

# Construction of Immune-gene-pairs and Prognostic Model and its Immune Cell Infiltration in Glioblastoma – A Study based on TCGA and GTEx

Yongxiang Wei

Medical School of Nanjing University

Jixin Shi (✉ [3149782906@qq.com](mailto:3149782906@qq.com))

Medical School of Nanjing University

---

## Research Article

**Keywords:** glioblastoma, lasso regression, immune infiltration, data mining

**Posted Date:** October 21st, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-960777/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** Gliomas are the most common primary intracranial tumor, containing 81% of brain malignant carcinoma. Glioblastoma is a primary brain tumor with a very poor prognosis. Existing treatment methods have many limitations, so more in-depth researches are urgently needed.

**Results:** To screen key genes and build models, we made full use of RNA sequencing (RNA-seq) data from TCGA and GTEx and perform immune infiltration analysis in multiple immune-related databases. Finally, we screened key genes and build the model. Meanwhile, we explored immune cell infiltration. KEGG and GO analysis prompted us COVID-2019 may have potential relation with GBM. Further selection of immune-related differential genes and gene-pairs were then explored. The prognosis model built by lasso regression was tested to have great effect. Key genes were all found to have quite diverse expression level compared with normal samples. Moreover, their co-expression analysis reminded us more attention could be paid to WNT and NOTCH signaling. Immune infiltration suggests that most immune cells are differentially expressed in tumor samples, and the interaction reminded us B cells may play a critical role.

**Conclusion:** This study is the first to build immune-gene-pairs and used lasso regression to construct the model and screen critical pairs. Key genes were then selected and observed their role in GBM.

## Introduction

Gliomas are the most common primary intracranial tumor, containing 81% of brain malignant carcinoma [1]. Among all types of this disease, glioblastoma (GBM) has the highest incidence rate and causes a great health burden [1]. It is believed to originate from neuroglial stem or progenitor cells and the most common type is isocitrate dehydrogenase (IDH)-wildtype tumor [2]. Although compared with certain cancers such as liver cancer or breast cancer, GBMs is relatively rare, its treatment prospects are not optimistic due to high strong invasive ability and high recurrence [3]. It remains as a difficultly cured cancer regardless of therapeutic options' improving, so more in-depth researches are urgently needed [4].

GBM exhibits high levels of not only inter-tumoral heterogeneity, but also intra-tumoral heterogeneity, and the major non-neoplastic cells in the microenvironment are so-called tumor-associated macrophages (TAMs), innate immune system cells [5]. In addition, these patients often have a distinctive feature—local or systemic immunosuppression, and PD-L1 on EVs maybe a potential target to suppress antitumor immunity [6]. As for treatment, immunotherapy also shows extraordinary potential. This treatment strategy with immune checkpoint inhibitors has been proved to obtain clinical improvements in advanced tumors for which conventional therapies have had limited success [7]. In recent studies, more attention is paid to interactions between tumor and immune cells. Based on immunohistochemistry approaches, infiltration of immune cells in different subtypes of GBM was researched, and CD68+ and CD163+ cells were the most familiar populations in all four subtypes [8]. In addition, bioinformatics analysis have been

widely used to simultaneously detect genes and therefore building risk prognosis models. Yang Kong et al. found TREM1 may be a novel immunotherapy target [9].

Since immune targets has shown broad prospects, we did related analysis. We screened out genes between normal samples and tumor ones and extracted their expression level. Afterwards we performed KEGG and GO pathways enrichment analysis. Furthermore, immune-related genes were selected among all differential genes, and immune-gene-pair were obtained through pairwise comparison. Then we filtered those relative to prognosis and used lasso regression to build a prognostic prediction model. We performed a comprehensive evaluation of this model and explored the infiltration of various immune cells. Besides, we explored key genes gene-pair included. All in all, this study aimed at constructing prognosis-related model in GBM and having a deep understanding of immune cells infiltration.

## Results

### KEGG and GO Pathway analysis of differential genes

RNA transcription data were downloaded and integrated from TCGA and GTEx. Afterwards, data cleansing was run to screen differentially expressed genes between GBM samples and normal ones. *KEGG* and *GO* pathway enrichment analysis were also performed and results was shown in Fig 1. Obviously, changes in *BP* were significantly enriched in translational initiation, establishment of protein localization to membrane, RNA catabolic process, mRNA catabolic process and protein targeting to ER. As for *CC*, they were mainly enriched in ribosomal subunit, focal adhesion, cell-substrate adherens junction, cell-substrate junction and mitochondrial protein complex. Changes of differential genes in *MF* were enriched in cell adhesion molecule binding, structural constituent of ribosome, cadherin binding, GTPase activity and ubiquitin-like protein ligase binding. Similarly, we can easily find changes in KEGG focused on Parkinson disease, ribosome, prion disease, huntington disease and coronavirus disease-COVID-19, diabetic cardiomyopathy and oxidative phosphorylation. Interestingly, coronavirus disease is spreading rapidly all over the world, and KEGG analysis provides clues between it with GBM. It also has a strong prompting effect for future research.

### The results of lasso regression

Next, we downloaded immune-related gene list from ImmPort and crossed it with differentially expressed genes to obtain immune-related differentially expressed genes which could be used to set up immune-gene-pair. Lasso regression was performed based on survival data. Results was shown in Fig 2. Meanwhile, risk score of each gene-pair and the most significant pairs were screened out. As was shown in Table 1, nine immune-related gene-pairs were believed to play a critical role in this model and worthy of further investigation. Besides, forest plots of these gene-pairs were displayed in Fig 3. All of the gene pairs have a *P* value of smaller than 0.001 and PTX3|BMP2 has the highest hazard ratio of 11.886.

### Most Significant gene-pairs screened out from lasso regression

ID	HR	HR.95	P-Value
AZGP1 BMP2	3.970	3.079-5.118	<0.001
HLA-DRB5 TNFRSF12A	0.214	0.161-0.285	<0.001
KLRC2 BIRC5	0.112	0.083-0.150	<0.001
PSMC2 NR2F1	5.576	4.236-7.341	<0.001
TMSB15A BMP2	4.815	3.731-6.214	<0.001
TMSB4XP8 PPP3CB	9.436	6.793-13.109	<0.001
WFDC2 JAG2	4.985	3.798-6.543	<0.001
TCF7L2 IKBKB	0.144	0.110-0.189	<0.001
PTX3 BMP2	11.886	8.949-15.785	<0.001

## Evaluation of the model

Evaluation of our model was also conducted. Firstly, we determined the critical value of risk score by ROC curve. As was shown in Fig 4A, the critical value of risk score was 3.672 and area under curve was 0.886. Next, we divided the patients into two groups according to risk score by setting cut-off value as 3.672 and carried out survival analysis. Kaplan-meier plot indicated that the survival between two groups were statistically significant ( $P < 0.001$ , Fig 4B). Also, we explored the AUC curve considering one-year, two-year and three-year survival data and found that the AUC were all greater than 0.85 (Fig 4C). Furthermore, in Fig 4D, there was a great divergence in the survival status on both sides of the cutoff value. Combining these results, our model was believed to be reliable.

## Clinically related analysis

Clinically related analysis were processed by integrating risk score data with clinical data. In Fig 5A, ROC curve indicated that risk score and age were important factors with AUC values greater than 0.800. Relationship between these factors and risk score was shown in Fig 5B-D which suggest age and grade were closely related to risk score. Heatmap also indicated that age and grade were significant clinical factors (Fig 5E). This conclusion could also be supported by multivariate independent prognostic analysis and single factor analysis. Age, grade and risk score were critical affecting factors according to Figure 5F-G.

## Further exploration of key genes

As was depicted before, noteworthy immune gene-pairs were got and these genes were regarded as key genes, including AZGP1, HLA-DRB5, TNFRSF12A, KLRC2, BIRC5, PSMC2, NR2F1, TMSB15A, TMSB4XP8, PPP3CB, WFDC2, JAG2, RBP1, RAC3, FURIN, CX3CL1, TCF7L2, IKBKB, PTX3 and BMP2. Results of violin plots were shown in Fig 6. Interestingly, expression level of these key genes between diverse risk groups were significantly different. Next, we conducted co-expression analysis. By analyzing samples with mutation and CNA data (273 patients/samples), the most common genetic alteration type was missense mutation and other types were also shown in Fig 7A. As for pathway analysis, WNT and NOTCH signal caught our attention (Fig 7B and Fig 7C), reminding us they may play unignorable role in the development process of GBM.

## Infiltration analysis of various types of immune cells

Next, we integrated immune databases from multiple sources and explored the infiltration situation of different types of immune cells. As shown in Fig 8, immune cell infiltration were significantly different between high and low risk group. Correlation between immune cell infiltration and risk score were also explored. As shown in Fig9, macrophage M2 have a high correlation coefficient in both Mcp-counter and CIBERSORT-ABS database. Cancer associated fibroblast from Mcp-counter and B cell plasma from CIBERSORT may have the most significant correlation, indicating their importance in immune infiltration in GBM.

## Discussion

This research aimed at constructing and assessing immune prognostic model and screened out key immune-gene-pairs and genes. Downloading and integrating RNA sequencing data from TCGA and GTEx, differentially expressed genes were screened out and performed KEGG and GO pathway analysis. Notably, In KEGG analysis, we found coronavirus disease (COVID-2019) was one of the enrichment pathways. No relevant studies before clarified relation between GBM and COVID-2019. However, in the review of Hyun Jee Han et al, patients tended to present worse situation when exposed to the virus [22]. It prompted us they may need further exploration since COVID-2019 was so serious that scientists all over the world were investigating its related information.

Then, we download immune-gene list from ImmPort and selected immune-related differentially expressed genes, which were used to construct immune-gene-pairs so as to adjust batch difference. After that, lasso regression was utilized to build model and nine most related gene-pairs were selected, including AZGP1|BMP2, HLA-DRB5|TNFRSF12A, KLRC2|BIRC5, PSMC2|NR2F1, TMSB15A|BMP2, TCF7L2|IKBKB, TMSB4XP8|PPP3CB, WFDC2|JAG2, and PTX3|BMP2. These genes were referred to as key genes. Some genes were proved to be relative to GBM or other cancers. For example, L Persano et al suggest reducing BMP expression level may be a great way to increase GBM responsiveness to chemotherapy [23].

Expression of AZGP1 in prostate cancer was especially higher compared with adjacent normal tissues [24]. More studies are needed to clarify in-depth mechanism. Furthermore, whether these genes related to GBM also correlate with COVID-2019 remain to be explored.

Assessment of the model was as important as building. Subsequently, through ROC curve we got the cutoff value – 3.672 and found area under one-year, two-year and three-year curve were all more than 0.75, reminding us this model was reliable. Similarly, there are great difference between high and low groups. Some clinical factors such as age, grade and sick score were all proved to be critical. Jigisha P Thakkar et al concluded younger age was related to better prognosis [25]. Moreover, key genes' expression level were higher in tumor samples; their co-expression level suggest WNT and NOTCH may be critical pathways and missense mutation was the most common genetic alteration type. Also, Nishani Rajakulendran et al reported that in orthotopic Wnt model, inhibition of Wnt/ $\beta$ -catenin and Notch signaling could lead to lower clonogenic potential [26]. More and more evidences unfolded NOTCH signaling was highly active in GBM stem cells, thus contributing to high resistance [27]. The influence of these genes and pathways in the prognosis of GBM patients and the underlying mechanism are worthy of exploration.

As for immune infiltration, most immune cells played an unignorable role especially B cell. Glioblastoma, having the poorest prognosis in all primary brain cancers, immunotherapy obviously provides new treatment approach [26]. We firstly constructed immune-gene-pairs and used lasso regression to build the model, and most importantly, it demonstrated great predictive ability and reliability. Christopher M Jackson et al discussed what role immunotherapy may play in GBM resistance model [28]. We also provided evidence for tumor immunotherapy and B cell was worthy of more attention. Catalina Lee-Chang et al found the role of B cells in GBM physiopathology and provides the mechanism [29].

In conclusion, we attempted to investigate GBM immune-related gene-pairs, key genes and screen potential biomarkers. Immune-model was built and evaluated, it had great predictive effect. Similarly, our research also has areas for improvement. Firstly, we lacked experimental projects and related data, all analysis was based on data downloaded from open databases. Combining the clinical project with this analysis will lead to more credible conclusions. Secondary, we did not pay more attention to exploration of signaling. Last but not least, the conclusions were based on retrospective data and more research is required to certificate these findings.

## Conclusions

This study is the first to build immune-gene-pairs and used lasso regression to construct the model and screen critical pairs. Key genes were then selected and observed their role in GBM.

## Methods

# Gene information and bioinformatics analysis

We downloaded gene expression data from TCGA (workflow type: HTSeqCounts) and GTEx (data format: fpkm). It contains 698 tumor samples and 1152 normal samples. Clinical data (1114 cases, data format: BCR XML) was downloaded from TCGA, which included survival time, survival state, age gender and grade. Additionally, the Ensembl database (<http://asia.ensembl.org/index.html>) was used to transform Ensembl IDs to corresponding gene symbols. Perl language was applied to integrate data and get data matrix, simultaneously, the listwise technology was a great way to deal with missing data. In terms of hypothesis test, we use Kruskal-Wallis test to determine the difference situation, especially clinical significantly relation.

## Identification of differentially immune-related expressed genes and immune-gene-pair.

After getting gene transcription data, we used "edgeR" R package (R version:3.6.1) to filter differentially expressed genes between GBM samples and normal samples. ImmPort (<http://www.immport.org>) is an open database which includes 17 categories, immune-related gene list was downloaded from this web [10]. Our study extracted differentially immune-related genes and their expression level. Afterwards, immune-gene-pair was constructed and it was combined with survival data. By the way, gene-pair meant comparing adjacent genes' expression levels and obtain results.

## KEGG and GO pathways analysis

The Kyoto Encyclopedia of Genes and Genomes (KEGG) is a database which is utilized to explore genes' biological functions [11]. Gene Ontology (GO) function analysis includes biological processes (BP), cellular components (CC) and molecular functions (MF), which could let us have a better understanding of biological processes [12]. By using "colorspace", "stringi", "ggplot2", "clusterProfiler" and "enrichplot" R package, we could perform GO functional annotations and KEGG pathway enrichment analysis of differential expressed genes we got before. Results could reveal pathways worthy of attention.

## Lasso regression

Lasso regression is a method that is more and more widely used in the field of bio-informatics. By constructing a penalty function, it compresses certain regression coefficients or making them zero, thereby establishing a more accurate model, especially for independent variables High-dimensional situation. In the model established in this study, survival time and status were taken as dependent variables, and immune gene pairs were taken as independent variables. We utilized the "glmnet" R language package (version 2.0-16) to fit the logistic LASSO regression model [13, 14]. As a result, risk score of each pair was calculated and summarized. Meanwhile, we obtained these immune-gene-pairs which have the most significant impact on prognosis. In addition, in order to verify the predictive effect of the model, our team drew ROC curve and got the cutoff value. Based on it, high and low risk groups were

dived and survival curve was built. Combing clinical information with risk file, our study performed univariate and multivariate independent prognostic analysis. At the same time, we drew ROC curve, forest map and boxplot curve for visual observation. R language and several packages were applied during it, including “survival”, “survivalROC”, “limma” and “ggpubr”.

## Key genes

Through lasso regression, most related immune-gene-pairs were selected and they were believed to play important roles in GBM patients. Similarly, gene-pairs composed of these genes were believed to play an important role in GBM patients. These genes were so-called key genes in this study. In addition to explore the role of gene-pairs, individual or co-expression of key genes in GBM patients were worthy of studying. By the way, violin program was used to visually represent the individual role and “\*\*\*” meant  $p$ -value < 0.01, since we compare certain genes’ expression level between patients with different prognosis.

## cBioPortal

The cBioPortal for Cancer Genomics (<http://cbioportal.org>) is an open and user-friendly dataset, which could explore and analyze multidimensional cancer genomics data [15]. By selecting GBM database and setting gene names, it is easy to interactively probe genetic alterations across samples, genes, and pathways. Due to its intuitiveness and simplicity of operation, the database has inspired many new discoveries in cancer bioinformatics. Through co-expression analysis, it suggests interesting pathways that maybe worth further exploration.

## Immune databases

xCell (<http://xCell.ucsf.edu/>) shows a novel method to present immune cells infiltration, which could explore 64 immune and stromal cell types [16]. Timer (<https://cistrome.shinyapps.io/timer/>) comprehensively collects and analyze molecular characterization, investigating expression levels of six immune-interaction subtypes in 10,897 tumors from 32 cancer types [17]. QuantIseq (<http://icbi.at/quantiseq>) is a fantastic way to quantify immune-infiltration, measuring the fractions of ten immune cell subtypes from a great number of RNA-sequencing data [18]. MCP-counter (Microenvironment Cell Populations-counter) focuses on exploring the infiltration of eight types of immune cells, thereby exploring the tumor microenvironment [19]. As a prospective cohort study, The European Prospective Investigation into Cancer and Nutrition (EPIC) aimed at investigating relationship between tumor and nutrition [20]. CIBERSORT (<http://cibersort.stanford.edu/>) could analyze RNA mixtures for exploring cellular biomarkers and therapeutic targets by sequencing samples from fresh, frozen and fixed tissues [21].

Although results and ways of these immune databases are not all the same, systematic integration of the immune infiltration in multiple immune cells can provide us with more convincing evidence. R packages “limma”, “scales”, “ggplot2” and “ggtext” were utilized to extract specific types of immune cell expression level and evaluate its correlation with risk score.

## **Declarations**

## **Acknowledgements**

We thank the researchers who gave their data for this analysis. We gratefully acknowledge their contributions.

## **Authors' contributions**

Yongxiang Wei and Jixin Shi designed the project, performed data analysis and wrote the manuscript together.

## **Funding**

This work is not supported by grants.

## **Availability of data and materials**

The datasets analyzed in the current study are available in the TCGA repository (<http://cancergenome.nih.gov/>) and GTEx (<http://www.broadinstitute.org/gtex>)

## **Ethics approval and consent to participate**

Not applicable.

## **Consent for publication**

Not applicable.

## **Conflicts of Interest**

The authors declare that they have no competing interests.

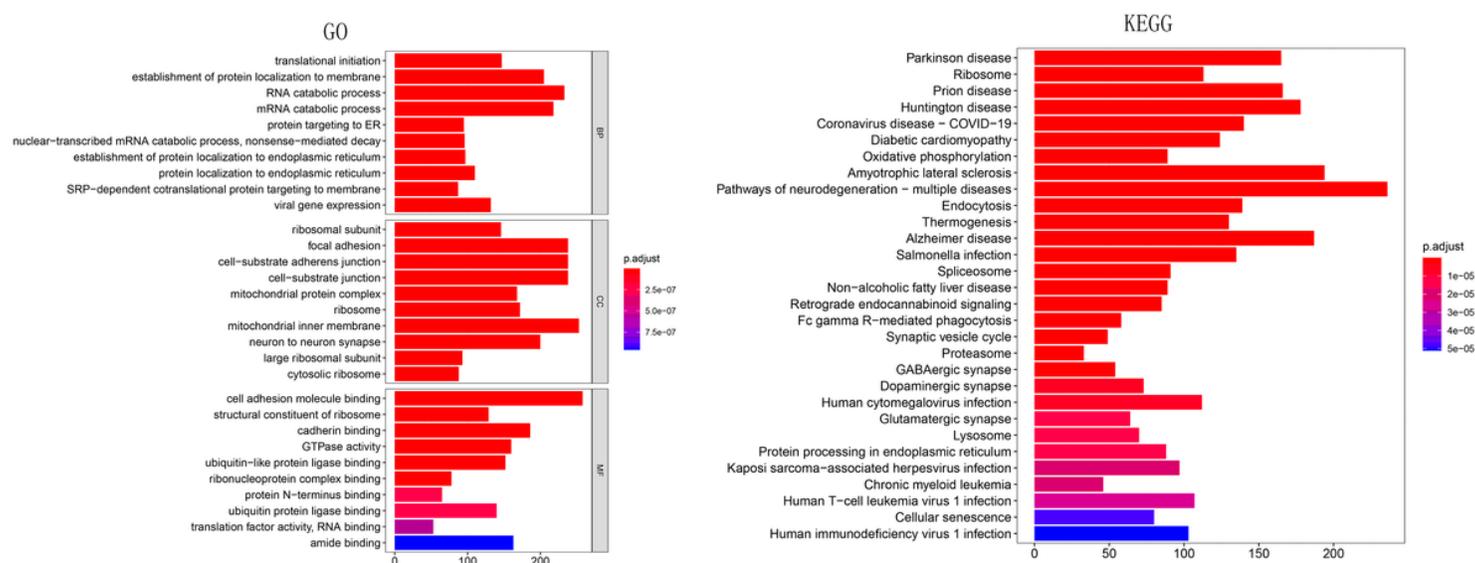
## References

1. Ostrom QT, Bauchet L, Davis FG, Deltour I, Fisher JL, Langer CE, Pekmezci M, Schwartzbaum JA, Turner MC, Walsh KM *et al*: The epidemiology of glioma in adults: a "state of the science" review. *Neuro Oncol* 2014;16(7):896–913. <https://doi.org/10.1093/neuonc/nou087>.
2. Le Rhun E, Preusser M, Roth P, Reardon DA, van den Bent M, Wen P, Reifenberger G, Weller M: Molecular targeted therapy of glioblastoma. *Cancer Treat Rev* 2019;80:101896. <https://doi.org/10.1016/j.ctrv.2019.101896>.
3. Campos B, Olsen LR, Urup T, Poulsen HS: A comprehensive profile of recurrent glioblastoma. *Oncogene* 2016;35(45):5819–5825. <https://doi.org/10.1038/onc.2016.85>.
4. Omuro A, DeAngelis LM: Glioblastoma and other malignant gliomas: a clinical review. *JAMA* 2013;310(17):1842–1850. <https://doi.org/10.1001/jama.2013.280319>.
5. Chen Z, Hambarzumyan D: Immune Microenvironment in Glioblastoma Subtypes. *Front Immunol* 2018;9:1004. <https://doi.org/10.3389/fimmu.2018.01004>.
6. Ricklefs FL, Alayo Q, Krenzlin H, Mahmoud AB, Speranza MC, Nakashima H, Hayes JL, Lee K, Balaj L, Passaro C *et al*: Immune evasion mediated by PD-L1 on glioblastoma-derived extracellular vesicles. *Sci Adv* 2018;4(3):eaar2766. <https://doi.org/10.1126/sciadv.aar2766>.
7. Preusser M, Lim M, Hafler DA, Reardon DA, Sampson JH: Prospects of immune checkpoint modulators in the treatment of glioblastoma. *Nat Rev Neurol* 2015;11(9):504–514. <https://doi.org/10.1038/nrneurol.2015.139>.
8. Martinez-Lage M, Lynch TM, Bi Y, Cocito C, Way GP, Pal S, Haller J, Yan RE, Ziober A, Nguyen A *et al*: Immune landscapes associated with different glioblastoma molecular subtypes. *Acta Neuropathol Commun* 2019;7(1):203. <https://doi.org/10.1186/s40478-019-0803-6>.
9. Kong Y, Feng ZC, Zhang YL, Liu XF, Ma Y, Zhao ZM, Huang B, Chen AJ, Zhang D, Thorsen F *et al*: Identification of Immune-Related Genes Contributing to the Development of Glioblastoma Using Weighted Gene Co-expression Network Analysis. *Front Immunol* 2020;11:1281. <https://doi.org/10.3389/fimmu.2020.01281>.
10. Shen S, Wang G, Zhang R, Zhao Y, Yu H, Wei Y, Chen F: Development and validation of an immune gene-set based Prognostic signature in ovarian cancer. *EBioMedicine* 2019;40:318–326. <https://doi.org/10.1016/j.ebiom.2018.12.054>.
11. Altermann E, Klaenhammer TR: PathwayVoyager: pathway mapping using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. *Bmc Genomics* 2005;6. <https://doi.org/10.1186/1471-2164-6-60>.
12. Ding J, Zhang Y: Analysis of Key GO Terms and KEGG Pathways Associated with Carcinogenic Chemicals. *Comb Chem High T Scr* 2017;20(10):861–871. <https://doi.org/10.2174/1386207321666171218120133>.
13. McEligot AJ, Poynor V, Sharma R, Panangadan A: Logistic LASSO Regression for Dietary Intakes and Breast Cancer. *Nutrients* 2020;12(9). <https://doi.org/10.3390/nu12092652>.

14. Lin Q, Zhao ZG, Liu JS: Sparse Sliced Inverse Regression via Lasso. *J Am Stat Assoc* 2019;114(528):1726–1739. <https://doi.org/10.1080/01621459.2018.1520115>.
15. Gao JJ, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun YC, Jacobsen A, Sinha R, Larsson E *et al*: Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. *Sci Signal* 2013;6(269). <https://doi.org/10.1126/scisignal.2004088>.
16. Aran D, Hu ZC, Butte AJ: xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol* 2017;18. <https://doi.org/10.1186/s13059-017-1349-1>.
17. Li TW, Fan JY, Wang BB, Traugh N, Chen QM, Liu JS, Li B, Liu XS: TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. *Cancer Res* 2017;77(21):E108-E110. <https://doi.org/10.1158/0008-5472.Can-17-0307>.
18. Finotello F, Mayer C, Plattner C, Laschober G, Rieder D, Hackl H, Krogsdam A, Loncova Z, Posch W, Wilflingseder D *et al*: Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data (vol 11, 34, 2019). *Genome Med* 2019;11. <https://doi.org/10.1186/s13073-019-0655-5>.
19. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, Selves J, Laurent-Puig P, Sautes-Fridman C, Fridman WH *et al*: Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression (vol 17, 218, 2016). *Genome Biol* 2016;17. <https://doi.org/10.1186/s13059-016-1113-y>.
20. Riboli E, Hunt KJ, Slimani N, Ferrari P, Norat T, Fahey M, Charrondiere UR, Hemon B, Casagrande C, Vignat J *et al*: European prospective investigation into cancer and nutrition (EPIC): study populations and data collection. *Public Health Nutr* 2002;5(6b):1113–1124. <https://doi.org/10.1079/Phn2002394>.
21. Newman AM, Liu CL, Green MR, Gentles AJ, Feng WG, Xu Y, Hoang CD, Diehn M, Alizadeh AA: Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* 2015;12(5):453–+. <https://doi.org/10.1038/Nmeth.3337>.
22. Han HJ, Nwagwu C, Anyim O, Ekweremadu C, Kim S: COVID-19 and cancer: From basic mechanisms to vaccine development using nanotechnology. *Int Immunopharmacol* 2021;90. <https://doi.org/10.1016/j.intimp.2020.107247>.
23. Persano L, Pistollato F, Rampazzo E, Della Puppa A, Abbadi S, Frasson C, Volpin F, Indraccolo S, Scienza R, Basso G: BMP2 sensitizes glioblastoma stem-like cells to Temozolomide by affecting HIF-1 alpha stability and MGMT expression. *Cell Death Dis* 2012;3. <https://doi.org/10.1038/cddis.2012.153>.
24. Cao RY, Ke M, Wu QX, Tian Q, Liu L, Dai Z, Lu S, Liu P: AZGP1 is androgen responsive and involved in AR-induced prostate cancer cell proliferation and metastasis. *J Cell Physiol* 2019;234(10):17444–17458. <https://doi.org/10.1002/jcp.28366>.
25. Thakkar JP, Dolecek TA, Horbinski C, Ostrom QT, Lightner DD, Barnholtz-Sloan JS, Villano JL: Epidemiologic and Molecular Prognostic Review of Glioblastoma. *Cancer Epidem Biomar* 2014;23(10):1985–1996. <https://doi.org/10.1158/1055-9965.Epi-14-0275>.

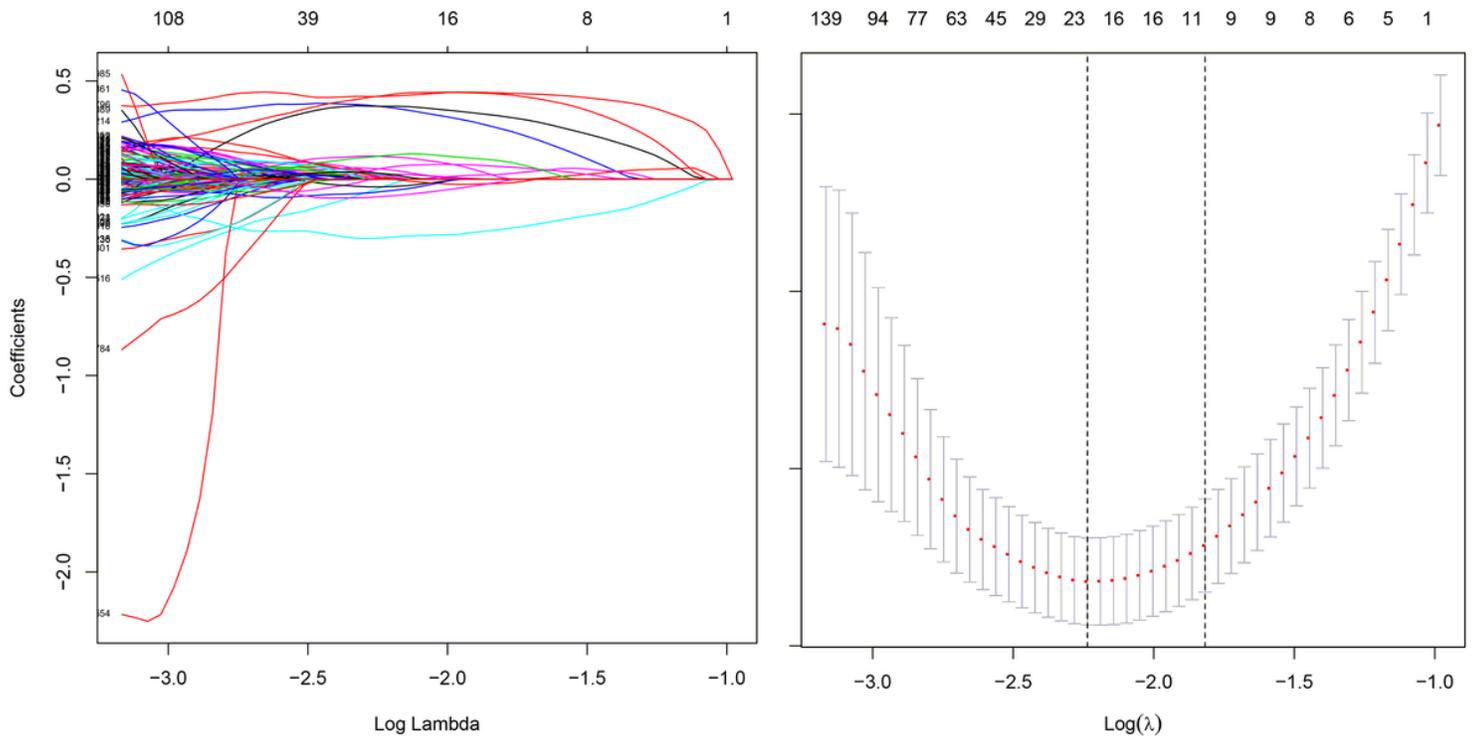
26. Rajakulendran N, Rowland KJ, Selvadurai HJ, Ahmadi M, Park NI, Naumenko S, Dolma S, Ward RJ, So M, Lee L *et al*: Wnt and Notch signaling govern self-renewal and differentiation in a subset of human glioblastoma stem cells. *Gene Dev* 2019;33(9-10):498–510. <https://doi.org/10.1101/gad.321968.118>.
27. Bazzoni R, Bentivegna A: Role of Notch Signaling Pathway in Glioblastoma Pathogenesis. *Cancers* 2019;11(3). <https://doi.org/10.3390/cancers11030292>.
28. Jackson CM, Choi J, Lim M: Mechanisms of immunotherapy resistance: lessons from glioblastoma. *Nat Immunol* 2019;20(9):1100–1109. <https://doi.org/10.1038/s41590-019-0433-y>.
29. Lee-Chang C, Rashidi A, Miska J, Zhang P, Pituch KC, Hou D, Xiao T, Fischietti M, Kang SJ, Appin CL *et al*: Myeloid-Derived Suppressive Cells Promote B cell-Mediated Immunosuppression via Transfer of PD-L1 in Glioblastoma. *Cancer Immunol Res* 2019;7(12):1928–1943. <https://doi.org/10.1158/2326-6066.Cir-19-0240>.

## Figures



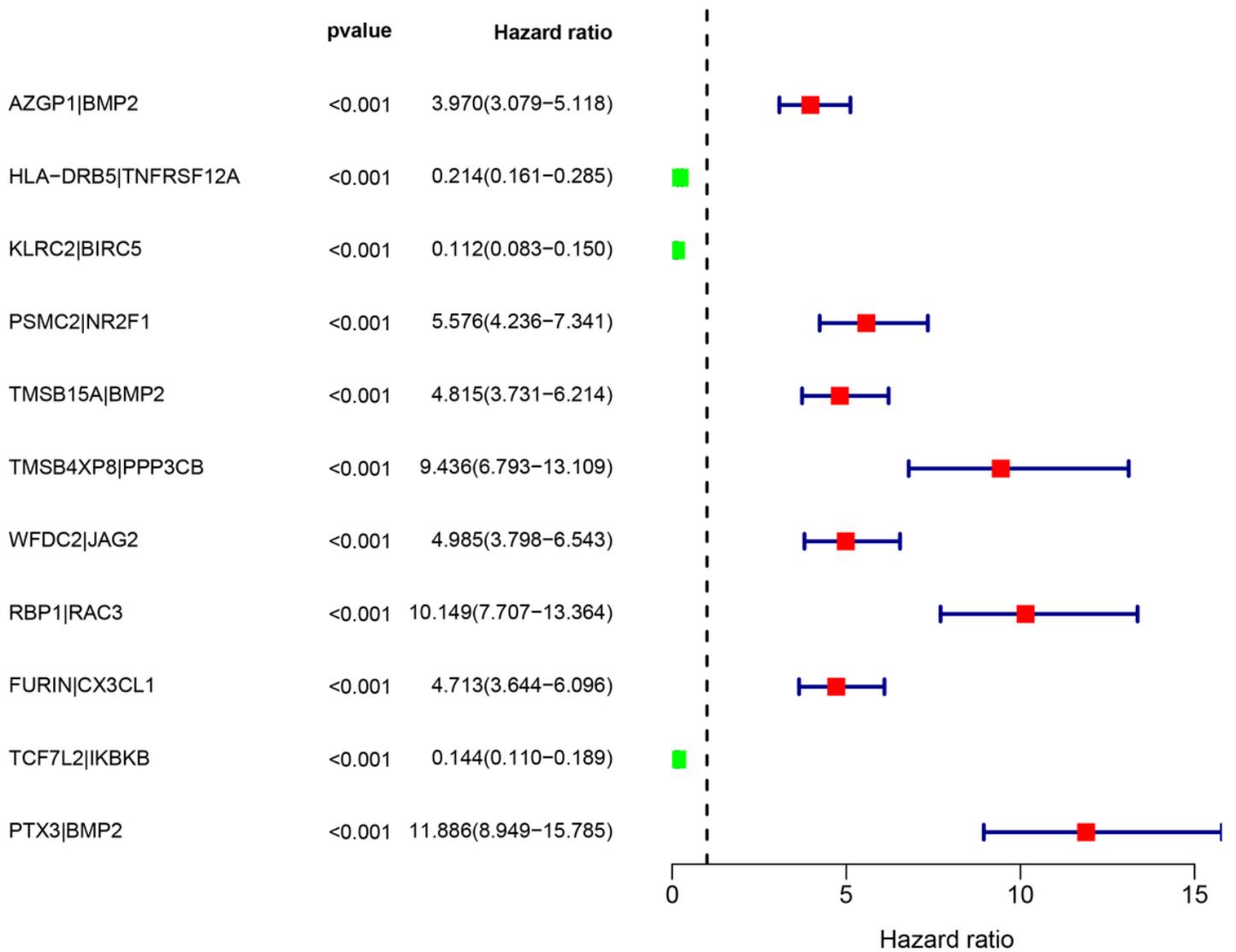
**Figure 1**

A. GO enrichment analysis predicted the functional roles of differentially expressed genes based on biological processes (BP), cellular components (CC), and molecular functions (MF). B. KEGG pathway analysis identified the enriched pathways.



**Figure 2**

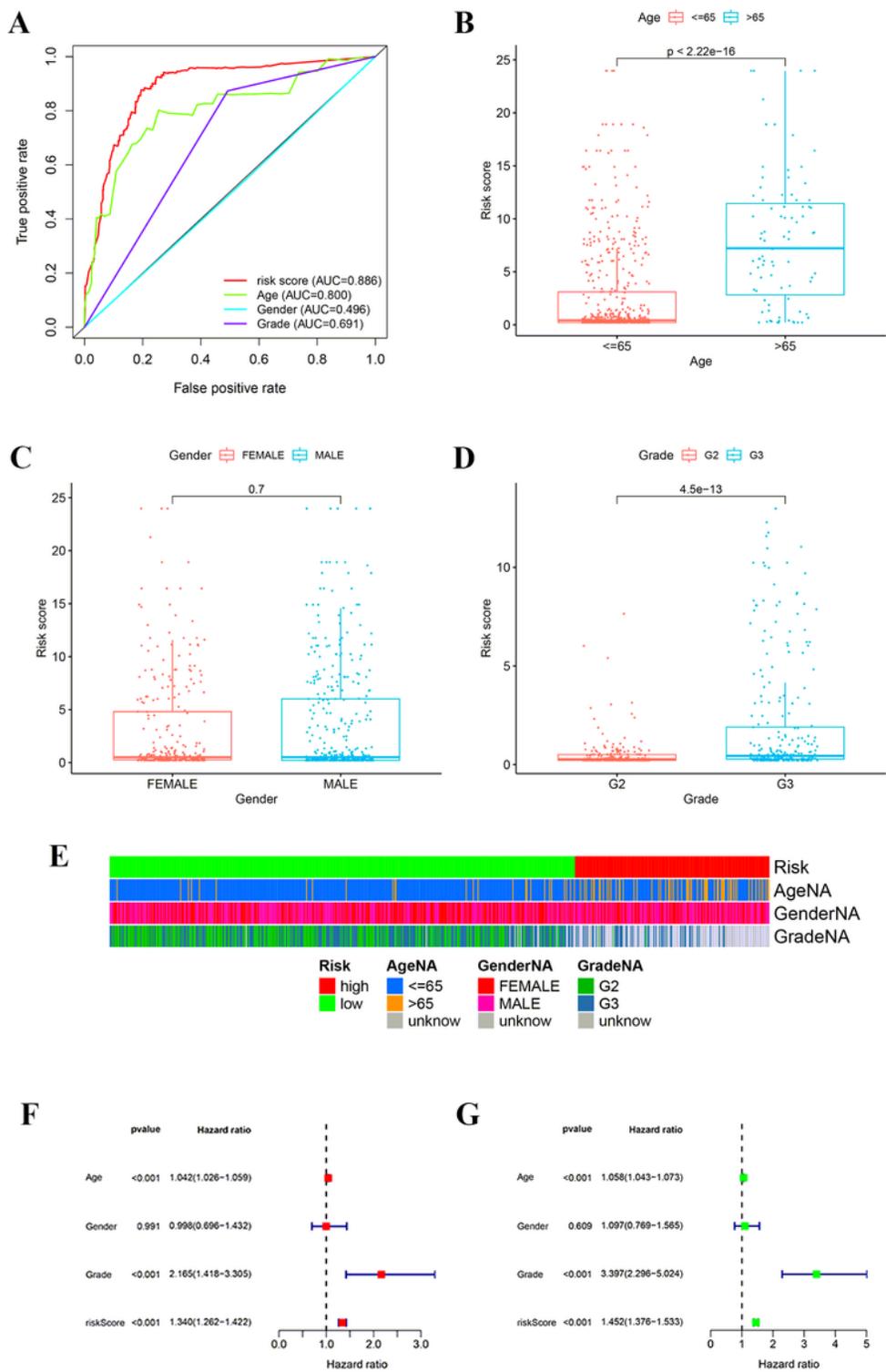
Results of lasso regression.



**Figure 3**

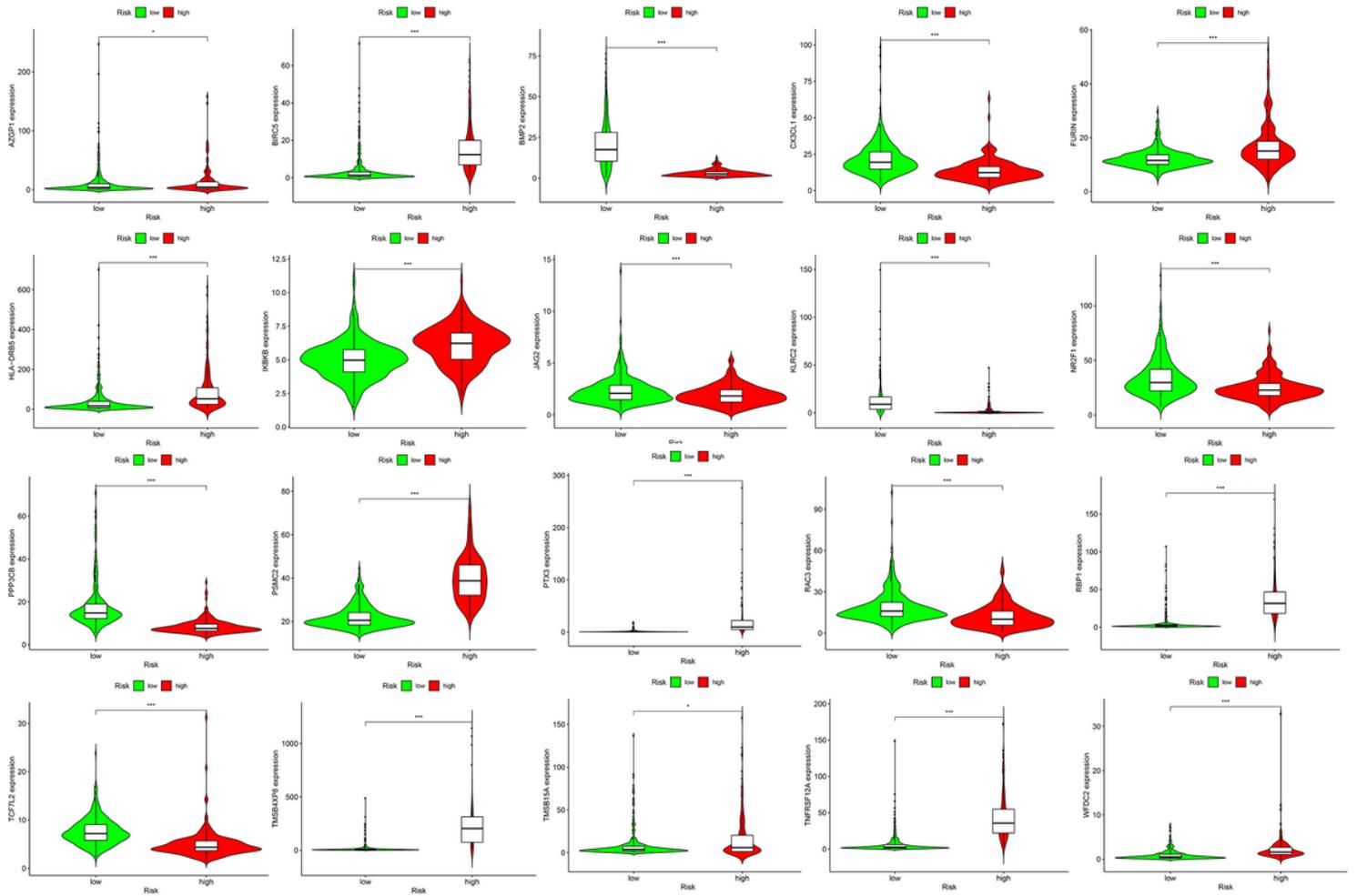
Forest picture of significantly expressed gene-pairs.





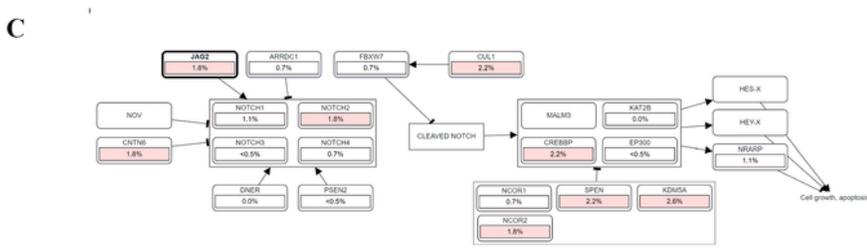
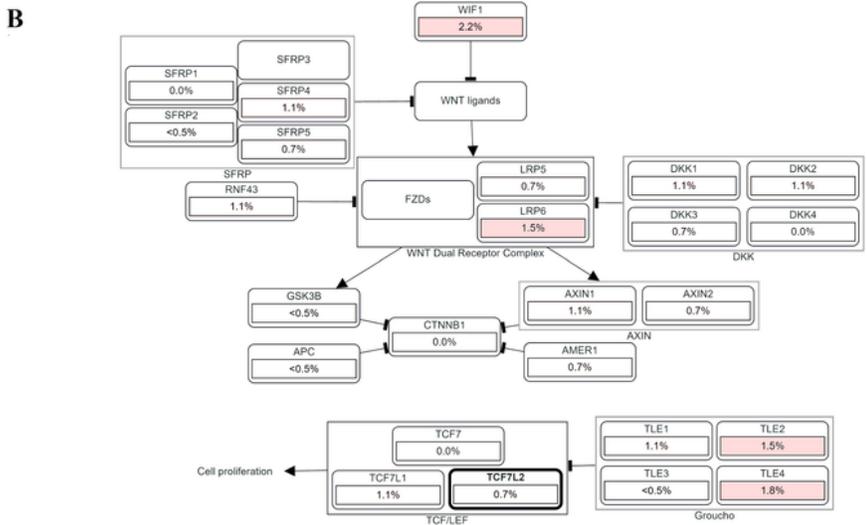
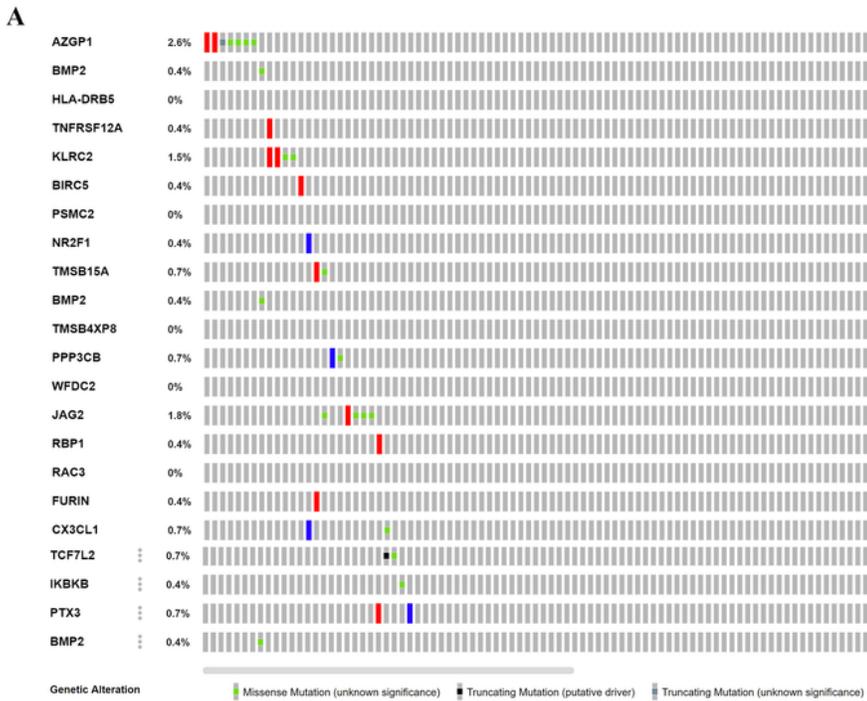
**Figure 5**

A. ROC curve of various clinical characteristics. B. Boxplot of age between different risk groups. C. Boxplot of gender between different risk groups. D. Boxplot of grade between different risk groups. E. Heatmap of several clinical characteristics. F. Forest picture of a few clinical characteristics (multi-factor analysis). G. Forest picture of a few clinical characteristics (univariate analysis).



**Figure 6**

Violin picture of key genes' expression between different risk groups.



**Figure 7**

A. Samples with mutation and CNA data (273 patients/samples, cBioportal). B.&C. Co-expression analysis of pathways (cBioportal).



Figure 8

Immune infiltration of most immune cells from several immune databases.



Figure 9

Correlation between immune cells and risk score.