

Identification of Ferroptosis Genes In Immune Infiltration and Prognosis In Thyroid Papillary Carcinoma Using Network Analysis

Ruoting Lin

Guangdong Pharmaceutical University

Conor E. Fogarty

University of the Sunshine Coast

Bowei Ma

The First Affiliated Hospital of Guangzhou University of Traditional Chinese Medicine

Hejie Li

University of the Sunshine Coast

Guoying Ni

Guangdong Pharmaceutical University

Xiaosong Liu

Guangdong Pharmaceutical University

Jianwei Yuan

Guangdong Pharmaceutical University

Tianfang Wang (✉ twang@usc.edu.au)

University of the Sunshine Coast

Research Article

Keywords: Thyroid papillary carcinoma, Immune-related genes, Tumour microenvironment, Ferroptosis genes, Prognostic model

Posted Date: April 9th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-389329/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 BMC Genomics on July 27th, 2021. See the published version at <https://doi.org/10.1186/s12864-021-07895-6>.

Abstract

Background: Papillary thyroid carcinoma (PTC) is the most common thyroid cancer. While many patients survive, a portion of PTC cases display high aggressiveness and even develop into refractory differentiated thyroid carcinoma. This may be alleviated by developing a novel model to predict the risk of recurrence. Ferroptosis is an iron-dependent form of regulated cell death (RCD) driven by lethal accumulation of lipid peroxides, is regulated by a set of genes and shows a variety of metabolic changes. To elucidate whether ferroptosis occurs in PTC, we analysed the gene expression profiles of the disease and established a new model for the correlation.

Methods: The thyroid carcinoma (THCA) datasets were downloaded from The Cancer Genome Atlas (TCGA), UCSC Xena and MisgDB, and included 502 tumour samples and 56 normal samples. A total of 60 ferroptosis related genes were summarised from MisgDB database. Gene set enrichment analysis (GSEA) and Gene set variation analysis (GSVA) were used to analyse pathways potentially involving PTC subtypes. Single sample GSEA (ssGSEA) algorithm was used to analyse the proportion of 28 kinds of immune cells in the tumour immune infiltration microenvironment in THCA and the hclust algorithm was used to conduct immune typing according to the proportion of immune cells. Spearman correlation analysis was performed on the ferroptosis gene expression and the correlation between immune infiltrating cells proportion. We established the WGCNA to identify genes modules that are highly correlated with the microenvironment of immune invasion. DEseq2 algorithm was further used for differential analysis of sequencing data to analyse the functions and pathways potentially involving hub genes. GO and KEGG enrichment analysis was performed using ClusterProfiler to explore the clinical efficacy of hub genes. Univariate Cox analysis was performed for hub genes combined with clinical prognostic data, and the results was included for lasso regression and constructed the risk regression model. ROC curve and survival curve were used for evaluating the model. Univariate Cox analysis and multivariate Cox analysis were performed in combination with the clinical data of THCA and the risk score value, the clinical efficacy of the model was further evaluated.

Results: We identified two subtypes in PTC, using *Braf* as the major mutation of subtype C1 and *NRAS* as the major mutation of subtype C2. The proportion of cluster 1 was significantly higher than cluster 2 in ferroptosis signature genes that are positively associated. Subsequent analyses of TME immune cells infiltration indicated cluster 1 was remarkably richer than cluster 2. The risk score of THCA in good performance evaluated by ROC curve and survival curve, in conjunction with univariate Cox analysis and multivariate Cox analysis results based on the clinical data showed that the risk score of the proposed model could be used as an independent prognostic indicator to predict the prognosis of patients with papillary thyroid cancer.

Conclusions: Our study found seven crucial genes, including *Ac008063.2*, *ApoE*, *Bcl3*, *Acap3*, *Alox5ap*, *Atxn2l* and *B2m*, and regulation of apoptosis by parathyroid hormone-related proteins significantly associated with ferroptosis and immune cells in PTC, and we constructed the risk score model which can be used as an independent prognostic index to predict the prognosis of patients with PTC.

Introduction

Thyroid cancer (TC) is the most common endocrine malignant tumour worldwide and its incidence has been increasing [1]. In 1990, there were an estimated 95,030 incident cases of TC and 22,070 deaths; this increased to 255,490 incident cases and 41,240 deaths in 2017 [2]. In 2020, TC has been found in 9th place for incidence among all cancers, with more than 586,000 cases diagnosed worldwide. In addition, there is a 3-fold higher incidence rate of TC in women, which is about 10.1 per 100,000 cases, representing one in every 20 cancers diagnosed among women [3]. TC shows the highest incidence rates in Northern America, Australia/New Zealand, Eastern Asia, and Southern

Europe. The incidence rate of TC is also increasing in China, and it has become one of the ten major cancers threatening the health of Chinese residents [4]. The rapid increase of thyroid cancer has been largely attributed to the use of progressively sensitive diagnostic imaging modalities [5, 6]. A study also found that obesity positively correlated with 16% of TC cases and 63% of large-size tumors diagnosed from 2013 to 2015 in the USA [7]. Differentiated thyroid cancer, which originates from thyroid follicular epithelial cells, accounts for more than 95% of TC, among which papillary thyroid carcinoma (PTC) constitutes more than 85% of cases [8, 9].

Traditional treatment methods of PTC include radical surgery, endocrine therapy and ^{131}I therapy, and most treatments display relatively good efficacy, e.g., patients who are disease-free after one course of ^{131}I ablation show a range of in 1–3% recurrence rate [10]. However, a small proportion of PTC displays high aggressiveness and even develops into RAI-refractory differentiated thyroid cancer (RAIR-DTC). This causes roughly 15% of the patients to experience recurrence within ten years after the initial treatment [11–13]. Accurate assessment of the prognosis of PTC is essential to ensure that high-risk and advanced patients receive appropriate treatment without over-treating low-risk patients. The American Thyroid Association (ATA) currently recommends the use of TNM staging to predict mortality and proposes a system to assess the risk of recurrence, which includes the size of the main tumour and whether it has grown into nearby areas (T), the extent of spread to nearby lymph nodes (N), and whether the cancer has metastasised to other organs of the body (M) [9, 14]. It has been addressed that tumour microenvironment (TME) possibly plays an important role in response to chemotherapy and antiangiogenic therapy [15, 16]. In the process of tumorigenesis, tumour cells interact with the surrounding microenvironment to promote immune tolerance, which then develops into immune escape that eventually results in tumour development and angiogenesis, invasion, metastasis, and chronic inflammation [10].

TME has been found to be a complex and continuously evolving matrix, containing such cells as stromal cells, fibroblasts, endothelial, innate and adaptive immune cells [17]. The transcriptome of PTC has been well-characterised and the major molecular events associated with most PTC cases, including the aberrant expression of *Brafv600e*, *Ras*, point mutations of *Tert* and RET/PTC rearrangements were elucidated [18–20]. However, reports on the interaction between PTC and stromal tissue, lymphocytic infiltrate, or normal thyrocytes and expression profiles of associated marker genes are limited. A study of 23 PTC patients using oligonucleotide microarrays identified 19 genes differentially expressed with reference to 10 patients with other thyroid disease, such as *Lrp4*, *Eva1*, *Tmprss4*, *Qpct*, and *Slc34a2* [21]. A recent study comparing the gene expression profiles of PTC and normal thyroid in both micro-dissected cells and whole tissue slides revealed crosstalk between cancer cells and TME possibly involving the functions of *Ptcsc*, *Ctgf*, *Tff3*, *Fn1*, *Mpped2* [22].

As a type of programmed cell death, ferroptosis is dependent on iron and induced by the accumulation of oxidatively damaged phospholipids, associated with the malfunction of glutathione-dependent antioxidant defences that are mediated by glutathione peroxidase 4 (*Gpx4*) via different pathways [23–25]. and, resulting in a large amount of ROS, which promotes ferroptosis. The declines in metabolism of lipid peroxides catalysed by *Gpx4* and glutathione (GSH) level intracellularly, lead to Fe^{2+} oxidising lipids in a Fenton-like manner, which enhances ferroptosis due to the elevation of lipid reactive oxygen species (ROS) in cells [24, 26]. In recent years, the induction of ferroptosis has been investigated as an alternative and/or joint therapeutic approach to trigger cancer cell death, with respect to other types of cell death, especially for the treatment of malignancies resistance issues in some cancers; the marker genes and modulator molecules of ferroptosis have been identified [13, 27, 28]. However, the role of ferroptosis in PTC currently remains elusive.

In this study, we integrated the THCA dataset from TCGA and the clinical databases of TCHA from UCSC Xena to identify reliable differentially expressed genes (DEGs) relevant to ferroptosis in PTC. Then, univariate Cox survival

analysis and lasso Cox regression analysis were performed to identify DEGs of PTC, and we proposed a prognostic prediction model using DEGs and clinical data from the TCGA-THCA and UCSC Xena dataset. We performed Multivariate Cox survival analysis on hub genes combined with clinical prognosis data and showed that the model can be used to predict PTC prognosis.

Methods

Data Processing

The papillary thyroid carcinoma (THCA) datasets were obtained from two platforms. THCA database level 3 count was searched in The Cancer Genome Atlas (TCGA) database (<https://portal.gdc.cancer.gov/repository>). The datasets in TCGA were transformed into Transcripts Per Million (TPM) values and the mutation data of The Single Nucleotide Polymorphism Database (SNP) were downloaded for THCA. The clinical data of THCA were obtained using the UCSC Xena browser (<https://xenabrowser.net/>). Only patients with primary tumours who did not receive neoadjuvant therapy were included in this retrospective analysis [13]. Their clinicopathological, genetic, epigenetic, and survival data were downloaded for a secondary analysis. In this study, the datasets of 502 tumour samples and 56 normal samples were used.

Identification of THCA subtypes based on related genes of ferroptosis

A set of 60 related genes of ferroptosis genes were summarised from the MisgDB database (<https://www.gsea-msigdb.org/gsea/msigdb>). The ConsensusClusterPlus package was used for consensus clustering and molecular subtype screening. In brief, k-means clustering was used, with 50 iterations (each using 80% of the samples) [29]. The best cluster number was determined by the clustering score for the cumulative distribution function (CDF) curve, and the relative changes in the area under the CDF curve were evaluated. Principal component analysis (PCA) and t-distributed Stochastic Neighbor Embedding (t-SNE), commonly used for dimensionality reduction, were used to verify the reliability of the consensus clusters.

Heatmap

The ssGSEA score x_i for each THCA sample i was converted to x_i' using the equation:

$$x_i' = (x_i - x_{min}) / (x_{max} - x_{min})$$

while x_{max} and x_{min} represent the maximum and minimum single-sample gene-set enrichment analysis (ssGSEA) scores for all samples in the THCA dataset, respectively. The relationship among subtypes, clinical parameters and ferroptosis genes gene expression were shown by thermography with R-package pheatmap.

Comparison of immune cell subgroups among molecular subtypes of THCA

We continued to use maftools to explore the difference between two subtypes and drew the waterfall of them. In order to investigate the biological difference and understand the different pathways involved by the subtypes, Gene Set Enrichment Analysis (GSEA) was used with the clusterProfiler package. Gene set variation analysis (GSVA) was used to further analyze the difference of pathways between the subtypes using GSVA package [30, 31].

Immune infiltration microenvironment in THCA

The difference between the subtypes may be due to the complexity of tumour microenvironment, so we used the ssGSEA algorithm to analyse the proportion of 28 kinds of immune cells in THCA, and used hclust algorithm to classify the immune cells according to the proportion of immune cells. Finally, a box-plot was constructed to contrast immune cell content between the subtypes. If the normal distribution and homogeneity of variance were met, T-test was used; if not, Wilcoxon test was used. The above operations were completed by using complexheatmap package, ssGSEA package, corrplot package and ggplot2 package.

Gene co-expression network analysis

To investigate characteristic immune gene distribution in each molecular subtype of THCA and identify genes or gene modules highly related to immune cell infiltration, the WGCNA R package was used to evaluate 502 immune-related genes comprising the expression matrix. WGCNA network was constructed and the significant modules were identified. Genes of the significant modules were selected as subtype related hub genes for follow-up analysis.

Data validation of the differentially expressed genes (DEGs)

To further analyse DEGs, DESeq2 algorithm was used for differential analysis of sequencing data (using adjusted P -value < 0.05 and $\log_2FC > 2$ as thresholds). Volcano plots and heatmap were applied to visualize the distribution of the overlapping DEGs between the training and test sets [32]. The above operation was done using the DESeq2 package.

Functional enrichment, pathway analysis of hub genes

In order to identify the possible functions and pathway of hub genes, Gene Ontology (GO) analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed using the clusterProfiler package [33]. The visualisation module of clusterProfiler was used for displaying analysis results. $P < 0.05$ was selected as the cutoff criterion.

Cox regression and survival analysis

To explore the clinical efficacy of hub genes, univariate Cox analysis was performed on hub genes combined with clinical prognosis data, and the results of univariate Cox regression ($P < 0.05$) were included in lasso regression. The lasso regression results were incorporated into multivariate Cox regression. The risk regression model was built according to the multivariate analysis results, and the formula of the model was as follows:

$$\text{risk score} = \sum_{i=1}^n \beta_i \times x_i$$

ROC curve and survival curve were used to evaluate the results of the model. Univariate Cox analysis and multivariate Cox analysis were performed by combining THCA clinical data and risk score values for further evaluation.

Statistical analysis

All the above analyses were completed by R software. Adjusted p -value < 0.05 , p -value < 0.05 and $FC > 2$ were used as statistical thresholds. The statistical methods and algorithms used are described in the corresponding steps.

Results

Identification of THCA subtypes based on immune gene sets

The 502 tumour samples of THCA were divided into k subtypes ($k = 2-9$) using the R package ConsensusClusterPlus. Consensus distributions for each k were displayed, which dissect the optimal k value at which the sample distribution was stable (**Figure 1A**). Based on the consensus score of the CDF curve, the optimal cluster number was determined as $k = 2$, and the relative changes in the area under the CDF curve were evaluated. These were finally divided into two distinct and nonoverlapping subtypes, i.e., cluster 1 and cluster 2. The consensus matrix heat map of these two clusters was shown in **Figure 1B**. A principal component analysis (PCA) plot (**Figure 1C**) and a two dimensional t-SNE analysis (**Figure 1D**) further verified the LUAD cohort and the reliability of the consensus clusters.

Association among subtypes, clinical parameters and ferroptosis signature genes

To explore the association among subtypes, clinical parameters and ferroptosis signature genes, we analysed gene cluster, tumour stage, survival status, pathologic stage, pharmaceutical, radiation, race, gender, age, and ferroptosis signature genes in the TCGA cohort. The relationship among subtypes, clinical parameters and ferroptosis signature genes is derived by thermography (**Figure 2**). On the left side of the heatmap, genes that were positively associated with ferroptosis were coded red (such as *Csf2*, *Hsbp1*, *Pebp1*, *Nfs1*, *Hmgcr*, *Sqle*, *Mt1g*, and *Pgd*), whereas blue colour was used to indicate genes that are negatively associated with ferroptosis (incl. *Tp53*, *Keap1*, *Rpl8*, *Nox1*, *Fth1*, *Hmox1*, *Acsl4*, *Sat1*, *Gpx4*, *Ptgs2*, *Gls2*, *Fancd2*, *Alox5*, *Slc1a5* and *Dpp4*). Gene cluster differed significantly between the two subtypes. In particular, the proportion of cluster 2 was significantly higher than cluster 1 in ferroptosis signature genes that are positively associated, yet the expression of these genes generally appeared higher in cluster 2, such as *Acsf2*, *Mt1g*, *Gclc* and *Aifm2*. Moreover, cluster 1 had remarkably larger proportion of genes that are negatively associated with ferroptosis, for instance, *Tp53*, *Dpp4*, *Slc1a5*, *Alox5* and *Ptgs2*.

In terms of the spread magnitude to nearby lymph nodes (stage_N), N1 was largely relevant to cluster 1, while N0 and NX were more present in cluster 2. The metastasis to distant sites (stage_M) showed an increased proportion of MX in cluster 2, while the extent of the tumour (stage_T) distributed across both cluster 1 and 2, with latter exhibited higher proportion of T2 (**Figure 2**). There were more pathological stage III and IV cases associated with cluster 1, whereas cluster 2 subtype showed more genes associated with stage II, which represents non-invasive tumour with no spread to lymph nodes and no metastasis.

Somatic mutation landscape of the two subtypes were identified by GSEA

We continued to explore the differences between the two subtypes using the maftools package to explore the mutation profile and drawing the waterfall map of them (**Figure 3**). It was found that the somatic mutation types of the two subtypes were missense mutations. The top 2 somatic mutation genes of subtype cluster 1 were *Braf* (88%) and *Tnn* (7%) (**Figure 3A**), with the other 47 genes with low mutation frequency less than 5%. Cluster 2 showed somatic mutations of *Nras*, *Hras* and *Tg* with more than 10% of mutation frequency (**Figure 3B**). This indicates biological differences between the two subtypes.

Gene set enrichment analysis

GSEA and GSVA were used to analyse the enriched pathways in each cluster to investigate the molecular differences in the phenotypes of these two subtypes. GSEA analysis suggested that subtype cluster 1 was mainly involved in bacterial invasion of epithelial cells, cell adhesion molecules, completion and coaggregation cascades and intestinal immune network for IgA production (**Figure 4A**). In subtype cluster 2 phenylalanine, tyrosine and tryptophan biosynthesis, thyroid hormone synthesis, tyrosine metabolism, ubiquinone and other terpenoid-quinone biosynthesis were enriched (**Figure 4B**). GSVA was used to further explore the differences in participating KEGG pathways between the two subtypes, and the results were shown in **Figure 4C** (adjust P -value <0.05 ; see **Supplementary Table S1** for more details). In addition, there were more immune response relevant pathways present in cluster 1 than cluster 2, such as antigen processing and presentation (P -value=1.27E-22), natural killer cell mediated cytotoxicity (P -value=1.36E-37), Fc gamma R-mediated phagocytosis (P -value=1.12E-23), cytokine-cytokine receptor interaction (P -value=7.18E-45), and toll-like receptor signalling pathway (P -value=1.17E-24). Cluster 2 was enriched with more metabolism pathways, for instance, metabolisms of glycine, serine and threonine (P -value=1.04E-30), porphyrin and chlorophyll (P -value=2.93E-33), and fatty acid (P -value=4.05E-41).

Immune characteristics of the two subtypes

The difference between two immune subtypes might be formed because of the complexity of the TME. To explore the biological behaviours between the two clusters, we performed ssGSEA algorithm to analyse the proportion of 28 kinds of immune cells in immune infiltration microenvironment in THCA. The results showed that the degree of immune infiltration in cluster 1 was largely higher than in cluster 2, suggested by the greater activation of immune response relevant cells, among which activated B cells, CD8 and CD4 T cells appeared to be the top three cell types (**Figure 5A**). Spearman correlation analysis was used to calculate the correlation between ferroptosis gene expression and the proportion of immune infiltrating cells (**Figure 5B**). This showed positive correlation of *Alox5*, *Fancd2*, *Dpp4*, *Hmox1*, *Ptgs2*, *Slc1a5*, *Sat1* and *Acs14* with most cell types with relatively high magnitudes. In contrast, the negative correlation of these cells with ferroptosis was indicated by the relative low expression of *Acsf2*, *Pebp1*, *Nfs1*, *Hsbp1*, *Hmgcr*, *Acs13* and *Gss*. The difference of immune cell contents between the two subtypes were compared (**Figure 5C** and **Supplementary Table S2**). Notably, cluster1 was remarkably enriched (P -value <0.0001) with nearly all innate immune cell types, except eosinophil and monocyte, compared to cluster 2.

Identification of DEGs distinct phenotypes derived from gene co-expression network

The distribution of genes expression profiles in the subtypes was investigated to identify genes or gene modules highly related to immune infiltration microenvironment. The WGCNA R-package was used to evaluate the expression matrix data from 502 tumour samples. Analysis of network topology for various soft-thresholding powers indicated that when the power value was equal to 5 ($\beta=5$), the predicted gene co-expression network exhibited scale-free topology by the fit index greater than 0.8, with inherent modular features (**Figure S1A**). The adjacency function was used to generate the adjacency matrix based on the β and gene expression matrix. The hierarchical clustering was built based on the TOM dissimilarity measure (**Figure 6A**). A total of 25 seven co-expression modules were detected. The module preservation statistics was employed to achieve reliable and preserved modules (**Figure S1B**). The co-expression network was examined by the NULMS Stanford dataset, with genes assigned to modules based on the

modules in the reference dataset. There are 16 modules strongly preserved (Z-summary more than 10), such as blue, green, turquoise yellow and red modules; while seven modules are moderately preserved ($5 < Z\text{-summary} < 10$).

The association between module eigengenes, subtypes and immunity were then computed through the Pearson's correlation coefficient, which evaluates the P -value was calculated for any given correlation. The blue module was most significantly correlated with subtypes, and the yellow module had the highest correlation with immune typing (**Figure S1C**). Therefore, the genes of these two modules were selected as subtype related hub genes for subsequent analysis, which included a total of 1,747 hub genes. The DEseq2 algorithm was used to further analyse the hub genes in the sequencing data, and significant differentially expression of 814 hub genes were found, among which 308 were up-regulated and 506 were down-regulated in tumour tissues (**Figure 6B**). The heat map comparing the expression of these 814 genes in \log_2FC values was displayed in **Figure 6C** (See **Supplementary Table S3** for the full gene list).

Gene ontology and KEGG pathway analysis showed functional enrichment in immune regulations

To further analyse the functions and pathways that 814 hub genes are potentially involved in, GO term and KEGG pathway enrichment analyses were performed using the clusterProfiler package. GO analysis showed that 814 hub genes were mainly involved in biological processes such as regulation of small GTPase mediated signal transduction, positive regulation of endocytosis and regulation of plasma lipoprotein particle levels. In terms of cellular components, these genes showed high relevance to membrane (**Figure 7A**). KEGG pathway analysis showed that 814 hub genes were mainly enriched in Salmonella infection and endocytosis pathways (**Figure 7B**). The enrichment analysis of Wikipathway revealed that regulation of apoptosis by parathyroid hormone-related protein, apoptosis modulation and signalling, and pathogenic *Escherichia coli* infection were the top 3 pathways (**Figure 7C**). Besides, several cancer relevant pathways were also enriched in the hub genes (**Supplementary Table S3**).

Establishment and assessment of the predict model

To explore the clinical efficacy of 814 hub genes, Univariate Cox analysis was performed with clinical prognosis data and the results of $P\text{-value} < 0.05$ were recorded in **Supplementary Table S3**. The results of Univariate Cox regression ($P\text{-value} < 0.05$) were included in lasso regression. Dimensionalisation was reduced according to the Lambda curve (**Figure 8A**) and the proportional hazards model curve (**Figure 8B**), which showed that the deviance was the smallest when the number of genes in the model was 19. The lasso regression results were incorporated into multivariate Cox regression, as seven genes were identified by LASSO regression (**Supplementary Table S3**), including *Ac008063.2*, *ApoE*, *Bcl3*, *Acap3*, *Alox5ap*, *Atxn2l* and *B2m*. The risk regression model diagram was displayed in **Figure 8C** according to the multivariate analysis results. In order to evaluate the specificity and sensitivity of the prognostic model, ROC curve and survival curve were used for evaluation. The AUC value of the ROC curve was 0.748 (**Figure 8D**) and the survival curve showed a significant difference between the high-risk and low-risk groups ($P\text{-value} = 6.329E-4$), indicating a good result of the model (**Figure 8E**).

The risk score predicted prognosis

To further evaluate the clinical efficacy of the model, univariate Cox analysis (**Figure 9A**) and multivariate Cox analysis (**Figure 9B**) were performed by combining THCA clinical data and risk score values, which implied that the

risk score value of the model could be used as an independent prognostic indicator to predict the prognosis of patients with thyroid papillary carcinoma (see **Supplementary Table S4** for full statistical results).

Discussion

The incidence of PTC is a heavy disease burden around the world, and has increased quite significantly over the past two decades [34–36]. Patients with recurrent PTC who undergo surgical treatment have a high risk of complications, including tracheal collapse, laryngeal edema and recurrent laryngeal nerve injury [9, 34, 37]. Furthermore, PTC is a highly heterogeneous disease and tumour progression involves a complex network comprised of multiple signalling pathways [38, 39]. Nomogram is widely used to evaluate clinical prognosis in oncology because it can integrate a variety of prognostic determinants, including molecular biological and clinicopathological parameters [8, 38, 40]. The combination of our prognostic gene signature with conventional clinical parameters may provide a better tool to predict PTC prognosis.

The discovery of distinct lethal subroutines of RCD has facilitated great progress in cancer treatment [41–43]. Ferroptosis was identified as a new form of RCD distinct from apoptosis, necrosis, autophagy, and other forms of cell death. It is an iron-dependent cell death process and caused by lipid peroxidation relying on reactive oxygen species (ROS) generation [24, 44, 45]. There has been growing interest in exploring the mechanisms underpinning ferroptosis in recent years, and several seminal discoveries have elucidated the process [42, 46, 47]. Ferroptosis is thought to only be inhibited by the phospholipid hydroperoxide-reducing enzyme glutathione peroxidase 4 and radical-trapping antioxidants [26, 42]. GSH-GPX4 and FSP1-CoQ10 are the two main ferroptosis resistance pathways [48, 49].

Growing evidence suggests that immune cells in TME play vital roles in tumorigenesis. These innate immune cells include macrophages, neutrophils, dendritic cells, innate lymphoid cells, myeloid-derived suppressor cells, and natural killer cells which potentially possess tumour-antagonising or tumour-promoting functions [50–52]. Cancer cells exhibited iron ion aggregation in the TME for active proliferation, thus from the perspective of iron homeostasis, regulating ferroptosis could effectively kill tumour cells. In addition, ferroptosis also plays an important immunological role, participating in the process of tumour surveillance by immune cells and tumour immunity [50, 53, 54]. The metabolism of immune cells affects their differentiation and function [55]. Given the complex interaction of environmental factors in TME as a profound impact on the metabolic activities of immunity, matrix and tumour cell types and ferroptosis [46, 56]. Thus, exploring immunophenotype in TME could enhance the significance of ferroptosis in cancer treatment. The cross study of ferroptosis and immune cells might lead to a discussion of clinical application.

In this study, we seek to investigate whether there was ferroptosis regulation via exploring the crucial genes and pathways of ferroptosis in PTC. A total of 502 PTC samples and 56 normal samples integrated from TCGA and UCSC Xena datasets were analysed, and 60 ferroptosis related genes were derived from MisgDB datasets. Two distinct subtypes (i.e. cluster 1, cluster 2) were identified in our study. The proportion of cluster 1 was significantly higher than cluster 2 in the signature genes that either positively or negatively associated with ferroptosis. The main mutation types of these two subtypes were missense mutation. The missense mutations of *Braf* have been detected in circulating DNA in the serum of some patients with PTC [57]. A close association of *Braf* mutations with extrathyroidal extension, lymph node metastasis, and stage III/IV of PTC has been strongly suggested [58]. The significance of missense mutations of *Nras* and *Hras* in PTC has also been identified [59–61].

The distinct enriched KEGG pathways revealed GSEA analysis of cluster 1 and cluster 2 implied biological differences between the two genotypes. The GSVA analysis showed that there were more immune response relevant pathways

enriched in cluster 1 (adjust *P*-value < 0.05, **Supplementary Table S1**). The difference between the two subtypes possibly resulted from the complexity of TME, since the degree of immune infiltration in cluster 1 was remarkably higher than that in cluster 2, which was reflected by the contents of many immune cell types. Subsequent analysis of TME cell infiltration indicated THCA cluster1 was remarkably rich in innate immune cell infiltration including T cells, natural killer cells, macrophage, eosinophil, mast cell, MDSC, plasmacytoid dendritic cell.

We subsequently analysed the related hub genes of the two modules with the highest correlation between molecular typing and immune typing in WGCNA network, and there were 814 DEGs including 308 up-regulated genes and 506 down-regulated genes in PTC tissues. Then, lasso regression was evaluated and incorporated into multivariate Cox regression, the crucial genes (i.e., *Ac008063.2*, *ApoE*, *Bcl3*, *Acap3*, *Alox5ap*, *Atxn2l* and *B2m*) were identified and the risk regression model was constructed according to the results of multivariate analysis. ROC curve and survival curve indicated a good result of the model. Based on these findings, the risk score of the model could be used as an independent prognostic indicator to predict the prognosis of patients with PTC.

All of these genes, with the exception of *Ac008063.2*, have been implicated in cancer cell proliferation. With the exception of *ApoE*, which is highly expressed in liver cancer, all of the genes have low cancer specificity, according to Human Protein Atlas (*Ac008063.2* could not be identified). *B2m* has been identified as a potential biomarker for thyroid cancer, kidney dysfunction and renal disease [62]. Higher rates of *B2m* mutation are correlated with lower patient survival rates, which is speculated to be due to increased immune evasion [63, 64]. *Atxn2l* is a stress response molecule which has been identified as a potential cancer prognosis gene for cancer, such as adrenocortical carcinoma [65, 66]. *Alox5ap* has been identified as specifically upregulated in leukemia stem cells and unchanged in hematopoietic stem cells, indicating specificity in cancer cell self-renewal and differentiation [67, 68]. There is relatively little oncological literature on *Acap3*, however it has been identified as highly down-regulated in late-stage liver cancer patients [69]. *Bcl3* up-regulation has been associated with poorer prognostic outcomes [70, 71]. It is classified as an oncogene and has been implicated in breast cancer cell migration [72, 73]. *ApoE* is associated with tumorigenesis in many cancers, including lung, gastric and thyroid cancer, and a higher risk of metastasis [74–76].

To our knowledge, a prognostic model based on ferroptosis genes and immune cells in tumour immune infiltration microenvironment has not been reported up to now. Compared with whole genome sequencing, our prediction model based on a limited number of gene expression levels might be more economical and practical. Further, the risk score of the model could predict the prognosis of patients with PTC. However, the external validation of the seven-gene signature and prognostic nomogram is needed in more independent cohorts. In addition, to clarify the molecular mechanism underlying the seven genes in relation to PTC, the spatial expression pattern and quantity of these genes at protein level warrants further investigation.

Conclusions

We demonstrated that ferroptosis was associated with immune cell infiltration in TME of patients with PTC, based on which a seven-gene signature and a prognostic model to predict the prognosis was established. The seven genes appeared closely related to the progression and prognosis of PTC and thus could be potential therapeutic targets. The predicted model proved to be reliable in predicting the prognosis of patients with PTC and might thus be beneficial for individualised treatment and medical decision making.

Abbreviations

PTC: Papillary thyroid carcinoma

RAI: Radioactive Iodine

THCA: thyroid carcinoma

TCGA: The Cancer Genome Atlas

GSEA: Gene set enrichment analysis

GSVA: Gene set variation analysis

GO: Gene ontology

KEGG: Kyoto Encyclopedia of Genes and Genomes

WGCNA: Weighted correlation network analysis

ROC: Receiver operating characteristic

AUC: Area under the ROC Curve

TME: Tumour microenvironment

TNM: Tumour-node-metastasis

FC: Fold-change

Declarations

Ethics approval and consent to participate

No specific permits were required, all data were downloaded from online data repository.

Consent for publication

Not applicable.

Availability of data and materials

The datasets generated and/or analysed during the current study are available in The Cancer Genome Atlas database (<https://portal.gdc.cancer.gov/projects/TCGA-THCA>, accession number phs000178) and the UCSC Xena ([https://xenabrowser.net/datapages/?cohort=TCGA%20Thyroid%20Cancer%20\(THCA\)&removeHub=https%3A%2F%2Fxena.treehouse.gi.ucsc.edu%3A443](https://xenabrowser.net/datapages/?cohort=TCGA%20Thyroid%20Cancer%20(THCA)&removeHub=https%3A%2F%2Fxena.treehouse.gi.ucsc.edu%3A443)).

Competing Interests

The authors declare that have no competing interests.

Funding

Foshan First People's Hospital Deng Feng project 2019A008, Foshan municipal Government (2015AG1003), Guangdong Provincial Government (2016A020213001) of China, MCRs project of Genecology Research Centre of University of the Sunshine Coast, and National Natural Science Foundation of China (31971355).

Authors' contributions

RL conceived and designed the study. RL and CEF analysed the data and drafted the manuscript. BM and TW performed the data analysis and generated the figures. HL and GN drafted and revised the manuscript. XL, JY and TW conceptualised and developed an outline for the manuscript and revised the manuscript. All authors read and approved the final manuscript.

Acknowledgments

Not applicable.

References

1. Siegel RL, Miller KD, Jemal A: **Cancer statistics, 2020**. *CA Cancer J Clin* 2020, **70**(1):7–30.
2. Deng Y, Li H, Wang M, Li N, Tian T, Wu Y, Xu P, Yang S, Zhai Z, Zhou L *et al*: **Global Burden of Thyroid Cancer From 1990 to 2017**. *JAMA network open* 2020, **3**(6):e208759.
3. Ferlay J, Ervik M, Lam F, Colombet M, Mery L, Piñeros M, Znaor A, Soerjomataram I, Bray F: **Global cancer observatory: cancer today**. International Agency for Research on Cancer, Lyon. In.; 2020.
4. Du L, Zhao Z, Zheng R, Li H, Zhang S, Li R, Wei W, He J: **Epidemiology of Thyroid Cancer: Incidence and Mortality in China, 2015**. *Front Oncol* 2020, **10**:1702.
5. Udelsman R, Zhang Y: **The epidemic of thyroid cancer in the United States: the role of endocrinologists and ultrasounds**. *Thyroid: official journal of the American Thyroid Association* 2014, **24**(3):472–479.
6. Brito JP, Morris JC, Montori VM: **Thyroid cancer: zealous imaging has increased detection and treatment of low risk tumours**. *BMJ (Clinical research ed)* 2013, **347**:f4706.
7. Kitahara CM, Pfeiffer RM, Sosa JA, Shiels MS: **Impact of Overweight and Obesity on US Papillary Thyroid Cancer Incidence Trends (1995–2015)**. *Journal of the National Cancer Institute* 2020, **112**(8):810–817.
8. Wu M, Yuan H, Li X, Liao Q, Liu Z: **Identification of a Five-Gene Signature and Establishment of a Prognostic Nomogram to Predict Progression-Free Interval of Papillary Thyroid Carcinoma**. *Front Endocrinol (Lausanne)* 2019, **10**:790.
9. Carling T, Udelsman R: **Thyroid cancer**. *Annu Rev Med* 2014, **65**:125–137.
10. Verburg FA, Van Santen HM, Luster M: **Pediatric papillary thyroid cancer: current management challenges**. *OncoTargets and therapy* 2017, **10**:165–175.
11. Rivera M, Ghossein RA, Schoder H, Gomez D, Larson SM, Tuttle RM: **Histopathologic characterization of radioactive iodine-refractory fluorodeoxyglucose-positron emission tomography-positive thyroid carcinoma**. *Cancer* 2008, **113**(1):48–56.
12. Riesco-Eizaguirre G, Galofre JC, Grande E, Zafon Llopis C, Ramon y Cajal Asensio T, Navarro Gonzalez E, Jimenez-Fonseca P, Santamaria Sandi J, Gomez Saez JM, Capdevila J: **Spanish consensus for the management of patients with advanced radioactive iodine refractory differentiated thyroid cancer**. *Endocrinol Nutr* 2016, **63**(4):e17-24.
13. Luo J, Zhang B, Cui L, Liu T, Gu Y: **FM01 gene expression independently predicts favorable recurrence-free survival of classical papillary thyroid cancer**. *Future Oncol* 2019, **15**(12):1303–1311.
14. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, Pacini F, Randolph GW, Sawka AM, Schlumberger M *et al*: **2015 American Thyroid Association Management Guidelines for Adult Patients with**

Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid* 2016, **26**(1):1–133.

15. Roma-Rodrigues C, Mendes R, Baptista PV, Fernandes AR: **Targeting Tumor Microenvironment for Cancer Therapy.** *International journal of molecular sciences* 2019, **20**(4).
16. De Palma M, Biziato D, Petrova TV: **Microenvironmental regulation of tumour angiogenesis.** *Nat Rev Cancer* 2017, **17**(8):457–474.
17. Denton AE, Roberts EW, Fearon DT: **Stromal Cells in the Tumor Microenvironment.** *Adv Exp Med Biol* 2018, **1060**:99–114.
18. **Integrated genomic characterization of papillary thyroid carcinoma.** *Cell* 2014, **159**(3):676–690.
19. Huang Y, Prasad M, Lemon WJ, Hampel H, Wright FA, Kornacker K, LiVolsi V, Frankel W, Kloos RT, Eng C *et al*: **Gene expression in papillary thyroid carcinoma reveals highly consistent profiles.** *Proceedings of the National Academy of Sciences of the United States of America* 2001, **98**(26):15044–15049.
20. Han J, Chen M, Wang Y, Gong B, Zhuang T, Liang L, Qiao H: **Identification of Biomarkers Based on Differentially Expressed Genes in Papillary Thyroid Carcinoma.** *Scientific reports* 2018, **8**(1):9912.
21. Jarzab B, Wiench M, Fujarewicz K, Simek K, Jarzab M, Oczko-Wojciechowska M, Wloch J, Czarniecka A, Chmielik E, Lange D *et al*: **Gene expression profile of papillary thyroid cancer: sources of variability and diagnostic implications.** *Cancer research* 2005, **65**(4):1587–1597.
22. Oczko-Wojciechowska M, Pfeifer A, Jarzab M, Swierniak M, Rusinek D, Tyszkiewicz T, Kowalska M, Chmielik E, Zembala-Nozynska E, Czarniecka A *et al*: **Impact of the Tumor Microenvironment on the Gene Expression Profile in Papillary Thyroid Cancer.** *Pathobiology: journal of immunopathology, molecular and cellular biology* 2020, **87**(2):143–154.
23. Stockwell BR, Friedmann Angeli JP, Bayir H, Bush AI, Conrad M, Dixon SJ, Fulda S, Gascon S, Hatzios SK, Kagan VE *et al*: **Ferroptosis: A Regulated Cell Death Nexus Linking Metabolism, Redox Biology, and Disease.** *Cell* 2017, **171**(2):273–285.
24. Dixon SJ, Lemberg KM, Lamprecht MR, Skouta R, Zaitsev EM, Gleason CE, Patel DN, Bauer AJ, Cantley AM, Yang WS *et al*: **Ferroptosis: an iron-dependent form of nonapoptotic cell death.** *Cell* 2012, **149**(5):1060–1072.
25. Galluzzi L, Vitale I, Aaronson SA, Abrams JM, Adam D, Agostinis P, Alnemri ES, Altucci L, Amelio I, Andrews DW *et al*: **Molecular mechanisms of cell death: recommendations of the Nomenclature Committee on Cell Death 2018.** *Cell Death Differ* 2018, **25**(3):486–541.
26. Li J, Cao F, Yin HL, Huang ZJ, Lin ZT, Mao N, Sun B, Wang G: **Ferroptosis: past, present and future.** *Cell Death Dis* 2020, **11**(2):88.
27. Elfarra AA, Krause RJ: **Potential roles of flavin-containing monooxygenases in sulfoxidation reactions of l-methionine, N-acetyl-l-methionine and peptides containing l-methionine.** *Biochim Biophys Acta* 2005, **1703**(2):183–189.
28. Krueger SK, Williams DE: **Mammalian flavin-containing monooxygenases: structure/function, genetic polymorphisms and role in drug metabolism.** *Pharmacol Ther* 2005, **106**(3):357–387.
29. Tang C, Ma J, Liu X, Liu Z: **Identification of Four Immune Subtypes in Bladder Cancer Based on Immune Gene Sets.** *Frontiers in Oncology* 2020, **10**.
30. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES *et al*: **Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.** *Proc Natl Acad Sci U S A* 2005, **102**(43):15545–15550.

31. Lei Z, Yu S, Ding Y, Liang J, Halifu Y, Xiang F, Zhang D, Wang H, Hu W, Li T *et al*: **Identification of key genes and pathways involved in vitiligo development based on integrated analysis.** *Medicine (Baltimore)* 2020, **99**(31):e21297.
32. Kang K, Xie F, Mao J, Bai Y, Wang X: **Significance of Tumor Mutation Burden in Immune Infiltration and Prognosis in Cutaneous Melanoma.** *Front Oncol* 2020, **10**:573141.
33. Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B: **WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs.** *Nucleic Acids Res* 2019, **47**(W1):W199-W205.
34. Cabanillas ME, McFadden DG, Durante C: **Thyroid cancer.** *The Lancet* 2016, **388**(10061):2783–2795.
35. Bergdorf K, Ferguson DC, Mehrad M, Ely K, Stricker T, Weiss VL: **Papillary thyroid carcinoma behavior: clues in the tumor microenvironment.** *Endocr Relat Cancer* 2019, **26**(6):601–614.
36. Fisher SB, Perrier ND: **The incidental thyroid nodule.** *CA Cancer J Clin* 2018, **68**(2):97–105.
37. Antonelli A, La Motta C: **Novel therapeutic clues in thyroid carcinomas: The role of targeting cancer stem cells.** *Med Res Rev* 2017, **37**(6):1299–1317.
38. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP: **Nomograms in oncology: more than meets the eye.** *The Lancet Oncology* 2015, **16**(4):e173-e180.
39. Zeiger MA, Schneider EB: **BRAF V600E mutation independently predicts central compartment lymph node metastasis in patients with papillary thyroid cancer.** *Ann Surg Oncol* 2013, **20**(1):3–4.
40. Dong S, Huo H, Mao Y, Li X, Dong L: **A risk score model for the prediction of osteosarcoma metastasis.** *FEBS Open Bio* 2019, **9**(3):519–526.
41. Hassannia B, Vandenabeele P, Vanden Berghe T: **Targeting Ferroptosis to Iron Out Cancer.** *Cancer Cell* 2019, **35**(6):830–849.
42. Doll S, Freitas FP, Shah R, Aldrovandi M, da Silva MC, Ingold I, Goya Grocin A, Xavier da Silva TN, Panzilius E, Scheel CH *et al*: **FSP1 is a glutathione-independent ferroptosis suppressor.** *Nature* 2019, **575**(7784):693–698.
43. Chen X, Kang R, Kroemer G, Tang D: **Broadening horizons: the role of ferroptosis in cancer.** *Nat Rev Clin Oncol* 2021.
44. Yan HF, Zou T, Tuo QZ, Xu S, Li H, Belaidi AA, Lei P: **Ferroptosis: mechanisms and links with diseases.** *Signal Transduct Target Ther* 2021, **6**(1):49.
45. Tang D, Chen X, Kang R, Kroemer G: **Ferroptosis: molecular mechanisms and health implications.** *Cell Res* 2021, **31**(2):107–125.
46. Angeli JPF, Shah R, Pratt DA, Conrad M: **Ferroptosis Inhibition: Mechanisms and Opportunities.** *Trends Pharmacol Sci* 2017, **38**(5):489–498.
47. Bersuker K, Hendricks JM, Li Z, Magtanong L, Ford B, Tang PH, Roberts MA, Tong B, Maimone TJ, Zoncu R *et al*: **The CoQ oxidoreductase FSP1 acts parallel to GPX4 to inhibit ferroptosis.** *Nature* 2019, **575**(7784):688–692.
48. Ubellacker JM, Tasdogan A, Ramesh V, Shen B, Mitchell EC, Martin-Sandoval MS, Gu Z, McCormick ML, Durham AB, Spitz DR *et al*: **Lymph protects metastasizing melanoma cells from ferroptosis.** *Nature* 2020, **585**(7823):113–118.
49. Maher P, Currais A, Schubert D: **Using the Oxytosis/Ferroptosis Pathway to Understand and Treat Age-Associated Neurodegenerative Diseases.** *Cell Chem Biol* 2020, **27**(12):1456–1471.
50. Lei X, Lei Y, Li JK, Du WX, Li RG, Yang J, Li J, Li F, Tan HB: **Immune cells within the tumor microenvironment: Biological functions and roles in cancer immunotherapy.** *Cancer Lett* 2020, **470**:126–133.

51. Wu F, Li F, Lin X, Xu F, Cui RR, Zhong JY, Zhu T, Shan SK, Liao XB, Yuan LQ *et al*: **Exosomes increased angiogenesis in papillary thyroid cancer microenvironment.** *Endocr Relat Cancer* 2019, **26**(5):525–538.
52. Tang C, Ma J, Liu X, Liu Z: **Identification of Four Immune Subtypes in Bladder Cancer Based on Immune Gene Sets.** *Front Oncol* 2020, **10**:544610.
53. La Rosa P, Petrillo S, Turchi R, Berardinelli F, Schirinzi T, Vasco G, Lettieri-Barbato D, Fiorenza MT, Bertini ES, Aquilano K *et al*: **The Nrf2 induction prevents ferroptosis in Friedreich's Ataxia.** *Redox Biol* 2021, **38**:101791.
54. Qiu Y, Cao Y, Cao W, Jia Y, Lu N: **The Application of Ferroptosis in Diseases.** *Pharmacol Res* 2020, **159**:104919.
55. Dong C: **Cytokine Regulation and Function in T Cells.** *Annu Rev Immunol* 2021.
56. Kaymak I, Williams KS, Cantor JR, Jones RG: **Immunometabolic Interplay in the Tumor Microenvironment.** *Cancer Cell* 2021, **39**(1):28–37.
57. Chuang TC, Chuang AY, Poeta L, Koch WM, Califano JA, Tufano RP: **Detectable BRAF mutation in serum DNA samples from patients with papillary thyroid carcinomas.** *Head & neck* 2010, **32**(2):229–234.
58. Xing M: **BRAF mutation in papillary thyroid cancer: pathogenic role, molecular bases, and clinical implications.** *Endocrine reviews* 2007, **28**(7):742–762.
59. Huang M, Yan C, Xiao J, Wang T, Ling R: **Relevance and clinicopathologic relationship of BRAF V600E, TERT and NRAS mutations for papillary thyroid carcinoma patients in Northwest China.** *Diagnostic pathology* 2019, **14**(1):74.
60. Fakhrudin N, Jabbour M, Novy M, Tamim H, Bahmad H, Farhat F, Zaatari G, Aridi T, Kriegshauser G, Oberkanins C *et al*: **BRAF and NRAS Mutations in Papillary Thyroid Carcinoma and Concordance in BRAF Mutations Between Primary and Corresponding Lymph Node Metastases.** *Scientific reports* 2017, **7**(1):4666.
61. Schulten HJ, Salama S, Al-Ahmadi A, Al-Mansouri Z, Mirza Z, Al-Ghamdi K, Al-Hamour OA, Huwait E, Gari M, Al-Qahtani MH *et al*: **Comprehensive survey of HRAS, KRAS, and NRAS mutations in proliferative thyroid lesions from an ethnically diverse population.** *Anticancer research* 2013, **33**(11):4779–4784.
62. Jasim A, Mohammed A, Ibrahim A: **Beta-2-Microglobulin as a Marker in Patients with**
63. **Thyroid Cancer.** *THE IRAQI POSTGRADUATE MEDICAL JOURNAL* 2019, **18**(1):6.
64. Middha S, Yaeger R, Shia J, Stadler ZK, King S, Guercio S, Paroder V, Bates DDB, Rana S, Diaz LA, Jr. *et al*: **Majority of B2M-Mutant and -Deficient Colorectal Carcinomas Achieve Clinical Benefit From Immune Checkpoint Inhibitor Therapy and Are Microsatellite Instability-High.** *JCO precision oncology* 2019, **3**.
65. Castro A, Ozturk K, Pyke RM, Xian S, Zanetti M, Carter H: **Elevated neoantigen levels in tumors with somatic mutations in the HLA-A, HLA-B, HLA-C and B2M genes.** *BMC medical genomics* 2019, **12**(Suppl 6):107.
66. Liang W, Sun F: **Prognostic Alternative mRNA Splicing in Adrenocortical Carcinoma.** 2021, **12**(177).
67. Lin L, Li X, Pan C, Lin W, Shao R, Liu Y, Zhang J, Luo Y, Qian K, Shi M *et al*: **ATXN2L upregulated by epidermal growth factor promotes gastric cancer cell invasiveness and oxaliplatin resistance.** *Cell death & disease* 2019, **10**(3):173.
68. Chen Y, Hu Y, Zhang H, Peng C, Li S: **Loss of the Alox5 gene impairs leukemia stem cells and prevents chronic myeloid leukemia.** *Nature genetics* 2009, **41**(7):783–792.
69. Chen Y, Li D, Li S: **The Alox5 gene is a novel therapeutic target in cancer stem cells of chronic myeloid leukemia.** *Cell cycle (Georgetown, Tex)* 2009, **8**(21):3488–3492.
70. Qi S, Su L, Li J, Zhang C, Ma Z, Liu G, Zhang Q, Jia G, Piao Y, Zhang S: **Arf6-driven endocytic recycling of CD147 determines HCC malignant phenotypes.** *Journal of experimental & clinical cancer research: CR* 2019, **38**(1):471.

71. Brenne AT, Fagerli UM, Shaughnessy JD, Jr., Våtsveen TK, Rø TB, Hella H, Zhan F, Barlogie B, Sundan A, Børset M *et al*: **High expression of BCL3 in human myeloma cells is associated with increased proliferation and inferior prognosis.** *European journal of haematology* 2009, **82**(5):354–363.
72. Legge DN, Chambers AC, Parker CT, Timms P, Collard TJ, Williams AC: **The role of B-Cell Lymphoma-3 (BCL-3) in enabling the hallmarks of cancer: implications for the treatment of colorectal carcinogenesis.** *Carcinogenesis* 2020, **41**(3):249–256.
73. Wakefield A, Soukupova J, Montagne A, Ranger J, French R, Muller WJ, Clarkson RW: **Bcl3 selectively promotes metastasis of ERBB2-driven mammary tumors.** *Cancer research* 2013, **73**(2):745–755.
74. Maldonado V, Melendez-Zajgla J: **Role of Bcl-3 in solid tumors.** *Molecular cancer* 2011, **10**:152.
75. Ostendorf BN, Bilanovic J, Adaku N, Tafreshian KN, Tavora B, Vaughan RD, Tavazoie SF: **Common germline variants of the human APOE gene modulate melanoma progression and survival.** *Nature medicine* 2020, **26**(7):1048–1053.
76. Chang NW, Chen DR, Wu CT, Aouizerat BE, Chen FN, Hung SJ, Wang SH, Wei MF, Chang CS: **Influences of apolipoprotein E polymorphism on the risk for breast cancer and HER2/neu status in Taiwan.** *Breast cancer research and treatment* 2005, **90**(3):257–261.
77. Zhao Z, Zou S, Guan X, Wang M, Jiang Z, Liu Z, Li C, Lin H, Liu X, Yang R *et al*: **Apolipoprotein E Overexpression Is Associated With Tumor Progression and Poor Survival in Colorectal Cancer.** *Frontiers in genetics* 2018, **9**:650.

Figures

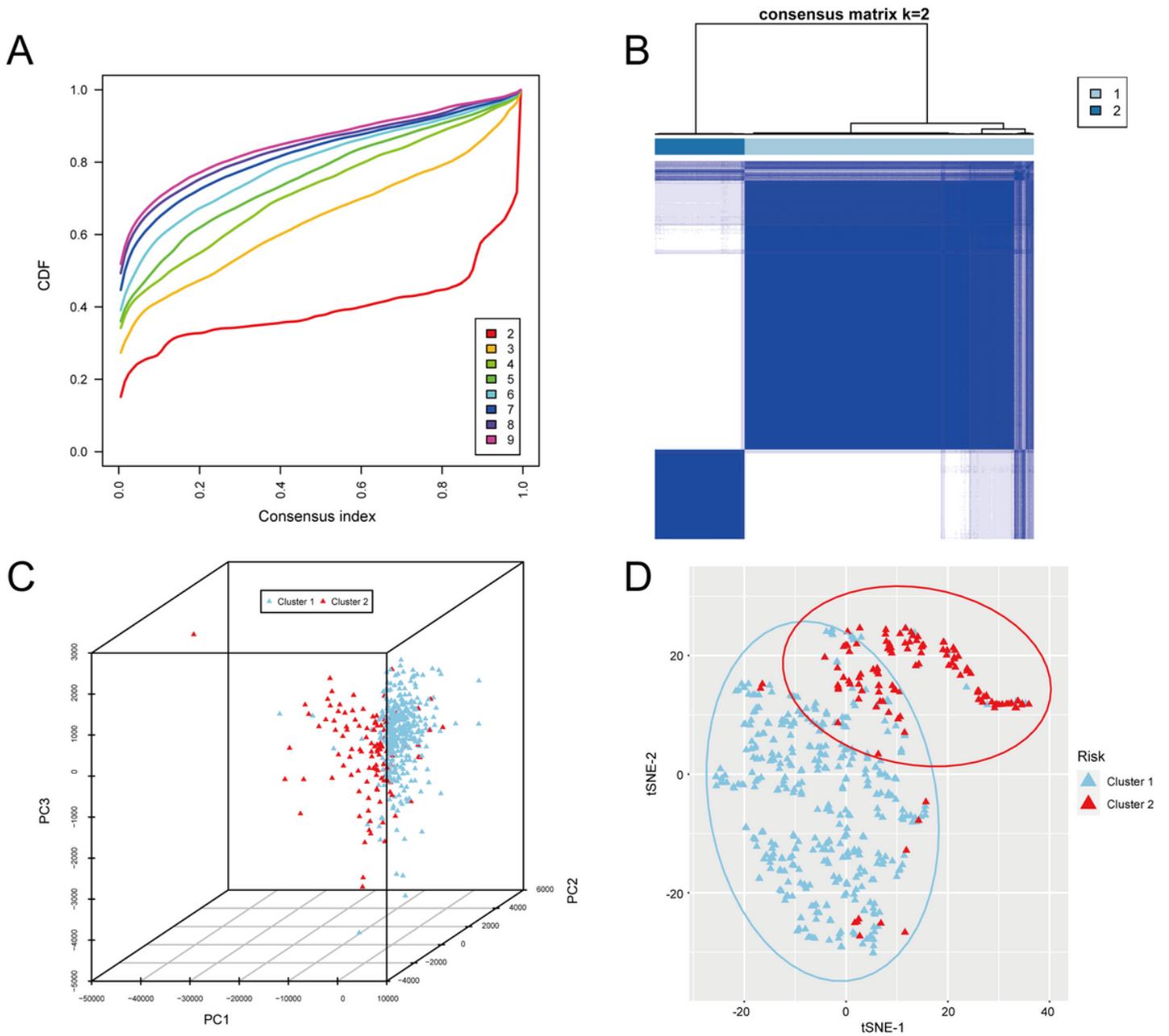


Figure 1

Consensus clustering of THCA TCGA cohorts. (A) CDF curve of the consistency score for different subtype numbers ($k = 2-9$). (B) The consensus score matrix for THCA samples when $k = 2$. A higher consensus score between two samples indicates that they are more likely to be assigned to the same cluster in different iterations. (C) Principal component analysis (PCA) plot of the TCGA cohort. (D) 2D t-SNE analysis of the TCGA cohort.

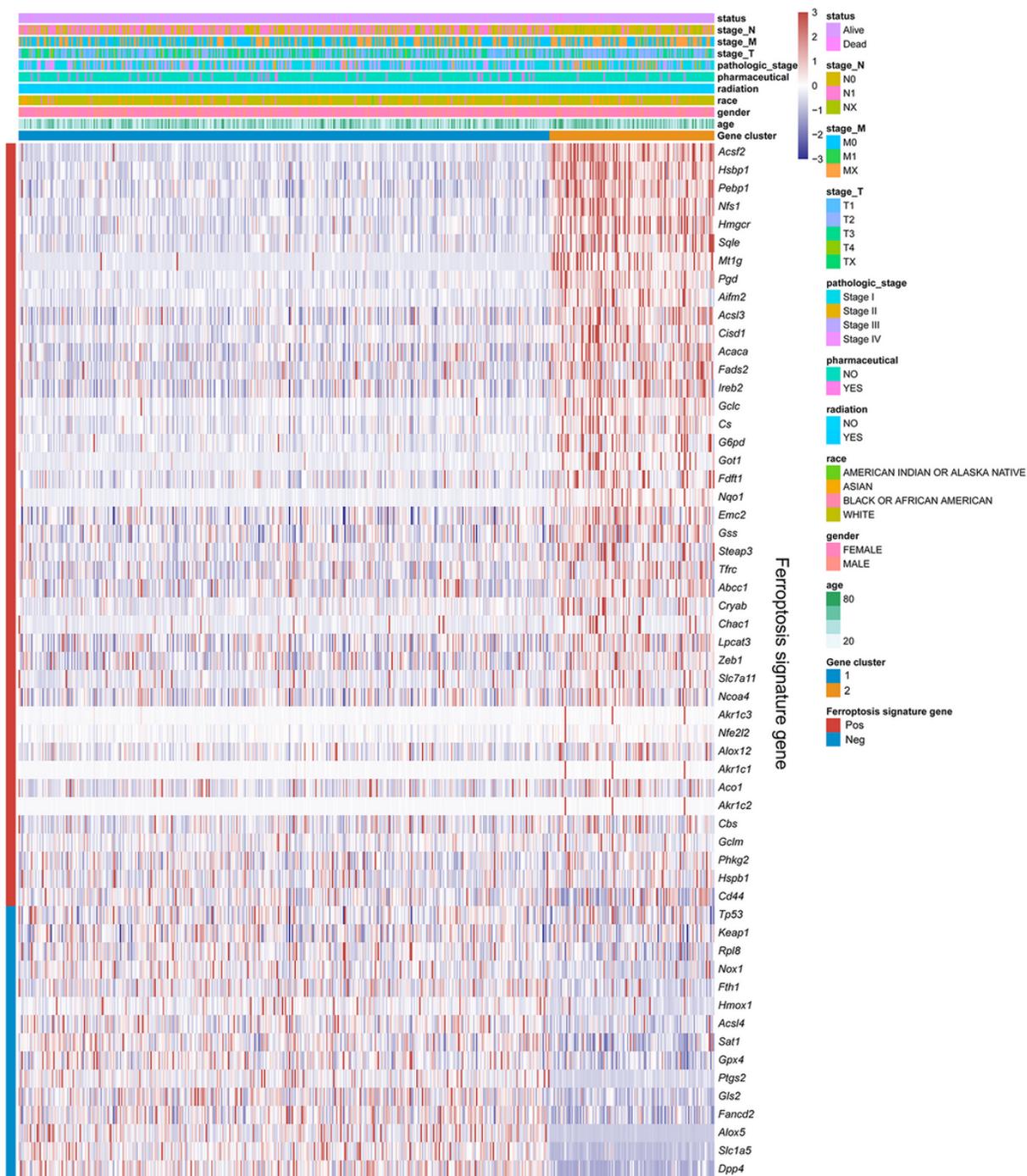


Figure 2

Unsupervised clustering of ferroptosis genes expression in the THCA cohort. The gene cluster, tumour stage, survival status, pathologic stage, pharmaceutical, radiation, race, gender, age, and ferroptosis signature genes were used as patient annotations. Red represented high expression of regulators and blue represented low expression.

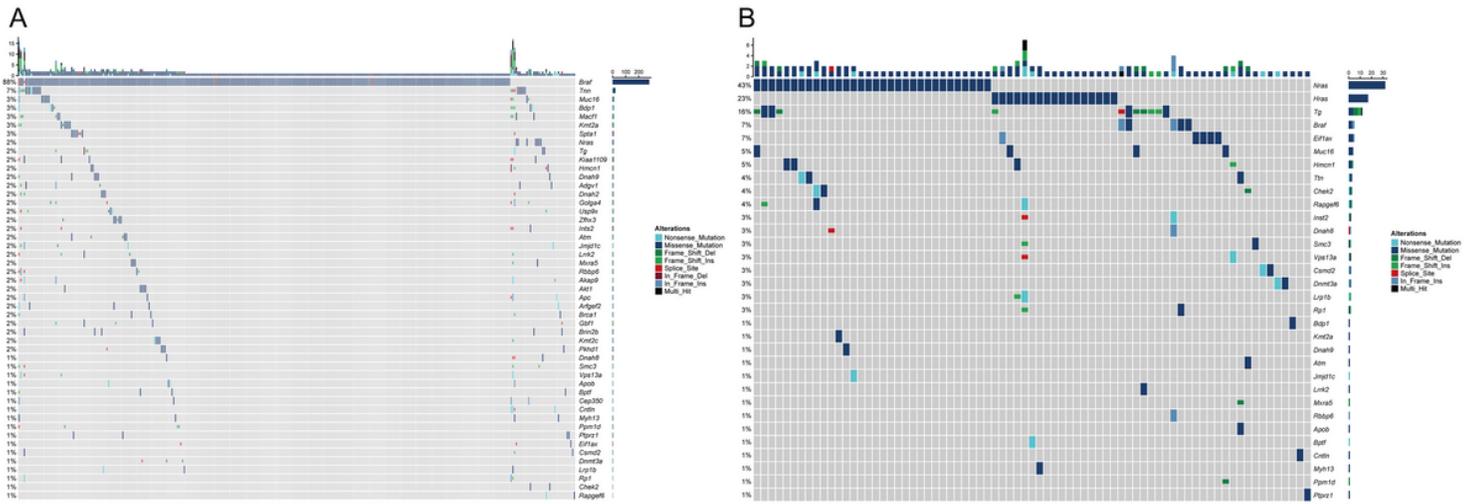


Figure 3

Somatic mutation landscapes of subtypes cluster 1 (A) and 2 (B) identified by GSEA in TCGA cohort. Landscape of mutation profiles in bladder cancer (BLCA) samples. A total of 34 genes with more than 10% of mutation frequency were chosen in the waterfall plot. Various colours of the waterfall plot with annotations at the bottom represent different mutation types. The barplot above the legend shows the mutation number.

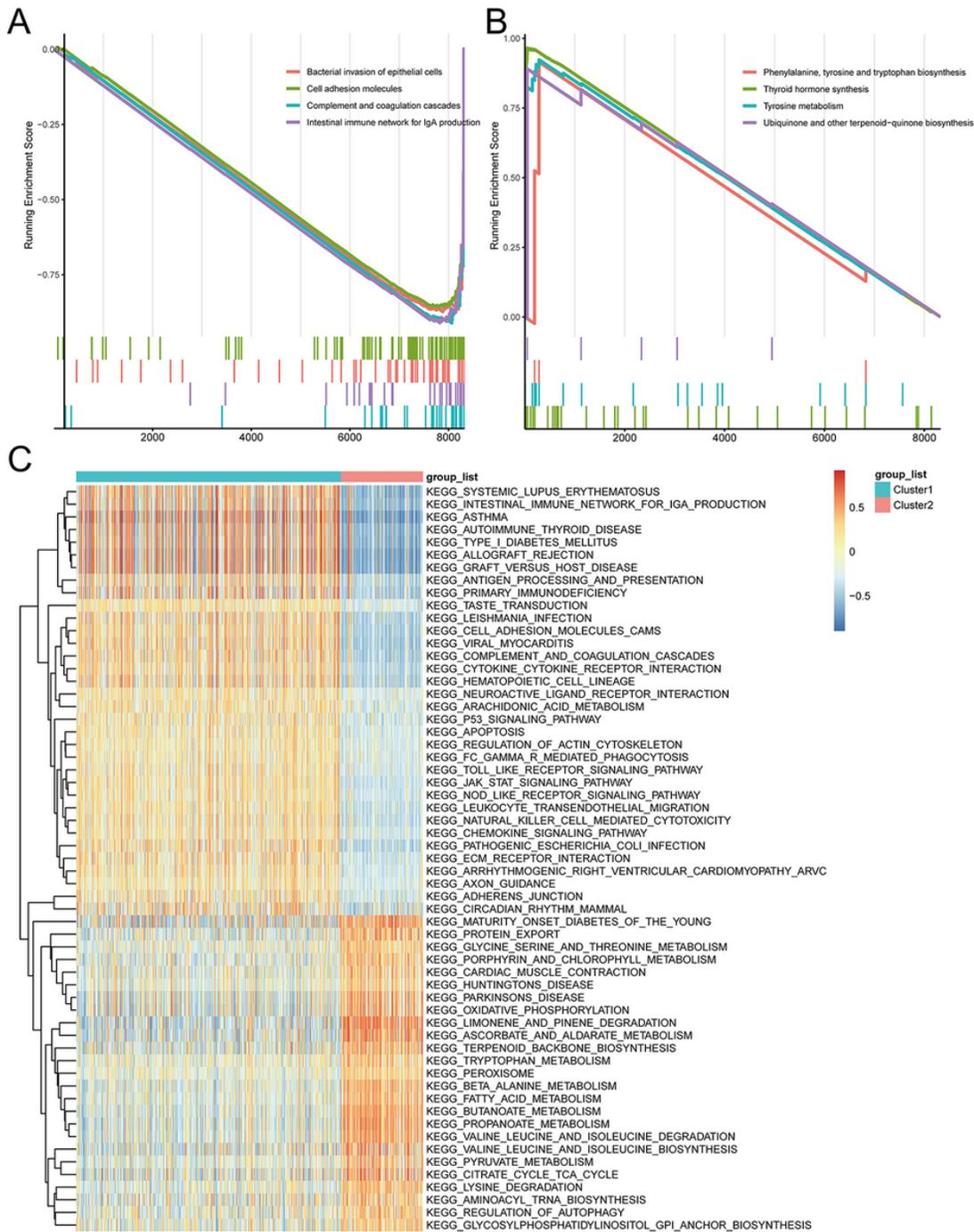


Figure 4

(A) The enriched KEGG pathways associated with DEGs of cluster 1 predicted by GSEA analysis. (B) The enriched KEGG pathways associated with DEGs of cluster 2 predicted by GSEA analysis. (C) Thermogram shows the activation state of KEGG pathways in different clusters after processing by GSVA. The red node represents up-regulation and the blue node represents down-regulation, P-value<0.05.

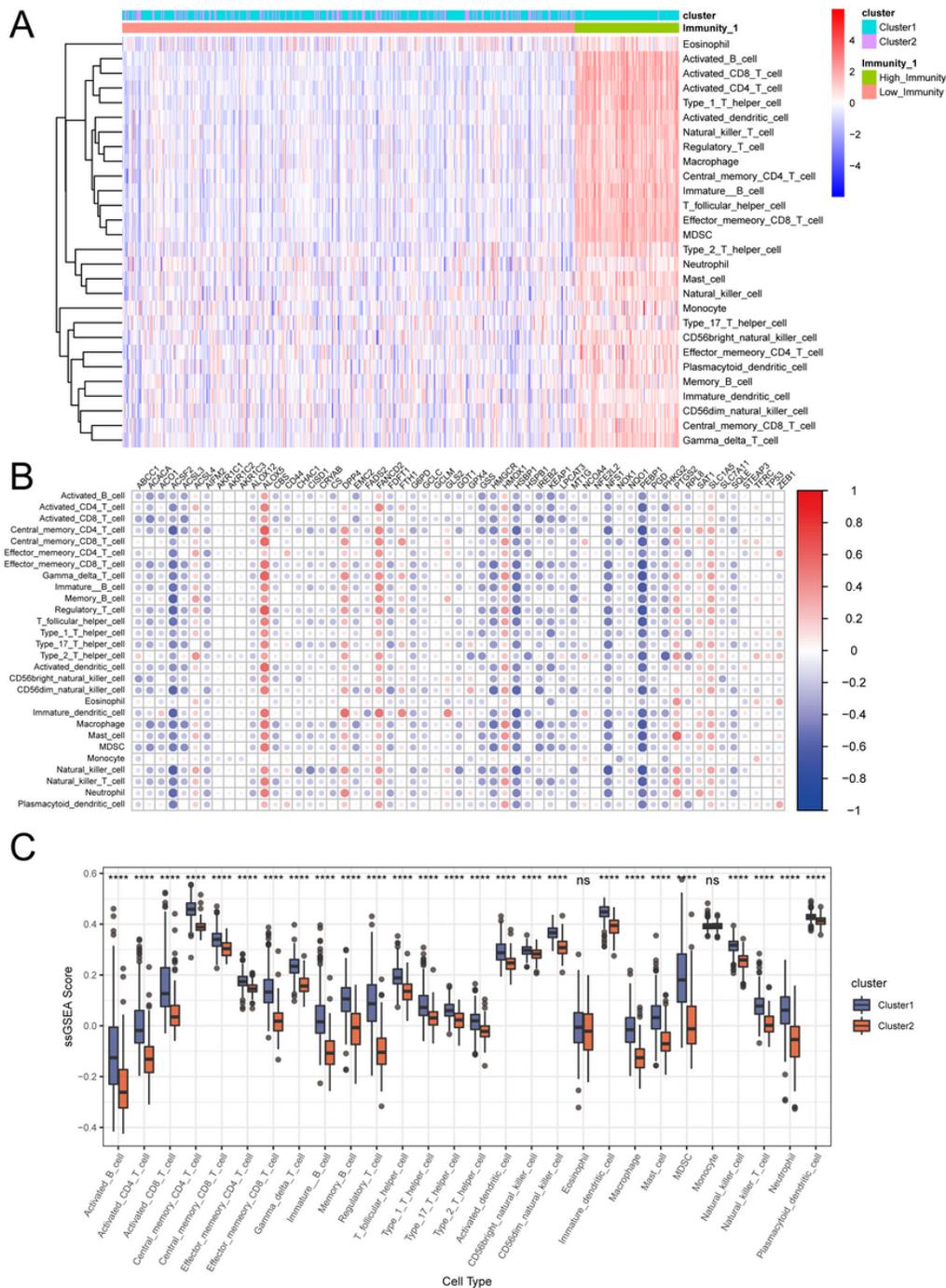


Figure 5

Identification of the two immune subtypes in the THCA cohort. (A) Heatmap of the two immune subtypes based on ssGSEA algorithm for 28 immune gene sets. (B) The correlation heatmap demonstrates the relationship between ferroptosis genes expressions and immune cells infiltration. Dots size shows the extent of their relationships and dots colour indicates if they are positive-related (red dots) or negative-related (blue dots). The number on the scale bar indicates the coefficients of correlation between genes expression and cells infiltration. Blank squares mean insignificance (P > 0.05). (C) The abundance of each TME infiltrating cell in THCA modification patterns. The upper and lower ends of the boxes represented interquartile range of values. The lines in the boxes represented median value, and black dots showed outliers. The asterisks represented the statistical P-value (*P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001).

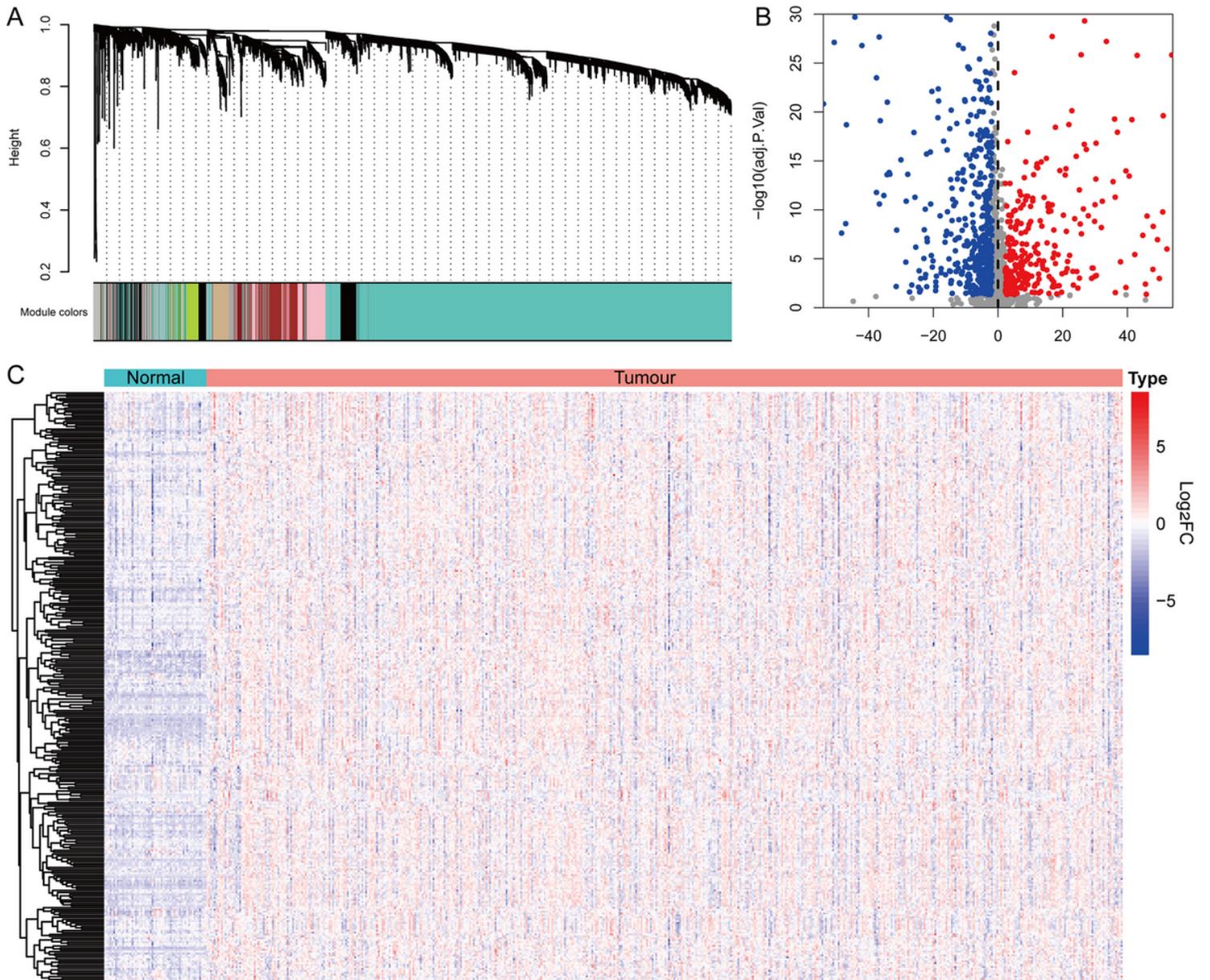


Figure 6

(A) Gene dendrogram and module colours of 502 PTC patients. (B) The volcano plot and (C) heatmap plot of 814 hub genes differentially expressed in tumour and normal tissues (including 308 up-regulated and 506 down-regulated genes).

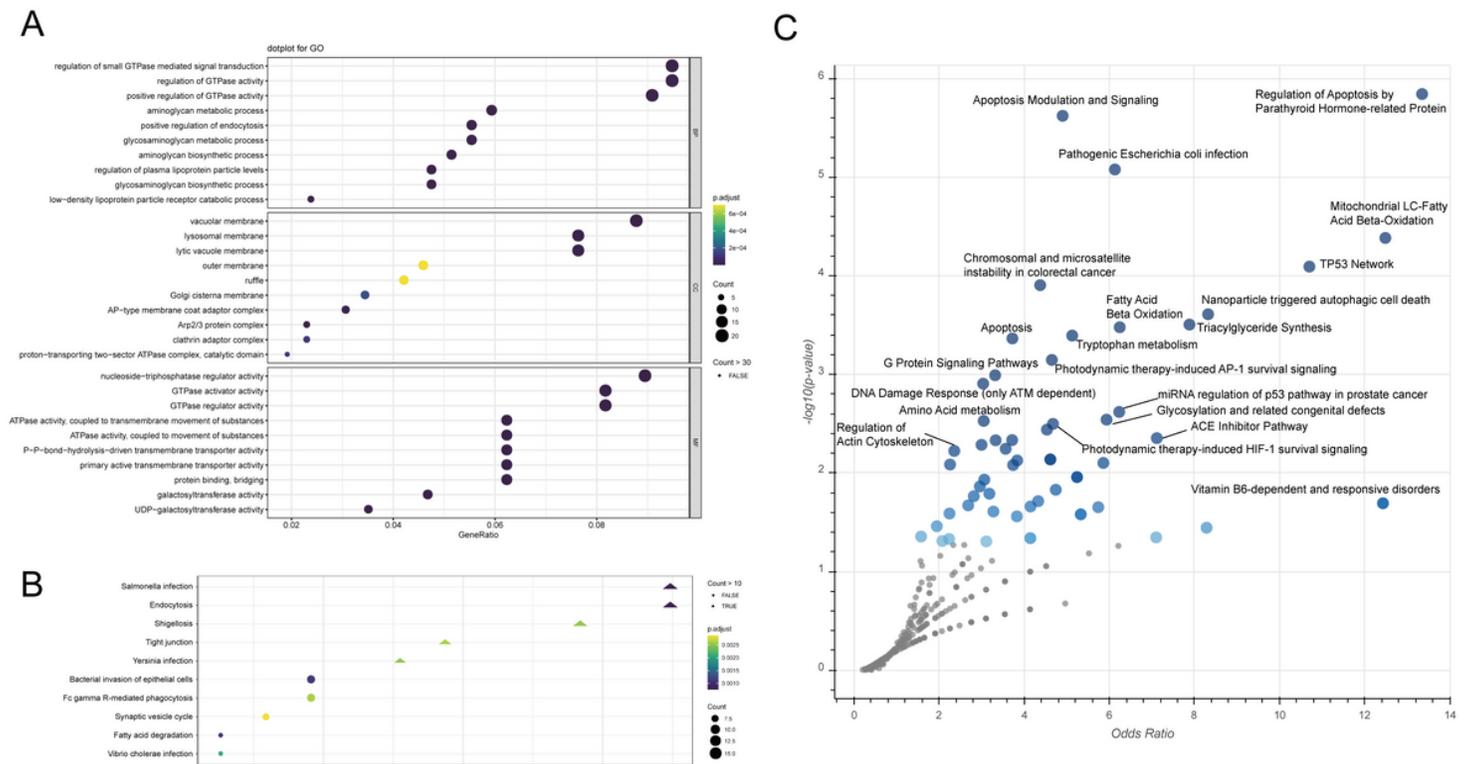


Figure 7

(A) The GO analysis and (B) the KEGG pathway enrichment analysis of 814 hub genes. The node colour changes gradually from yellow to black in ascending order according to the adjusted P-values. The size of the node represents the number of counts. (C) The Wikipathway enrichment analysis. Larger blue points represent significant terms (P -value < 0.05); smaller grey points represent non-significant terms. The darker the blue colour of a point, the more significant it is.

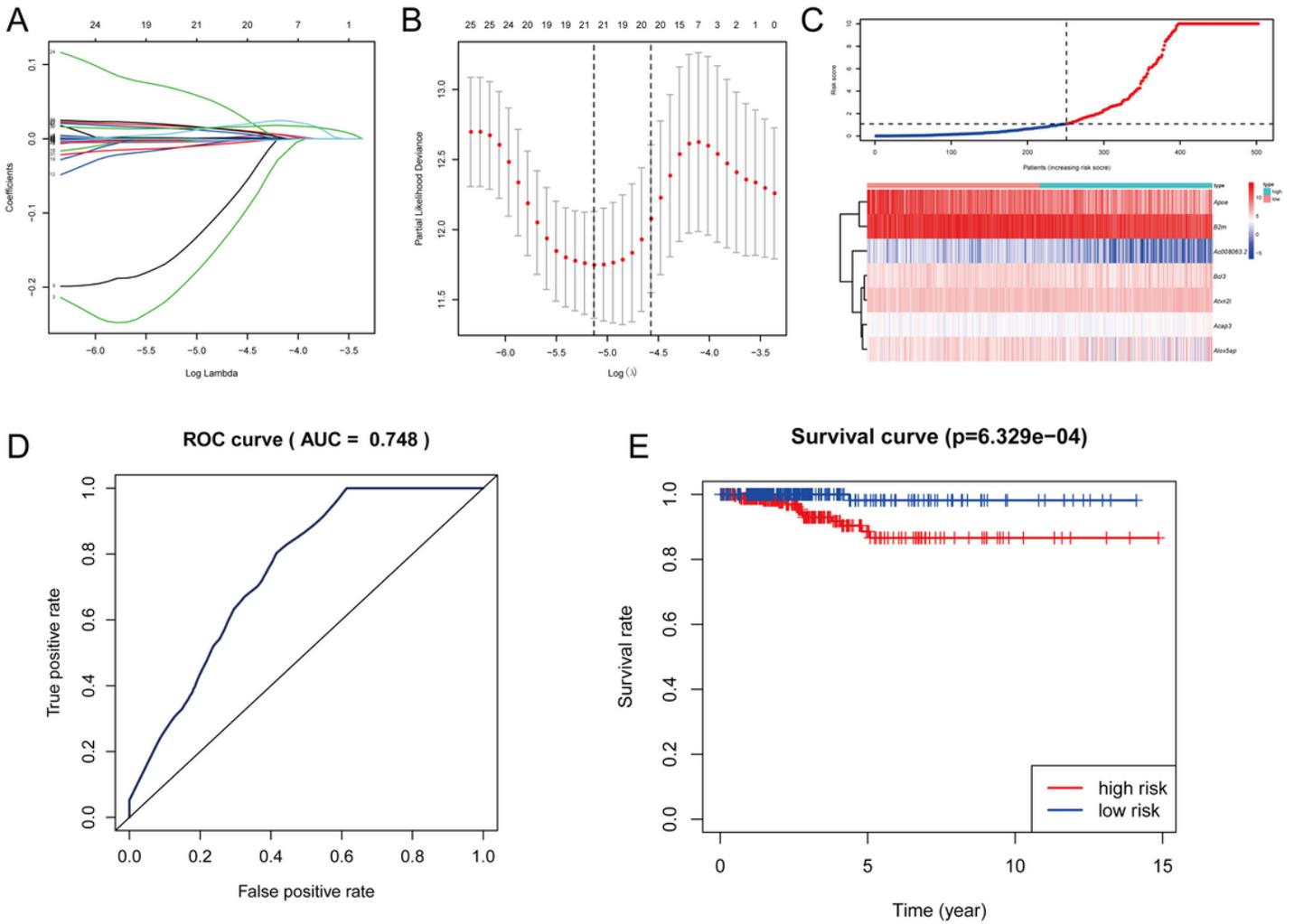


Figure 8

(A) Log (Lambda) value of the 24 genes in LASSO model. (B) Cross-validation for tuning parameter selection in the proportional hazards model, dotted vertical lines were drawn at the optimal values, and seven genes were identified by LASSO regression. (C) The distribution of risk score and gene expression levels among patients. (D) The ROC curve for assessing the reliability of the predict model. (E) The Kaplan-Meier curve of the predicted model.

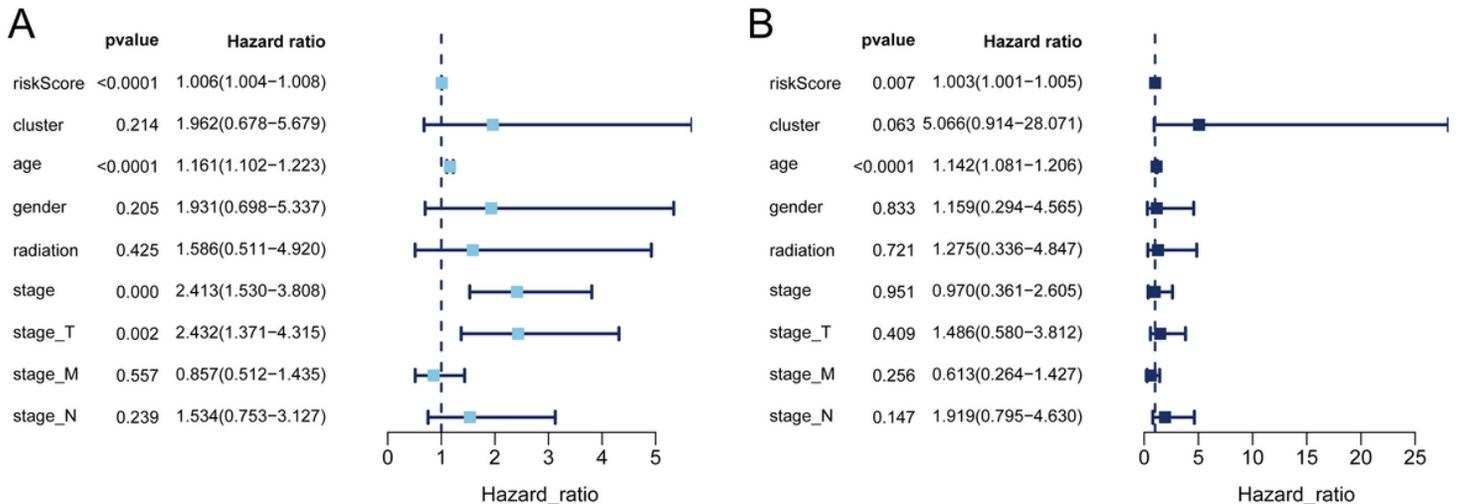


Figure 9

Univariate and Multivariate Cox analysis analysis of risk score, cluster, age, gender, radiation and TNM stage. (A) Prognostic value detection of the gene signature via univariate survival-related analysis in PTC. (B) Prognostic value detection of the gene signature via multivariate survival-related analysis in PTC.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FigureS1.png](#)
- [TableS1KEGGpathways.xlsx](#)
- [TableS2Immunecells.xlsx](#)
- [TableS3Genes.xlsx](#)
- [TableS4Unicox.xlsx](#)