk and low-risk patients (1-year AUC, 0.74; 2-year AUC, 0.64) (Fig. 7C). The core genes of the C2 subtype include CDF3, TDGF1, KRT18, CD24, FGF4, DNMT3B, SOX2, ZIC3, and the higher the expression of these genes, the higher the Riskscore score (Fig. 7E). KM survival curve analysis found that patients with high expression of C2 subtype core genes had significantly lower disease-free survival, and these core gene signatures could be used to distinguish high-risk and low-risk patients (1-year AUC, 0.72; 2-year AUC, 0.71) (Fig. 7F).

4. Discussion

In this study, combined the sequencing data of the TGCT cohort in the TCGA database with our previous lncRNA sequencing data, TGCTs were divided into three types: C1, C2, and C3 according to differentially expressed lncRNAs. The C3 subtype has a better prognosis, while the C1 and C2 subtypes had poor prognosis due to high TNM grades, a high proportion of metastatic recurrence, and no response to radiotherapy.

The C1, C2 and C3 subtypes had 58 %, 77.8 % and 76.9 % of genetic mutation samples, respectively. The gene with the highest mutation frequency of C1 subtype was PCLO, and the C2 subtype was MUC5B. The biological functions and molecular mechanisms of PCLO and MUC5B in TGCT are still not reported. PCLO has been reported to regulate synaptic vesicle trafficking and release of monoamine neurotransmitters in presynaptic active regions and was often mutated in poorly differentiated hepatocellular carcinoma [25,26]. MUC5B is a mucin gene. Abnormal expression of mucin-related genes is an important feature of many epithelial cell-derived tumors and is associated with poor prognosis [27]. Germline mutations in MUC5B are closely associated with a variety of tumorigenesis [28]. These studies suggest that PCLO and MUC5B may play key roles in the occurrence and development of TGCT. The frequently mutated molecules of the C3 subtype were KIT and KRAS. Through gene enrichment of these mutant genes, it was found that the C3 subtype has high activity in receptor tyrosine kinases, disease of signal transduction by growth factor receptors and second messengers, and NOTCH3 intracellular domain regulates transcription pathway-related genes, while C1 and C2 subtypes are low. The proto-oncogene c-Kit belongs to the type 3 tyrosine kinase receptor family. The ligand of Kit is a stem cell factor (SCF). The SCF-kit pathway regulates the differentiation of germ cells and is of great significance to the survival of primordial germ cells [29]. When the proto-oncogene c-Kit is mutated, the Kit protein show an activated state without binding to the ligand SCF, thereby stimulating the continuous proliferation of tumor cells and inducing the uncontrolled anti-apoptotic signal, resulting in the cancerization [30]. It was reported that compared with normal testicular tissue, 88 % (22/25) of seminoma and 44.4 % (4/9) of non-seminoma expressed kit protein, of which 9 cases expressed kit protein seminoma detected c-kit gene mutation (40.9 %, 9/22) [31]. In addition, Cheng et al. reported that among 22 patients with germ cell tumors, 6 cases (27 %) were detected with c-Kit gene point mutation [32]. Mutation of kirsten rat sarcoma viral oncogene (KRAS) is a common biological event in many tumors, such as colon cancer, pancreatic cancer, and

non-small cell lung cancer, and is also one of the main mechanisms of many tumors [33]. KRAS gene mutation can cause the body to lose GTP hydrolase activity, and its carcinogenic effect mainly affects the growth and proliferation of cells [34]. These phenomena suggest that the C3 subtype TGCT may be sensitive to KRAS and KIT mutation inhibitors.

Although TGCT has high chemotherapy sensitivity, 20–30 % of TGCT patients are resistant to conventional chemotherapy [35]. In addition, radiotherapy and chemotherapy are harmful to patients' reproductive function, leading to a significant decrease in sperm quality and event permanent infertility [36]. Therefore, it is crucial to find new options that can supplement or even replace the traditional treatment methods of TGCT [37]. The advent of immunotherapy has provided new directions for the treatment of TGCT patients. It was reported that the infiltration of activated T cells was strongly associated with a positive prognosis in patients with seminoma [38]. PD-L1 inhibitors have also been shown to be useful in patients with TGCT who are refractory to chemotherapy or radiation therapy [39]. Further analysis of the immunological characteristics of the C1, C2, and C3 subtypes showed that the C3 subtype had a large amount of immune cell infiltration, while the C1 and C2 subtypes had lower immune scores, suggesting that the C3 subtype would benefit from immune checkpoint inhibitor therapy.

Cancer stem cell stemness affects the sensitivity of tumor cells to chemotherapeutic drugs [40]. Our analysis showed that the C1 subtype had the highest stemness, the C2 subtype had the lowest, the C1 subtype had the lowest IC50 for cisplatin, bleomycin, and vinblastine and the C2 subtype had the highest IC50 for these three chemotherapeutic drugs, suggesting that the C1 subtype had the lowest IC50 for chemotherapy. These results suggest that C1 subtype was sensitive to these drugs, while the C2 subgroup is not, which may be resulted by cancer stem cell stemness. We further analyzed the characteristics and phenotypic mechanisms of the three subtypes of TGCT, and found that the C1 subtype was significantly enriched in epidermal cell proliferation and differentiation-related pathways, the C2 subtype was significantly enriched in mesenchymal cell proliferation and differentiation-related signaling pathways, and the C3 subtype was significantly enriched in T-cell activation-related pathways. Therefore, we defined the C1 subtype as epidermal progression, the C2 subtype as mesenchymal progression, and the C3 subtype as T cell activation. These definitions tightly link the three subtypes to cellular origin and the tumor immune microenvironment. Survival curve analysis also confirmed that the disease-free survival of patients with high expression of C1 and C2 subtype core genes was significantly reduced. The above results suggest that the C3 subtype is the least malignant and has the greatest benefit after receiving immunotherapy, the C1 subtype is the most sensitive to chemotherapy drugs, and the C2 subtype is not sensitive to both immunotherapy and chemotherapy.

5. Limitations

This paper also has certain limitations: 1. The results based on lncRNA sequencing lead to the fact that patients of the same subtype may have different heterogeneity due to different characteristics on other omics data platforms; 2. The study was based on a retrospective design and prospective studies should be performed to verify the results; 3. The biological functions and molecular mechanisms of the key genes should be further studied to promote their clinical application in TGCT.

6. Conclusion

This paper proposes three subtypes of TGCT based on differentially expressed lncRNAs: C1, epidermal cell progressive type; C2, mesenchymal cell progressive type; C3, T cell activation type. The C3 subtype has the lowest degree of malignancy and the greatest benefit from immunotherapy, the C1 subtype is the most sensitive to chemotherapy drugs, and the C2 subtype is not sensitive to both immunotherapy and chemotherapy and has the worst prognosis. Subgrouping based on differential gene expression and immunological characteristics is helpful for the precise treatment of TGCT, reducing the side effects of treatment and improving the quality of life of patients.

Funding

This work was supported by Research Project of Hunan Provincial Health Commission (202204055545) and Natural Science Foundation of Hunan Province (2021JJ41091).

Ethics approval

Not applicable.

Consent to participate

Not applicable.

Consent for publication

Not applicable.

Data availability statement

The datasets used or analyzed during the current study are available from the corresponding author on reasonable request. R software packages were used to conduct analysis. The code was available by reasonable request on corresponding author.

CRediT authorship contribution statement

Jian Cao: Writing – review & editing, Writing – original draft, Methodology, Data curation. **Zhizhong Liu:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation. **Junbin Yuan:** Writing – review & editing, Writing – original draft, Validation, Methodology, Data curation. **Yanwei Luo:** Writing – review & editing, Writing – original draft, Resources, Investigation, Conceptualization. **Jinrong Wang:** Writing – review & editing, Writing – original draft, Methodology, Data curation. **Jianye Liu:** Writing – review & editing, Writing – original draft. **Hao Bo:** Writing – review & editing, Writing – original draft, Supervision, Funding acquisition, Conceptualization. **Jie Guo:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis, Data curation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] C.C. Guo, B. Czerniak, Somatic-type malignancies in testicular germ cell tumors, Hum. Pathol. 127 (2022) 123–135.
[2] M. Soleimani, C. Kollmannsberger, L. Nappi, Emerging role of biomarkers in testicular germ cell tumors, Curr. Oncol. Rep. 24 (2022) 437–442.
[3] M.J. Murray, R.A. Huddart, N. Coleman, The present and future of serum diagnostic tests for testicular germ cell tumours, Nat. Rev. Urol. 13 (2016) 715–725.
[4] L. Cheng, S.A. Mann, A. Lopez-Beltran, M. Chovanec, M. Santoni, M. Wang, et al., Molecular Characterization of testicular germ cell tumors using tissue microdissection, Methods Mol. Biol. 2195 (2021) 31–47.
[5] D. Ilijazi, S.F. Shariat, M.R. Hassler, U. Lemberger, I.E. Ertl, Epigenetic alterations of testicular germ cell tumours, Curr. Opin. Urol. 30 (2020) 264–270.
[6] P. Chieffi, M. De Martino, F. Esposito, Further insights into testicular germ cell tumor oncogenesis: potential therapeutic targets, Expert Rev. Anticancer Ther. 20 (2020) 189–195.
[7] H. Goldberg, Z. Klaassen, T. Chandrasekar, N. Fleshner, R.J. Hamilton, M.A.S. Jewett, Germ cell testicular tumors-contemporary diagnosis, staging and management of localized and advanced disease, Urology 125 (2019) 8–19.
[8] M. Barchi, P. Bielli, S. Dolci, P. Rossi, P. Grimaldi, Non-coding RNAs and splicing activity in testicular germ cell tumors, Life (2021) 11.
[9] K. Kalavska, L. Kucerova, S. Schmidtova, M. Chovanec, M. Mego, Cancer stem cell niche and immune-active tumor microenvironment in testicular germ cell tumors, Adv. Exp. Med. Biol. 1226 (2020) 111–121.
[10] M. Chovanec, J. Mardiak, M. Mego, Immune mechanisms and possible immune therapy in testicular germ cell tumours, Andrology 7 (2019) 479–486.
[11] H. Shen, J. Shih, D.P. Hollern, L. Wang, R. Bowlby, S.K. Tickoo, et al., Integrated molecular characterization of testicular germ cell tumors, Cell Rep. 23 (2018) 3392–3406.
[12] Y. Song, X. Qi, J. Kang, X. Wang, N. Ou, J. Zhu, et al., Identification of new biomarkers in immune microenvironment of testicular germ cell tumour, Andrologia 53 (2021) e13986.
[13] H. Bo, F. Zhu, Z. Liu, Q. Deng, G. Liu, R. Li, et al., Integrated analysis of high-throughput sequencing data reveals the key role of LINC00467 in the invasion and metastasis of testicular germ cell tumors, Cell Death Discov 7 (2021) 206.
[14] J. Gliozzo, M. Mesiti, M. Notaro, A. Petrini, A. Patak, A. Puertas-Gallardo, et al., Heterogeneous data integration methods for patient similarity networks, Brief Bioinform 23 (2022).
[15] A. Mayakonda, D.C. Lin, Y. Assenov, C. Plass, H.P. Koeffler, Maftools: efficient and comprehensive analysis of somatic variants in cancer, Genome Res. 28 (2018) 1747–1756.
[16] Y. Zhou, B. Zhou, L. Pache, M. Chang, A.H. Khodabakhshi, O. Tanaseichuk, et al., Metascape provides a biologist-oriented resource for the analysis of systems-level datasets, Nat. Commun. 10 (2019) 1523.
[17] M.D. Wilkerson, D.N. Hayes, ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking, Bioinformatics 26 (2010) 1572–1573.
[18] M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, et al., Limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res. 43 (2015) e47.
[19] G. Yu, L.G. Wang, Y. Han, Q.Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters, OMICS 16 (2012) 284–287.
[20] Z. Tang, C. Li, B. Kang, G. Gao, C. Li, Z. Zhang, GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses, Nucleic Acids Res. 45 (2017) W98–W102.
[21] W. Yang, J. Soares, P. Greninger, E.J. Edelman, H. Lightfoot, S. Forbes, et al., Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells, Nucleic Acids Res. 41 (2013) D955–D961.
[22] G. Sturm, F. Finotello, F. Petitprez, J.D. Zhang, J. Baumbach, W.H. Fridman, et al., Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology, Bioinformatics 35 (2019) i436–i445.
[23] T. Li, J. Fu, Z. Zeng, D. Cohen, J. Li, Q. Chen, et al., TIMER2.0 for analysis of tumor-infiltrating immune cells, Nucleic Acids Res. 48 (2020) W509–W514.
[24] T.M. Malta, A. Sokolov, A.J. Gentles, T. Burzykowski, L. Poisson, J.N. Weinstein, et al., Machine learning identifies stemness features associated with oncogenic dedifferentiation, Cell 173 (2018) 338–354 e15.
[25] A. Minelli, C. Scassellati, C.R. Cloninger, E. Tessari, M. Bortolomasi, C. Bonvicini, et al., PCLO gene: its role in vulnerability to major depressive disorder, J. Affect. Disord. 139 (2012) 250–255.
[26] L. Yin, L. Zhou, R. Xu, Identification of tumor mutation burden and immune infiltrates in hepatocellular carcinoma based on multi-omics analysis, Front. Mol. Biosci. 7 (2020) 599142.
[27] J.A. Espinoza, I. Riquelme, E.A. Sagredo, L. Rosa, P. Garcia, C. Bizama, et al., Mucin 5B, carbonic anhydrase 9 and claudin 18 are potential theranostic markers of gallbladder carcinoma, Histopathology 74 (2019) 597–607.
[28] K. Guda, H. Moinova, J. He, O. Jamison, L. Ravi, L. Natale, et al., Inactivating germ-line and somatic mutations in polypeptide N-acetylgalactosaminyltransferase 12 in human colon cancers, Proc Natl Acad Sci U S A 106 (2009) 12921–12925.
[29] E. Kim, L. Cai, S.H. Hyun, Effects of stem cell factor/c-kit signaling on in vitro maturation of porcine oocytes and subsequent developmental competence after fertilization, Front. Vet. Sci. 8 (2021) 745488.
[30] A. Kaushik, S. Anand, D. Bhartiya, Altered biology of testicular VSELs and SSCs by neonatal endocrine disruption results in defective spermatogenesis, reduced fertility and tumor initiation in adult mice, Stem Cell Rev Rep 16 (2020) 893–908.

[31] D.A. Landero-Huerta, R.M. Vigueras-Villasenor, E. Yokoyama-Rebollar, F. Garcia-Andrade, J.C. Rojas-Castaneda, L.A. Herrera-Montalvo, et al., Cryptorchidism and testicular tumor: comprehensive analysis of common clinical features and search of SNVs in the KIT and AR genes, Front. Cell Dev. Biol. 8 (2020) 762.
[32] M.T. Sung, G.T. Maclennan, A. Lopez-Beltran, S. Zhang, R. Montironi, L. Cheng, Primary mediastinal seminoma: a comprehensive assessment integrated with histology, immunohistochemistry, and fluorescence in situ hybridization for chromosome 12p abnormalities in 23 cases, Am. J. Surg. Pathol. 32 (2008) 146–155.
[33] L. Huang, Z. Guo, F. Wang, L. Fu, KRAS mutation: from undruggable to druggable in cancer, Signal Transduct Target Ther 6 (2021) 386.
[34] D. Tang, G. Kroemer, R. Kang, Oncogenic KRAS blockade therapy: renewed enthusiasm and persistent challenges, Mol. Cancer 20 (2021) 128.
[35] G. de Vries, X. Rosas-Plaza, M. van Vugt, J.A. Gietema, S. de Jong, Testicular cancer: determinants of cisplatin sensitivity and novel therapeutic opportunities, Cancer Treat Rev. 88 (2020) 102054.
[36] C. Grasso, M. Popovic, E. Isaevska, F. Lazzarato, V. Fiano, D. Zugna, et al., Association study between polymorphisms in DNA methylation-related genes and testicular germ cell tumor risk, Cancer Epidemiol. Biomarkers Prev. 31 (2022) 1769–1779.
[37] J. Guo, S. Wang, Z. Jiang, L. Tang, Z. Liu, J. Cao, et al., Long non-coding RNA RFPL3S functions as a biomarker of prognostic and immunotherapeutic prediction in testicular germ cell tumor, Front. Immunol. 13 (2022) 859730.
[38] R. Boldrini, M.D. De Pasquale, O. Melaiu, M. Chierici, G. Jurman, M.C. Benedetti, et al., Tumor-infiltrating T cells and PD-L1 expression in childhood malignant extracranial germ-cell tumors, OncoImmunology 8 (2019) e1542245.
[39] S. Zschabitz, F. Lasitschka, B. Hadaschik, R.D. Hofheinz, K. Jentsch-Ullrich, M. Gruner, et al., Response to anti-programmed cell death protein-1 antibodies in men treated for platinum refractory germ cell cancer relapsed after high-dose chemotherapy and stem cell transplantation, Eur. J. Cancer 76 (2017) 1–7.
[40] J.L. Boormans, J. Mayor de Castro, L. Marconi, Y. Yuan, M.P. Laguna Pes, C. Bokemeyer, et al., Testicular tumour size and rete testis invasion as prognostic factors for the risk of relapse of clinical stage I seminoma testis patients under surveillance: a systematic review by the testicular cancer guidelines panel, Eur. Urol. 73 (2018) 394–405.