



OPEN ACCESS

EDITED BY

Ming Jun Zheng,
Ludwig Maximilian University of Munich,
Germany

REVIEWED BY

Zhe Wang,
Third Military Medical University, China
Dandan Yuan,
The Second Affiliated Hospital of Harbin
Medical University, China

*CORRESPONDENCE

Xiaoxue Li,
✉ lixiaoxue168211069@csu.edu.cn

[†]These authors have contributed equally
to this work and share first authorship

SPECIALTY SECTION

This article was submitted to
Computational Genomics,
a section of the journal
Frontiers in Genetics

RECEIVED 05 January 2023

ACCEPTED 09 March 2023

PUBLISHED 17 March 2023

CITATION

Hu D, Zhang Z, Zhang Y, Huang K and Li X
(2023), Identification of immune related
molecular subtypes and prognosis model
for predicting prognosis, drug resistance
in cervical squamous cell carcinoma.
Front. Genet. 14:1137995.
doi: 10.3389/fgene.2023.1137995

COPYRIGHT

© 2023 Hu, Zhang, Zhang, Huang and Li.
This is an open-access article distributed
under the terms of the [Creative
Commons Attribution License \(CC BY\)](#).
The use, distribution or reproduction in
other forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Identification of immune related molecular subtypes and prognosis model for predicting prognosis, drug resistance in cervical squamous cell carcinoma

Dongzhi Hu^{1†}, Zijian Zhang^{2†}, Yongjing Zhang³, Kangni Huang¹
and Xiaoxue Li^{3*}

¹Department of Obstetrics and Gynecology, Yiyang Central Hospital, Yiyang, China, ²Department of General Surgery, The Second Xiangya Hospital of Central South University, Changsha, China, ³Department of Obstetrics and Gynecology, The Second Xiangya Hospital of Central South University, Changsha, China

Background: One of the features of tumor immunity is the immunosuppressive tumor microenvironment (TME). In this study, TME gene signatures were used to define the characteristics of Cervical squamous cell carcinoma (CESC) immune subtypes and construct a new prognostic model.

Methods: Single sample gene set enrichment analysis (ssGSEA) was used to quantify pathway activity. RNA-seq of 291 CESC were obtained from the Cancer Genome Atlas (TCGA) database as a training set. Microarray-based data of 400 cases of CESC were obtained from the Gene Expression Compilation (GEO) database as an independent validation set. 29 TME related gene signatures were consulted from previous study. Consensus Cluster Plus was employed to identify molecular subtype. Univariate cox regression analysis and random survival forest (RSF) were used to establish the immune-related gene risk model based on the TCGA data set of CESC, and the accuracy of prognosis prediction was verified by GEO data set. ESTIMATE algorithm was used to perform immune and matrix scores on the data set.

Results: three molecular subtypes (C1, C2, C3) were screened in TCGA-CESC on account of 29 TME gene signatures. Among, C3 with better survival outcome had higher immune related gene signatures, while C1 with worse prognosis time had enhanced matrix related features. Increased immune infiltration, inhibition of tumor related pathways, widespread genomic mutations and prone immunotherapy were observed in C3. Furthermore, a five immune genes signature was constructed and predicted overall survival for CESC, which successfully validated in GSE44001 dataset. A positive phenomenon was observed between five hub genes expressions and methylation. Similarly, high group enriched in matrix related features, while immune related gene signatures were enriched in low group. Immune cell, immune checkpoints genes expression levels were negatively, while most TME gene signatures were positively correlated with Risk Score. In addition, high group was more sensitive to drug resistance.

Conclusion: This work identified three distinct immune subtypes and a five genes signature for predicting prognosis in CESC patients, which provided a promising treatment strategy for CESC.

KEYWORDS

tumor microenvironment, endocervical adenocarcinoma, immunotherapy, genomic mutations, prognosis

1 Introduction

Globally, there are more than 500,000 new cases of cervical cancer every year, and about 300,000 deaths from cervical cancer, and its incidence and mortality rank the fourth place in female malignant tumors (Bray et al., 2018). Large-scale promotion of HPV vaccination and early screening and diagnosis of cervical cancer has reduced the disease burden of patients to some extent. The traditional treatment, mainly surgery and supplemented by chemoradiotherapy, has a good effect on the treatment of early cervical cancer, but the 5-year survival rate of advanced, metastatic, and recurrent cervical cancer is less than 20% (Pfaendler and Tewari, 2016; Tewari et al., 2017).

Tumor microenvironment (TME) is the cellular environment in which tumor cells reside, which is composed of immune cells, mesenchymal cells, endothelial cells, inflammatory mediators, and extracellular matrix (ECM) (Hanahan and Weinberg, 2011; Hanahan and Coussens, 2012). The cells and molecules in TME are in a dynamic process that reflects the evolutionary nature of cancer and work together to promote immune escape, growth, and metastasis of tumors (Jiang et al., 2019; Ren et al., 2020). Immune cells and stromal cells are two major types of non-tumor components, which are considered to have important value in the diagnosis and prognosis of tumors (Zhu et al., 2021). In recent years, immunotherapy is a new means of tumor treatment. Its mechanism is to significantly improve the survival time by reactivating the anti-tumor immune system to strongly and continuously kill tumor cells. Currently, the most comprehensive immunotherapy is immune checkpoint inhibitor, whose representative drug is programmed death protein 1 (PD-1) inhibitor (Pembrolizumab), which has been proved effective in a variety of cancers, but the overall objective effective rate is only 20%–30% (Iwai et al., 2017). Currently, the molecular targets used to guide immunotherapy are mainly limited to the expression level of programmed death protein ligand 1 (PD-L1), high microsatellite instability (MSI-H) (Le et al., 2015), mismatch fixes system defects (dMMR) (Le et al., 2017), Tumor mutation burden (TMB) (Goodman et al., 2017; Yarchoan et al., 2017). TMB to predict the inaccurate treatment response of immunosuppressive agents in some cancer patients. Therefore, it is particularly important to screen more reasonable molecular markers to guide immunotherapy through comprehensive analysis of tumor microenvironment.

In view of this, this study obtained CESC expression profile data through The Cancer Genome Atlas (TCGA) database, and analyzed the relationship between immune pathway score and survival prognosis of patients with CESC by ssGSEA algorithm. Combined with the data set from the Gene Expression Omnibus (GEO) database (GSE44001), differentially expressed genes, (DEGs) analysis, functional enrichment and survival analysis were performed to screen out hub genes to construct prognostic models, and to explore the relevance of prognostic models in predicting the prognosis of patients with CESC and

immunotherapy, so as to provide references for the research of biomarkers related to CESC immunity and immunotherapy.

2 Materials and methods

2.1 Data acquisition and preprocessing

Using “CESC”, “transcriptome profiling (transcripts per million (TPM))”, and “Gene Expression Quantification” as search terms, the results can be obtained from the TCGA database to download a sequence dataset containing 291 CESC tissues and corresponding clinical information. Using “cervical cancer” as a keyword in the GEO database. The gene-chip dataset GSE44001 contains 300 CESC tissues was downloaded.

For TCGA-CESC, the sample with clinical information, survival time greater than 0 and Status (alive and death) is retained and Ensembl is converted into Gene symbol, and the expression with multiple Gene Symbol is the median value. For the GSE44001 dataset, probes are mapped to genes based on annotation information, and probes that match one probe to multiple genes are removed. When multiple probes matched a gene, the mean value was taken as the expression value of the gene.

2.2 ssGSEA analysis

Twenty nine TME related gene signatures, covering known cellular and functional TME properties, were extracted from previously study (Bagaev et al., 2021). A total 257 genes were found in 29 gene signatures and ssGSEA using GSVA package (Yi et al., 2020) was employed to quantitate TME score.

2.3 Sample cluster analysis

ConsensusClusterPlus (Wilkerson and Hayes, 2010) was employed to construct consistency matrix for TCGA-CESC samples clustering on account on 29 TME gene signatures scores. 80% samples were carried out 500 bootstraps using km algorithm and distance of 1-pearson correlation. Number of Clusters was set as 2–10 and optimal cluster number was determined in terms of consistency matrix and cumulative distribution function. Principal component analysis (PCA) was also performed to test rationality of molecular subtype distribution.

2.4 Evaluation of immune infiltration

CIBERSORT algorithm (<https://cibersort.stanford.edu/>) was used to quantify the relative abundance of 22 types of immune cells in CESC. At the same time, the ESTIMATE (Yoshihara et al., 2013) software was used to calculate the proportion of immune cells.

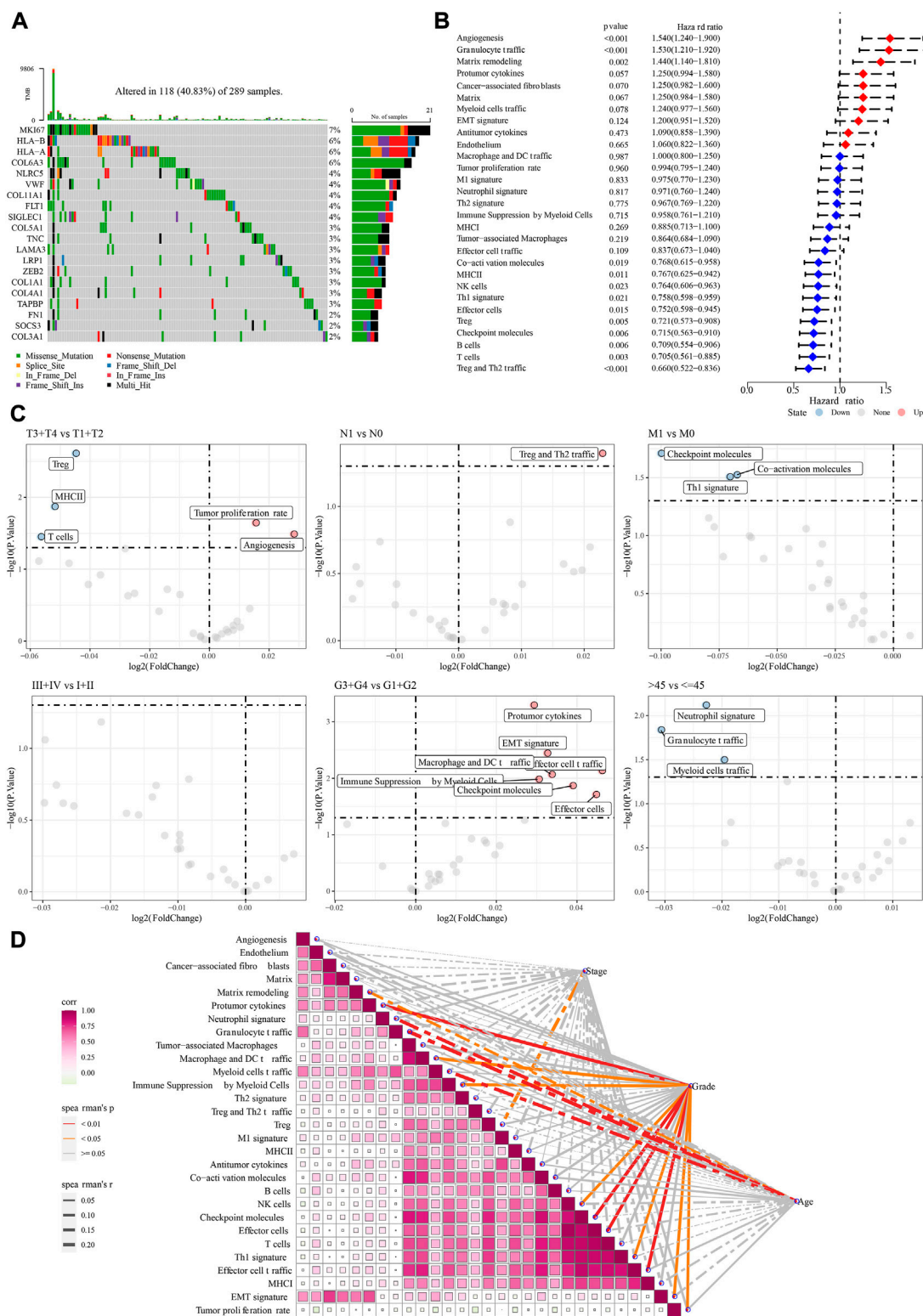


FIGURE 1

The association between 29 TME gene signatures and clinical characteristics in TCGA-CESC patients. **(A)**: Top20 TME related genes mutations in TCGA-CESC dataset. **(B)**: Univariate cox regression analysis of 29 TME gene signatures. **(C)**: The differences of 29 TME gene signatures among clinical feature grouping. **(D)**: Correlation between 29 TME gene signatures with each other as well as stage, grade, and age. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$; ns: no significance.

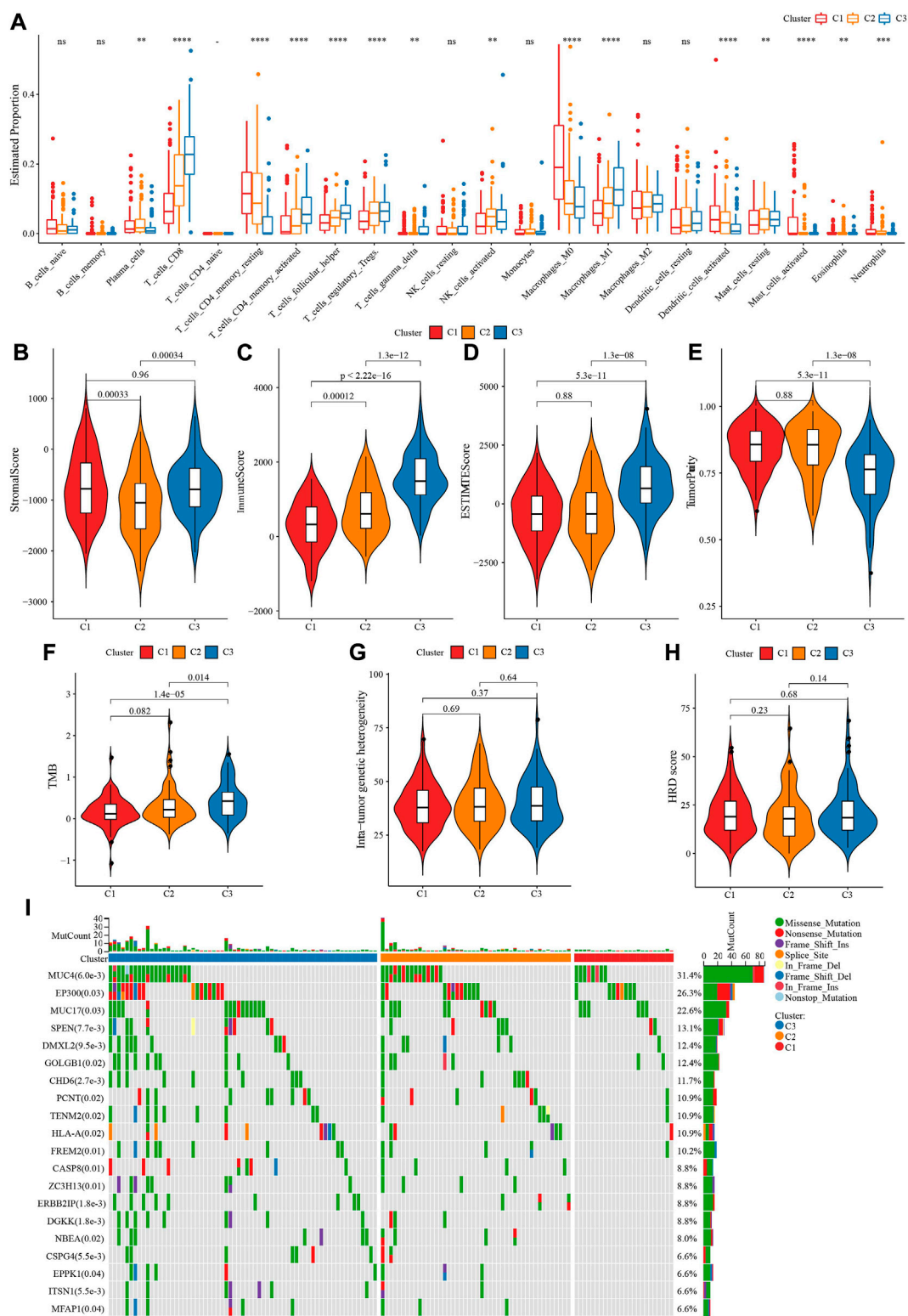
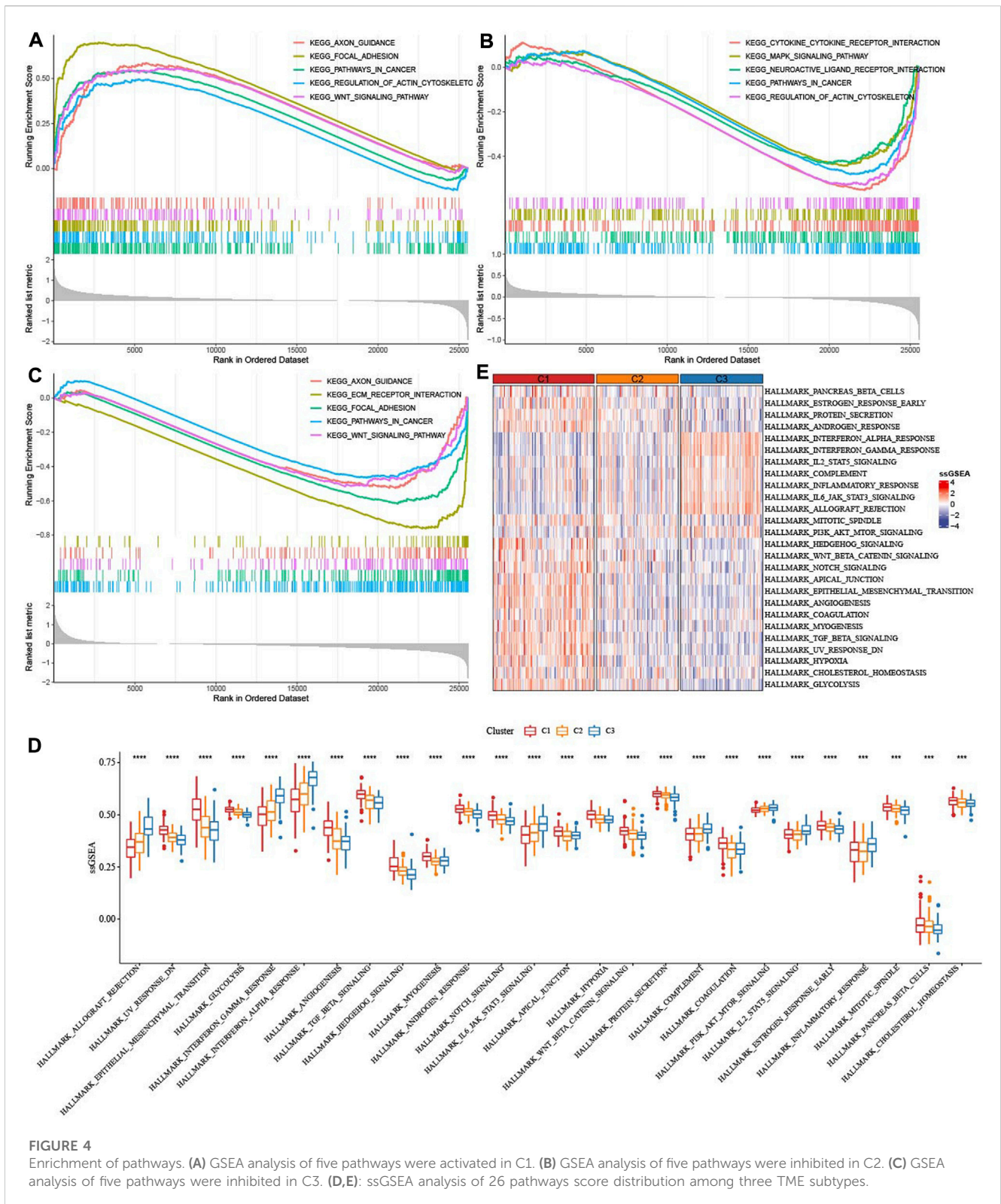


FIGURE 3

Immune infiltration analysis among three TME subtypes. **(A)** CIBERSORT analysis of 22 immune cells distribution among three TME subtypes. **(B)** The difference of StromalScore among three TME subtypes. **(C)** The difference of ImmuneScore among three TME subtypes. **(D)** The difference of ESTIMATEScore among three TME subtypes. **(E)** The difference of TumorPurity among three TME subtypes. **(F)** The difference of TMB among three TME subtypes. **(G)** The difference of Intra-tumor genetic heterogeneity among three TME subtypes. **(H)** The difference of HRD score among three TME subtypes. **(I)** Somatic cell mutation among three TME subtypes. *** $p < 0.001$; **** $p < 0.0001$.



TIDE (Xiao et al., 2018; Fu et al., 2020) software was used to evaluate the potential clinical effects of immunotherapy included dysfunction of tumor infiltration cytotoxic T lymphocytes (CTL) (Dysfunction) and exclusion of CTL (Exclusion), M2 subtype of

tumor-associated fibroblasts (CAF), tumor-associated macrophages (TAM), myeloid-derived suppressor cells (MDSCs), a higher TIDE predictive score indicates a greater likelihood of immune escape, suggesting that patients are less likely to benefit from immunotherapy.

pRRophetic (Geeleher et al., 2014) was used to predict the sensitivity of traditional medicines to half maximal inhibitory concentration (IC50).

2.7 Construction and validation of prognosis model

Among molecular subtypes, limma analysis (Ritchie et al., 2015) and univariate cox regression analysis were implemented to screen genes affecting CESC prognosis ($p < 0.05$). Random Forest SRC package was introduced to construct a random forest model and the most highly predictive variables were screened when variable importance (VIMP) value > 0.4 . Finally, the optimal genes were used to construct Risk score using stepAIC method in MASS package.

$$\text{Risk score} = \sum \text{coef}_i * \text{Exp}_i$$

Exp_i is the expression level of genes, and coef_i is the regression correlation coefficient.

Survminer package was conducted to determine optimal cutoff to divided CESC samples into high group and low group. KM survival and ROC analysis using timeROC package were used to predict performance of Risk score. TCGA-CSEC was a training dataset and GSE44001 dataset was acted as independently validate dataset.

2.8 Statistical analysis

R (4.0.2) software was used for statistical analysis. WebGestaltR package (Yu et al., 2012) was used to carry out functional enrichment analysis. Genetic mutations were determined by maftools. Wilcoxon non-parametric rank sum test was used to analyze the differences. $p < 0.05$ was considered to be statistically significant. Sangerbox was used for analysis (Shen et al., 2022).

3 Results

3.1 29 TME gene signatures was association with clinical characteristics for TCGA-CESC samples

As we know, somatic mutations could lead to carcinogenesis. 199 of 257 genes (from 29 TME gene signature) were mutated and Top20 genes mutation rate were showed, among, MKI67 had highest mutation rate (7%) (Figure 1A). Univariate cox regression analysis of 29 TME gene signatures found 13 TME gene signatures affecting prognosis of CESC samples (Figure 1B). The differences of TME gene signatures scores in clinical features indicated that Tumor proliferation rate, Angiogenesis scores were increased in T3 + T4 stage, Protumor cytokines, Macrophage and DC traffic, Effector cell traffic, Immune Suppression by Myeloid Cells, Effector cells scores were enhanced in G3 + G4 stage (Figure 1C). TME gene characteristics were positively correlated with each other and with Grade (Figure 1D).

3.2 Identification of three molecular subtypes

Based on 29 TME gene signature, TCGA-CESC samples were divided into three molecular subtypes when $k = 3$ on account of CDF and CDF delta area (Supplementary Figure S1A–C). KM survival curve showed that the overall survival and progression-free survival (PFS) in C3 had longest time, followed by C2 and C1 (Figures 2A, B). PCA suggested that the three molecular subtypes have distinct regional divisions (Figure 2C). The distribution of TME gene signatures among three molecular subtypes indicated that immune related gene signatures, such as Treg and Th2 traffic, Antitumor cytokines, were enriched in C3, while Matrix related gene signatures, such as Angiogenesis, Endothelium, Cancer-associated fibroblasts, Matrix, Matrix remodeling, were enriched in C1 (Figure 2D). TNM stage also had distribution differences among three molecular subtypes (Figure 2E).

3.3 Differences of infiltration of immune cells and somatic cell mutation among three molecular subtypes

In TCGA-CESC dataset, CIBERSORT analysis on 22 immune cells showed 17 of which had statistical significance among three molecular subtypes, such as T_cells_CD8, T_cells_CD4_memory_activated were involved in C3, while C1 enriched in T_cells_CD4_memory_resting, Macrophages_M0, Dendritic_cells_activated (Figure 3A). ESTIMATE analysis demonstrated that C3 had enhanced StromalScore, ImmuneScore, ESTIMATEScore, while TumprPurity was lowest (Figures 3B–E). Changes in genome among three molecular subtypes were explored, and we found that C3 harbored a significantly higher TMB (Figure 3F). No statistical significance of Mutant-allele tumor heterogeneity and HRD score among subtypes were observed (Figures 3G, H). In addition, MUC4, EP300, MUC17 genes had a wide range of somatic mutations in CESC (Figure 3I).

3.4 Functional characterization of three molecular subtypes

GSEA analysis indicated that axon guidance, focal adhesion, pathways in cancer, regulation of actin cytoskeleton, WNT signaling pathway were activated in C1 (Figure 4A), Cytokine-cytokine receptor interaction, MAPK signaling pathway, Neuroactive ligand receptor interaction, pathway in cancer were inhibited in C2 (Figure 4B), Axon guidance, ECM receptor interaction, pathway in cancer and WNT signaling pathway were inhibited in C3 (Figure 4C). ssGSEA analysis of 26 pathways scores had difference among three molecular subtypes. EMT-related pathways such as HALLMARK_WNT_BETA_CATENIN_SIGNALING were enriched in C1, in addition, the C3 subtype is significantly enriched in some immune-related pathways such as HALLMARK_INTERFERON_ALPHA_RESPONSE and HALLMARK_INTERFERON_GAMMA_RESPONSE (Figures 4D, E). Those data suggested that C3 presented immunoinfiltration state, and Cell growth-related pathways were activated in C1.

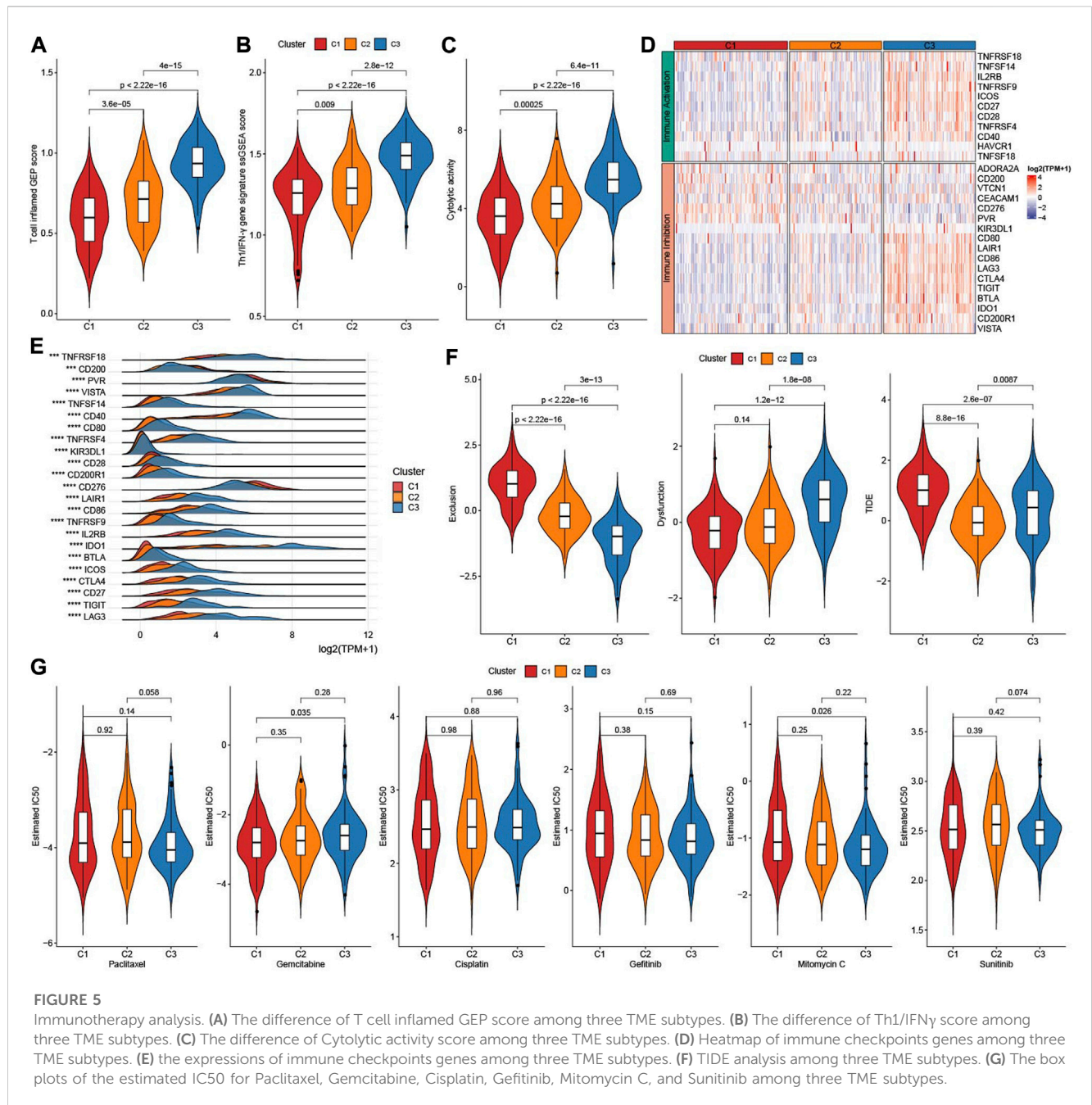


FIGURE 5

Immunotherapy analysis. (A) The difference of T cell inflamed GEP score among three TME subtypes. (B) The difference of Th1/IFN γ score among three TME subtypes. (C) The difference of Cytolytic activity score among three TME subtypes. (D) Heatmap of immune checkpoints genes among three TME subtypes. (E) the expressions of immune checkpoints genes among three TME subtypes. (F) TIDE analysis among three TME subtypes. (G) The box plots of the estimated IC50 for Paclitaxel, Gemcitabine, Cisplatin, Gefitinib, Mitomycin C, and Sunitinib among three TME subtypes.

3.5 Analysis of immunotherapy and chemotherapy among three molecular subtypes

As showed in Figures 5A, B, Figure 3 factors (T cell inflamed GEP score, Th1/IFN γ gene signature score, and Cytolytic activity score) that predict immunotherapy effect were all elevated in subtype C3 (Figures 5A–C). Given that immune checkpoint blockade (ICB) is a key factor for cancer immunotherapy, we evaluated a few representative genes. Most immune inhibition genes and activation genes were upregulated in C3 (Figure 5D). Moreover, 23 immune checkpoint genes had highest expressions in C3 (Figure 5E). Exclusion score and TIDE score were significantly

highest in C1, while Dysfunction score was highest in C3 (Figure 5F). Sensitivity analysis of molecular subtypes to traditional chemotherapy drugs showed C3 was more sensitive to Paclitaxel, Mitomycin C, and C1 maybe benefit from Gemcitabine (Figure 5G).

3.6 Construction and validation of risk model

Firstly, DEGs were screened among three molecular subtypes, which 165 upregulated genes and 96 downregulated genes in C1 (Supplementary Figure S2A), 216 increased genes and 106 decreased

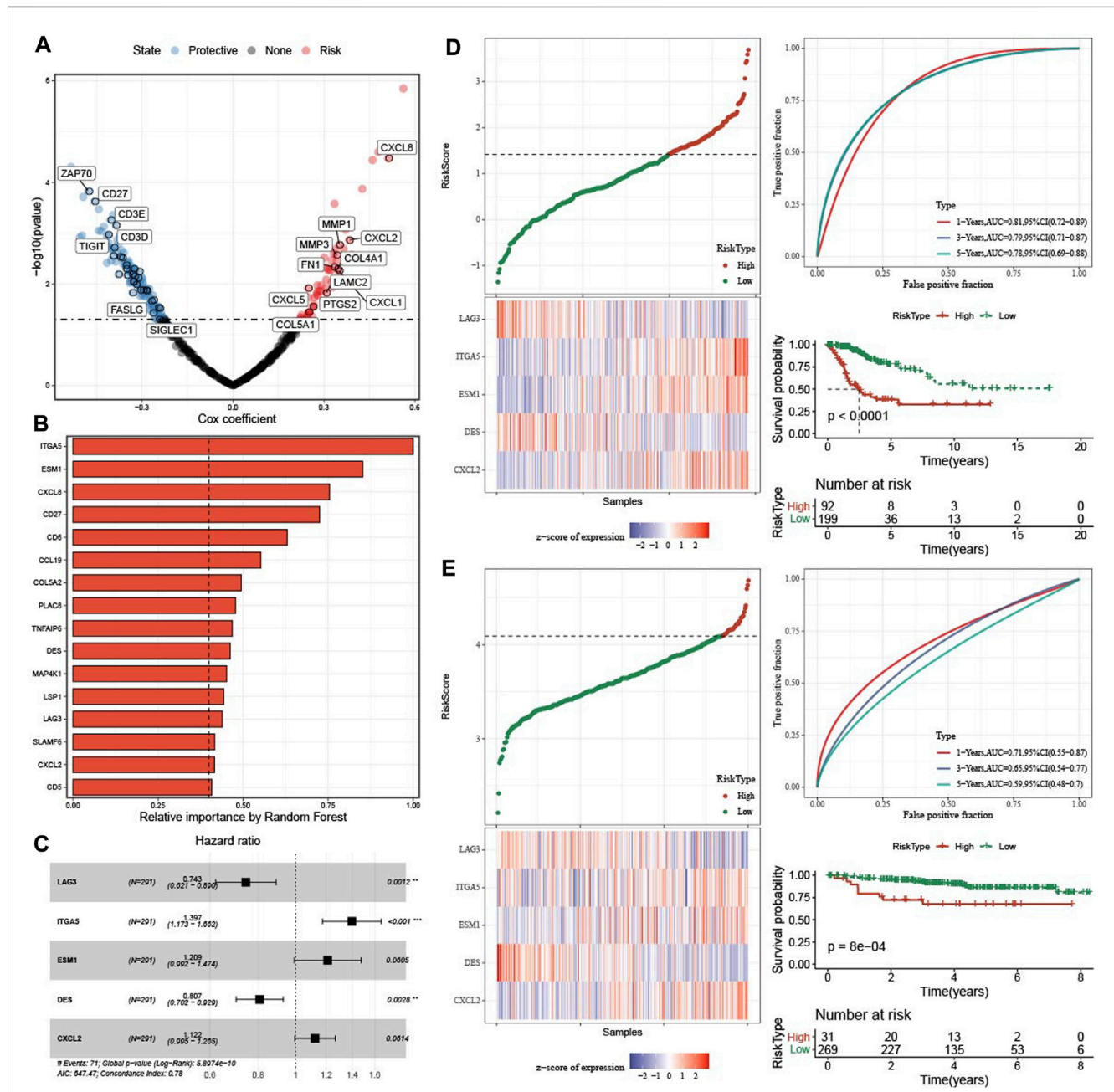


FIGURE 6

Construction and validation of prognosis model. (A) Univariate cox regression analysis of TME related genes. (B) The 16 most predictive genes selected by random survival forest. (C) Univariate cox regression analysis of five hub genes. (D) The distribution of RiskScore, expression of five hub genes in TCGA-CESC dataset. ROC analysis and AUC of RiskScore in TCGA-CESC dataset. KM survival curve of high group and low group in TCGA-CESC dataset. (E) The distribution of RiskScore, expression of five hub genes in GSE44001 dataset. ROC analysis and AUC of RiskScore in GSE44001 dataset. KM survival curve of high group and low group in GSE44001 dataset.

genes in C3 (Supplementary Figure S2B). Finally, 429 DEGs were found among three molecular subtypes (Supplementary Figure S2C). 186 genes affecting prognosis of CESC samples were screened from 429 genes (Figure 6A). 186 genes were reduced to 16 genes using a random forest model (Figure 6B). Finally, five hub gene were determined from 16 genes by stepAIC method (Figure 6C). RiskScore = $-0.297 \times \text{LAG3} + 0.334 \times \text{ITGA5} + 0.19 \times \text{ESM1} - 0.214 \times \text{DES} + 0.115 \times \text{CXCL2}$. The five

hub genes expressions were positively correlated with methylation levels (Supplementary Figure S3).

In TCGA-CESC dataset, the distribution of RiskScore and five genes expression were showed. 1-, 3-, five- year AUC was 0.81, 0.79, and 0.78 respectively, and patients in high group had worse survival time (Figure 6D). In GSE44001 queue, the 1-, three-, and five- year AUC was 0.71, 0.65, and 0.59, respectively, and samples in high group also had poor survival time (Figure 6E).

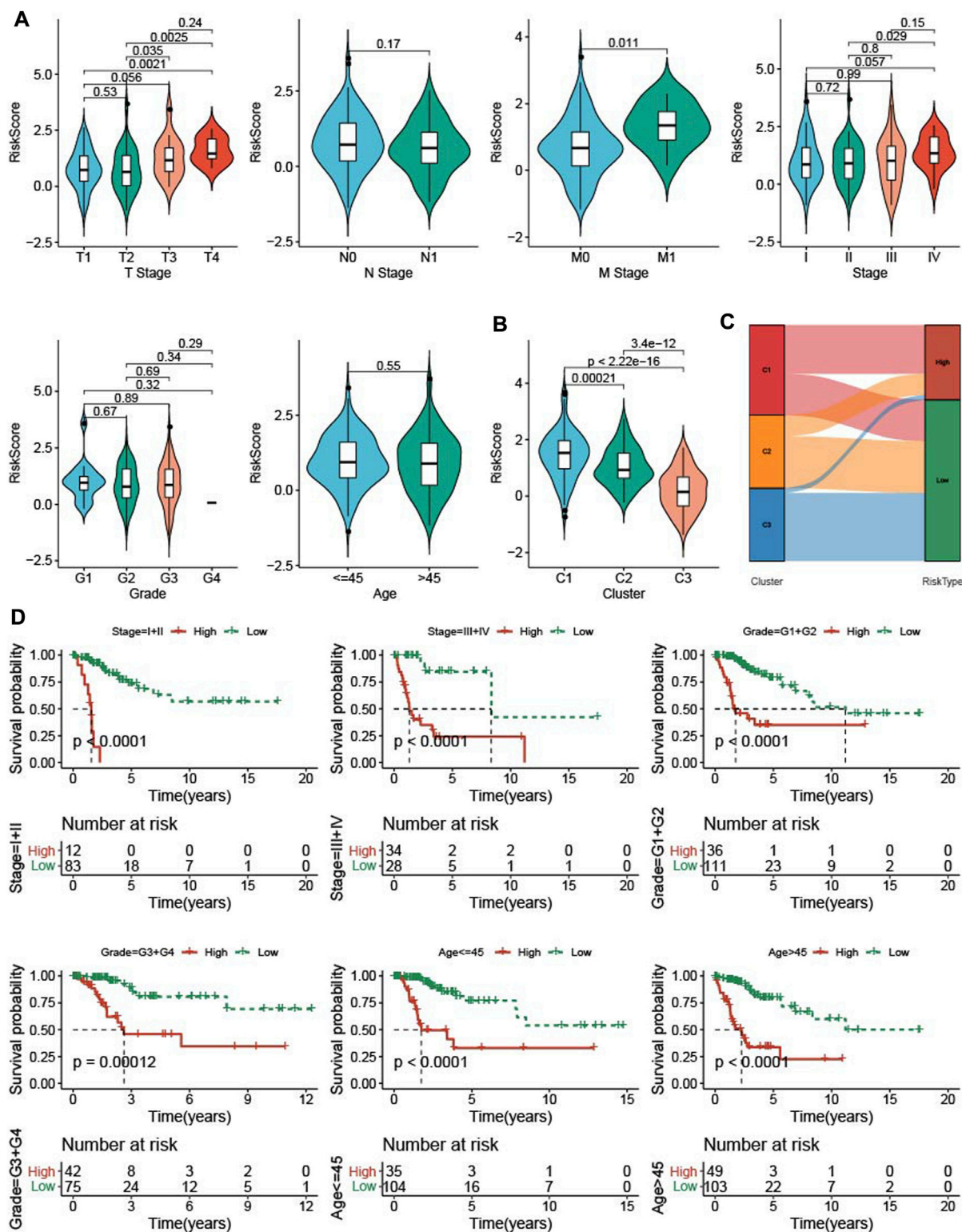
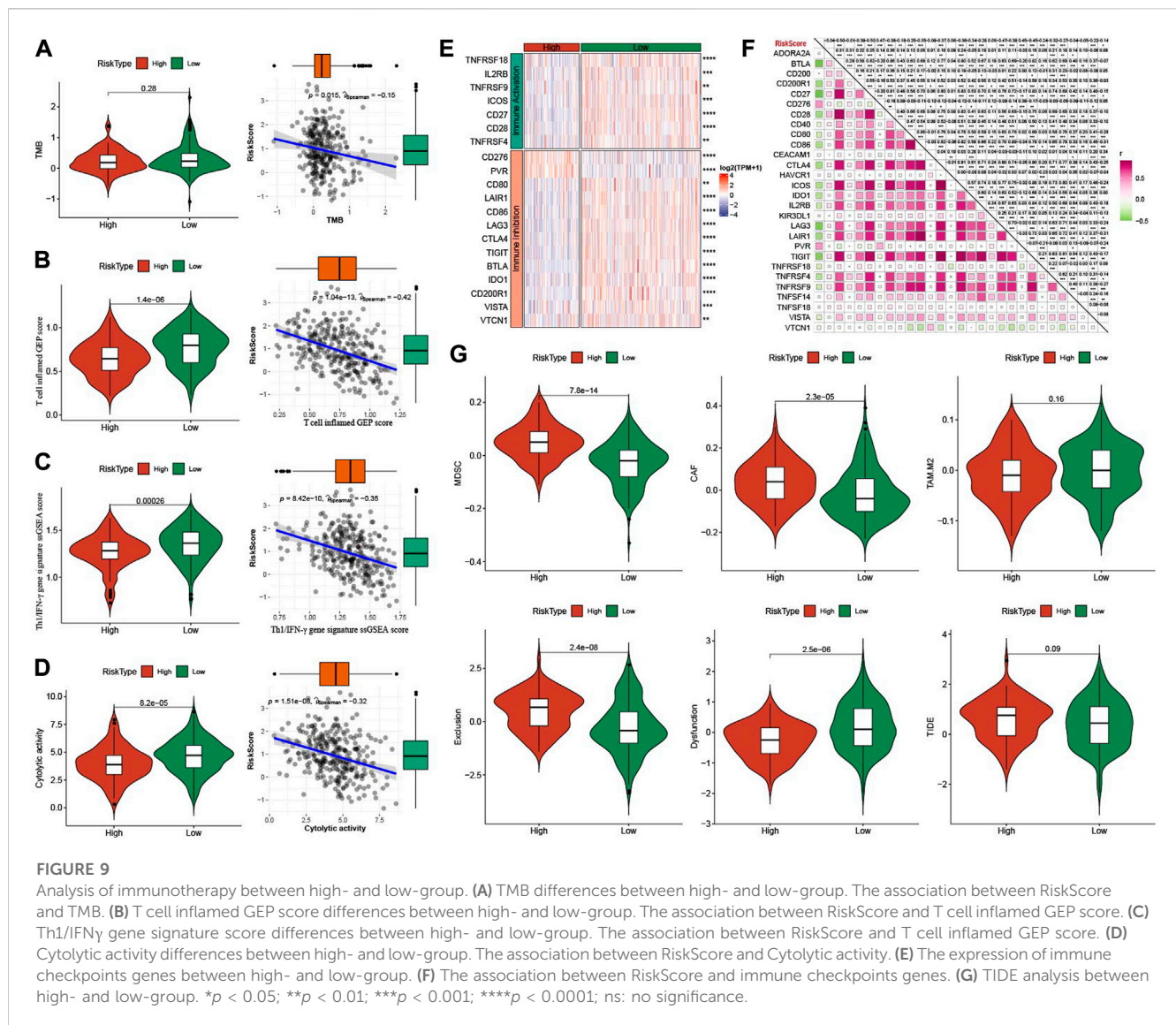


FIGURE 7

The RiskScore differences on samples with clinical features. (A) the differences of RiskScore in patients with various clinical features including T stage, N stage, M stage, stage, grade, and age. (B) the differences of RiskScore among three TME subtypes. (C) Matches of two subtypes and high- and low-groups. (D) KM survival of patients in high group and low group with various clinical features divided by RiskScore.



T_cells_follicular_helper, and Macrophages_M1. However, there were significant positive correlations with Angiogenesis, Matrix, Matrix remodeling, Protumor cytokines, Myeloid cells traffic (Figure 8E).

3.9 Immunotherapy response of identified CESE subtypes

Firstly, we compared the TMB in high group and low group, which it had no significance between the two groups ($p = 0.28$), but there was a negatively association between RiskScore and TMB (Figure 9A). T cell inflamed GEP score, TH1/IFN γ gene signature score and Cytolytic activity score were all enhanced in low group, and all them were negatively correlated with RiskScore (Figures 9B–D). Immune checkpoint genes were higher expressions in low group and negatively correlated with RiskScore (Figures 9E, F). MDSC, CAF, TAM.M2, and Exclusion were decreased, while

Dysfunction was increased in low group in comparison to high group (Figure 9G).

4 Discussion

Studies have shown that CESC interstitium has a large number of immune cell infiltration (Dossus et al., 2013; Liu et al., 2020), Immune cell infiltration is believed to play an important role in the development of various malignant tumors (Hanahan and Coussens, 2012; Yoshihara et al., 2013), and immunotherapy has made great progress in the field of anti-tumor. In this study, it was found that in the 29 TME gene signatures, the higher the CESC pathological grade, the higher the infiltration of some TME gene signatures, and the infiltration abundance is related to the patient's prognosis. Targeted therapy targeting these immune cells is expected to improve the patient's prognosis.

In this study, based on 29 TME gene signatures, the TCGA-CESC cohort samples were divided into three immune subtypes (C1, C2, C3), which showed significant differences in prognosis, immune characteristics, pathway enrichment, gene mutation, and immunotherapy. C3 with good prognosis presented immunoinfiltration state, and cell growth-related pathways were activated in C1 accompanied by poor prognosis. Based on the three immune subtypes, the risk model was constructed by univariate Cox regression analysis and random survival forest model. We found that patients in the low-risk group had longer survival than those in the high-risk group, and there were significant differences in immunoinfiltration and immunotherapy.

In recent years, immune system therapies such as immune checkpoint inhibitors have shown remarkable effects in the field of anti-tumor. Studies have shown that highly mutated tumor genes can induce the production of a large number of neoantigens, which can activate immune cells and lead to a tumor-suppressing immune response (Büttner et al., 2019). MSI is closely related to the efficacy of tumor immunotherapy (Baretti and Le, 2018). Multiple studies have demonstrated that TMB, T cell inflamed GEP, TH1/IFN- γ , TIDE are emerging biomarkers for predicting the efficacy of tumor immunotherapy (Samstein et al., 2019). This study found that T cell inflamed GEP and TH1/IFN- γ scores were negatively correlated with RiskScore, and the low-risk group had a lower TIDE score. We speculated that patients in the low-risk group may benefit from immunotherapy.

Among the five key genes, ITGA5, ESM1, and CXCL2 were risk factors for the prognosis of CESC, while LAG3, and DES were protective factors. Multiple studies have shown that increased ITGA5 expression predicts poor prognosis of tumors, such as ovarian cancer (Gong et al., 2016), breast cancer (Xiao et al., 2018), and lung cancer (Zheng et al., 2016). CXCL2 expression level was closely related to lymph node metastasis and prognosis of cervical cancer patients (Zhang et al., 2018; Yang et al., 2021). Patients with high levels of LAG-3 peripheral t cells may suppress the antitumor response in a way that PD-1 or CTLA-4 blockers cannot overcome. LAG-3 has shown promise as a target in preclinical models, and drugs targeting LAG-3 are already in the early stages of clinical development, showing modest activity in unselected patient populations (Grosso et al., 2007; Brignone et al., 2009; Kraman et al., 2020). Two other genes including ESM1, and DES were little studied and their involvement in CESC remains largely unexplored, and more basic researches are needed to reveal their biological function in CESC.

There are some limitations in this study. First, it is necessary to verify the significance of hub genes in cancer tissues through experiments, such as RT-qPCR, IHC, and Western blot. Second, although our results show good predictive potential and clinical value of the five gene prognostic signature, prospective studies are needed to demonstrate the clinical application and prognostic value of this model in patients.

In this study, based on 29 TME gene signatures, we not only identified three subtypes and constructed a 5-key genes prognostic signature of CESC, which had a potential prognostic value. Those findings maybe provided prognosis prediction and precision treatment for clinicians.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

Conceptualization, DH and XL; Methodology, ZZ; Software, ZZ; Validation, XL, DH, and ZZ; Formal analysis, ZZ; Investigation, YZ and KH; Data curation, YZ; Writing—original draft preparation, XL; Writing—review and editing, DH and ZZ; Visualization, XL and DH; Supervision, YZ; Project administration, XL and ZZ; Funding acquisition, DH and XL. All authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by Clinical Medical Technology Innovation Guide Project of Hunan Province, grant number 2021SK52601.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2023.1137995/full#supplementary-material>

SUPPLEMENTARY FIGURE S1

Cluster analysis. A: Cumulative distribution function. B: Delta area of Cumulative distribution function. C: sample clustering in TCGA-CESC dataset when $k=3$.

SUPPLEMENTARY FIGURE S2

Identification of Differentially expressed gene. A: Differentially expressed gene in C1 vs C2/C3. B: Differentially expressed gene in C3 vs C1/C2. C: Differentially expressed gene among 3 molecular subtypes.

SUPPLEMENTARY FIGURE S3

The association of 5 hub genes expressions and methylation level.

References

- Ayers, M., Lunceford, J., Nebozhyn, M., Murphy, E., Loboda, A., Kaufman, D. R., et al. (2017). IFN- γ -related mRNA profile predicts clinical response to PD-1 blockade. *J. Clin. Invest.* 127, 2930–2940. doi:10.1172/JCI91190
- Bagaev, A., Kotlov, N., Nomic, K., Svekolkin, V., Gafurov, A., Isaeva, O., et al. (2021). Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell.* 39, 845–865. e7. doi:10.1016/j.ccell.2021.04.014
- Baretti, M., and Le, D. T. (2018). DNA mismatch repair in cancer. *Pharmacol. Ther.* 189, 45–62. doi:10.1016/j.pharmthera.2018.04.004
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68, 394–424. doi:10.3322/caac.21492
- Brignone, C., Escudier, B., Grygar, C., Marcu, M., and Triebel, F. (2009). A phase I pharmacokinetic and biological correlative study of IMP321, a novel MHC class II agonist, in patients with advanced renal cell carcinoma. *Clin. Cancer Res.* 15, 6225–6231. doi:10.1158/1078-0432.CCR-09-0068
- Büttner, R., Longshore, J. W., López-Ríos, F., Merkelbach-Bruse, S., Normanno, N., Rouleau, E., et al. (2019). Implementing TMB measurement in clinical practice: Considerations on assay requirements. *ESMO Open* 4, e000442. doi:10.1136/esmoopen-2018-000442
- Daniilova, L., Ho, W. J., Zhu, Q., Vithayathil, T., De Jesus-Acosta, A., Azad, N. S., et al. (2019). Programmed cell death ligand-1 (PD-L1) and CD8 expression profiling identify an immunologic subtype of pancreatic ductal adenocarcinomas with favorable survival. *Cancer Immunol. Res.* 7, 886–895. doi:10.1158/2326-6066.CIR-18-0822
- Dossus, L., Lukanova, A., Rinaldi, S., Allen, N., Cust, A. E., Becker, S., et al. (2013). Hormonal, metabolic, and inflammatory profiles and endometrial cancer risk within the EPIC cohort—a factor analysis. *Am. J. Epidemiol.* 177, 787–799. doi:10.1093/aje/kws309
- Fu, J., Li, K., Zhang, W., Wan, C., Zhang, J., Jiang, P., et al. (2020). Large-scale public data reuse to model immunotherapy response and resistance. *Genome Med.* 12, 21. doi:10.1186/s13073-020-0721-z
- Gao, Z., Tao, Y., Lai, Y., Wang, Q., Li, Z., Peng, S., et al. (2020). Immune cytolytic activity as an indicator of immune checkpoint inhibitors treatment for prostate cancer. *Front. Bioeng. Biotechnol.* 8, 930. doi:10.3389/fbioe.2020.00930
- Geeleher, P., Cox, N., and Huang, R. S. (2014). pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PLoS One* 9, e107468. doi:10.1371/journal.pone.0107468
- Gong, C., Yang, Z., Wu, F., Han, L., Liu, Y., and Gong, W. (2016). miR-17 inhibits ovarian cancer cell peritoneal metastasis by targeting ITGA5 and ITGB1. *Oncol. Rep.* 36, 2177–2183. doi:10.3892/or.2016.4985
- Goodman, A. M., Kato, S., Bazhenova, L., Patel, S. P., Frampton, G. M., Miller, V., et al. (2017). Tumor mutational burden as an independent predictor of response to immunotherapy in diverse cancers. *Mol. Cancer Ther.* 16, 2598–2608. doi:10.1158/1535-7163.MCT-17-0386
- Grosso, J. F., Kelleher, C. C., Harris, T. J., Maris, C. H., Hipkiss, E. L., De Marzo, A., et al. (2007). LAG-3 regulates CD8+ T cell accumulation and effector function in murine self- and tumor-tolerance systems. *J. Clin. Invest.* 117, 3383–3392. doi:10.1172/JCI31184
- Hanahan, D., and Coussens, L. M. (2012). Accessories to the crime: Functions of cells recruited to the tumor microenvironment. *Cancer Cell.* 21, 309–322. doi:10.1016/j.ccr.2012.02.022
- Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell.* 144, 646–674. doi:10.1016/j.cell.2011.02.013
- Iwai, Y., Hamanishi, J., Chamoto, K., and Honjo, T. (2017). Cancer immunotherapies targeting the PD-1 signaling pathway. *J. Biomed. Sci.* 24, 26. doi:10.1186/s12929-017-0329-9
- Jiang, X., Wang, J., Deng, X., Xiong, F., Ge, J., Xiang, B., et al. (2019). Role of the tumor microenvironment in PD-L1/PD-1-mediated tumor immune escape. *Mol. Cancer* 18, 10. doi:10.1186/s12943-018-0928-4
- Kraman, M., Farouqi, M., Allen, N. L., Kmieciak, K., Gliddon, D., Seal, C., et al. (2020). FS118, a bispecific antibody targeting LAG-3 and PD-L1, enhances T-cell activation resulting in potent antitumor activity. *Clin. Cancer Res.* 26, 3333–3344. doi:10.1158/1078-0432.CCR-19-3548
- Le, D. T., Durham, J. N., Smith, K. N., Wang, H., Bartlett, B. R., Aulakh, L. K., et al. (2017). Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 357, 409–413. doi:10.1126/science.aan6733
- Le, D. T., Uram, J. N., Wang, H., Bartlett, B. R., Kemberling, H., Eyring, A. D., et al. (2015). PD-1 blockade in tumors with mismatch-repair deficiency. *N. Engl. J. Med.* 372, 2509–2520. doi:10.1056/NEJMoa1500596
- Liu, J., Nie, S., Wu, Z., Jiang, Y., Wan, Y., Li, S., et al. (2020). Exploration of a novel prognostic risk signatures and immune checkpoint molecules in endometrial carcinoma microenvironment. *Genomics* 112, 3117–3134. doi:10.1016/j.ygeno.2020.05.022
- Pfaendler, K. S., and Tewari, K. S. (2016). Changing paradigms in the systemic treatment of advanced cervical cancer. *Am. J. Obstet. Gynecol.* 214, 22–30. doi:10.1016/j.ajog.2015.07.022
- Ren, Q., Zhu, P., Zhang, H., Ye, T., Liu, D., Gong, Z., et al. (2020). Identification and validation of stromal-tumor microenvironment-based subtypes tightly associated with PD-1/PD-L1 immunotherapy and outcomes in patients with gastric cancer. *Cancer Cell. Int.* 20, 92. doi:10.1186/s12935-020-01173-3
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-seq and microarray studies. *Nucleic Acids Res.* 43, e47. doi:10.1093/nar/gkv007
- Samstein, R. M., Lee, C. H., Shoushtari, A. N., Hellmann, M. D., Shen, R., Janjigian, Y. Y., et al. (2019). Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat. Genet.* 51, 202–206. doi:10.1038/s41588-018-0312-8
- Shen, W., Song, Z., Xiao, Z., Huang, M., Shen, D., Gao, P., et al. (2022). Sangerbox: A comprehensive, interaction-friendly clinical bioinformatics analysis platform. *iMeta* 1, e36. doi:10.1002/imt2.36
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102, 15545–15550. doi:10.1073/pnas.0506580102
- Tewari, K. S., Sill, M. W., Penson, R. T., Huang, H., Ramondetta, L. M., Landrum, L. M., et al. (2017). Bevacizumab for advanced cervical cancer: Final overall survival and adverse event analysis of a randomised, controlled, open-label, phase 3 trial (gynecologic oncology group 240). *Lancet* 390, 1654–1663. doi:10.1016/S0140-6736(17)31607-0
- Wilkerson, M. D., and Hayes, D. N. (2010). ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinformatics* 26, 1572–1573. doi:10.1093/bioinformatics/btq170
- Xiao, Y., Li, Y., Tao, H., Humphries, B., Jiang, Y., Li, A., et al. (2018). Integrin $\alpha 5$ down-regulation by miR-205 suppresses triple negative breast cancer stemness and metastasis by inhibiting the Src/Vav2/Rac1 pathway. *Cancer Lett.* 433, 199–209. doi:10.1016/j.canlet.2018.06.037
- Yang, P., Ruan, Y., Yan, Z., Gao, Y., Yang, H., and Wang, S. (2021). Comprehensive analysis of lymph nodes metastasis associated genes in cervical cancer and its significance in treatment and prognosis. *BMC Cancer* 21, 1230. doi:10.1186/s12885-021-08945-8
- Yarchoan, M., Hopkins, A., and Jaffee, E. M. (2017). Tumor mutational burden and response rate to PD-1 inhibition. *N. Engl. J. Med.* 377, 2500–2501. doi:10.1056/NEJMcl1713444
- Yi, M., Nissley, D. V., McCormick, F., and Stephens, R. M. (2020). ssGSEA score-based Ras dependency indexes derived from gene expression data reveal potential Ras addiction mechanisms with possible clinical implications. *Sci. Rep.* 10, 10258. doi:10.1038/s41598-020-66986-8
- Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-García, W., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* 4, 2612. doi:10.1038/ncomms3612
- Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omic* 16, 284–287. doi:10.1089/omi.2011.0118
- Zhang, W., Wu, Q., Wang, C., Yang, L., Liu, P., and Ma, C. (2018). AKIP1 promotes angiogenesis and tumor growth by upregulating CXC-chemokines in cervical cancer cells. *Mol. Cell. Biochem.* 448, 311–320. doi:10.1007/s11010-018-3335-7
- Zheng, W., Jiang, C., and Li, R. (2016). Integrin and gene network analysis reveals that ITGA5 and ITGB1 are prognostic in non-small-cell lung cancer. *Oncotargets Ther.* 9, 2317–2327. doi:10.2147/OTT.S91796
- Zhu, Y., Zhou, Y., Jiang, H., Chen, Z., and Lu, B. (2021). Analysis of core genes for colorectal cancer prognosis based on immune and stromal scores. *PeerJ* 9, e12452. doi:10.7717/peerj.12452