

# Construction of ceRNA Network Based on TCGA and UCSC Xena Official Website to Provide Certain Help for Immunotherapy of Thyroid Cancer.

**Junwei Du**

First Affiliated Hospital of Zhengzhou University

**Xinguang Qiu** (✉ [qxg2000@126.com](mailto:qxg2000@126.com))

First Affiliated Hospital of Zhengzhou University

**Guoquan Li**

First Affiliated Hospital of Zhengzhou University

**Yangsen Li**

First Affiliated Hospital of Zhengzhou University

---

## Research Article

**Keywords:** Thyroid cancer, TCGA, UCSC Xena

**Posted Date:** February 8th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-153862/v1>

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

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Scientific Reports on August 10th, 2021. See the published version at <https://doi.org/10.1038/s41598-021-90538-3>.

# Abstract

**Background:** Thyroid cancer is the most common endocrine tumor, and papillary thyroid carcinoma is the most common. Commonly used treatment methods include surgery and radiotherapy. As an emerging treatment option in recent years, immunotherapy has played an essential role in the treatment of many tumors, but there are few studies in thyroid cancer.

**Method:** We first downloaded the mRNA and lncRNA data of thyroid cancer patients from the TCGA official website <https://portal.gdc.cancer.gov/>, and downloaded the miRNA of thyroid cancer patients from UCSC Xena <https://xenabrowser.net/> data. Through the tumor microenvironment's immune score, tumor samples are divided into cold tumors and hot tumors. We obtained the differential genes of DEGs, lncRNA and miRNA, and jointly constructed a ceRNA network through differential analysis of the mRNA data of cold and hot tumours.

**Results:** We first assessed the immune infiltration of the tumor microenvironment of patients with thyroid cancer, and divided the samples into cold tumors and hot tumors based on the immune score. A total of 568 up-regulated DEGs and 412 down-regulated DEGs were screened by analysing the difference between hot and cold tumours. Then we conducted a differential analysis of lncRNA and miRNA and obtained 629 differential genes of lncRNA, and 114 differential genes of miRNA. Finally, we constructed a ceRNA network for the acquired differential genes. Five miRNA-centric hubnets, hsa-mir-204, hsa-mir-128, hsa-mir-214, hsa-mir-150 and hsa-mir-338, have been identified.

**Conclusion:** Our research has identified the immune-related mRNA, lncRNA and miRNA DEGs of thyroid cancer, and constructed a ceRNA network. This result is of great significance for exploring immune-related mechanisms of thyroid cancer and provides certain help for immunotherapy of thyroid cancer.

## Background

In 1863, Rudolf Virchow, the father of modern cytopathology, proposed the relationship between microinflammation and subsequent cancer development[8]. In 1909, Paul Ehrlich proposed using the immune system to control cancer[9]. Tumor microenvironment, especially the immune system, plays an essential role in regulating tumor progression and tumor response to treatment. It mainly stimulates tumor specific immune response by inducing the immunogenic death of tumor cells or participating in immune response mechanism[10]. According to the immune infiltration of tumor microenvironment, tumor is divided into "cold tumor" and "hot tumor". In the tumor microenvironment of "hot" tumor, the degree of immune invasion and immune effect is relatively higher, with strong antigen presentation ability and T cell activation. Such immune infiltration will lead to the production of tumor specific CD8 + T cells, which can clear cancer cells and generate systemic tumor specific immunity, forming a long-term anti-tumor memory response[11, 12]. While "cold" tumors have no immune cell infiltration in the tumor microenvironment or are mainly infiltrated by suppressive regulatory cell subtypes (including regulatory T cells (Tregs), regulatory B cells (bregs) and myeloid suppressor cells (MDSCs)) [13-15]. The growth of

cancer is not inhibited in immunology. In thyroid cancer, we divided tumor samples into cold tumor group and hot tumor group by the degree of immune invasion, and further screened the differential genes of lncRNA, miRNA and mRNA between the two groups to screen out the immune mechanism related to thyroid cancer.

Cancer Genome Atlas (TCGA) database can be used for large-scale global gene expression profile analysis and database mining to find the potential correlation between genes and the overall survival rate of various malignant tumors[16]. In this study, we first from TCGA official website <https://portal.gdc.cancer.gov/> The fpkm data of mRNA and lncRNA of thyroid cancer patients were downloaded and transformed into TPM from the official website of ucscxena <https://xenabrowser.net/> . The clinical information corresponding to the miRNA data set of patients with thyroid cancer was downloaded. We used ssgsea, MCP counter, cibersort, xcell four packages to evaluate the immune cells of thyroid cancer tumor samples and normal samples. We used xcell's own algorithm and estimate package to evaluate the immune and matrix scores. In addition, thyroid cancer was divided into four subtypes by congruent clustering to understand the differences in immune cell types, immune related molecules, tumor size distribution and grading among the four subtypes. According to the difference of immune score and matrix score, cluster 3 and Cluster 4 were defined as a hot tumor, and cluster 1 and cluster 2 were defined as cold tumors. Then, differential analysis of cold and hot tumors, enrichment analysis of differential genes, and protein interaction network construction were carried out.

Similarly, we compared the expression of lncRNA and miRNA in cold and hot tumors. Finally, the differentially expressed lncRNA, miRNA and mRNA obtained by us were used to construct the immune related Cerna network. Finally, we identified five immune connected CERN networks in thyroid cancer, which is of great significance to understand the mechanism of immune invasion and immunotherapy.

The purpose of this study was to construct a Cerna regulatory network using microarray data collected from a public database, and preliminarily identify the regulatory mechanism mediated by a novel lncRNA miRNA mRNA Cerna in thyroid cancer. This study may provide a target for the development of new therapeutic strategies for thyroid cancer.

## Methods

### Data download and immune infiltration assessment

We are from TCGA official website <https://portal.gdc.cancer.gov/> The mRNA and lncRNA data of thyroid cancer patients were downloaded from ucscxena database <https://xenabrowser.net/> To download miRNA data from. Among them, there were 57 normal samples and 511 tumor samples. In tumor microenvironment, there are four commonly used methods to evaluate immune cell infiltration: single sample gene set enrichment analysis (ssgsea), microenvironment cell populations (MCP) - counter, cibersort and xcell[17-20]. In order to reduce the error, we put four methods into the study. Inclusion criteria:  $P < 0.05$  (the p value of cibersort was  $p < 0.1$  to obtain sufficient samples).

## **Immune and matrix evaluation of tumor and paracancerous samples**

In the tumor microenvironment, the immune and matrix scores were calculated by xcell's own calculation method and estimate package, Immune and stromal cells are two major types of non tumor components and have been proposed to be valuable for diagnostic and diagnostic assessment of tumors. There were significant differences in the distribution of immune cells and immune scores between tumor tissues and adjacent tissues.

### **Correlation of immune cells**

We used the results of four evaluation methods to calculate the correlation between different immune cells. Among them, blue represents positive correlation and red represents negative correlation. Through the correlation analysis of thyroid cancer cells and adjacent immune cells, we know that the interaction between adjacent and cancer cells is significantly different.

### **Cell consistent clustering**

Using the method of cell consistent clustering, we divided thyroid cancer samples into different subtypes according to the results of immune infiltration and drew the cumulative distribution function (CDF) to identify the optimal number of subsets. Finally, we identified four different subtypes and attracted the heat map to compare the immune matrix score and the distribution of immune cells among different subtypes.

### **Differences of immune matrix score, immune related molecules and tumor size distribution and grading among different subtypes**

The progress and metastasis of cancer depend on the two-way interaction between cancer cells and their environment, forming the tumor microenvironment (TME)[21].Tumor microenvironment is different in different tumour progression stages, which can promote tumor formation and inhibit tumor formation. We have known that immune cells can be activated to promote tumor formation and progression[22]. The immune matrix score can predict the immune invasion of tumor microenvironment. We compared the immune matrix score, immune-related molecules' expression, tumor size distribution, and grade of four subtypes of immune cell infiltration.

### **Analysis of the difference between cold and hot tumors**

Cancer progression requires tumor cells to produce immune tolerance. Although the immune tolerance mechanism is involved, the tumor can be divided into two subtypes according to T lymphocyte infiltration[23, 24]. The different proportion of effector T cells and regulatory T cells in thermal tumors reflects the different degree of immunosuppression and affects the progress of tumor[25]. We compared the results of immune matrix score among different subtypes. Cluster 3 and 4 with higher immune matrix scores were classified as hot tumors, and cluster 1 and 2 were cold tumors. The lists of differentially expressed genes (DEGs) between controls and patients with DTC were generated using the LIMMA

method[26], where statistical significance was set as  $|\log\text{-fold change (logFC)}| > 1$  and Benjamini and Hochberg-corrected false discovery rates (FDR)  $< 0.05$ . A hierarchical cluster heatmap based on Euclidean distance was generated using the pheatmap R package (Version: 1.0.12) and represents DEGs' expression intensity and direction.

### **Enrichment analysis of differentially expressed genes and construction of protein interaction network of differentially expressed genes**

We use R software package "clusterprofile" (version 3.7.0; <http://cytoscape.org/>) Gene Ontology (go) and the Kyoto Encyclopedia of genes and genomes (KEGG) pathway enrichment analysis were performed. The selection criteria were  $\log_{2}fc > 1$  or  $< -0.5$ ,  $adjpvalue < 0.05$ [27]. DEGs-encoded proteins and PPI network information data were obtained using the Search Tool for the Retrieval of Interacting Genes (STRING) database (<http://string-db.org>). Cytoscape software (version 3.7.0; <http://cytoscape.org/>) was used for visualization of the interactions among the candidate DEGs[28].

### **Objective to search for tumor related regulatory molecules and enrich them**

Based on the above research, we obtained the differential genes between tumor and adjacent tumor, and then crossed the genes up-regulated in cold tumor ( $ADJP < 0.05$ ) with the genes over expressed in tumor. These genes are related to cancer and participate in negative immune regulation. Similarly, the up-regulated genes in hot tumors are crossed with the low expressed genes in cancer, which are positive immune regulatory genes. Finally, we used go and KEGG enrichment analysis to obtain the immune regulatory molecules.

### **Differences between lncRNA and miRNA in cold and hot tumors**

We performed the differential analysis (up:  $|\log_{2}fc| > 1$ , down:  $|\log_{2}fc| > 1$ , adjusted p-value  $< 0.05$ ) by comparing cold and hot tumors in the R computing environment using limma package.

### **Construction of hubcemanetwork**

To ensure the functions of the lncRNAs, miRNAs, and mRNAs in the ceRNA network and to further improve the ceRNA network reliability, a coexpression network of DElncRNAs, DEmiRNAs, and DEmRNAs was constructed; the ggalluvial R package (Version: 0.9.1) was used to visualize the ceRNA network[29]. MiRNAs are 19-23nt short RNAs transcribed from endogenous transcriptome and distributed throughout the cell[30]. Long non coding RNAs (lncRNAs) are involved in a variety of cellular functions, most of which require interaction with one or more RNA binding proteins (RBPs)[31]. Competitive endogenous RNA (Cerna) includes protein coding RNA, tRNA, rRNA, long non coding RNA (lncRNA), pseudogene RNA and circular RNA [32]. DEmRNAs targeted by the DEmiRNAs were retrieved from the miRDB (Version 5.0; <http://mirdb.org>), miRTarBase (Version 7.0; <http://mirtarbase.mbc.nctu.edu.tw/>), and TargetScan (Version 7.2; [http://www.targetscan.org/vert\\_72/](http://www.targetscan.org/vert_72/)) databases[33-35]. Besides, Cytoscape software (<https://cytoscape.org/>) was used to visualize the relationship of ceRNA network. MiRNA target genes were predicted by mirdb, targetscan, and mirtarbase. We constructed an immune related Cerna network

with different lncRNA, miRNA and mRNA, and defined the first five networks as hubnet through the cytohub module.

## Results

### **Data download, evaluation of immune invasion and comparison of immune cells between cancer and adjacent tumor**

In this study, we first from TCGA official website <https://portal.gdc.cancer.gov/> We downloaded the mRNA and lncRNA data of osteosarcoma patients, miRNA and corresponding clinical data are UCSC xena <https://xenabrowser.net/> It is downloaded directly from the database. We used four methods: single sample gene set environment analysis (ssgsea), microenvironment cell populations (MCP) - counter, cibersort and xcell to compare the differences of immune cells between cancer and adjacent tissues. The four methods included different numbers of immune cells (cibersort: 22, ssgsea: 28, MCP counter: 10, xcell: 67). In order to reduce possible errors, we included all four evaluation methods into the study. We used four methods to compare the difference of immune cells between cancer and paracancerous. The results showed that the number of most immune cells in cancer was significantly lower than that in paracancerous (Figure 1A-D), and the immune and matrix scores in cancer were lower than that in paracancerous (Figure 1E).

### **Correlation of immune cells**

We analyzed the correlation of immune cells in tumor and paracancerous samples (Figure 2A-H). The results showed that the correlation of immune cells in tumor and paracancerous samples was significantly different, which may be due to the synergistic infiltration of immune cells activated by cancer antigen. In tumor tissue, the synergistic effect of different immune cells constitutes the tumor immune microenvironment, and plays an important role in tumor invasion and development.

### **Cell consensus clustering**

We classified thyroid cancer samples into different subtypes according to the results of different immune infiltration. The cumulative distribution function (CDF) is drawn to identify the number of optimal subgroups (Figure 3A, B & C). Finally, we identified four different subtypes. We compared the distribution of cells and the expression of immune matrix score between different subtypes by mapping heat map. Among the four subtypes, we compared the differences of tumor purity, immune score and stromal score in innate and adaptive immune mechanisms. We observed that the tumor purity of cluster3 and 4 was lower than that of the other two subtypes, while the immunescore and stromalscore were higher than those of the other two subtypes (Figure 3D).

### **Immune matrix scores and immune-related molecules among different subtypes**

Immune infiltration in the tumor microenvironment can be assessed by the immune matrix score. We respectively compared the immune matrix scores and the expression levels of immune-related molecules

in 4 different immune infiltrating subtypes (Figure 4A-F). According to the ESTIMATE algorithm, the matrix score is between -600 and 1700, and the immune score is between -800 and 3500. Among them, the immune scores of clusters 3 and 4 are significantly higher than those of clusters 1 and 2. This means that immune and matrix scores are meaningful in the classification of subtypes.

### **Differences in tumor size distribution and grade**

The American Joint Committee on Cancer/International Cancer Control (AJCC/UICC) released the 8th edition of the TNM staging system[36]. Where T represents the size of the cancer, N represents the lymph node metastasis or not, and M represents whether there is distant metastasis. However, the TNM staging system alone is not enough to accurately reflect the disease stage of the tumor, and other tumor-related factors have also been paid more and more attention, so the different grades of tumors (stage1-4) are named[37]. We draw a pie chart to understand the difference in tumor size distribution and grading between different subtypes of thyroid cancer (Figure 5A-H). Among them, the ratio of cluster4 at T1 and stage1 is significantly higher than that of other clusters, which suggests that hot tumors have a lower pathological stage than cold tumors.

### **Analysis of differences between cold and hot tumors**

We define clusters 3 and 4 with higher immune matrix scores as hot tumors, and clusters 1 and 2 with lower immune matrix scores as cold tumors. We drew a heat map and found that the immune cell types of thyroid cancer and adjacent samples were significantly different (Figure 6A). We used limma package ( $|\text{Log}_2\text{FC}| > 2$ , adjusted P value  $< 0.05$ ) to identify differentially expressed genes, and drew a volcano map to compare the molecular differences in gene expression between cold and hot tumors (Figure 6B).

### **Enrichment analysis of differential genes and construction of protein interaction network of differential genes**

Through screening, we found 568 up-regulated differential genes and 412 down-regulated differential genes. We use the clusterProfiler package to perform GO enrichment and KEGG pathway enrichment analysis on differential genes related to cold and hot tumors[38]. The up-regulated differential genes are mainly enriched in regulation of leukocyte activation (GO:0002694), T cell activation (GO:0042110), regulation of lymphocyte activation (GO:0051249) and leukocyte cell-cell adhesion (GO:0007159); down-regulated The genes are mainly enriched in hormone metabolic process (GO:0042445), thyroid hormone metabolic process (GO:0042403), hormone biosynthetic process (GO:0042446) and thyroid hormone generation (GO:0006590) (Figure 7A-D). KEGG pathway enrichment analysis showed that the up-regulated differential genes were mainly enriched in Hematopoietic cell lineage (hsa04640), Cytokine-cytokine receptor interaction (hsa04060) and Viral protein interaction with cytokine and cytokine receptor (hsa04061), and the down-regulated differential genes were mainly enriched in Thyroid hormone synthesis (hsa04918), Rap1 signaling pathway (hsa04015) and Cortisol synthesis and secretion (hsa04927) (Figure 7E-H). We used the STRING online tool to construct a protein interaction network (PPI) of differential genes related to cold and hot tumors. Among the up-regulated differential genes, we chose

the node connection greater than 20 as the hub gene, and in the down-regulated differential gene we chose the node connection Those greater than 10 are defined as hub genes (Figure 7I: down-regulated differential gene protein interaction network; Figure 7J: up-regulated differential gene protein interaction network). We have 19 up-regulated hub genes and 9 down-regulated hub genes. In the protein interaction network, we have observed a wide range of links between the markers related to the matrix in the differential gene and the markers related to the immune in the differential gene, which may be related to the higher scores of immunity and matrix in hot tumors .

### **Looking for tumor-related regulatory molecules and enrichment analysis of tumor-related molecules**

T cell-based cancer immunotherapy, such as checkpoint suppression or adoptive cell therapy, has greatly changed the way cancer is treated[39]. We know that immunity plays a very important role in the occurrence and development of tumors. We first calculated the differential genes between cancer and adjacent cancers, and then intersected the genes up-regulated in cold tumors ( $\text{adjpvalue} < 0.05$ ) with the highly expressed genes in the tumors. A total of 717 genes were negatively regulated immune genes related to cancer. So these genes may be therapeutic targets. Similarly, the genes that are up-regulated in hot tumors ( $\text{adjpvalue} < 0.05$ ) are crossed with genes that are low-expressed in cancer. There are 1246 genes in total. These genes are positively regulated immunity, so agonists need to be added during the treatment. We enrich these differential genes. In GO enrichment analysis, the up-regulated differential genes are mainly enriched in: T cell activation (GO:0042110), regulation of T cell activation (GO:0050863), regulation of lymphocyte activation (GO:0051249), external side of plasma membrane (GO:0009897), cytokine activity (GO:0005125), etc. (Figure 8A&B). In the enrichment analysis of the KEGG pathway, the up-regulated differential genes were mainly enriched in: Cytokine-cytokine receptor interaction (hsa04060) and Viral protein interaction with cytokine and cytokine receptor (hsa04061) (Figure 8C&D). The down-regulated differential genes are not enriched in related pathways.

### **Differences between lncRNA and miRNA in cold and hot tumors**

We will compare the expression differences of the downloaded lncRNA and miRNA between the two groups of cold and hot tumors (Figure 9A-D). We used limma package to identify differentially expressed genes, and drawn volcano maps to compare the differentially expressed genes between cold and hot tumors. A total of 629 differential lncRNAs ( $|\text{Log}_2\text{FC}| > 1$ , adjusted P value  $< 0.05$ ) and 114 differential miRNAs ( $|\text{Log}_2\text{FC}| > 0.5$ , adjusted P value  $< 0.05$ ) were obtained.

To better understand the effect of lncRNAs on mRNAs mediated by combination with miRNAs in DTC, we built a ceRNA network based on the above mentioned data and used the ggalluvial R package (Version: 0.9.1) to visualize the network. miRNA target gene prediction passed miRDB, TargetsCan, and miRTarBase take the intersection. The circle represents lncRNA ( $\text{adj} < 0.05$ ), the diamond represents mRNA ( $\text{adj} < 0.05$ ,  $\log\text{FC} > 1$ ,  $\log\text{FC} < 0.05$ ), and the triangle represents miRNA ( $\text{adj} < 0.05$ ). Finally, the cytohub module calculates the first 5 networks defined as hubnet (Figure 10). Five miRNAs, hsa-mir-204, hsa-mir-128, hsa-mir-214, hsa-mir-150 and hsa-mir-338, are located in the central area of the network, and are of great significance to the immunity of thyroid cancer.

## Discussion

Thyroid cancer is one of the malignant tumors, the screening of RNAs transcripts was facilitated in the past 20 years, lncRNAs and miRNAs were tested to be an emerging factor which strongly associated with tumorigenesis and metastasis in thyroid carcinoma[40-42]. Interactions between tumor cells and various components of TME are significant and contribute to all the hallmarks of cancer[43]. TME can affect how a tumor grows and spreads. Therefore, identification of key genes in thyroid cancer microenvironment aids appropriate management and treatment of thyroid cancer. The immune infiltration of TME is very important for the immune-related treatment of thyroid cancer. Our research aims to identify immune-related mRNAs, lncRNAs and miRNAs, and further explore the relationship between them.

More and more studies indicate that lncRNAs play a vital role in biological functions through multiple levels of regulation, which involve transcriptional, posttranscriptional, and epigenetic regulation [44, 45]. A large number of studies have shown that there is a complex and closely related regulatory network between miRNA and lncRNA. In triple negative breast cancer[46], the relationship between miRNA and lncRNA has been determined. The ceRNA hypothesis was proposed to explain the mechanism of tumorigenesis; this hypothesis provides a novel guiding theory and suggests valuable strategies and research directions for the diagnosis and treatment of malignancies [47], lncRNAs with sequences similar to their target miRNA can regulate mRNA expression by acting as a sponge of miRNA.

At present, there are few studies on ceRNA networks related to thyroid cancer. In this study, we downloaded mRNA, lncRNA and miRNA of patients with thyroid cancer from TCGA. We first divide the sample into cold tumors and hot tumors through immune evaluation[48]. In the tumor microenvironment of "hot" tumors, the degree of immune infiltration and immune effect is relatively higher, and it has strong antigen presentation ability and T cell activation. Such immune infiltration will lead to the production of tumor-specific CD8 + T cells, which can eliminate cancer cells and generate systemic tumor-specific immunity, forming a long-term anti-tumor memory response[11, 12]. However, "cold" tumors have no immune cell infiltration in the tumor microenvironment or are mainly infiltrated by inhibitory regulatory cell subtypes (including regulatory T cells (Tregs), regulatory B cells (Bregs) and myeloid suppressor cells (MDSCs)) [13-15], the growth of cancer is immunologically uninhibited. We used DEGs to construct ceRNA, and finally identified a hubnet centered on five miRNAs: hsa-mir-204, hsa-mir-128, hsa-mir-214, hsa-mir-150 and hsa-mir-338. miRNAs are approximately 22 nucleotides long RNA molecules that bind to the 3'-untranslated region (3'-UTR) of their respective target genes, and exert their influence on gene expression by inhibiting protein translation or degrading mRNA[49]. We already know that miRNA plays a very important role in the occurrence and development of cancer. hsa-mir-204 has been determined to be related to many cancer malignancies, including: melanoma [50], breast cancer [51] and liver cancer [52]. hsa-mir-128-3p can increase the chemical sensitivity of colorectal cancer to chemotherapy drugs [53]. hsa-mir-128-3p has an important relationship with the drug resistance of chemotherapeutics, but there are few related reports in thyroid cancer. It is worth noting that hsa-mir-214 has been found to play a very important role in regulating the proliferation and metastasis of papillary thyroid cancer cells [54]. hsa-mir-150 and hsa-mir-338 are also related to tumor proliferation and cell invasion, such as colorectal

cancer [55], non-small cell carcinoma [56] and cervical cancer [57]. In thyroid cancer, these related hub miRNAs and related lncRNA and mRNA jointly affect the tumor microenvironment. It may be closely related to immune cell infiltration and tumor invasion. Because the number of deaths from thyroid cancer samples is relatively small, in this study, hubnet was not analyzed for survival.

Inevitably, our research had some innate limitations which need to be addressed. Although this bioinformatics study was designed well, the major drawback of our study was the lack of in vivo and vitro validation. Our research has identified 5 hubnets related to thyroid cancer, which may be related to cancer immune infiltration and tumor invasion. This result has important significance for understanding the tumor invasion and mechanism of action in patients with thyroid cancer, and provides possible markers for treatment.

## Conclusions

In this study, we identified the hubnet centered on five miRNAs: hsa-mir-204, hsa-mir-128, hsa-mir-214, hsa-mir-150 and hsa-mir-338. In the ceRNA network we established, the relationship between miRNA, lncRNA and immune-related mRNA was determined. The understanding of the molecular role of thyroid cancer tumor microenvironment is of great significance.

## Abbreviations

TCGA: The Cancer Genome Atlas

UCSC: University of California Santa Cruz

KEGG: Kyoto Encyclopedia of Genes and Genomes

GO: Gene Ontology

DEGs: differentially expressed genes

## Declarations

### Authors' contributions

DJW and QXG contributed to the conception of the study. LGQ performed the data analyses. LYS contributed significantly to process data and wrote the manuscript. All of the authors read and approved the final manuscript.

### Ethical Statement

The data of this study are from TCGA and UCSC, and do not involve animal experiments and human specimens, no ethics-related issues.

## Consent for publication

Not Applicable

## Availability of data and material

The data of this study are from GEO database.

## Competing interests

The authors declare that they have no conflicts of interest.

## Funding

The present study was supported by the key projects of high learning institutions in Henan province (Grant No. 18A320008)

## Acknowledgments

We thank these researchers who gave their data for this analysis. It is cheerful to acknowledge their contribution.

## References

1. Siegel RL, Miller KD, Jemal A. **Cancer Statistics, 2017.** *CA Cancer J Clin* 2017; 67(1):7-30.
2. Mirian C, Gronhoj C, Jensen DH, Jakobsen KK, Karnov K, Jensen JS, et al. **Trends in thyroid cancer: Retrospective analysis of incidence and survival in Denmark 1980-2014.** *Cancer Epidemiol* 2018; 55:81-87.
3. Stjepanovic N, Capdevila J. **Multikinase inhibitors in the treatment of thyroid cancer: specific role of lenvatinib.** *Biologics* 2014; 8:129-139.
4. Raue F, Frank-Raue K. **Thyroid Cancer: Risk-Stratified Management and Individualized Therapy.** *Clin Cancer Res* 2016; 22(20):5012-5021.
5. Steven A, Fisher SA, Robinson BW. **Immunotherapy for lung cancer.** *Respirology* 2016; 21(5):821-833.
6. Morrison AH, Byrne KT, Vonderheide RH. **Immunotherapy and Prevention of Pancreatic Cancer.** *Trends Cancer* 2018; 4(6):418-428.
7. Sugie T. **Immunotherapy for metastatic breast cancer.** *Chin Clin Oncol* 2018; 7(3):28.
8. Balkwill F, Mantovani A. **Inflammation and cancer: back to Virchow?** *Lancet* 2001; 357(9255):539-545.
9. Dunn GP, Bruce AT, Ikeda H, Old LJ, Schreiber RD. **Cancer immunoediting: from immunosurveillance to tumor escape.** *Nat Immunol* 2002; 3(11):991-998.
10. Galluzzi L, Senovilla L, Zitvogel L, Kroemer G. **The secret ally: immunostimulation by anticancer drugs.** *Nat Rev Drug Discov* 2012; 11(3):215-233.

11. Ayers M, Lunceford J, Nebozhyn M, Murphy E, Loboda A, Kaufman DR, et al. **IFN-gamma-related mRNA profile predicts clinical response to PD-1 blockade.** *J Clin Invest* 2017; 127(8):2930-2940.
12. Martinez-Lostao L, Anel A, Pardo J. **How Do Cytotoxic Lymphocytes Kill Cancer Cells?** *Clin Cancer Res* 2015; 21(22):5047-5056.
13. Kather JN, Suarez-Carmona M, Charoentong P, Weis CA, Hirsch D, Bankhead P, et al. **Topography of cancer-associated immune cells in human solid tumors.** *Elife* 2018; 7.
14. Mauri C, Menon M. **Human regulatory B cells in health and disease: therapeutic potential.** *J Clin Invest* 2017; 127(3):772-779.
15. Lindau D, Gielen P, Kroesen M, Wesseling P, Adema GJ. **The immunosuppressive tumour network: myeloid-derived suppressor cells, regulatory T cells and natural killer T cells.** *Immunology* 2013; 138(2):105-115.
16. Jia D, Li S, Li D, Xue H, Yang D, Liu Y. **Mining TCGA database for genes of prognostic value in glioblastoma microenvironment.** *Aging (Albany NY)* 2018; 10(4):592-605.
17. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. **Robust enumeration of cell subsets from tissue expression profiles.** *Nat Methods* 2015; 12(5):453-457.
18. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, et al. **Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1.** *Nature* 2009; 462(7269):108-112.
19. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. **Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression.** *Genome Biol* 2016; 17(1):218.
20. Aran D, Hu Z, Butte AJ. **xCell: digitally portraying the tissue cellular heterogeneity landscape.** *Genome Biol* 2017; 18(1):220.
21. Meurette O, Mehlen P. **Notch Signaling in the Tumor Microenvironment.** *Cancer Cell* 2018; 34(4):536-548.
22. Kim J, Bae JS. **Tumor-Associated Macrophages and Neutrophils in Tumor Microenvironment.** *Mediators Inflamm* 2016; 2016:6058147.
23. Gajewski TF. **The Next Hurdle in Cancer Immunotherapy: Overcoming the Non-T-Cell-Inflamed Tumor Microenvironment.** *Semin Oncol* 2015; 42(4):663-671.
24. Sharma P, Allison JP. **The future of immune checkpoint therapy.** *Science* 2015; 348(6230):56-61.
25. Fridman WH, Pages F, Sautes-Fridman C, Galon J. **The immune contexture in human tumours: impact on clinical outcome.** *Nat Rev Cancer* 2012; 12(4):298-306.
26. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. **limma powers differential expression analyses for RNA-sequencing and microarray studies.** *Nucleic Acids Res* 2015; 43(7):e47.
27. Gong J, Jiang H, Shu C, Hu MQ, Huang Y, Liu Q, et al. **Integrated analysis of circular RNA-associated ceRNA network in cervical cancer: Observational Study.** *Medicine (Baltimore)* 2019; 98(34):e16922.
28. Kohl M, Wiese S, Warscheid B. **Cytoscape: software for visualization and analysis of biological networks.** *Methods Mol Biol* 2011; 696:291-303.

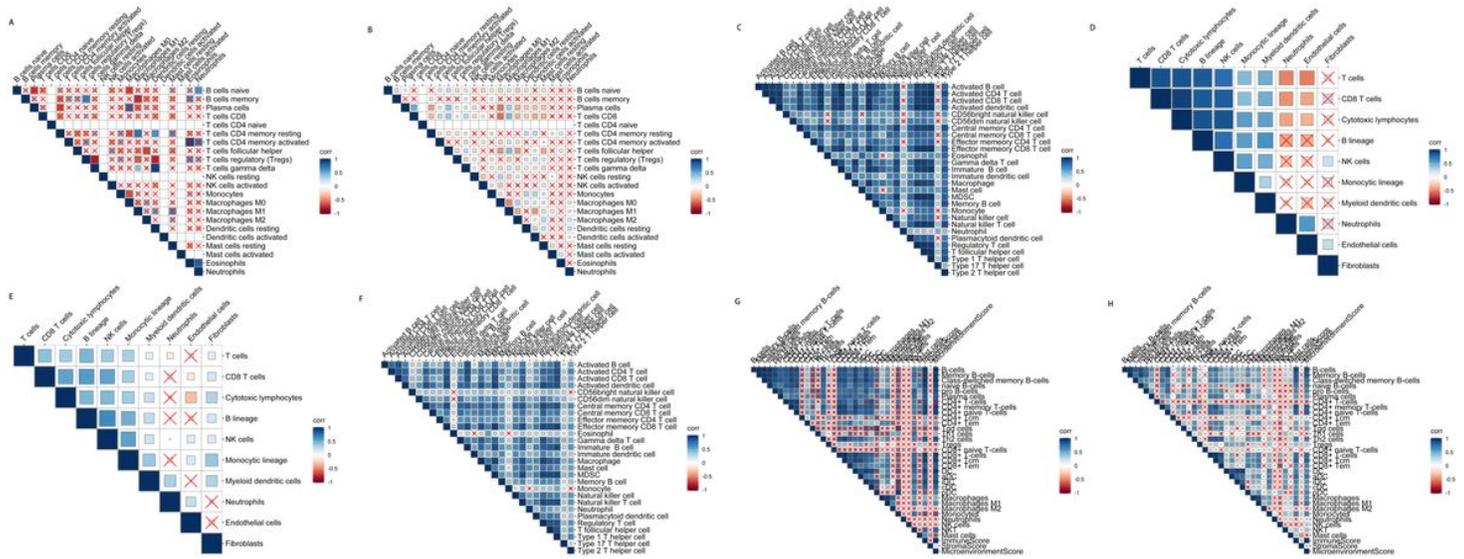
29. Rosvall M, Bergstrom CT. **Mapping change in large networks.** *PLoS One* 2010; 5(1):e8694.
30. Lundstrom K. **Micro-RNA in disease and gene therapy.** *Curr Drug Discov Technol* 2011; 8(2):76-86.
31. Ferre F, Colantoni A, Helmer-Citterich M. **Revealing protein-lncRNA interaction.** *Brief Bioinform* 2016; 17(1):106-116.
32. Qu S, Yang X, Li X, Wang J, Gao Y, Shang R, et al. **Circular RNA: A new star of noncoding RNAs.** *Cancer Lett* 2015; 365(2):141-148.
33. Wong N, Wang X. **miRDB: an online resource for microRNA target prediction and functional annotations.** *Nucleic Acids Res* 2015; 43(Database issue):D146-152.
34. Chou CH, Shrestha S, Yang CD, Chang NW, Lin YL, Liao KW, et al. **miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions.** *Nucleic Acids Res* 2018; 46(D1):D296-D302.
35. Agarwal V, Bell GW, Nam JW, Bartel DP. **Predicting effective microRNA target sites in mammalian mRNAs.** *Elife* 2015; 4.
36. Lee J, Lee SG, Kim K, Yim SH, Ryu H, Lee CR, et al. **Clinical Value of Lymph Node Ratio Integration with the 8(th) Edition of the UICC TNM Classification and 2015 ATA Risk Stratification Systems for Recurrence Prediction in Papillary Thyroid Cancer.** *Sci Rep* 2019; 9(1):13361.
37. Perrier ND, Brierley JD, Tuttle RM. **Differentiated and anaplastic thyroid carcinoma: Major changes in the American Joint Committee on Cancer eighth edition cancer staging manual.** *CA Cancer J Clin* 2018; 68(1):55-63.
38. Yu G, Wang LG, Han Y, He QY. **clusterProfiler: an R package for comparing biological themes among gene clusters.** *OMICS* 2012; 16(5):284-287.
39. Medina-Echeverz J, Hinterberger M, Testori M, Geiger M, Giessel R, Bathke B, et al. **Synergistic cancer immunotherapy combines MVA-CD40L induced innate and adaptive immunity with tumor targeting antibodies.** *Nat Commun* 2019; 10(1):5041.
40. Li HM, Yang H, Wen DY, Luo YH, Liang CY, Pan DH, et al. **Overexpression of LncRNA HOTAIR is Associated with Poor Prognosis in Thyroid Carcinoma: A Study Based on TCGA and GEO Data.** *Horm Metab Res* 2017; 49(5):388-399.
41. Wang Y, He H, Li W, Phay J, Shen R, Yu L, et al. **MYH9 binds to lncRNA gene PTCS2 and regulates FOXE1 in the 9q22 thyroid cancer risk locus.** *Proc Natl Acad Sci U S A* 2017; 114(3):474-479.
42. Zhu H, Lv Z, An C, Shi M, Pan W, Zhou L, et al. **Onco-lncRNA HOTAIR and its functional genetic variants in papillary thyroid carcinoma.** *Sci Rep* 2016; 6:31969.
43. Hanahan D, Weinberg RA. **Hallmarks of cancer: the next generation.** *Cell* 2011; 144(5):646-674.
44. Muers M. **RNA: Genome-wide views of long non-coding RNAs.** *Nat Rev Genet* 2011; 12(11):742.
45. Caley DP, Pink RC, Trujillano D, Carter DR. **Long noncoding RNAs, chromatin, and development.** *ScientificWorldJournal* 2010; 10:90-102.
46. Augoff K, McCue B, Plow EF, Sossey-Alaoui K. **miR-31 and its host gene lncRNA LOC554202 are regulated by promoter hypermethylation in triple-negative breast cancer.** *Mol Cancer* 2012; 11:5.

47. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. **A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language?** *Cell* 2011; 146(3):353-358.
48. Galon J, Bruni D. **Approaches to treat immune hot, altered and cold tumours with combination immunotherapies.** *Nat Rev Drug Discov* 2019; 18(3):197-218.
49. Patel N, Garikapati KR, Makani VKK, Nair AD, Vangara N, Bhadra U, et al. **Regulating BMI1 expression via miRNAs promote Mesenchymal to Epithelial Transition (MET) and sensitizes breast cancer cell to chemotherapeutic drug.** *PLoS One* 2018; 13(2):e0190245.
50. Diaz-Martinez M, Benito-Jardon L, Alonso L, Koetz-Ploch L, Hernando E, Teixido J. **miR-204-5p and miR-211-5p Contribute to BRAF Inhibitor Resistance in Melanoma.** *Cancer Res* 2018; 78(4):1017-1030.
51. Muller V, Oliveira-Ferrer L, Steinbach B, Pantel K, Schwarzenbach H. **Interplay of lncRNA H19/miR-675 and lncRNA NEAT1/miR-204 in breast cancer.** *Mol Oncol* 2019; 13(5):1137-1149.
52. Yu Y, Wang Y, Xiao X, Cheng W, Hu L, Yao W, et al. **MiR-204 inhibits hepatocellular cancer drug resistance and metastasis through targeting NUAK1.** *Biochem Cell Biol* 2019; 97(5):563-570.
53. Liu T, Zhang X, Du L, Wang Y, Liu X, Tian H, et al. **Exosome-transmitted miR-128-3p increase chemosensitivity of oxaliplatin-resistant colorectal cancer.** *Mol Cancer* 2019; 18(1):43.
54. Liu F, Lou K, Zhao X, Zhang J, Chen W, Qian Y, et al. **miR-214 regulates papillary thyroid carcinoma cell proliferation and metastasis by targeting PSMD10.** *Int J Mol Med* 2018; 42(6):3027-3036.
55. Fan H, Liu X, Zheng WW, Zhuang ZH, Wang CD. **MiR-150 alleviates EMT and cell invasion of colorectal cancer through targeting Gli1.** *Eur Rev Med Pharmacol Sci* 2017; 21(21):4853-4859.
56. Lu W, Zhang H, Niu Y, Wu Y, Sun W, Li H, et al. **Long non-coding RNA linc00673 regulated non-small cell lung cancer proliferation, migration, invasion and epithelial mesenchymal transition by sponging miR-150-5p.** *Mol Cancer* 2017; 16(1):118.
57. Luan X, Wang Y. **LncRNA XLOC\_006390 facilitates cervical cancer tumorigenesis and metastasis as a ceRNA against miR-331-3p and miR-338-3p.** *J Gynecol Oncol* 2018; 29(6):e95.

## Figures

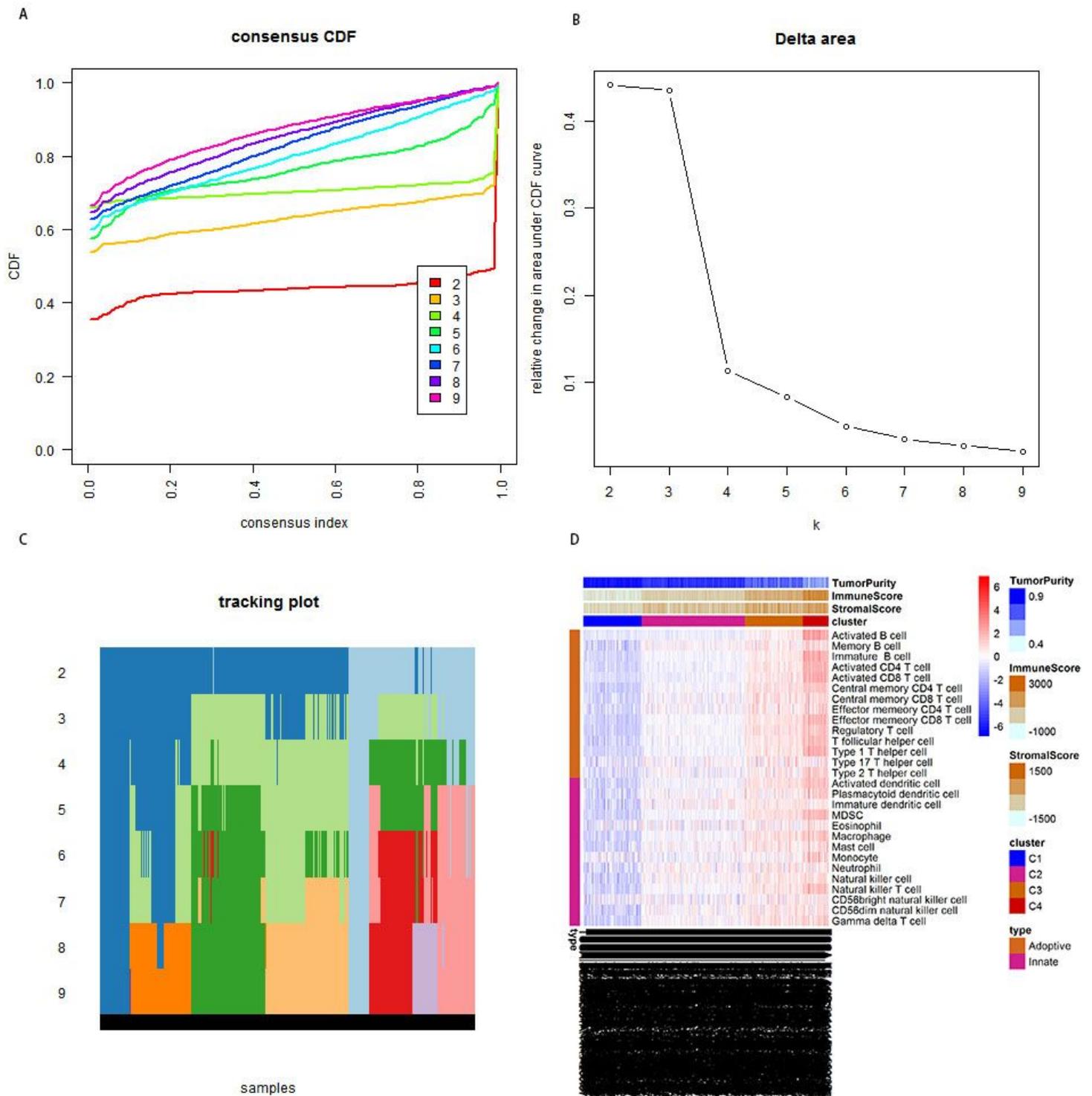


cancer is higher than that in adjacent cancer. (C)ssgsea evaluation method shows that the number of most T cells and B cells in cancer is lower than that in adjacent cancer, while the number of mast cells, monocytic lineage and neutrophils in cancer is higher than that in adjacent cancer. (D)X-cell evaluation method shows that the number of B cells, T-cells and monocyte in cancer is less than that in adjacent cancer, while the number of dendritic cells, mast cells and natural killer cells is higher than that in adjacent cancer.(E)the immune and matrix scores in carcinoma were lower than those in adjacent carcinoma.



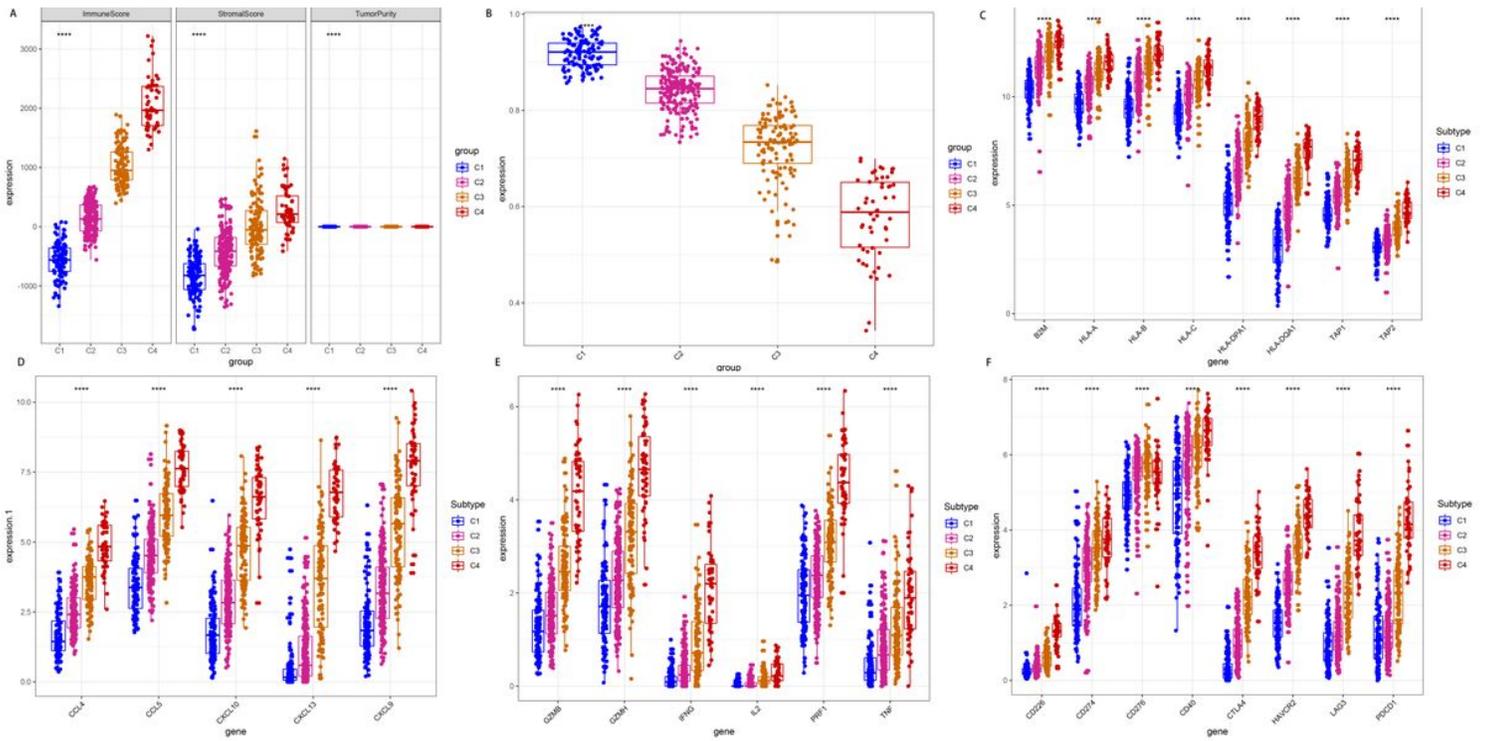
**Figure 2**

correlation of immune cells. (A-H)The correlation of immune cells measured by four different methods is shown in the figure,in which blue represents positive correlation and red represents negative correlation.



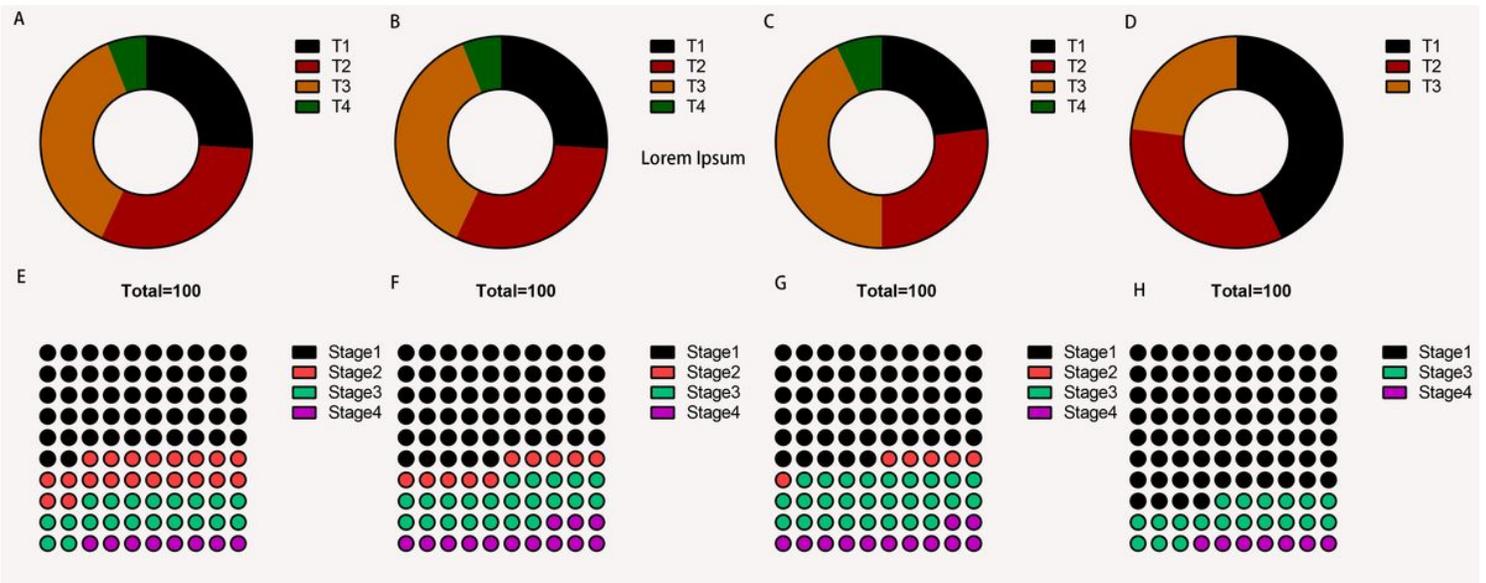
**Figure 3**

Consistent cluster map of immune cells. (A, B & C) CDF (cumulative distribution function) shows that when k is 4, it is the optimal number of subgroups. (D) A heat map is drawn to show the difference of TumorPurity, ImmuneScore and StromalScore among different subtypes in the innate and adaptive immune mechanisms.



**Figure 4**

Box plot of immune scores and expression levels of immune-related molecules between different subtypes. (A&B) The immune score and immune matrix score difference between the four clusters, clusters 3 and 4 have higher immune scores than clusters. (C-F) Among the four clusters, the expression levels of multiple immune molecules are compared, and clusters 3 and 4 show higher expression levels of immune molecules. \*,  $P < 0.05$ . \*\*,  $P < 0.01$ . \*\*\*,  $P < 0.001$ . \*\*\*\*,  $P < 0.0001$ .



**Figure 5**

Tumor size distribution and grading differences between different subtypes. (A-H) shows the distribution ratio and grading of thyroid cancer patients' tumor sizes between different clusters. Among them, the

ratio of cluster4 in T1 and stage1 is significantly higher than that of other clusters.

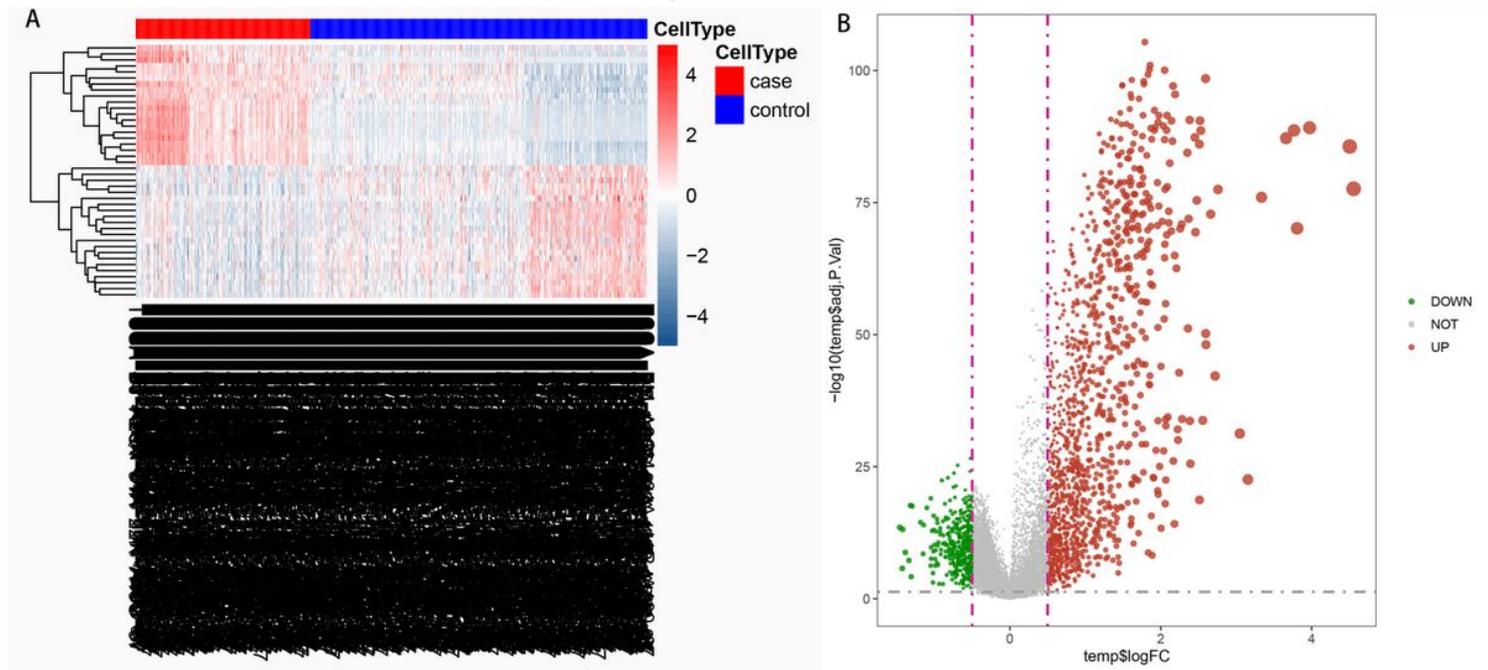
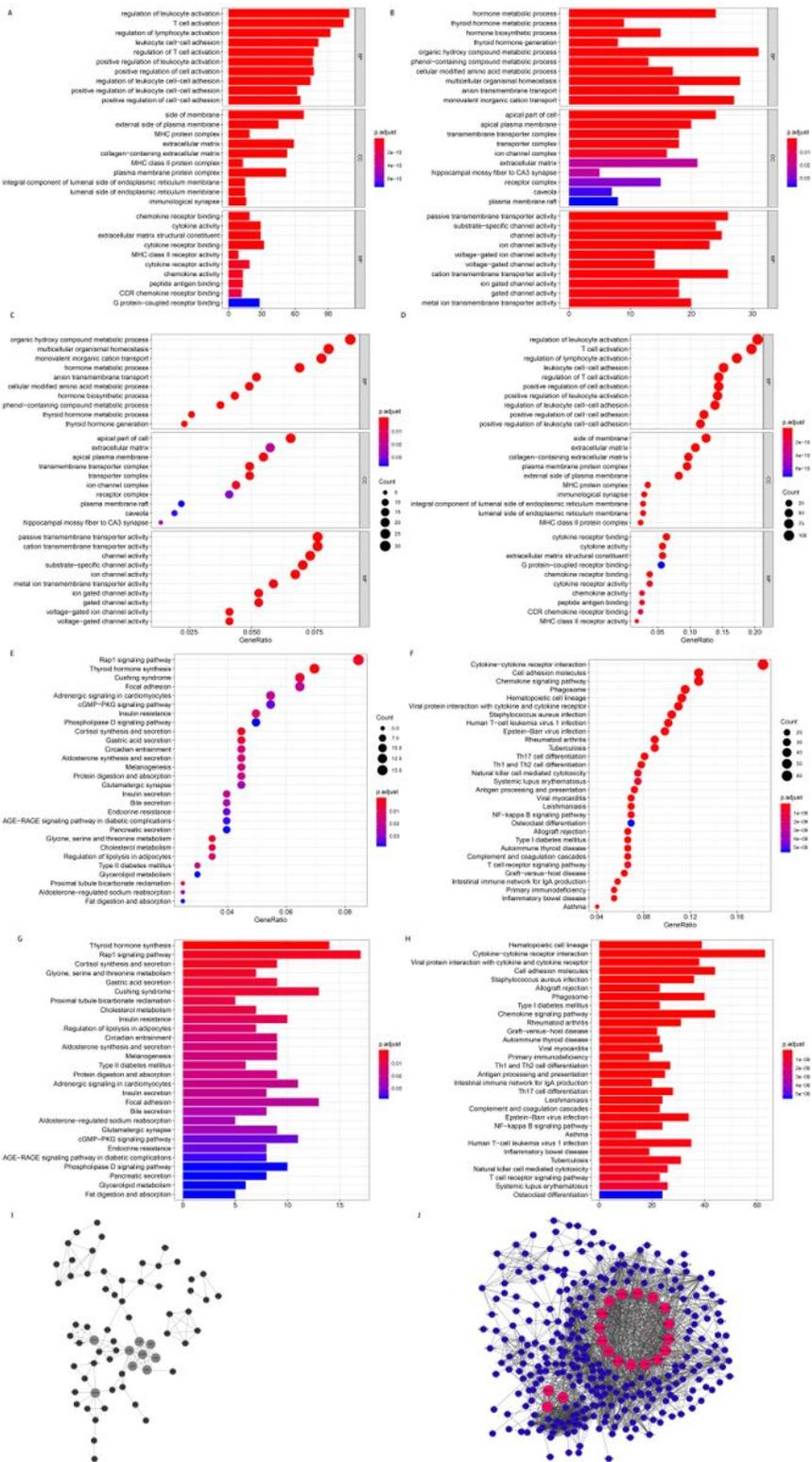


Figure 6

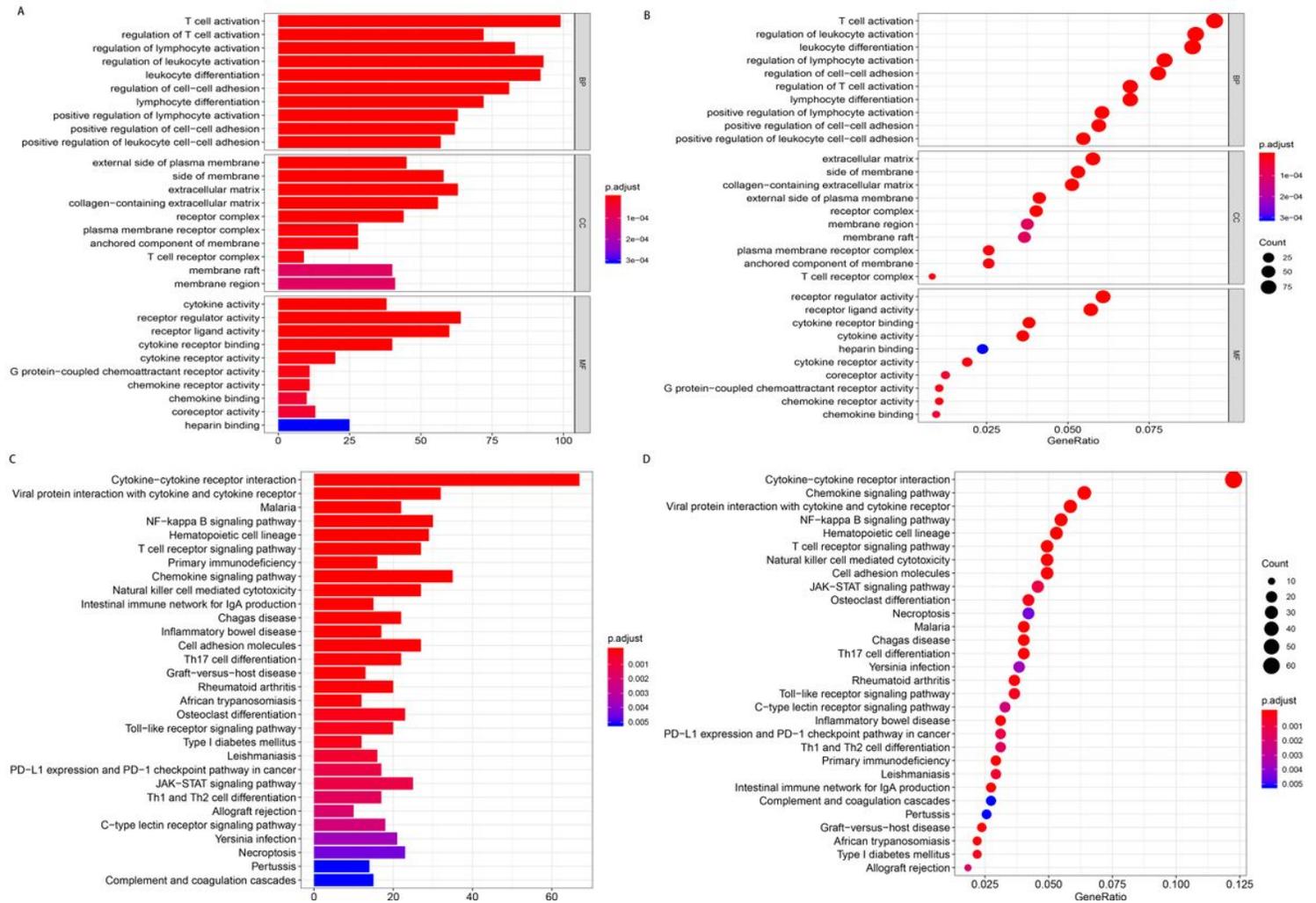
Differences in immune cell types in different samples and differentially expressed genes in cold and hot tumors. (A) shows the difference in immune cell types between cancer and adjacent samples. (B) shows the up-regulation and down-regulation of differentially expressed genes in cold and hot tumors.



**Figure 7**

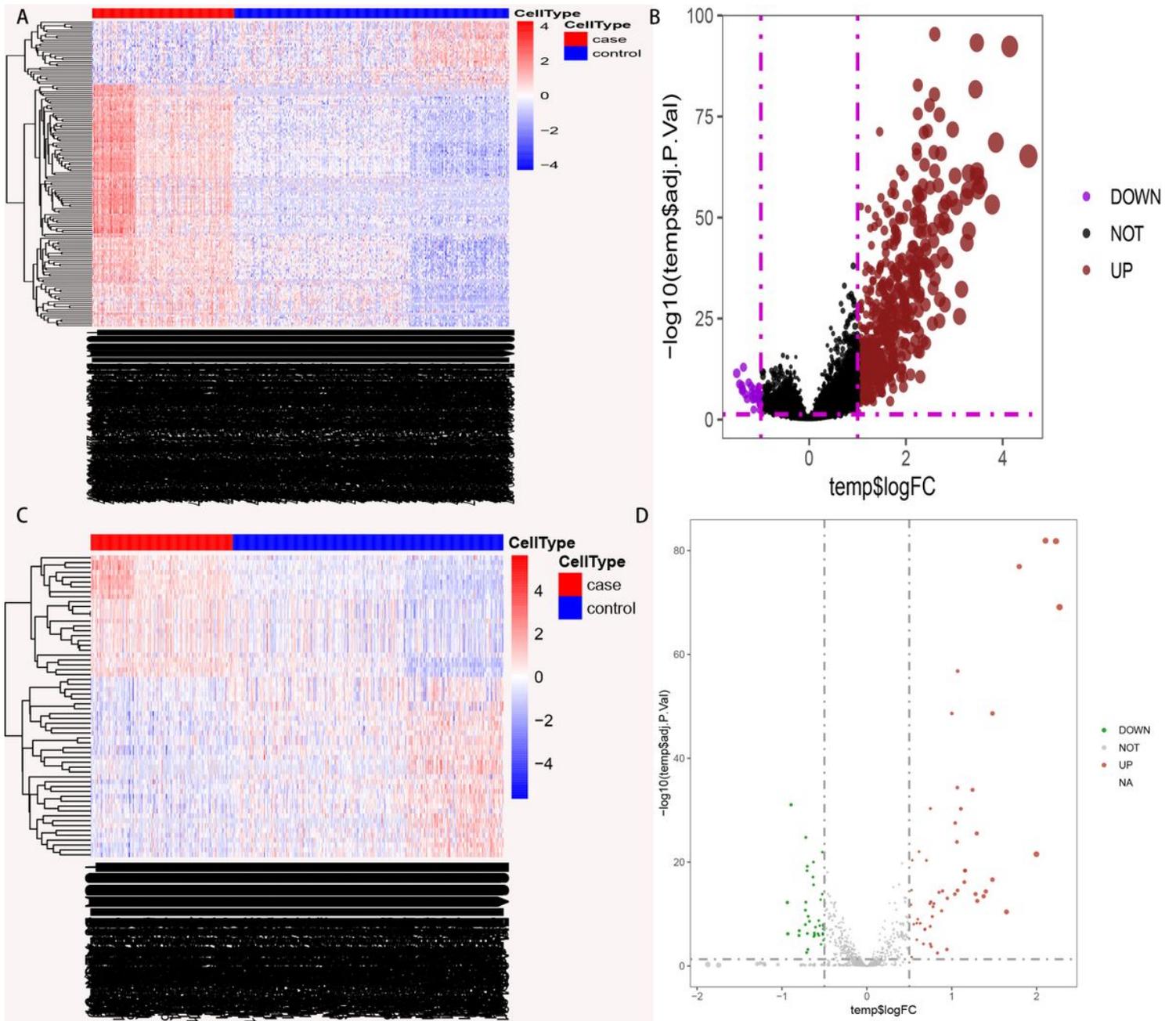
The GO, KEGG and PPI analysis of DEGs related to cold and hot tumors. (A-D) The GO analysis of DEGs related to cold and hot tumors. (E-H) The KEGG pathway analysis of DEGs related to cold and hot tumors. (I&J) The PPI analysis of DEGs related to cold and hot tumors. Among them, (I) shows down-regulation and (J) shows up-regulation. BP, Biological Process; CC, Cellular Component; MF, Molecular Function; GO,

Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; PPI, Protein–Protein Interaction; DEGs, Differentially expressed genes.



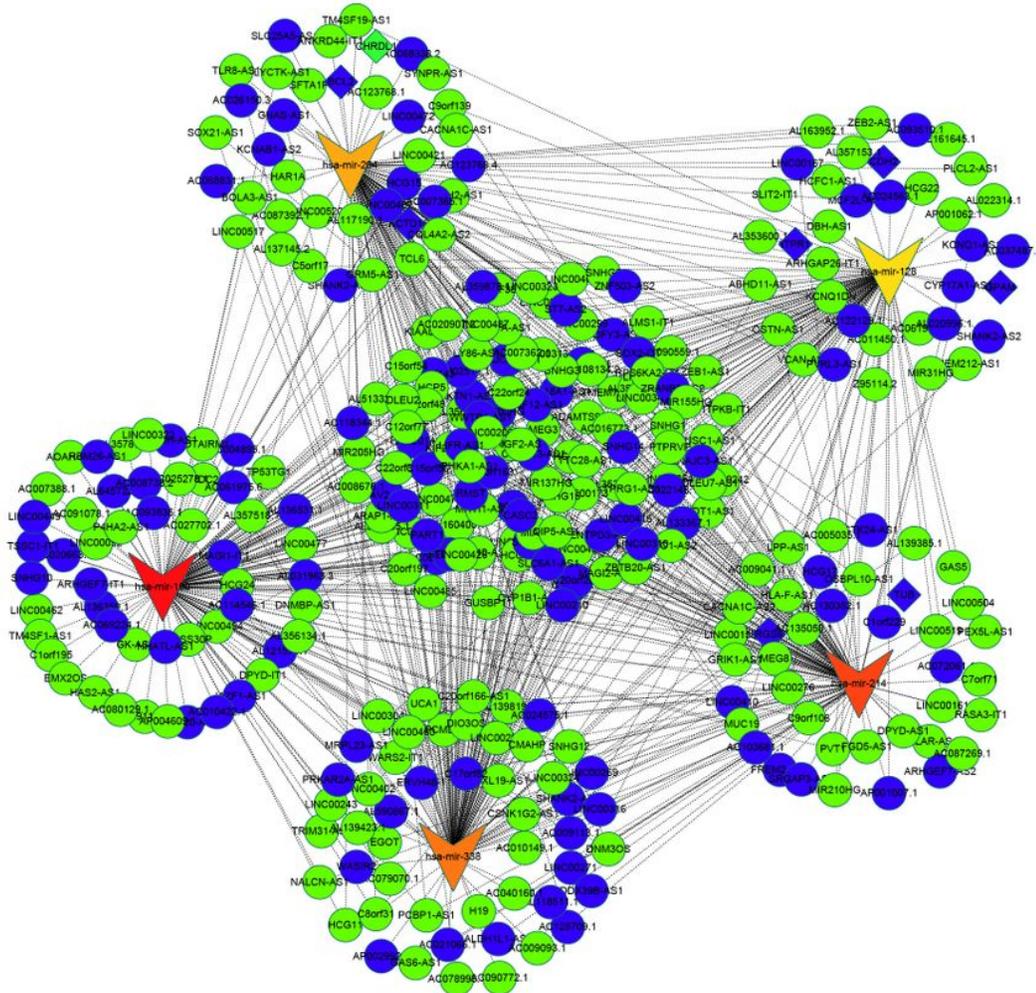
**Figure 8**

GO and KEGG enrichment analysis of immune-related genes. (A&B) The up-regulated DEGs are functionally enriched in GO. (C&D) The main pathway of up-regulated DEGs is enriched in KEGG.



**Figure 9**

Differences in the expression of lncRNA and miRNA in cold and hot tumors. (A&B) The differences in the expression of lncRNA in cold and hot tumors. (C&D) The differences in the expression of miRNA in cold and hot tumors.



**Figure 10**

Construction of ceRNA network. A ceRNA network consisting of 629 lncRNAs, 114 miRNAs and 980 mRNAs was constructed. We select the first five networks and define them as hubnet. The circle represents lncRNA ( $adj < 0.05$ ), the diamond represents mRNA ( $adj < 0.05$ ,  $\log FC > 1$ ,  $\log FC < 0.05$ ), and the triangle represents miRNA ( $adj < 0.05$ ).