



Identification of Immune-Related Hub Genes in Thymoma: Defects in CD247 and Characteristics of Paraneoplastic Syndrome

Lin-Fang Deng^{1,2*}

¹College of Sciences, Shanghai University, Shanghai, China, ²College of Medicine, Shanghai University, Shanghai, China

Background: Thymomas (Ts) and thymic carcinomas (TCs) are rare primary tumors of the mediastinum. Paraneoplastic syndrome (PNS) is an important feature of thymoma, which presents great challenges to clinicians.

Methods: The present study uses the weighted gene co-expression network analysis (WGCNA) to identify possible immunologic mechanisms of thymoma. RNA sequencing data from thymoma samples were downloaded from the TCGA. Core genes were taken from the module that is closely related to the WHO's stage of classification. Enhanced analysis using the online database "Metascape" and an overall survival (OS) analysis were carried out via the Kaplan–Meier method. The hub genes were obtained from the protein–protein interaction (PPI) network. In addition, we jointly analyzed multiple sets of PNS data related to thymomas from other sources to verify the correlation between thymomas and PNS. The impact of hub genes on the prognosis of PNS was evaluated via the ROC curve, with simultaneous analysis of immune infiltration by CIBERSORT.

Findings: The 14 immune hub genes closely related to thymomas were found to be jointly involved in the T-cell receptor signaling pathway. Compared to the normal thymus and type B1/B2 thymoma, there is a lower number of T-cells in type A/B3 thymoma and thymic carcinoma. The expression of genes related to the T-cell receptor signaling pathway appeared defective. The low expression of CD247 and the decrease in the number of mature T-cells are common features among thymomas, specific pulmonary fibrosis, rheumatoid arthritis, and systemic lupus erythematosus.

Keywords: Thymomas, immune, WGCNA, paraneoplastic syndrome (PNS), CD247

INTRODUCTION

Thymomas (Ts) and thymic carcinomas (TCs) are rare primary tumors of the mediastinum, originating from the thymic epithelium (Scorsetti et al., 2016). The World Health Organization (WHO) divides thymoma into type A, AB, B1, B2, B3, and C according to the morphology of thymoma epithelial cells and the ratio of lymphocytes to epithelial cells in tissues (Marx et al., 2015; Travis et al., 2015). Type C is a thymic carcinoma, and others are thymomas. Thymomas are relatively common primary anterior mediastinal mass, while thymic carcinoma is rare. These two can be synchronous. Thymoma sometimes can develop into carcinomas, but it usually takes 10 to 14 years (Ettinger et al., 2013). In addition to the WHO classification system, the Masaoka staging

OPEN ACCESS

Edited by:

Tao Huang,
Shanghai Institute of Nutrition and
Health (CAS), China

Reviewed by:

Xin Du,
Capital Medical University, China
Jiankang Yang,
Dali University, China

*Correspondence:

Lin-Fang Deng
lfdeng@shu.edu.cn

Specialty section:

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

Received: 14 March 2022

Accepted: 06 May 2022

Published: 14 June 2022

Citation:

Deng L-F (2022) Identification of
Immune-Related Hub Genes in
Thymoma: Defects in CD247 and
Characteristics of
Paraneoplastic Syndrome.
Front. Genet. 13:895587.
doi: 10.3389/fgene.2022.895587

system and TNM staging system are also commonly used in clinical thymoma (Ried et al., 2018). Masaoka classifies thymomas into stages I, II (IIa and IIb), III, and IV (IVa and IVb) according to their comprehensive capsule infiltration and tumor metastasis (Masaoka, 2010). However, the TNM staging system includes lymph nodes in the evaluation criteria, which is of great significance for thymoma. In general, there appears to be a clear correlation and correspondence between several staging methods.

Due to the insidious onset and slow progression of thymoma, the first diagnosis is usually found by chance in physical examination, or physical discomfort caused by paraneoplastic syndrome (PNS). PNS is an important feature of thymoma, bringing great challenges to clinicians (Blum et al., 2020). The common PNS of thymoma are myasthenia gravis, total red cell aplasia, polymyositis, systemic lupus erythematosus, rheumatoid arthritis, Cushing syndrome, and syndrome of inappropriate antidiuretic hormone secretion (Rajan and Zhao, 2019; Rubinstein et al., 2019; Hashimoto et al., 2020; Sideris and Huang, 2020). A sizeable percentage (25–40%) of thymoma patients with myasthenia gravis, and more than 15% of patients with fever and immunodeficiency such as syndromes (Zhao et al., 2020; Liao et al., 2021). Thus, deciphering the relationship between this particular clinical symptom and tumor is crucial for the treatment of thymoma. The treatment and management of thymomas include chemoradiotherapy, the use of corticosteroids, immunotherapy, tyrosine kinase inhibitors, and surgical resection. However, the complications, caused by surgery and radiotherapy, such as radiation pericarditis, radiation pneumonia, and pulmonary fibrosis can hasten the patients' death (Ettinger et al., 2013; Jeffrey Yang et al., 2020; Tian et al., 2020; Yu et al., 2020).

In recent years, with the development of systems biology, an increasing number of researchers have explored and predicted the prognostic targets of tumors from the perspectives of gene expression and molecular interaction. However, the etiology of thymoma is still unclear, and the association between PNS and thymoma is also ambiguous. Therefore, this study explores the biochemical mechanism of thymoma from the perspective of systems biology. Weighted correlation network analysis (WGCNA) is a high-quality method for finding clusters (modules) of highly correlated genes (Langfelder and Horvath, 2008; Langfelder and Horvath, 2012), which is a method of data reduction and unsupervised classification. This method can form a module of co-expressed genes to simplify the complex data matrix. Moreover, we can further explore the correlation between gene network and phenotype of concern and explore the hub genes in the network. This method has been widely used since its development. For example, Tian et al. used WGCNA to identify two gene co-expression modules involved in the process of lung squamous cell carcinoma metastasis and suggested that CFTR, SCTR, and FIGF genes could be used as a potential prognostic biomarker (Tian et al., 2017). Magdalena Niemira et al. (2019) applied WGCNA for exploring molecular networks associated with clinical traits such as tumor size, SUV max, BMI, smoking status, recurrence-free survival, and disease-free survival (Niemira et al., 2019). In addition, WGCNA has been applied

to breast cancer, liver cancer, colon adenocarcinoma, and other diseases (Yin et al., 2018; Zhou et al., 2018; Wang et al., 2019).

This study analyzes the immunologic mechanism of thymoma by WGCNA. RNA sequencing data from thymoma samples were downloaded from the TCGA. The thymomas datasets GSE177522, GSE57892, and GSE29695 from GEO were used as the validation set to verify the reliability of the results. In addition, we jointly analyze multiple sets of PNS data related to thymomas (GSE33566, GSE93272, and GSE138458) to explore the internal relationship between thymomas and PNS.

MATERIALS AND METHODS

Data Sources and Pre-Processing

RNA sequencing data and clinical information for thymomas were downloaded using TCGAAbiolinks, a third-party tool officially recommended by GDC in R language (Colaprico et al., 2016; Silva et al., 2016; Mounir et al., 2019). The TCGAAbiolinks tool downloads data through GDC official API. There were 119 cases of thymomas (including 36 cases of type AB thymoma, 15 of type A thymoma, 14 of type B1 thymoma, 30 of type B2 thymoma, 13 of type B3 thymoma, and 11 of thymic carcinomas) and two cases of normal thymus tissue. The validation sets GSE177522, GSE57892, and GSE29695 were downloaded from the GEO official website (<https://www.ncbi.nlm.nih.gov/geo/>). Some common PNS data were also downloaded from the GEO official website, including rheumatoid arthritis (GSE93272), systemic lupus erythematosus (GSE138458), and pulmonary fibrosis (GSE33566).

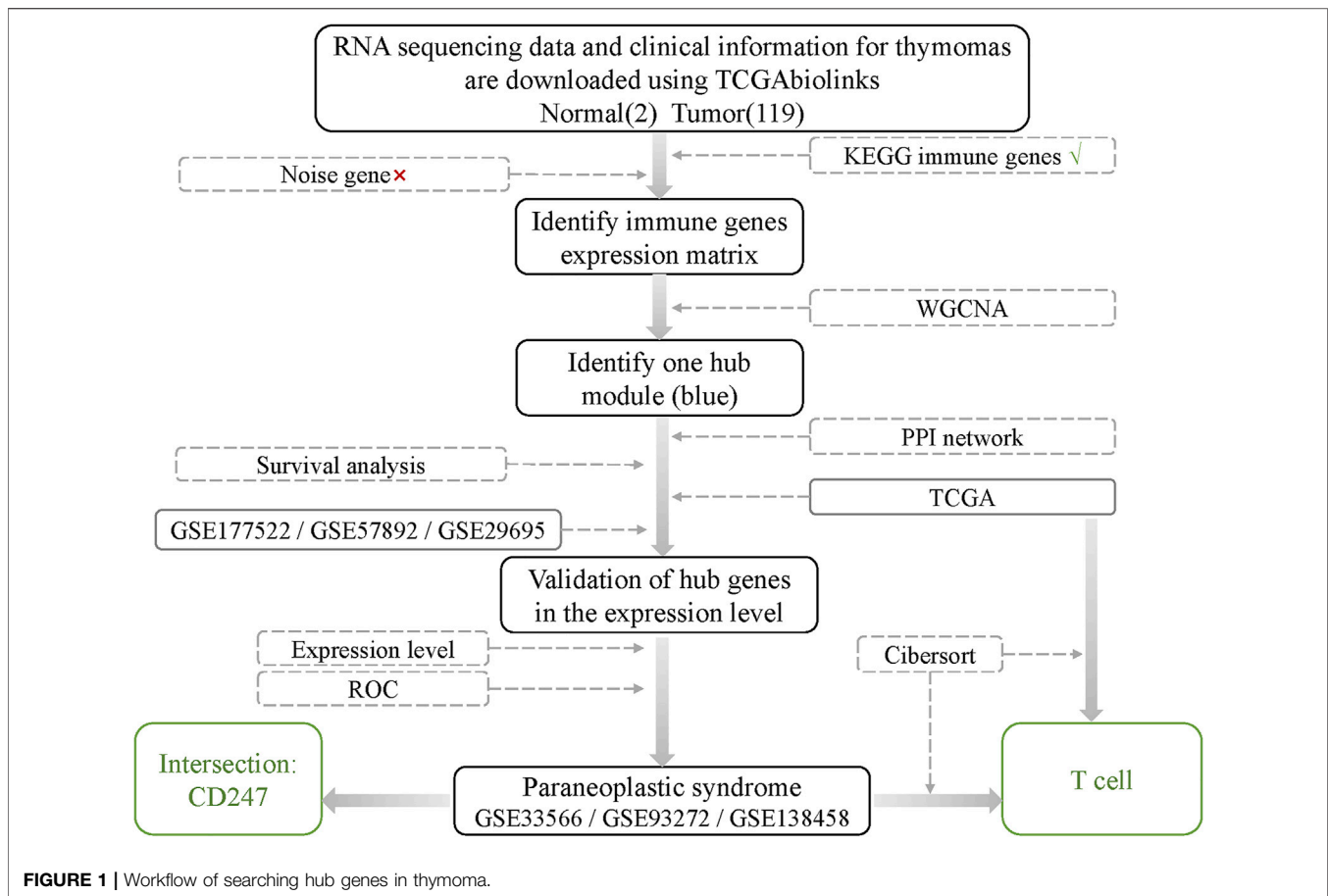
The workflow of the study is shown in **Figure 1**.

Screening of Immune Gene and Noise Gene

The purpose of this study is to find out the relationship between thymoma and immunity. Therefore, a total of 2,381 genes in all immune-related pathways were collected from the KEGG website (<https://www.genome.jp/kegg/>). First, we extract the expression profiles of these immune genes from thymoma data. For details of the gene list, cf. Table 1.XLSX of the **Supplementary Materials** (SM). Second, noise genes were screened by the correlation method. Specifically, the Spearman correlation coefficient matrix of immune genes was calculated. It was determined that the two genes are not correlated when the correlation coefficient r is in the range of $(-0.2, 0.2)$. When a gene is not correlated with 70 % of the remaining genes, it is determined to be a noise gene (cf. SI-1 of the Data Sheet 1.PDF). The data matrix of the optimized immune-related genes is 2023×121 .

Weighted Gene Co-Expression Network Analysis

This study uses WGCNA (version: 1.70–3), downloaded by BiocManager (version: 1.30.10) in the R suite, to construct the immune co-expression network (Langfelder and Horvath, 2008; Langfelder and Horvath, 2012). After screening and testing the immune expression matrix, the Pearson correlation coefficient matrix is calculated and a suitable soft threshold is selected.



$$c_{ij} = \frac{\text{cov}(i, j)}{\sigma_i \sigma_j},$$

where $\text{cov}(i, j)$ refers to the covariance of genes i and j , σ represents the standard deviation of gene, and c_{ij} is the Pearson correlation coefficient between genes i and j . Then, the adjacency matrix, a matrix of weighted correlation between genes, was constructed using the power function. This method strengthens the strong correlation and weakens the weak correlation or negative correlation, which makes the correlation value more in line with the scale-free network characteristics and more biological significance.

$$a_{ij} = |c_{ij}|^\beta,$$

where a_{ij} is adjacency between those two genes. Then, the topological overlap matrix (TOM) was constructed using the adjacency function to reduce noise and false correlation.

$$\text{TOM}_{i,j} = \frac{I_{ij} + a_{ij}}{\min(k_i + k_j) + 1 - a_{ij}},$$

where I_{ij} is the product's sum of the adjacency coefficients of the nodes connected by genes i and j , and k refers to the sum of the adjacency coefficients of the given gene with all other genes in the weighted network. The TOM is a method to quantitatively

describe the similarity in nodes by comparing the weighted correlation between two nodes and other nodes.

Identification of Clinically Significant Modules and Immune Hub Genes in Thymoma

The co-expression module is a collection of genes with high topological overlap similarity. First, principal component analysis (PCA) was used to find the first principal component of each module (module eigengene, ME) to represent the expression pattern of the module. Then, the correlation between these modules and clinical data was calculated to determine the concerned clinical information and gene modules. Module membership (MM) and gene significance (GS) were used to describe the reliability of a gene in the module. The intramodular connectivity may be interpreted as a measure of MM. Genes with intramodular connectivity greater than 0.8 were selected as hub genes (highly connected genes).

Survival Analysis of Hub Genes

The ability of hub genes to predict survival is based on Kaplan–Meier analysis by using the “survival” package (version: 3.2–7) in the R suite. First, the expression profile of hub genes was extracted from TCGA data. Second, the median

expression value of each hub gene was determined. Third, the tumor samples were divided into high-expression groups and low-expression groups with the median of each gene as the boundary. The median sample is divided into high-expression groups. Finally, differences in survival between high- or low-expression groups were assessed for significance by the log-rank test. If $p < 0.05$, we consider the gene to be a validated hub gene.

Functional Enrichment of Hub Genes

In order to analyze the biological functions and signaling pathways of differentially expressed genes in thymoma-related modules, we perform enrichment analysis using the online database “Metascape” (<https://metascape.org/gp/index.html#/main/step1>) (Zhou et al., 2019). At the same time, protein-protein interaction (PPI) network can be given by Metascape, which uses the molecular complex detection (MCODE) algorithm.

CIBERSORT

CIBERSORT (Newman et al., 2015) is an R package/webpage tool for deconvolution of the expression matrix of human immune cell subtypes based on the principle of linear support vector regression. The proportions of the 22 tumor-infiltrating immune cells (TIICs) from each sample were determined by using the “CIBERSORT” (R package). CIBERSORT was used to analyze the relative expression levels of 547 genes in individual tissue samples according to their gene expression profiles, to predict the proportion of 22 types of TIICs in each tissue. CIBERSORT derived a p -value for the deconvolution of each sample, which provided a measure of confidence in the results, and $p < 0.05$ was considered accurate. Significant results ($p < 0.05$) were selected for subsequent analysis.

The CIBERSORT results of TCGA thymoma data were downloaded from the GDC website (<https://gdc.cancer.gov/about-data/publications/panimmune>) (Thorsson et al., 2018). The CIBERSORT result of GEO data was computed using the “cibersort” package of R (the number of permutations: perm = 1,000). For details of the list, see **Supplementary Material: results_cibersort.xlsx**.

RESULTS

Immune Gene Expression Profile Data in Thymoma

A total of 2,381 genes related to immune pathways were obtained from the KEGG website. The data matrix of the immune gene expression profile of thymoma was obtained by taking intersection with TCGA thymoma expression profile data, and the matrix dimension was 2247 * 121. Then the noise genes were screened by the correlation method, and the optimized immune gene expression data matrix of thymoma was obtained, with the matrix dimension of 2023 * 121. The optimized matrix avoided the interference of noise in the analysis, and there was a strong correlation between the 2,023 immune genes that remained. The correlation heat map of the immune expression data before and after noise reduction is shown in **Figure 2**.

Construction of Weighted Gene Co-Expression Network

WGCNA was used to construct a network based on the expression matrix of 2023 immune genes and clinical data from 121 thymoma samples.

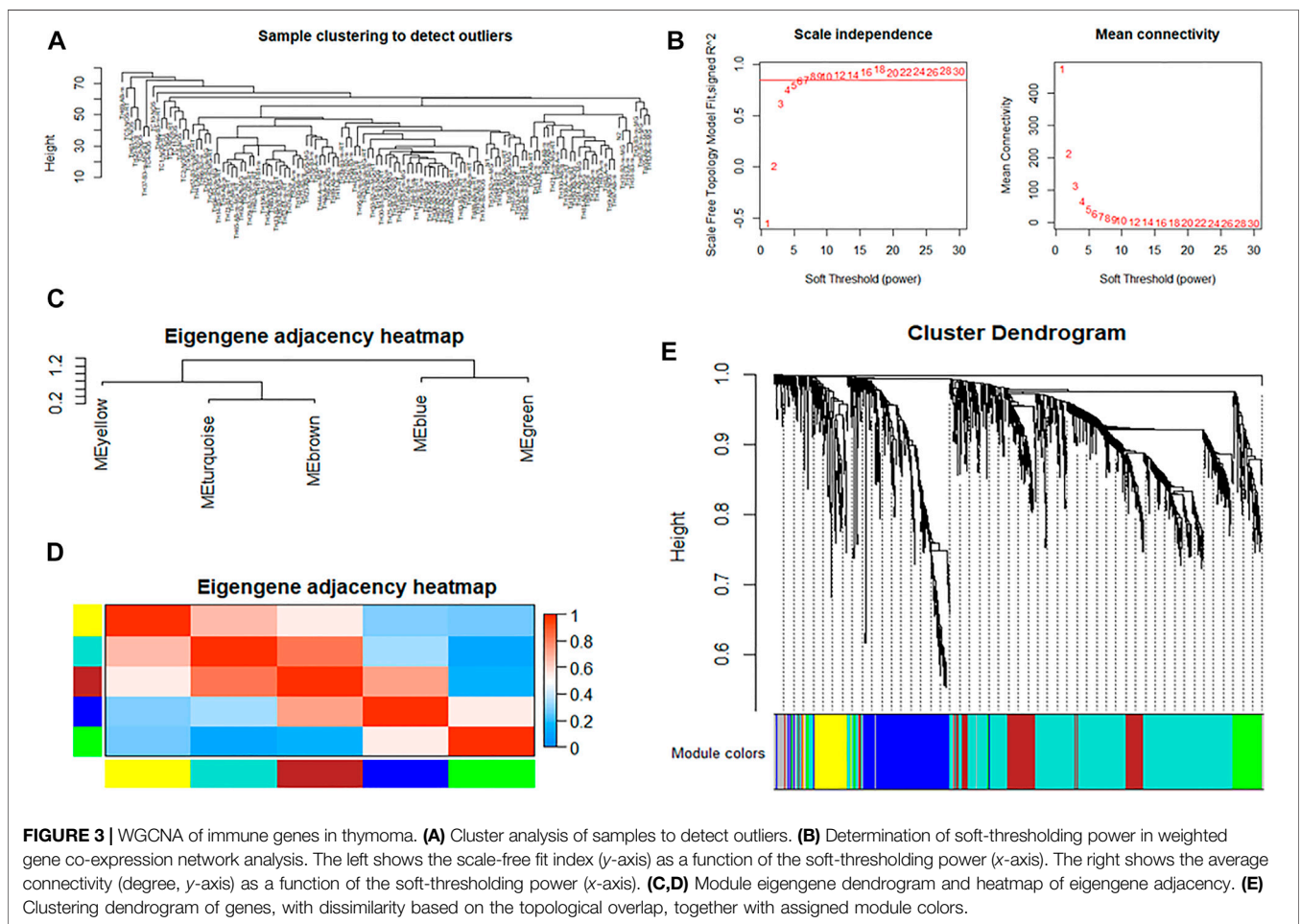
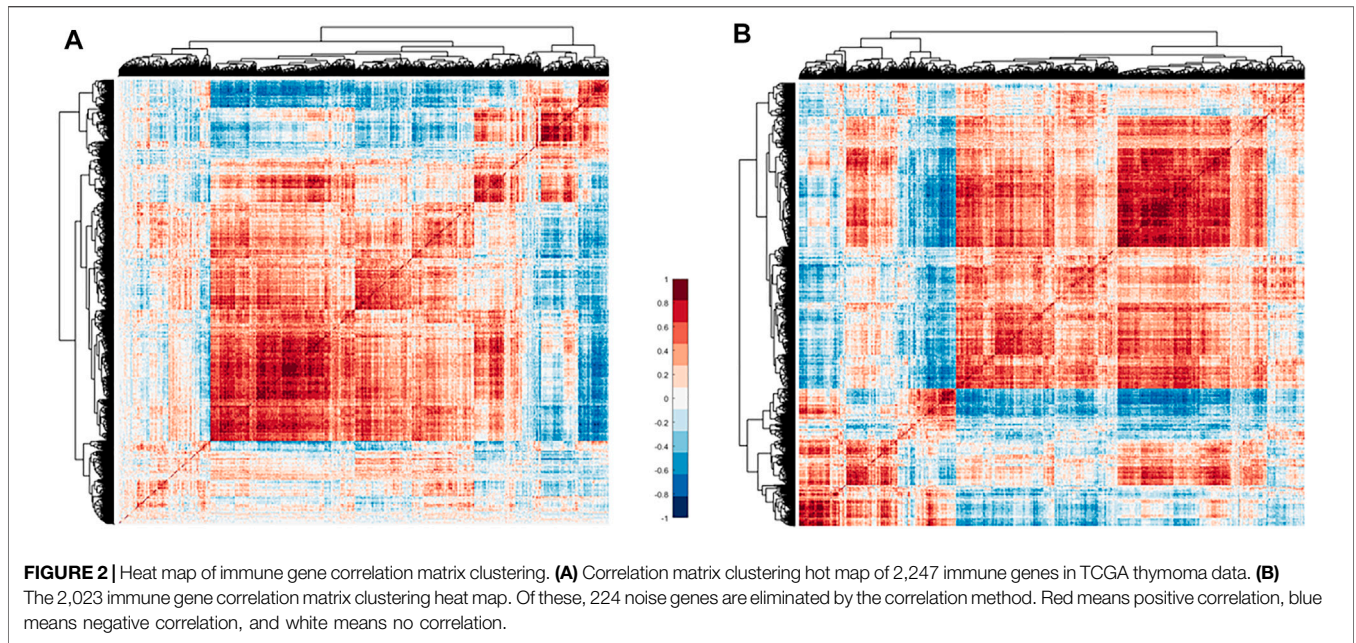
A dendrogram of samples was clustered by the average linkage method and Pearson’s correlation method to check the quality of the data from the 121 samples, and no outliers were identified for removal (**Figure 3A**). To construct a scale-free network, we set the soft threshold power β to 8, and the independence degree to 0.9 (**Figure 3B**). The result showed that there were a total of five co-expression modules; the gray module contains genes that could not be incorporated into any other module (**Figures 3C–E**).

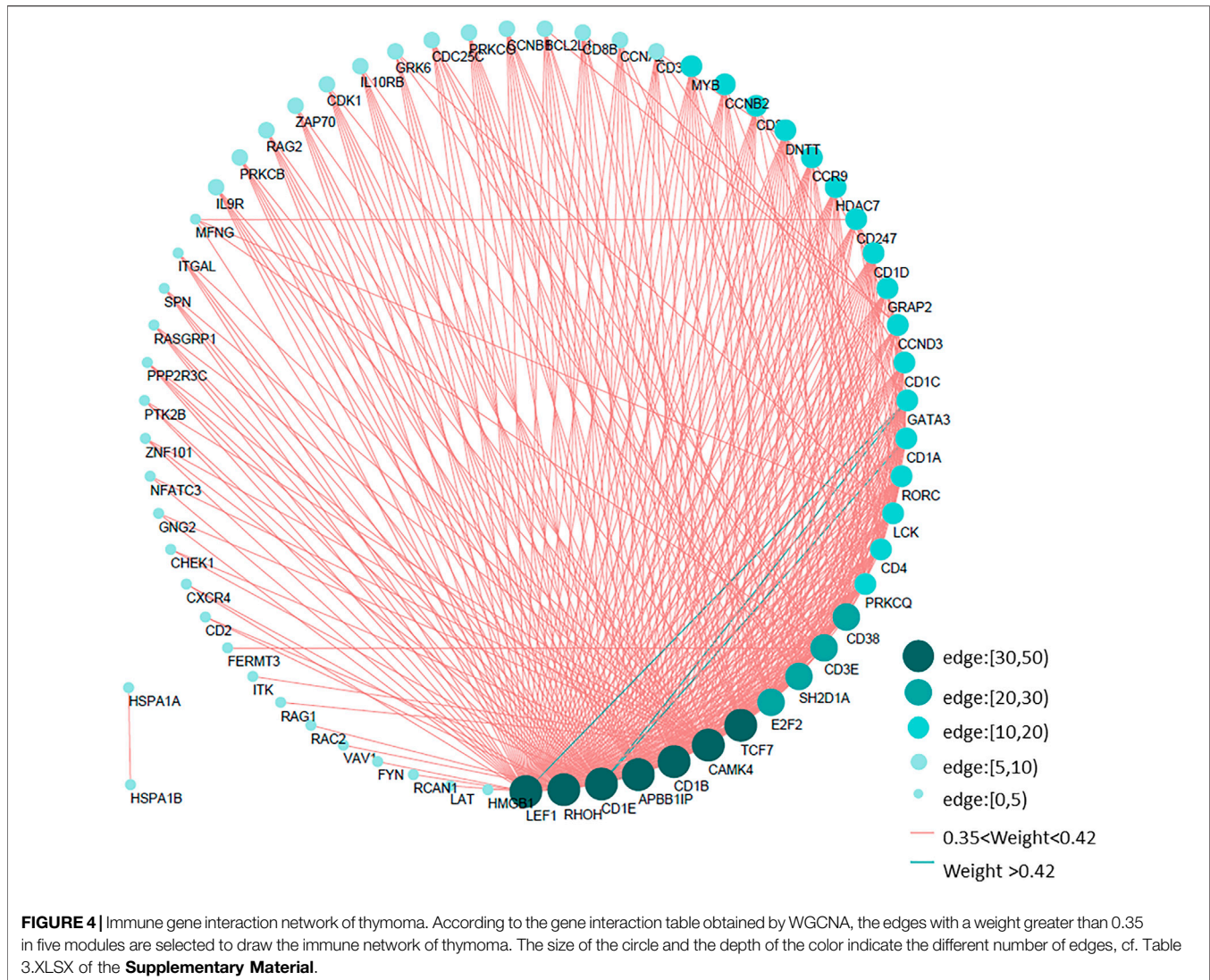
According to the topological overlap matrix (TOM), the connection relationship between genes in each module could be obtained. Different colors indicate that the weight in different modules was greater than 0.35, and the immune gene interaction network of thymoma was visualized by Cytoscape. Among them, LEF1, RHOH, APBB1IP, CD1B, CAMK4, and TCF7 interact with more than 30 immune genes, which are the core genes of the immune network of thymoma, shown as **Figure 4**.

Association of Modules With Clinical Traits and Determination of Core Genes

In addition, we calculated the correlation between module genes and clinical traits of thymoma. For each module, the gene co-expression was summarized by the eigengene, and the correlations of each eigengene with clinical traits were calculated, such as history myasthenia gravis, radiation therapy, gender, OS, OS time, age, tumor, tissue or organ, Masaoka stage, and WHO stage. We found that the blue module had the highest correlation with the WHO stage ($cor = 0.72$, $p = 8e - 21$) (**Figures 5A,B**).

Therefore, the blue module was analyzed for core genes. The first 87 intersection genes of the blue module with the highest correlation (intramodular connectivity >0.7) were selected as the core genes for subsequent study and enrichment analysis. For details of the core gene list, cf. Table 4.XLSX of the **Supplementary Material**. The results of functional enrichment analysis were obtained by the online database Metascape; terms with a p -value < 0.01, a minimum count of 3, and an enrichment factor >1.5 (the enrichment factor is the ratio between the observed counts and the counts expected by chance) were collected and grouped into clusters based on their membership similarities. We selected the top 20 clusters with their representative enriched terms, mainly involving T-cell activation (GO:0042110), T-cell receptor (TCR) signaling pathway (WP69), human immunodeficiency virus 1 infection (hsa05170), hematopoietic cell lineage (hsa04640), primary immunodeficiency (hsa05340), hemostasis (R-HSA-109582), adaptive immune response (GO:0002250), human T-cell leukemia virus 1 infection (hsa05166), and leukocyte activation involved in immune response (GO:0002366) (**Figure 5C**). Metascape combines DisGeNET, a comprehensive platform integrating information on human disease-associated genes





and variants, and gives the disease information related to these core genes (**Figure 5D**), most of which are autoimmune diseases or T-cell-related diseases, such as precursor T-Cell lymphoblastic leukemia-lymphoma, T-cell lymphoma, combined immunodeficiency, and lupus erythematosus. Metascape also gives a summary of enrichment analysis in PaGenBase. PaGenBase is a novel repository for the collection of tissue- and time-specific pattern genes. The results show that most of these core genes were specifically expressed in the thymus (**Figure 5E**), which proved the correctness of these core genes to some extent.

At the same time, the protein-protein interaction (PPI) network of core genes was given by Metascape (**Figure 5F**), which applied a mature complex identification algorithm called MCODE to automatically extract protein complexes embedded in such a large network. Where the T-cell receptor signaling pathway was the core cluster in this PPI network ($\log_{10}(p) = -37.4$), the hub genes of this cluster include

CD247, CD8A, CD8B, PRKCQ, CD3E, CD3G, GRAP2, VAV1, CD4, LCK, ZAP70, LCP2, ITK, FYN, and LAT.

Validation of Hub Genes in the Expression Level

In the 87 core genes screened by WGCNA, which have a strong correlation with WHO's classification of thymomas, a core cluster, the T-cell receptor signaling pathway, was obtained by PPI enrichment analysis, which contained 15 interrelated hub genes. We observed the expression levels of these 15 genes in TCGA data using a thermal map (**Figure 6A**), and found that these hub genes that were lowly expressed in type A and type B3 thymomas and thymic carcinomas are highly expressed in type B1 and type B2 thymomas. There is no consistent expression pattern in AB thymomas, but thymoma patients with MG have relatively low-expression in AB (cf. SI-2 of the Data Sheet 1.PDF). To verify these results, three groups of thymic tumor data were

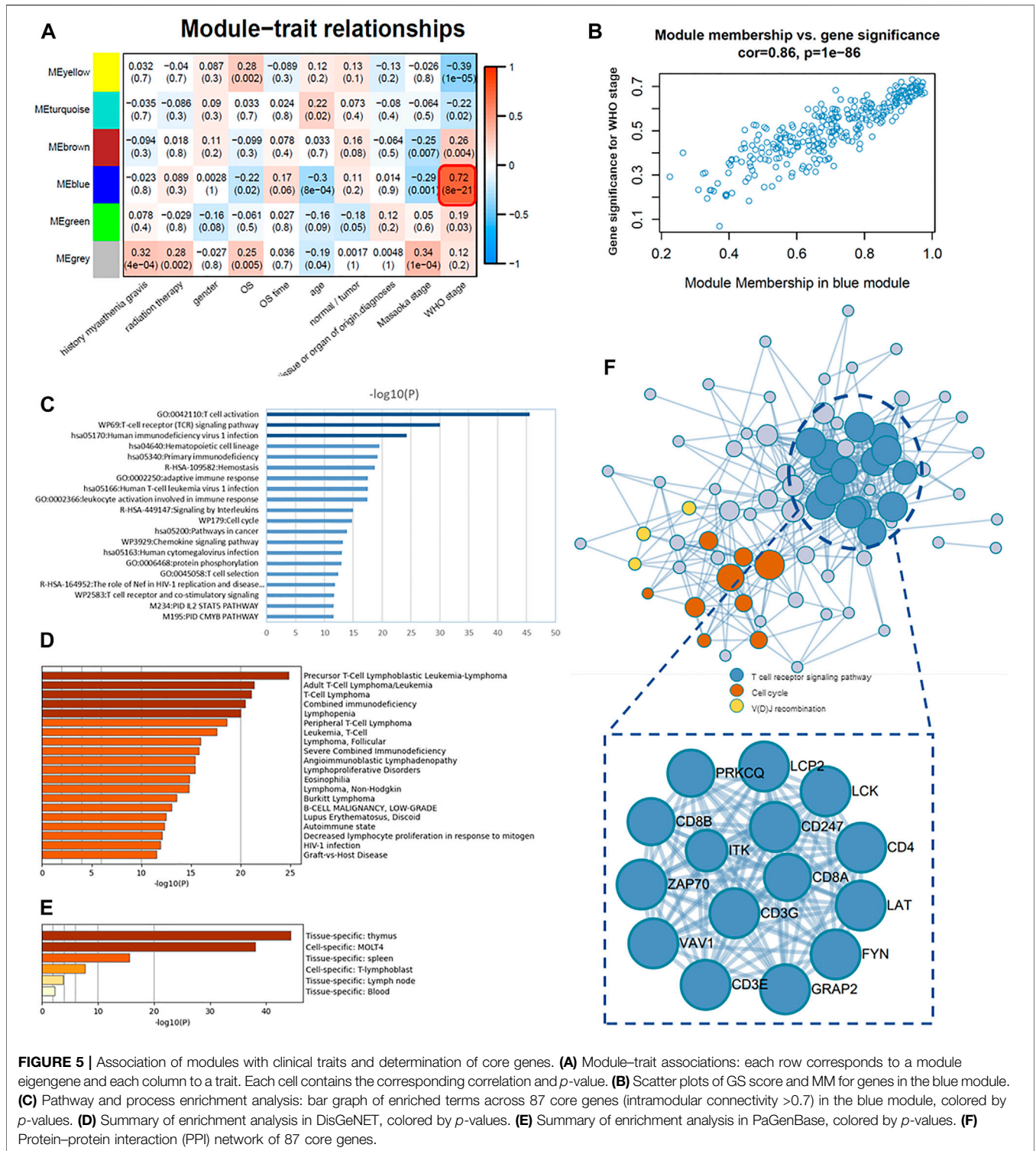


FIGURE 5 | Association of modules with clinical traits and determination of core genes. **(A)** Module-trait associations: each row corresponds to a module eigengene and each column to a trait. Each cell contains the corresponding correlation and *p*-value. **(B)** Scatter plots of GS score and MM for genes in the blue module. **(C)** Pathway and process enrichment analysis: bar graph of enriched terms across 87 core genes (intramodular connectivity >0.7) in the blue module, colored by *p*-values. **(D)** Summary of enrichment analysis in DisGeNET, colored by *p*-values. **(E)** Summary of enrichment analysis in PaGenBase, colored by *p*-values. **(F)** Protein-protein interaction (PPI) network of 87 core genes.

used as validation data, and the data processing method was consistent with TCGA data, which were standardized. In GSE57892, type A and B3 thymoma and thymic carcinoma has almost low expression in these 15 hub genes, while type B2 thymoma has a large expression level (Figure 6C). The data GSE177522 also verify the result that the 15 hub genes were

almost low expressed in thymic carcinoma (Figure 6D). The data GSE29695 clearly show that B3 thymoma cannot be expressed in these 15 hub genes, and the relative expression levels of B1 and B2 were higher (Figure 6E).

The WHO stage is mainly based on the morphology of thymoma epithelial cells and the ratio of lymphocytes to

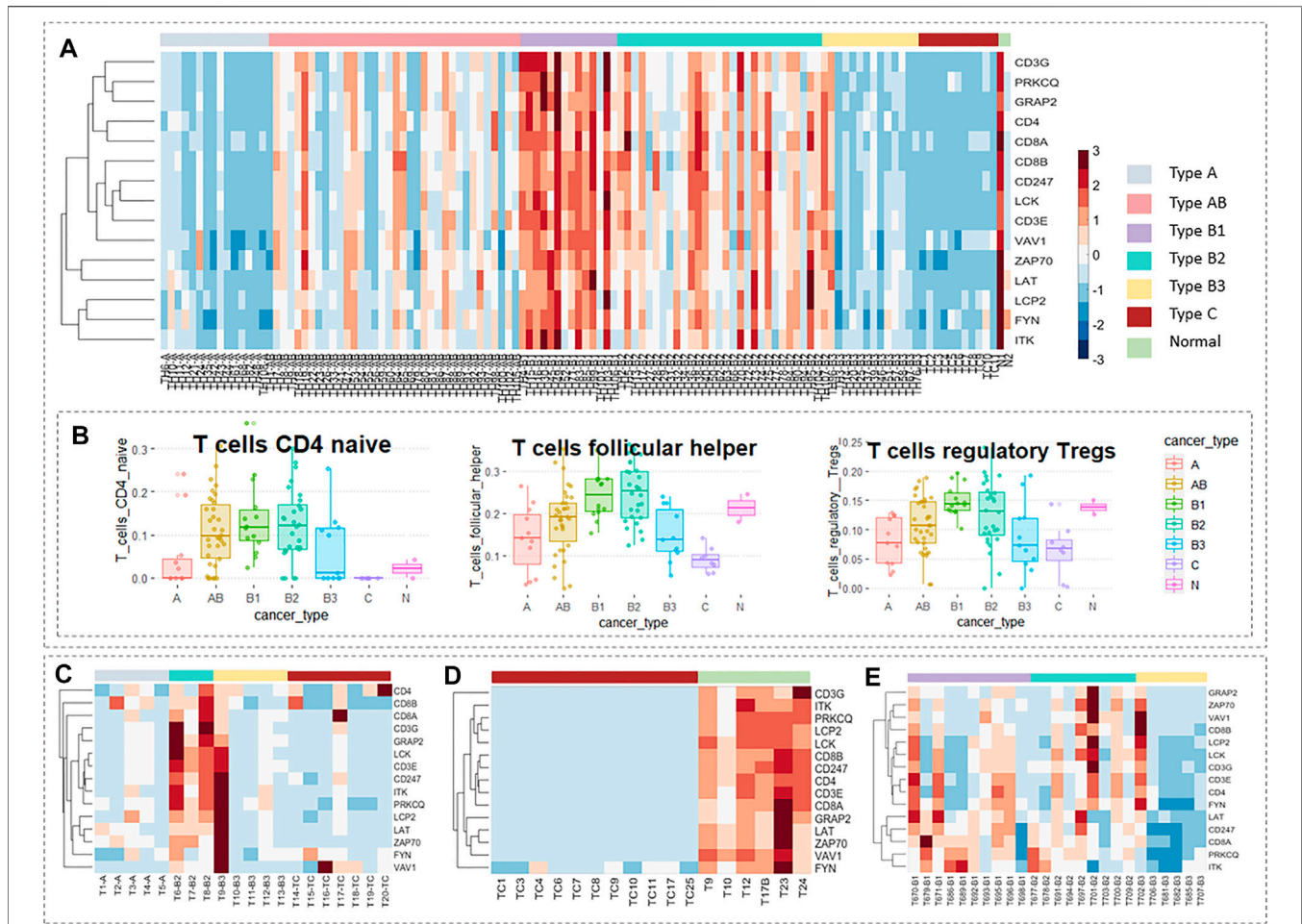


FIGURE 6 | Validation of hub genes in the expression level. Expression levels of 15 hub genes in TCGA data. **(B)** The proportion of T-cells of each thymic tumor subtype in TCGA data obtained by the CIBERSORT algorithm. **(C)** Expression levels of 15 hub genes in GSE57892 data. There are 25 samples in total, including two cases of type AB thymoma, five of type A thymoma, three of type B2 thymoma, five of type B3 thymoma, seven of thymic carcinomas, and three of cell line. **(D)** Expression levels of 15 hub genes in GSE177522 data. There are 19 samples in total, including 11 cases of thymic carcinoids, two cases of thymoma, and six cases of the normal thymus. **(E)** Expression levels of 15 hub genes in GSE29695 data. There are 41 samples in total, including nine cases of type AB thymoma, one of type A thymoma, 10 of type B1 thymoma, nine of type B2 thymoma, six of type B3 thymoma, one of type A/B thymoma, one of type B1/B2 thymoma, and four of the cell line. In particular, cell lines and subtypes of less than three samples were not included in the study.

epithelial cells in tissues. We know that these 15 genes which have a strong correlation with WHO classification belong to the T-cell receptor signaling pathway. Interestingly, the variation trend of the proportions of three T-cells in different thymic tumor subtypes obtained from the CIBERSORT algorithm in **Figure 6B** was very consistent with the variation trend of the expression levels of these 15 hub genes (cf. Table 2.XLSX of the **Supplementary Materials**). A similar trend of expression spectrum and immune infiltration can be used as an inspiration for the phenomenological hypothesis in this group of hub genes.

Certainly, for sake of comprehending whether the 15 hub genes are related to the prognosis of patients, we performed a Kaplan–Meier analysis of these genes according to the clinical data of 119 cases of thymomas in TCGA. It was found that 14

genes were associated with prognosis except FYN ($p < 0.05$). The Kaplan–Meier survival curves of 15 hub genes are given in SI-6 of the Data Sheet 1.PDF.

Taken together, these validation analyses confirm that the more severe the subtype of thymomas is, the less the number of T-cells is, and the expression of related genes in the T-cell receptor signaling pathway is defective.

DISCUSSION

In this study, we obtained 14 hub immune genes of thymomas through statistical analysis. The common correlation pathway of these 14 genes was the T-cell receptor signaling pathway, and their expression was closely related to the WHO stage.

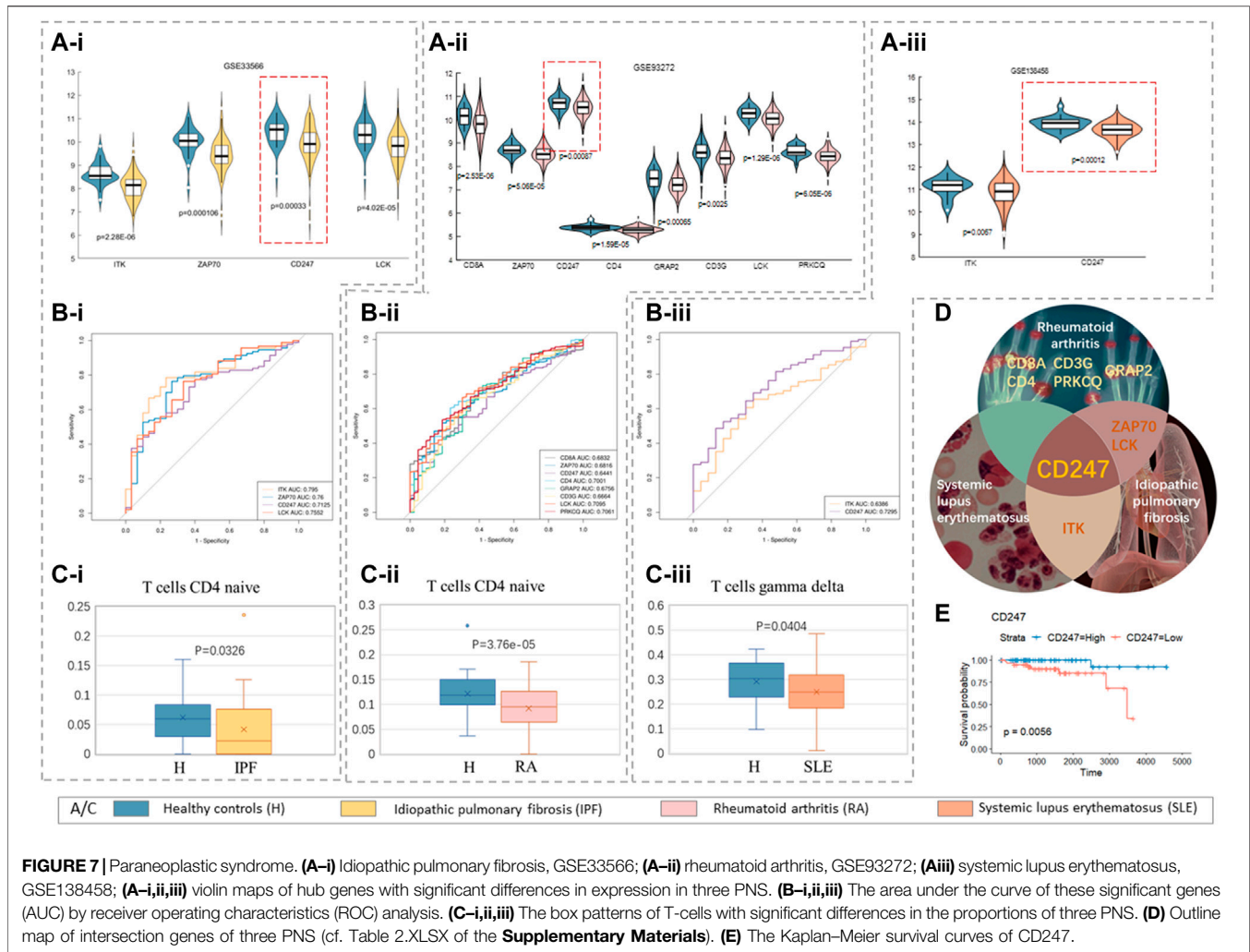


FIGURE 7 | Paraneoplastic syndrome. **(A–i)** Idiopathic pulmonary fibrosis, GSE33566; **(A–ii)** rheumatoid arthritis, GSE93272; **(A–iii)** systemic lupus erythematosus, GSE138458; **(A–i,ii,iii)** violin maps of hub genes with significant differences in expression in three PNS. **(B–i,ii,iii)** The area under the curve of these significant genes (AUC) by receiver operating characteristics (ROC) analysis. **(C–i,ii,iii)** The box patterns of T-cells with significant differences in the proportions of three PNS. **(D)** Outline map of intersection genes of three PNS (cf. Table 2.XLSX of the **Supplementary Materials**). **(E)** The Kaplan–Meier survival curves of CD247.

It is well known that there are generally no specific symptoms for patients with thymomas, but they may have nonspecific symptoms such as chest pain, chest tightness, palpitation, fatigue, and cough. Therefore, they are often ignored by patients who miss the best opportunity for early diagnosis. The first diagnosis is usually found by chance in physical examination or because of the physical discomfort caused by PNS. Thymoma is closely related to autoimmune disorders, and most of its concomitant PNS are autoimmune diseases. Therefore, in this study, we selected two kinds of autoimmune PNS, namely, systemic lupus erythematosus (GSE138458) and rheumatoid arthritis (GSE93272), as the objects of discussion. At the same time, specific pulmonary fibrosis (GSE33566), a severe complication that often occurs after thymic tumor surgery, was also selected. There were 330 samples in dataset GSE138458, including 307 cases of systemic lupus erythematosus (SLE) and 23 cases of healthy control. For dataset GSE93272, there were 232 cases of rheumatoid arthritis (RA) and 43 cases of healthy control.

The dataset GSE33566 includes 93 cases of idiopathic pulmonary fibrosis (IPF) and 30 cases of healthy control.

Considering the condition that the p -value of survival analysis was less than 0.01 and the p -value of expression difference was less than 0.01, we found that ITK, ZAP70, CD247, and LCK were lower expressed in patients with specific pulmonary fibrosis than in healthy controls (**Figure 7A–i**). The expression of CD8 A, ZAP70, CD247, CD4, GRAP2, CD3 G, LCK, and PRKCQ in patients with rheumatoid arthritis was lower than that in healthy controls (**Figure 7A–ii**). Compared with the healthy control group, the expression levels of ITK and CD247 in patients with systemic lupus erythematosus decreases (**Figure 7A–iii**). All hub genes expression levels were significantly different between the highest and lowest quartiles, and the area under the curve (AUCs) of these genes was higher than 0.6 (**Figures 7B–i,ii,iii**), confirming the authenticity of the differences in these genes in their respective PNS. We calculated the proportion of immune cells in patients with three types of by CIBERSORT concomitantly, and found that the proportion of T-cells that were

CD4 naive in patients with specific pulmonary fibrosis and rheumatoid arthritis was significantly different, while the number of T-cells gamma delta in patients with systemic lupus erythematosus was significantly different.

In summary, the decreased proportion of T-cells and the lack of hub gene expression are the common links between thymomas and specific pulmonary fibrosis, rheumatoid arthritis, and systemic lupus erythematosus. The lack of expression of different core genes may be the reason why different patients tend to have different PNS.

Obviously, CD247 was at a low expression level in all three PNS (**Figure 7D**), and the Kaplan–Meier survival curves of CD247 in TCGA data ($p = 0.0056$) were given in (**Figure 7E**). The protein encoded by CD247 (also called CD3 ζ) is T-cell receptor zeta. The zeta chain plays an important role in coupling antigen recognition to several intracellular signal-transduction pathways. Low expression of the antigen results in impaired immune response (Call et al., 2006). Moreover, this gene plays an important role in intrathymic T-cell differentiation, and its lack of expression may lead to the reduction of mature T-cells. Petros Christopoulos et al. (2015) proposed a novel thymoma-associated immunodeficiency in 2015. Its characteristics are an accumulation of CD247-deficient, hyporesponsive naive $\gamma\delta$ and $\alpha\beta$ T-cells and an increased susceptibility to infections (Christopoulos et al., 2015). In 2018, his team further suggested that deficient CD247 expression was a typical histopathological characteristic of thymomas with cortical features (Christopoulos et al., 2018). Recent evidence has demonstrated that CD247 is a potential T-cell-derived disease severity and prognostic biomarker in patients with idiopathic pulmonary fibrosis (Li et al., 2021). Some studies have shown that CD3 ζ plays a vital role in multiple autoimmune diseases, such as the gene expression assays showing that CD3 ζ mRNA levels are downregulated in PBMCs of patients with RA when compared to healthy controls (Li et al., 2016). Moreover, available evidence suggests that SLE is associated with a deficiency in a cluster of differentiation 247 (Takeuchi and Suzuki, 2013). CD247 is shared by various autoimmune disorders and supports a common T-cell-mediated mechanism. A classic T-cell phenotype in SLE is the downregulation and replacement of the CD3 ζ chain that alters T-cell receptor signaling (Martins et al., 2015).

In summary, the common characteristics of thymomas and these three PNS are the low expression of CD247 and the inhibition of T-cell differentiation.

REFERENCES

- Blum, T. G., Misch, D., Kollmeier, J., Thiel, S., and Bauer, T. T. (2020). Autoimmune Disorders and Paraneoplastic Syndromes in Thymoma. *J. Thorac. Dis.* 12 (12), 7571–7590. doi:10.21037/jtd-2019-thym-10
- Call, M. E., Schnell, J. R., Xu, C., Lutz, R. A., Chou, J. J., and Wucherpfennig, K. W. (2006). The Structure of the $\zeta\zeta$ Transmembrane Dimer Reveals Features Essential for its Assembly with the T Cell Receptor. *Cell* 127 (2), 355–368. doi:10.1016/j.cell.2006.08.044
- Christopoulos, P., Chung, I., Bozorgmehr, F., Muley, T., Meister, M., Kobinger, S., et al. (2018). Deficient CD247 Expression Is a Typical Histopathological

CONCLUSION

In this study, 14 hub immune genes closely related to thymomas, jointly involved in the T-cell receptor signaling pathway, were found by analysis of the expression data of immune genes in thymomas. Compared with normal thymus and type B1/ B2 thymoma, the number of T-cells in type A/B3 thymoma and thymic carcinoma is less, and the expression of genes related to T-cell receptor signaling pathway is low. Then, we also discussed the expression of these 14 core genes in three PNS, and found that CD247 has not only minimal expression in multiple subtypes of thymomas, but also low expression in specific pulmonary fibrosis, rheumatoid arthritis, and systemic lupus erythematosus. The low expression of CD247 and the decrease in the number of mature T-cells are the common features of thymomas and these three PNS.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. These data can be found here: TCGA-THYM <https://portal.gdc.cancer.gov/GSE177522>, GSE57892, GSE29695, GSE93272, GSE138458, and GSE33566 <https://www.ncbi.nlm.nih.gov/geo/>.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor to this work and has approved it for publication.

ACKNOWLEDGMENTS

I thank Prof. Yong-Cong Chen for his critical reading and suggestions on various parts of the work. I thank Dr. Zhang He-Yao for his support and encouragement. I am grateful to Dr Shi Chun-Xiao, Dr Cheng Run-Tan, and Dr Yao Meng-Chao for valuable discussions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.895587/full#supplementary-material>

Characteristic of Thymomas with Cortical Features. *Histopathology* 73 (6), 1040–1043. doi:10.1111/his.13724

Christopoulos, P., Dopfer, E. P., Malkovsky, M., Esser, P. R., Schaefer, H.-E., Marx, A., et al. (2015). A Novel Thymoma-Associated Immunodeficiency with Increased Naive T Cells and Reduced CD247 Expression. *J. Immunol.* 194 (7), 3045–3053. doi:10.4049/jimmunol.1402805

Colaprico, A., Silva, T. C., Olsen, C., Garofano, L., Cava, C., Garolini, D., et al. (2016). TCGAAbiolinks: an R/Bioconductor Package for Integrative Analysis of TCGA Data. *Nucleic Acids Res.* 44 (8), e71. doi:10.1093/nar/gkv1507

Ettinger, D. S., Riely, G. J., Akerley, W., Borghaei, H., Chang, A. C., Cheney, R. T., et al. (2013). Thymomas and Thymic Carcinomas: Clinical Practice Guidelines

- in Oncology. *J. Natl. Compr. Canc Netw.* 11 (5), 562–576. doi:10.6004/jncn.2013.0072
- Hashimoto, S., Hayasaka, K., Suzuki, K., Endoh, M., Yanagawa, N., and Shiono, S. (2020). Thymic Small Cell Carcinoma Associated with Lambert-Eaton Myasthenic Syndrome. *Ann. Thorac. Surg.* 109 (5), e347–e348. doi:10.1016/j.athoracsur.2019.08.080
- Jeffrey Yang, C.-F., HurdYang, J., Shah, S. A., Liou, D., Wang, H., Backhus, L. M., et al. (2020). A National Analysis of Open versus Minimally Invasive Thymectomy for Stage I to III Thymoma. *J. Thorac. Cardiovasc. Surg.* 160 (2), 555–e15. doi:10.1016/j.jtcvs.2019.11.114
- Langfelder, P., and Horvath, S. (2012). Fast R Functions for Robust Correlations and Hierarchical Clustering. *J. Stat. Softw.* 46 (11), i11. doi:10.18637/jss.v046.i11
- Langfelder, P., and Horvath, S. (2008). WGCNA: an R Package for Weighted Correlation Network Analysis. *BMC Bioinforma.* 9, 559. doi:10.1186/1471-2105-9-559
- Li, P., Wang, X., Zhao, M.-Q., Li, L.-J., Zhang, C., Li, B.-Z., et al. (2016). TCR-CD3 ζ Gene Polymorphisms and Expression Profile in Rheumatoid Arthritis. *Autoimmunity* 49 (7), 466–471. doi:10.1080/08916934.2016.1174855
- Li, Y., Chen, S., Li, X., Wang, X., Li, H., Ning, S., et al. (2021). CD247, a Potential T Cell-Derived Disease Severity and Prognostic Biomarker in Patients with Idiopathic Pulmonary Fibrosis. *Front. Immunol.* 12, 762594. doi:10.3389/fimmu.2021.762594
- Liao, C.-H., Lyu, S.-Y., Chen, H.-C., Chang, D.-M., and Lu, C.-C. (2021). Thymoma-Related Paraneoplastic Syndrome Mimicking Reactive Arthritis. *Medicina* 57 (9), 932. doi:10.3390/medicina57090932
- Martins, M., Williams, A. H., Comeau, M., Marion, M., Ziegler, J. T., Freedman, B. I., et al. (2015). Genetic Association of CD247 (CD3 ζ) with SLE in a Large-Scale Multiethnic Study. *Genes Immun.* 16 (2), 142–150. doi:10.1038/gene.2014.73
- Marx, A., Chan, J. K. C., Coindre, J.-M., Detterbeck, F., Girard, N., Harris, N. L., et al. (2015). The 2015 World Health Organization Classification of Tumors of the Thymus: Continuity and Changes. *J. Thorac. Oncol.* 10 (10), 1383–1395. doi:10.1097/JTO.0000000000000654
- Masaoka, A. (2010). Staging System of Thymoma. *J. Thorac. Oncol.* 5 (10 Suppl. 4), S304–S312. doi:10.1097/JTO.0b013e3181f20c05
- Mounir, M., Lucchetta, M., Silva, T. C., Olsen, C., Bontempi, G., Chen, X., et al. (2019). New Functionalities in the TCGAAbiolinks Package for the Study and Integration of Cancer Data from GDC and GTEx. *PLoS Comput. Biol.* 15 (3), e1006701. doi:10.1371/journal.pcbi.1006701
- Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust Enumeration of Cell Subsets from Tissue Expression Profiles. *Nat. Methods* 12 (5), 453–457. doi:10.1038/nmeth.3337
- Niemira, M., Collin, F., Szalkowska, A., Bielska, A., Chwialkowska, K., Reszec, J., et al. (2019). Molecular Signature of Subtypes of Non-small-cell Lung Cancer by Large-Scale Transcriptional Profiling: Identification of Key Modules and Genes by Weighted Gene Co-expression Network Analysis (WGCNA). *Cancers* 12 (1), 37. doi:10.3390/cancers12010037
- Rajan, A., and Zhao, C. (2019). Deciphering the Biology of Thymic Epithelial Tumors. *Mediastinum* 3, 36. doi:10.21037/med.2019.08.03
- Ried, M., Eicher, M.-M., Neu, R., Kraus, D., Inderhees, S., Marx, A., et al. (2018). Vergleich der Masaoka-Koga-Klassifikation mit der neuen TNM-Klassifikation der IASLC/ITMIG für Thymome und Thymuskarzinome. *Zentralbl Chir.* 143 (S 01), S44–S50. doi:10.1055/a-0606-5603
- Rubinstein, M. M., Goss, C., Avencia, S., Dubé, G. P., Riely, G. J., and Mones, J. V. (2019). Management of Thymoma-associated Pure Red Cell Aplasia: A Novel Use of Blood Substitute HBOC-201 in a Jehovah's Witness. *Clin. Case Rep.* 8 (2), 289–292. doi:10.1002/ccr3.2626
- Scorsetti, M., Leo, F., Trama, A., D'Angelillo, R., Serpico, D., Macerelli, M., et al. (2016). Thymoma and Thymic Carcinomas. *Crit. Rev. Oncology/Hematology* 99, 332–350. doi:10.1016/j.critrevonc.2016.01.012
- Sideris, A., and Huang, J. (2020). Commentary: Toward a Better Understanding of the Natural History of Surgically Treated Thymoma Associated with Paraneoplastic Syndromes. *J. Thorac. Cardiovasc. Surg.* 160 (1), 316–317. doi:10.1016/j.jtcvs.2019.12.059
- Silva, T. C., Colaprico, A., Olsen, C., D'Angelo, F., Bontempi, G., Ceccarelli, M., et al. (2016). TCGA Workflow: Analyze Cancer Genomics and Epigenomics Data Using Bioconductor Packages. *F1000Res* 5, 1542. doi:10.12688/f1000research.8923.2
- Takeuchi, T., and Suzuki, K. (2013). CD247 Variants and Single-Nucleotide Polymorphisms Observed in Systemic Lupus Erythematosus Patients. *Rheumatology* 52 (9), 1551–1555. doi:10.1093/rheumatology/ket119
- Thorsson, V., Gibbs, D. L., Brown, S., Wolf, D., Bortone, D. S., Ou Yang, T. H., et al. (2018). The Immune Landscape of Cancer. *Immunity* 48 (4), 812–e14. doi:10.1016/j.immuni.2018.03.023
- Tian, F., Zhao, J., Fan, X., and Kang, Z. (2017). Weighted Gene Co-expression Network Analysis in Identification of Metastasis-Related Genes of Lung Squamous Cell Carcinoma Based on the Cancer Genome Atlas Database. *J. Thorac. Dis.* 9 (1), 42–53. doi:10.21037/jtd.2017.01.04
- Tian, W., Li, X., Tong, H., Weng, W., Yang, F., Jiang, G., et al. (2020). Surgical Effect and Prognostic Factors of Myasthenia Gravis with Thymomas. *Thorac. Cancer* 11 (5), 1288–1296. doi:10.1111/1759-7714.13396
- Travis, W. D., Brambilla, E., Burke, A. P., Marx, A., and Nicholson, A. G. (2015). Introduction to the 2015 World Health Organization Classification of Tumors of the Lung, Pleura, Thymus, and Heart. *J. Thorac. Oncol.* 10 (9), 1240–1242. doi:10.1097/JTO.0000000000000663
- Wang, C. N., Li, C. Y., Cai, J.-H., Sheu, P., Tsai, J., Wu, M., et al. (2019). Identification of Prognostic Candidate Genes in Breast Cancer by Integrated Bioinformatic Analysis. *Jcm* 8 (8), 1160. doi:10.3390/jcm8081160
- Yin, L., Cai, Z., Zhu, B., and Xu, C. (2018). Identification of Key Pathways and Genes in the Dynamic Progression of HCC Based on WGCNA. *Genes* 9 (2), 92. doi:10.3390/genes9020092
- Yu, L., Zhang, B. X., Du, X., Yu, Z., Yang, X. G., and Jiang, Y. X. (2020). Evaluating the Effectiveness of Chemotherapy for Thymic Epithelial Tumors Using the CD-DST Method. *Thorac. Cancer* 11 (5), 1160–1169. doi:10.1111/1759-7714.13362
- Zhao, J., Bhatnagar, V., Ding, L., Atay, S. M., David, E. A., McFadden, P. M., et al. (2020). A Systematic Review of Paraneoplastic Syndromes Associated with Thymoma: Treatment Modalities, Recurrence, and Outcomes in Resected Cases. *J. Thorac. Cardiovasc. Surg.* 160 (1), 306–314. doi:10.1016/j.jtcvs.2019.11.052
- Zhou, X.-g., Huang, X.-l., Liang, S.-y., Tang, S.-m., Wu, S.-k., Huang, T.-t., et al. (2018). Identifying miRNA and Gene Modules of Colon Cancer Associated with Pathological Stage by Weighted Gene Co-expression Network Analysis. *Ott* 11, 2815–2830. doi:10.2147/OTT.S163891
- Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., et al. (2019). Metascape Provides a Biologist-Oriented Resource for the Analysis of Systems-Level Datasets. *Nat. Commun.* 10 (1), 1523. doi:10.1038/s41467-019-09234-6

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

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

Copyright © 2022 Deng. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.