

Original Research

Comparative Proteomic and Metabolomic Analyses of Plasma Reveal the Novel Biomarker Panels for Thyroid Dysfunction

Haodong Xia^{1,2,3}, Yaohan Li², Shengzhi Liu², Haote Han¹, Chaoting Zhou⁴, Luyang Wang⁵, Qiang Jin⁴, Jingkui Tian¹, Xiangmin Tong³, Ying Wang^{3,*}, Wei Zhu^{1,*}

¹The Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital), Institute of Basic Medicine and Cancer (IBMC), Chinese Academy of Sciences, 310002 Hangzhou, Zhejiang, China

²College of Biomedical Engineering and Instrument Science, Zhejiang University, 310027 Hangzhou, Zhejiang, China

³Phase I Clinical Research Center, Zhejiang Provincial People's Hospital, 310014 Hangzhou, Zhejiang, China

⁴School of Pharmacy, Zhejiang University of Technology, 310014 Hangzhou, Zhejiang, China

⁵School of Pharmacy, Hangzhou Medical College, 310059 Hangzhou, Zhejiang, China

*Correspondence: zhuwei@ibmc.ac.cn (Wei Zhu); nancywangying@163.com (Ying Wang)

Academic Editor: François Chevalier

Submitted: 22 December 2021 Revised: 13 February 2022 Accepted: 23 February 2022 Published: 8 March 2022

Abstract

Background: Thyroid dysfunction, including hypothyroidism (THO) and hyperthyroidism (THE), commonly arise from pathological processes in the thyroid gland. The current diagnosis of thyroid dysfunction varies because of the age and sex of the patients. The aim of this study was to explore novel candidate biomarker panels for hypothyroidism and hyperthyroidism screening with mass spectrometry and bioinformatics. **Methods:** Plasma samples were collected from 15 THE patients, 9 THO patients, and 15 healthy controls. Data Independent Acquisition (DIA)-based proteomic and untargeted metabolomic analyses were performed to identify novel biomarker panels for THO and THE patients. Finally, three candidate biomarkers were verified by ELISA in 34 samples. **Results:** A total of 2738 proteins and 6103 metabolites were identified, and 173 proteins and 2487 metabolites were found to be differentially expressed among the THE, THO and control groups. The results of the ensemble feature selection, K-means clustering and least absolute shrinkage and selection operator (LASSO) regression model showed that two proteins (C4-A and C3/C5 convertase) combined with two metabolites (L-arginine and L-proline), and proteins (APOL1 and ITIH4) combined with metabolites (cortisol, and cortisone) identified by plasma proteomics and metabolomics could help distinguish THO and THE patients from healthy controls, respectively. **Conclusions:** This study identified and verified two pairs of biomarker panels that can be used to distinguish THE and THO patients regardless of age and sex. Consequently, our findings represent a comprehensive analysis of thyroid dysfunction plasma, which is significant for clinical diagnosis.

Keywords: hypothyroidism; hyperthyroidism; plasma proteomics; plasma metabolomics; biomarkers

1. Introduction

Thyroid hormones are essential for growth, development, and energy metabolism [1]. Thyroid dysfunction, including hyperthyroidism (THE) and hypothyroidism (THO), is a global high-risk disease that seriously affects human health [2]. For patients with thyroid dysfunction, if not treated in time, it can lead to serious and even life-threatening complications: diabetes and cardiovascular disease [3]. In addition, women with thyroid dysfunction during pregnancy might have a high incidence of miscarriage, placental abruption, preeclampsia, premature delivery, and decreased intelligence in their offspring [4]. The current diagnosis of THE and THO is mainly dependent on the levels of thyroid-stimulating hormone, but the reference range varies according to patients' age and sex [5,6]. Therefore, it is necessary to identify the potential biomarkers for the diagnosis of thyroid dysfunction.

Multimomics techniques, such as proteomics and metabolomics, can be a powerful tool to discover biomarkers related to thyroid disease at the protein and metabolic

levels [7,8]. Recently, the plasma proteomics analysis of THE and euthyroid groups showed that 20 differentially abundant proteins related to the NF- κ B and MAPK pathways were identified, which was served as clinical markers for the early detection of side effects in patients [9]. Another comparison of plasma proteomics study between THO and euthyroid states revealed the changes in circulating protein levels to characterize changes in thyroid hormone status [10]. Nuclear magnetic resonance-based metabolomics analysis of serum determined the metabolic changes in hypothyroid patients before and after levothyroxine treatment, which contributed to integrate the hormone assays and the diagnosis of euthyroid status [11]. And the metabolomic study of hyperthyroidism patients before and after antithyroid drug treatment showed that their metabolomic characteristics were deeply affected by thyroid hormone levels, and this change persisted after its normalization [12]. Furthermore, serum metabolomics analyses of patients with autoimmune thyroid disease showed that the significantly changed metabolites, including 22



metabolites from hyperthyroidism and 17 metabolites from hypothyroidism were involved in amino acid metabolism and aminoacyl-transfer ribonucleic acid biosynthesis [13]. However, unlike single omics techniques, multiomics could provide a more valuable reference for disease prediction and diagnosis at multidimensional levels [14].

In the present study, to explore the potential markers on thyroid dysfunction for clinical diagnosis and treatment, the integration of metabolomics and proteomics analyses of plasma samples among the normal, THE, and THO groups were performed. Perform Receiver Operating Characteristic (ROC) analysis on the identified and quantified proteins and metabolites, and combine the expression of proteins and metabolites in the sample group to find potential biomarkers. This was followed by an integrated analysis of proteomic and metabolomic correlations and LASSO analysis in order to explore the performance of the combined markers. Finally, independent samples were used to validate biomarkers using ELISA kits. This will provide more reference information for early diagnosis strategies of THE and THO. These results will inform strategies for the early diagnosis of THE and THO.

2. Materials and Methods

2.1 Patients and Samples

Plasma samples were obtained from 39 participants in Zhejiang Provincial People's Hospital. These 39 participants included 15 patients with hyperthyroid (THE), 9 patients with hypothyroid (THO), and 15 healthy controls (N), the thyroid hormone levels of patients were shown in **Supplementary Table 1**. All the participants were female and signed written informed consent forms. This study was approved by the ethics committee for clinical studies of Zhejiang Provincial People's Hospital (No.2021QT355). All pathological diagnoses of hyperthyroidism and hypothyroidism were confirmed by experts based on clinical examination standards. Plasma was collected from the peripheral vein from the participants after overnight fasting. EDTA blood samples were centrifuged at $1000 \times g$ for 10 min within four hours after collection, and the supernatant was collected and stored at -80°C until further analysis (**Supplementary Fig. 1**).

2.2 Proteomic Analysis of Plasma Specimens

Protein extraction and digestion from plasma samples were performed as previously described with minor modification [15]. SDS free lysate was added to 100 μL of plasma samples, making up a total volume of 1 mL. The proteins were reduced with 10 mM dithiothreitol (DTT) for 30 minutes at 37°C , then alkylated with 55 mM iodoacetamide (IAA) in the dark for 30 minutes at room temperature. Protein enrichment was performed using a solid phase extraction (SPE) C18 column. The eluate was collected and freeze-dried. The dried proteins were resolved in 20 μL of 50 mM ammonium bicarbonate and quanti-

fied by Pierce Quantitative Fluorometric Peptide Assay. Trypsin was added to the protein solution with an enzyme-to-substrate ratio of 1:20. The mixture was incubated at 37°C for 14–16 h.

Then, the mixed peptides were analyzed by LC-MS/MS in data-dependent acquisition (DDA) mode for library construction. For data-independent acquisition (DIA) analysis, each digested peptide sample was ionized by a nanoESI source and injected into a tandem mass spectrometer Q-Exactive HF X (Thermo Fisher Scientific, USA) in DIA detection mode (the specific steps of protein extraction, digestion and the settings of high-performance liquid chromatography and tandem mass spectrometry are detailed in the supplementary methods). For library construction, the DDA data were analyzed by MaxQuant [16] (version 1.5.3.30) and matched against the UniProtKB database (homo sapiens with 172,419 entries, downloaded 2020.07.20). Oxidation of methionine was set as a variable modification and carbamidomethylation of cysteine was set as a fixed modification. The mass spectra library was constructed using Spectronaut with FDR $< 1\%$. Differentially abundance proteins (DAPs) were analyzed by MSstats with linear mixed-effects models [17], and significant DAPs were considered with the conditions of fold change > 2 and p value < 0.05 . STRING (<http://string-db.org/>, version: 11.0) software was used to analyze functional protein association networks [18].

2.3 Untargeted Metabolomics Assay

After thawing the plasma sample, the extract and internal standard were added, and the metabolites were extracted after mixing and centrifugation. Chromatographic separation was carried out using an ultra-performance liquid chromatography system (Waters, Milford, MA, USA). The sample was injected on a Waters BEH C18 column. Mass data acquisition was performed using a Thermo Q-Exactive (Thermo Fisher Scientific, San Jose, CA, USA) equipped with an electrospray ionization source. Detected metabolites were identified using multiple databases, including the BGI library, mzCloud database, Chemspider database, HMDB, KEGG database, and Lipidmaps database (The preparation of mass spectrometry samples and the setting of mass spectrometry conditions are detailed in the supplementary methods.). Compound Discoverer 3.1 software (San Jose, CA, USA) from Thermo Fisher Scientific was used for data processing as previously described [15].

2.4 Statistical and Bioinformatic Analyses

The NAGuideR package was used to interpolate the missing values. Ensemble feature selection (EFS) analysis was performed to reduce the biases of any individual feature selection. Statistical significance was evaluated by Student's t test when only two groups were compared and one-way ANOVA with Tukey's test when multiple groups were compared. K-means clustering, DAVID,

and Cytoscape analyses were performed using R v3.6.1. The optimal clustering number was determined by elbow method and “ggplot2” R package was used for data visualization. Metabolic pathway enrichment analysis was performed based on the KEGG database through MetaboAnalyst. LASSO regression algorithm was used to select the minimum set of features to classify the sample groups. Fast missing value imputation is achieved by the chained random forests of the R software package in each separate dataset. The caret and “glmnet” R software package were used for regular machine learning to train, test, and evaluate the LASSO logical classification model. The Pearson correlation coefficient was calculated between the expression profiles of proteins and metabolites to identify potential biomarkers. The protein-metabolite coexpression network was drawn with the Pearson correlation coefficient ± 0.7 as the boundary.

3. Results

3.1 Demographic Characteristics of the Study Population

We used a total of 24 plasma samples from patients with abnormal thyroid function and 15 healthy control plasma samples matched for age, sex, and the levels of thyroid hormone (**Supplementary Table 1**). 15 of the 24 patients were hyperthyroidism, and 9 were hypothyroidism.

3.2 Global Proteomic Profiling of Hyperthyroid and Hypothyroid

In DDA mode, a total of 20,322 peptides and 2738 proteins were identified for the library construction. Among them, 2622 proteins contained at least one unique peptide, accounting for 95.76% of the total. A total of 2171 proteins with a sequence coverage of at least 20% accounted for 79.29%. In DIA mode, a total of 757 proteins with at least one unique peptide and a false discovery rate (FDR) $< 1\%$ were identified in the plasma samples from all 39 samples. Principal component analysis of protein quantitative values from each sample showed a clear distinction between THE and THO groups (**Supplementary Fig. 2A**). To deepen the comprehensive molecular features related to thyroid dysfunction, the expanded cohort was analyzed through standard proteomic workflows as well as metabolomic approaches and the potential biomarkers were validated by the ELSIA approach. Since the initial proteome matrix displayed obvious missing values, the NAGuideR R-package was used to interpolate the missing values, and proteins with more than half of missing values among three groups were removed. Finally, 173 differentially abundance proteins (DAPs) obtained by one-way ANOVA analysis ($p < 0.05$) were used for the subsequent analysis.

3.2.1 Definition of the Potential Protein Markers between the Thyroid Dysfunction and Healthy Groups

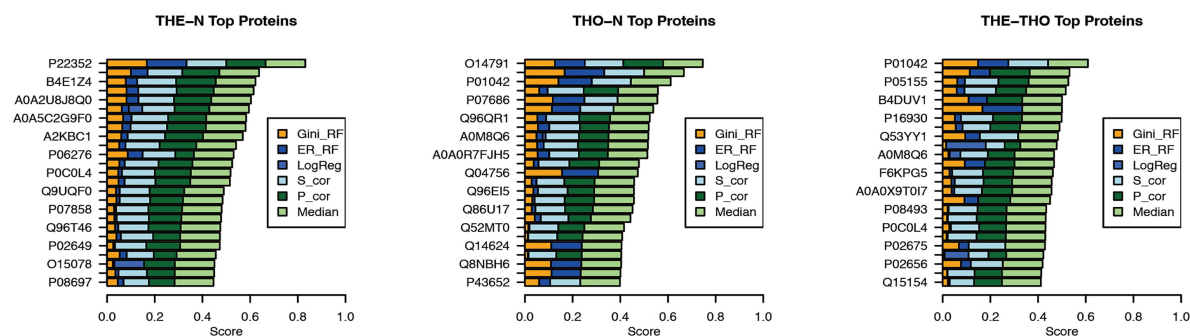
To identify reliable protein markers in the plasma among the THO, THE, and N groups, the ensemble feature

selection (EFS) approach and student's t test were used to aggregate and rank the results [19] (Fig. 1A). The top 13 EFS proteins among THE, THO, and N groups are listed in Fig. 1A. Importantly, the biomarker ranks by EFS and t tests were concordant among the three groups (Fig. 1B). The highest ranked protein markers of the participants in THE/N and THO/N groups were glutathione peroxidase 3 and apolipoprotein L1 (proteins Uniprot ID: P22352 and O14791), respectively. The EFS approach ensures that the top-ranked biomarkers are not correlated with one another [20]. According to the average results of the EFS ranking and the t test, the top-ranked proteins were used for the further definition of the potential protein markers. ROC analyses were performed on these proteins, and 16 proteins with AUC values greater than 0.7 were recognized as the potential markers (**Supplementary Table 2**). The relative quantitative comparison and ROC analyses of 16 proteins are shown in **Supplementary Figs. 3,4**. Among those 16 proteins, four proteins, including complement C4-A, complement C3/C5 convertase, apolipoprotein L1 (APOL1), and interalpha-trypsin inhibitor heavy chain H4 (ITIH4), showed the best distinguishing effect between the thyroid dysfunction groups (THE or THO) and the normal group. For example, the relative abundance of complement C4-A and C3/C5 convertase in the THE group showed an increasing and decreasing trend compared to the N group (Fig. 1C), while the relative abundance of apolipoprotein L1 and interalpha-trypsin inhibitor heavy chain H4 in the THO group were significantly increased compared to the THE and N groups (Fig. 1D). ROC analysis also showed high AUC values (more than 0.8) of four proteins (Fig. 1C,D).

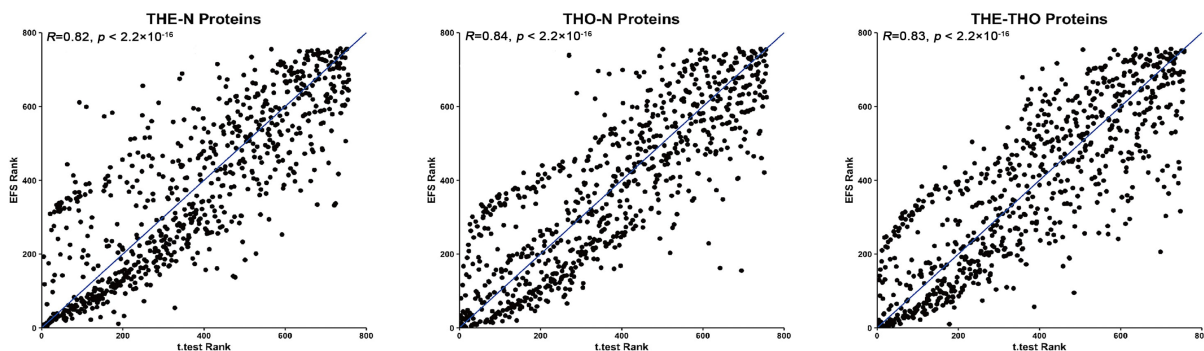
3.2.2 Unbiased Clustering of Thyroid Dysfunction Module Analysis

Beyond defining biomarkers to predict thyroid dysfunction, we performed k-means clustering and elbow clustering using the proteomic dataset. It was grouped into 5 clusters (Fig. 2A,B), revealing expression profiles of interest. Clusters 1 and 3 showed decreases and increases in the THO group, which was largely different from the results in the THE and N groups. Clusters 2 and 4 were largely different in the thyroid dysfunction groups (THE and THO groups) compared to the normal group. To examine the crosstalk of proteins between clusters, protein-protein interaction networks of DAPs among the three groups ($p < 0.05$) were obtained using STRING based on conserved genomic neighborhoods, gene fusion, coexpression and cooccurrence of genes across genomes, known metabolic pathways from databases and experimental events. The confidence score was set at a high level (> 0.7). Then, the output table from STRING was visualized by Cytoscape (Fig. 2C) and R for connections among different clusters, indicating that Cluster 2 was strongly connected with Cluster 5 (Fig. 2D). To define the functional roles, GO enrichment analysis of proteins in each cluster was performed

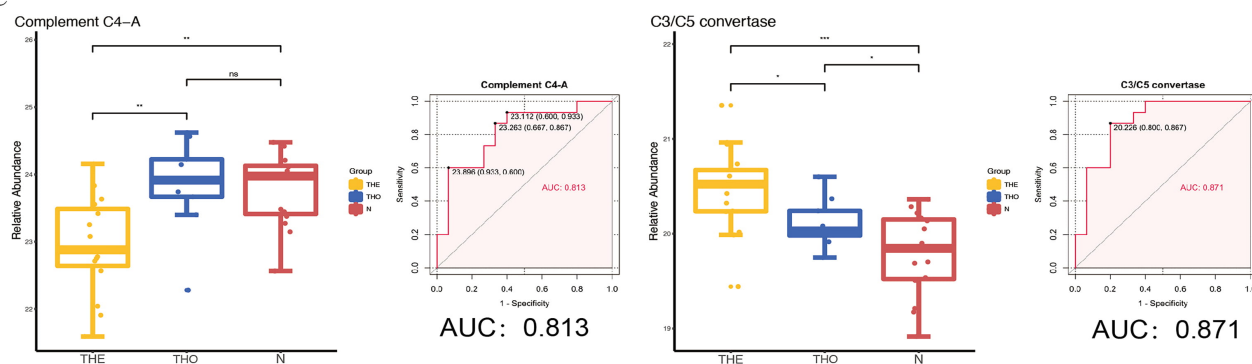
A



B



C



D

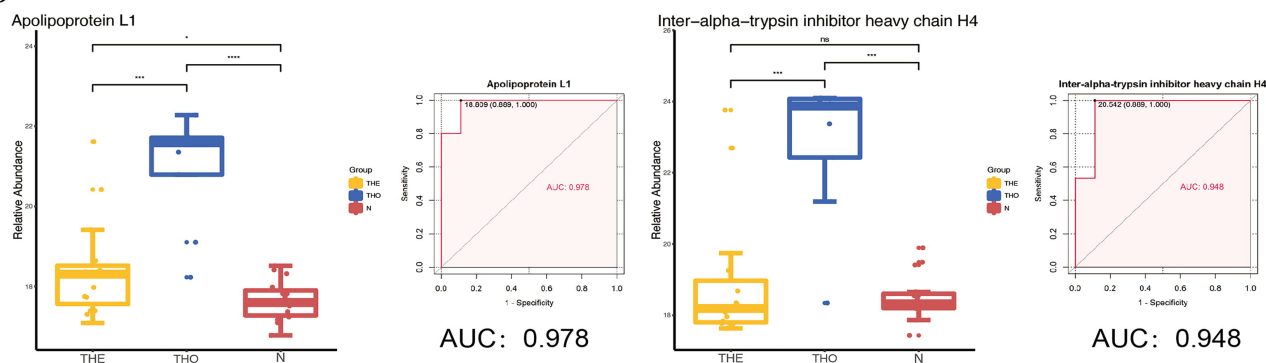


Fig. 1. Ensemble feature selection analyses of proteomics data and ROC analyses of potential biomarkers among N, THE, and THO groups. (A) EFS analysis for Group THE/N, THO/N and THE/THO. (B) Comparison of EFS and *t*-test statistical analysis of THE/N, THO/N and THE/THO. (C) The relative abundance and ROC analyses of the potential protein biomarkers between THE and N groups. (D) The relative abundance and ROC analyses of the potential protein biomarkers between THO and N groups.

(Fig. 2E). The results showed that Cluster 1 and Cluster 3 were commonly dominated by extracellular exosomes and space. Cluster 3 was dominated by extracellular regions, blood microparticles, and serine-type endopeptidase inhibitor activity. The top three enriched GO functions in Clusters 2 and 5 were the same: extracellular space, region, and exosome. For Cluster 4, the top three enriched GO functions were hydrogen peroxide catabolic process, hemoglobin complex, and hemoglobin binding, suggesting a potential difference between thyroid dysfunction and normal (Fig. 2E).

3.3 Metabolomics Analysis of the Thyroid Dysfunction and Healthy Groups

To identify the differential plasma metabolite profiles related to thyroid dysfunction, we performed metabolomic analysis using UPLC Orbitrap/MS. A total of 6311 features including 4435 in positive ion mode and 1876 in negative ion mode were detected from 39 plasma samples. Among them, 3121 metabolites were identified. Based on one-way ANOVA analysis, 2487 metabolites with $p < 0.05$ were recognized as the significantly changed metabolites (SCMs) among the THE, THO, and N groups. Principal component analysis of metabolites from each sample showed a clear distinction between the THE and THO groups (**Supplementary Fig. 2B**). To identify potential metabolic biomarkers in hyperthyroidism and hypothyroidism patients, the EFS approach was used to further rank the results and applied to the three primary datasets to rank the top biomarkers (Fig. 3A). In addition, Student's t test was also performed to rank the variables between two groups, which was concordant with the EFS results (Fig. 3B). In the hyperthyroid group, the contents of L-arginine and L-proline were significantly different from those in the other two groups. Moreover, the contents of cortisone and cortisol in the hypothyroid group were also significantly different from those in the other two groups. Subsequently, the top metabolites from the average results of EFS ranking and t test were used for ROC analyses, and 16 metabolites with AUC values greater than 0.7 were selected as the potential metabolic biomarkers (**Supplementary Table 3**). Among them, L-arginine and L-proline showed a significant increase in the THE group compared to the N group (Fig. 3C). Cortisol and cortisone were significantly increased in the THO group compared to the THE and N groups (Fig. 3D). ROC analyses of these four metabolites showed high AUC values, suggesting that they might be potential metabolic biomarkers (Fig. 3C,D). The remaining 12 metabolites were also analyzed by ROC (**Supplementary Figs. 5,6**).

For metabolomics analysis, the SCMs were used for k-means cluster analysis and grouped into 8 clusters (Fig. 4A). The violin chart intuitively indicates the relationship between the clusters and thyroid dysfunction. Clusters 1, 2, and 5 showed increases in the THO group compared

to the N group, while Clusters 3, 7, and 8 showed decreases compared to the N groups. Clusters 1, 4, 5, and 6 showed differences in the THE group compared to the N group, while Clusters 2, 3, 7, and 8 showed large differences in the THE group compared to the THO group (Fig. 4A). Furthermore, kinship variables were divided into 8 groups by the elbow method (Fig. 4B). To evaluate the importance of metabolic pathways affected by thyroid dysfunction, metabolic pathway enrichment analysis was performed based on the KEGG database through MetaboAnalyst (Fig. 4C). The top 25 enriched pathways from the metabolite sets are shown in Fig. 4C, of which 5 pathways, arginine biosynthesis, steroid hormone biosynthesis, butanoate metabolism, D-glutamine and glutamate metabolism, and the citrate cycle, were significantly enriched with p values less than 0.05 (Fig. 4C).

3.4 Potential Biomarkers of Thyroid Dysfunction Based on Integrated Omics Analyses

To gain a holistic view of the changes in protein and metabolite profiles between the normal and thyroid dysfunction groups, the proteomic and metabolomic data were integrated for comprehensive analysis. Based on the above analysis, proteins (complement C4-A and C3/C5 convertase) (Fig. 1C) and metabolites (L-arginine and L-proline) (Fig. 3C) were the largest difference biomarker between THE and N groups. While proteins (apolipoprotein L1 and interalpha-trypsin inhibitor heavy chain H4) (Fig. 1D) and metabolites (cortisol and cortisone) (Fig. 3D) were the largest difference biomarkers between THO and N groups. They were selected for the dual-omic logistic regression model analysis (Fig. 1C,D and Fig. 3C,D). The results showed that the mean AUC of the dual-omic model was 0.978 between the THE and N groups and 0.963 between the THO and N groups, which were significantly higher than those of the individual biomarkers (Fig. 5A). Furthermore, correlation analysis of proteins and metabolites in the plasma landscape was performed. The links with a Pearson correlation coefficient >0.5 were exported and made into a network colored with K-means clustering results in Cytoscape (Fig. 5B). The results indicated that GPX3 and C4-A in protein Cluster 5 were surrounded by features in metabolite Cluster 3 and Cluster 6 between the THE and N groups. APOL1 and KNG1 in protein Cluster 2 were surrounded by features in metabolite Cluster 1 between the THO and N groups (Fig. 5B).

In addition, least absolute shrinkage and selection operator (LASSO) regression was used to validate the reliability of the combined biomarkers. All the identified proteins and metabolites from the THE, THO, and N groups were merged for LASSO and logistic regression algorithm analyses, and then the minimum features required for classification were obtained. For participants in the THE-N groups, a set of 126 features (**Supplementary Table 4**) was screened through the “varImp” function in the caret R package, in-

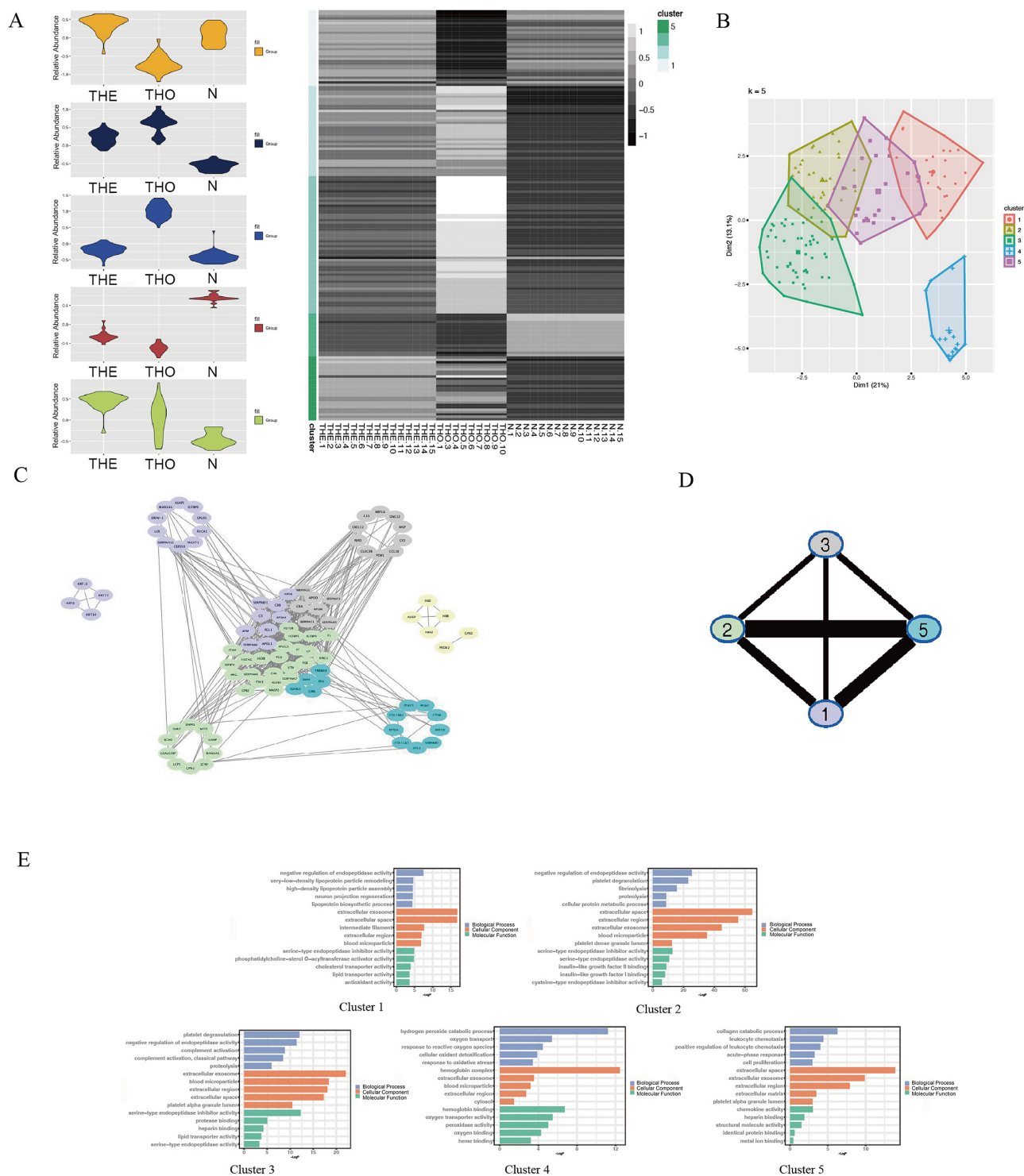
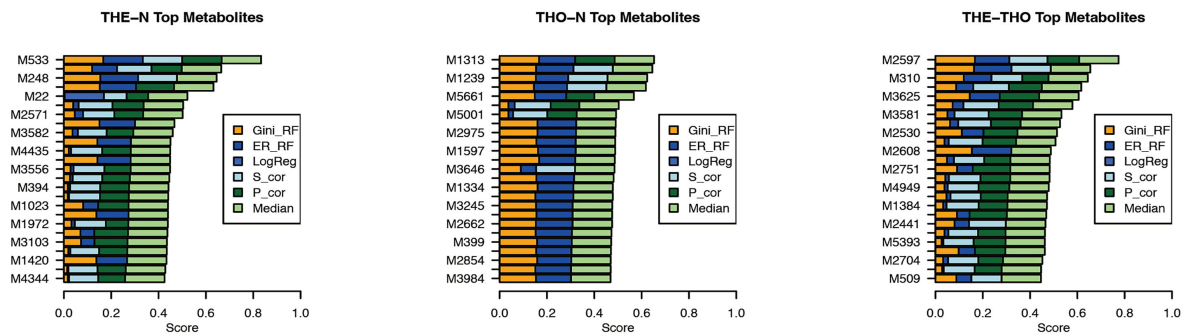


Fig. 2. Cluster-based modular analysis of differential proteins in thyroid dysfunction. (A) K-means clustered heatmap and violin diagram. (B) Elbow clustering of identified proteins. (C) Protein-protein interaction network. (D) Statistics of the number of protein connections among clusters. (E) GO enrichment for each cluster.

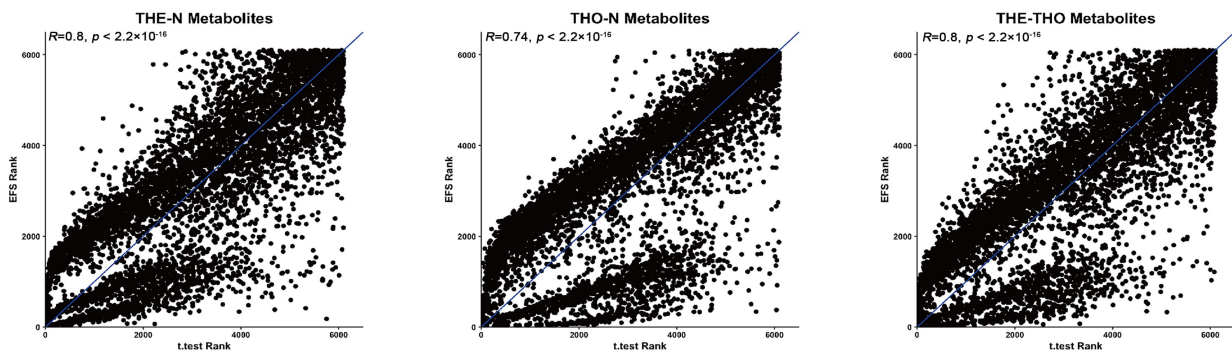
cluding 8 proteins and 118 metabolites and 238 features for participants in the THO-N groups (**Supplementary Table 5**). From the heatmap, these features can obviously group the samples (**Supplementary Fig. 7**). Due to the small size of the sample queue, we selected the first six important

variables to build a logistic regression model through the Sklearn package in Python (version 3.7.3, Python Software Foundation, <http://www.python.org>). LOOCV (leave-one-out-cross-validation) was used to enhance the robustness of the model. As we expected, the model with an average ROC

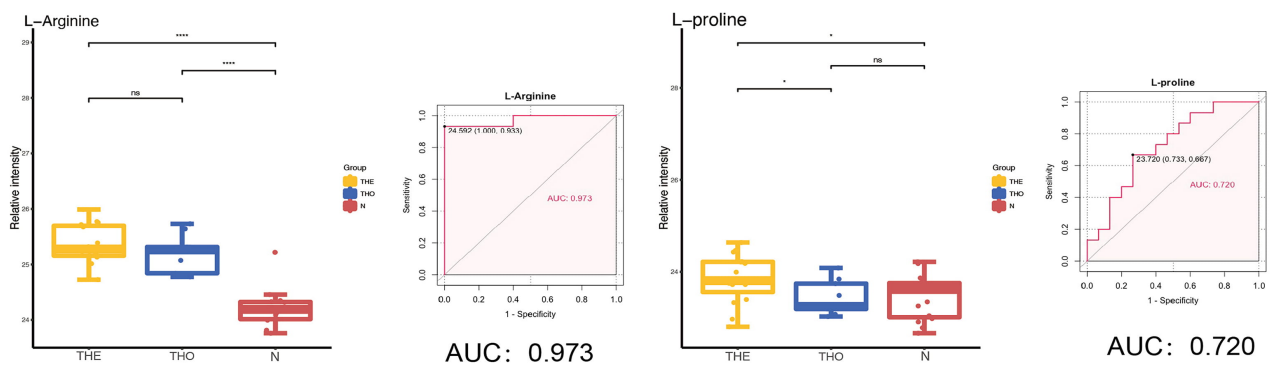
A



B



C



D

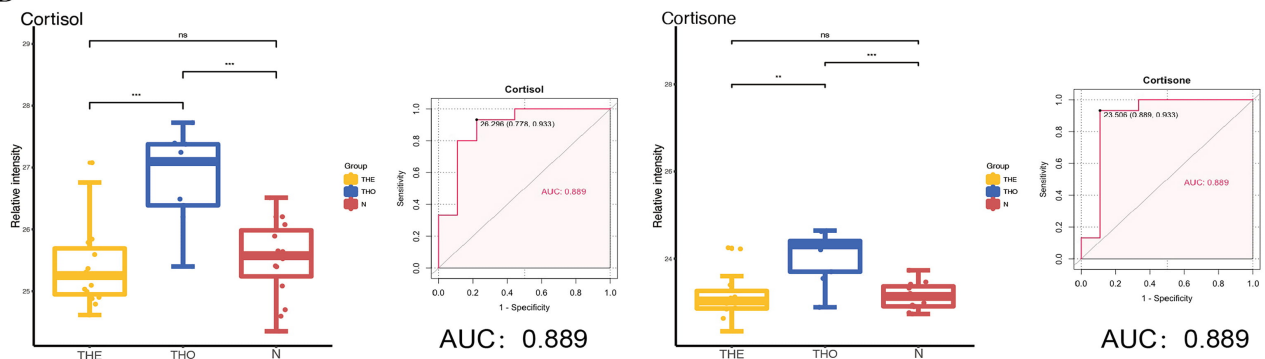


Fig. 3. Ensemble feature selection analyses of metabolomics data and ROC analyses of potential metabolite biomarkers among N, THE, and THO groups. (A) EFS analysis for Group THE/N, THO/N and THE/THO. (B) Comparison of EFS and *t*-test statistical analysis of THE/N, THO/N and THE/THO. (C) The relative abundance and ROC analyses of the potential metabolite biomarkers between THE and N groups. (D) The relative abundance and ROC analyses of the potential metabolite biomarkers between THO and N groups.

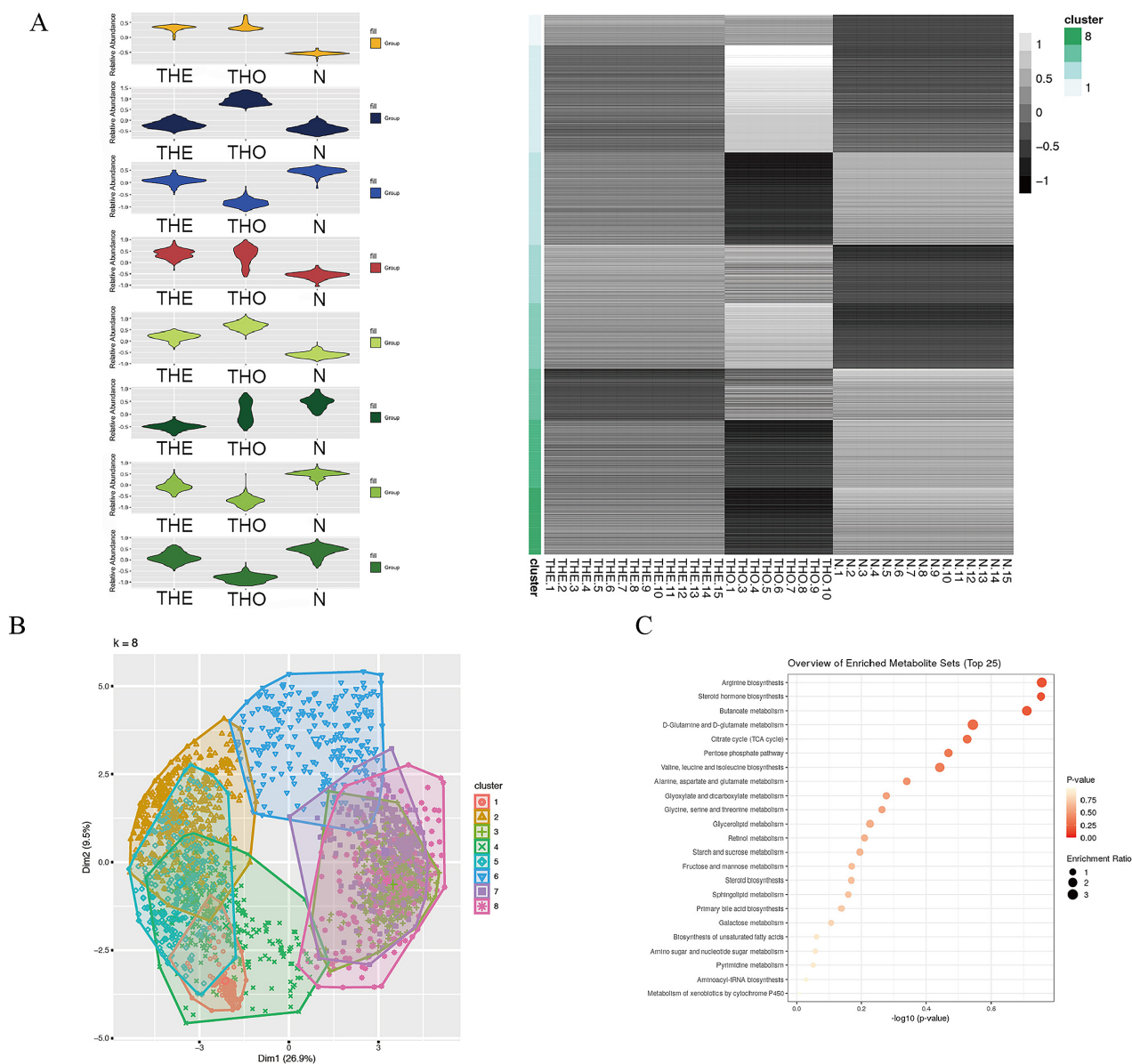


Fig. 4. Cluster-based modular analysis of differential metabolites in thyroid dysfunction. (A) K-means clustered heatmap and violin diagram. (B) Metabolites elbow clustering. (C) Dot plot of the top 25 enriched metabolite sets using MSEA analysis based on KEGG database.

AUC of 0.991 and 0.993 was more effective than the former combined biomarkers (Fig. 5C).

3.5 Validation of Potential Biomarkers for Thyroid Dysfunction

To verify the distinguishing performance of the biomarkers obtained by omics analysis in patients with dysthyroid function and normal people, independent batch plasma samples containing 9 THO, 10 THE, and 15 normal groups were used for the analysis. The contents of three potential markers, apolipoprotein L1, complement C4-A, L-arginine and cortisol, were measured by ELISA kits. The results showed that in the control group, apolipoprotein L1 was 143.45 ng/mL, complement C4-A was 12.07 ng/mL,

and cortisol was 1.90 $\mu\text{g}/\mu\text{L}$, L-arginine was 0.4471 ng/mL; in the THE group, apolipoprotein L1 was 141.34 ng/mL, complement C4-A was 8.56 ng/mL and cortisol was 1.92 $\mu\text{g}/\mu\text{L}$, L-arginine was 0.4536 ng/mL; and in the THO group, apolipoprotein L1 was 147.76 ng/mL, complement C4-A was 12.04 ng/mL, and cortisol was 2.22 $\mu\text{g}/\mu\text{L}$, L-arginine was 0.4479 ng/mL. Apolipoprotein L1 was significantly higher in the THO group than N group, while complement C4-A was significantly lower in the THO group than N group, and cortisol was significantly higher in the THO group than N group (Fig. 6A–C). This results were consistent with the omics results. The dispersion analyses of these biomarkers were shown in **Supplementary Table 6**.

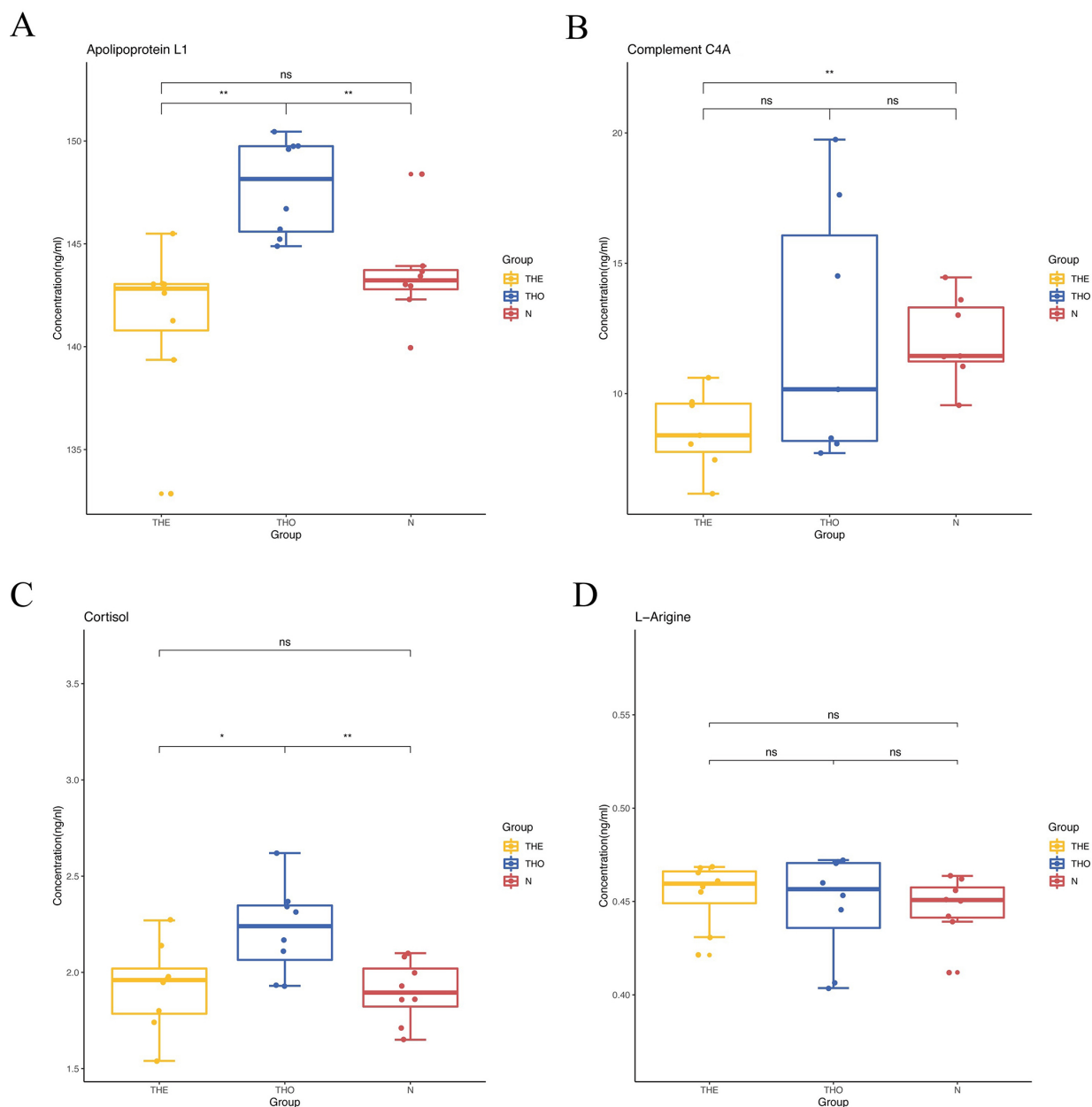


Fig. 6. Validation of the levels of potential biomarkers among N, THE, and THO groups using ELISA assays. (A) The levels of Apolipoprotein L1. (B) The levels of Complement C4-A. (C) The levels of cortisol. (D) The levels of L-arginine.

4. Discussion

Clinically, thyroid function tests, including thyroxine (T4), triiodothyronine (T3), thyroid stimulating hormone (TSH), free T3, and free T4, are general indicators for the diagnosis of thyroid dysfunction. For participants in the THE group, one of the most common causes was Graves' disease, and the other causes were toxic multinodular goiter and toxic adenoma [21]. Generally, the TSH level was used for preliminary screening; if the results were uncertain, then the method of radionuclide uptake was used for the definite diagnosis. For participants in the THO group, there are no uniform clinical diagnostic criteria because the reference level of TSH differs based on age, weight, area

and medication history. In addition, THO is the most common drug-induced thyroid dysfunction [22]. In the present study, the integrated proteomics and metabolomics analyses revealed the potential biomarkers for thyroid dysfunction.

For hyperthyroid, four molecules, complement C4-A, complement C3/C5 convertase, L-arginine and L-proline, were identified as potential biomarker panels to predict THE (Fig. 1C). Complement C4-A is a short-term fragment of complement C4 that plays an important role in the function of the lectin complement pathway [23], and is related to other autoimmune inflammation, infectious diseases and neurological diseases. Alfadda *et al.* [10] found that the

expression of complement C4-A in patients with abnormal thyroid function was increased. In addition to B lymphocytes, complement proteins can also be synthesized by thyroid cells [24], and the metabolic function of thyroid cells is impaired by complement activation. C3/C5 convertase was a protein involved in the process of complement activation that can cleave the complement component C3 into C3a and C3b and the complement component C5 into C5a and C5b [25]. Jafarzadeh *et al.* [26] found that the expression of complement C3 was increased in the patients with hypothyroidism. In this study, the expression levels of C3/C5 convertase were increased in patients with thyroid dysfunction, especially in the patients with THE which was much higher than that in the patients with hypothyroidism. It suggests that C3/C5 convertase could be a potential marker for distinguishing THE and normal groups. L-arginine is an essential amino acid in the human body, which participates in many biological processes such as the normal functions of the cardiovascular and immune systems [27]. Rodríguez-Gómez *et al.* [28] found that the abundance of arginase I was increased in the aorta, heart, and kidney of hyperthyroid rats, whereas the abundance of arginase I was decreased in the kidney and aorta of hypothyroid rats, and arginase II in hyperthyroidism increased in the aorta and kidney of rats and remained unchanged in all organs of hypothyroid rats. In the present study, the abundance of L-arginine in thyroid dysfunction patients, especially for the participants in THE group, was increased (Fig. 3C), which means that L-arginine could be one of the potential biomarkers for THE. L-Proline is the active product of OAT, with L-ornithine as the substrate, and the highest content in the kidneys. The level of L-proline is positively regulated by thyroid hormones in the liver [28]. In this study, the level of L-proline in patients with hyperthyroidism had an upward trend, which indicated a close relationship between L-proline and thyroid hormone. Additionally, the above four markers were integrated for LASSO and logistic regression analysis. The results showed that the combination of multiple markers was more effective than a single marker (Fig. 5D). The AUC value of multiple markers combined with ROC was 0.991, which was significantly higher than that of a single marker.

For the THO, the four potential biomarkers including APOL1, ITIH4, cortisone and cortisol can also significantly distinguish the THO and normal groups. APOL1 is a minor HDL3-related apolipoprotein related to lipid transport and metabolism, apoptosis, autophagic cell death, and cell lysis caused by membrane pore formation [29]. Masood *et al.* [9] found that the expression level of APOL1 was increased in the patients with thyroid dysfunction. It is known that the increased level of APOL1 is positively correlated with hyperglycemia and plasma triglycerides in patients with coronary artery disease with high-density lipoprotein (HDL), and it is a potential factor for premature cardiovascular disease [30]. Previous studies found that HDL participated

in the diffusion of thyroid hormone through the cell membrane and inner nuclear membrane [31], which indicated that the level of APOL1 may affect the diffusion of thyroid hormone. In this study, the expression level of APOL1 in patients with THO was increased significantly (Fig. 1D). It was consistent with the independent validation experiments (Fig. 6). Previous studies showed that ITIH4 was a potential diagnostic and prognostic marker for several diseases, such as acute ischemic stroke, ovarian cancer, interstitial cystitis, and liver fibrosis [32]. And its level in the plasma of thyroid cancer patients was increased compared to healthy controls [33]. Similarly, in this study, ITIH4 has a higher expression in patients with hypothyroidism. Cortisol, also known as hydrocortisone, is the adrenal cortex hormone with the strongest effect on carbohydrate metabolism. Nobumasa *et al.* [34] found that one patient with corticotropin deficiency can treat primary hypothyroidism with cortisol. In this study, the levels of cortisol and cortisone in THO group were significantly higher than the normal group, suggesting they might be the potential biomarkers of THO. Additionally, in this study, the above four markers were integrated for LASSO and logistic regression analysis. The results showed that the combination of multiple markers is more effective than a single marker (Fig. 5D). The AUC value of multiple markers combined with ROC was 0.991, which was significantly higher than that of a single marker. Furthermore, among these markers of thyroid dysfunction, two proteins and one metabolite were validated using the independent samples by ELISA assays (Fig. 6), which are consistent with the omics results.

To further explore the mechanism of thyroid dysfunction, integrated omics of clustering and network analyses were performed. K-means cluster analyses showed that proteins related to thyroid dysfunction were highly correlated with negative regulation of endopeptidase activity, hydrogen peroxide catabolic process and platelet degranulation (Fig. 2). Mousa *et al.* [35] found that platelet function was regulated by L-thyroxine (T4), while T4 induced platelet aggregation and degranulation. Therefore, in the present study, complement activation and platelet degranulation may affect the normal thyroid function by affecting the metabolism of thyroid cells. Furthermore, previous study has shown that the levels of glycine and L-serine in patients with hyperthyroidism are reduced, and the serum glutamine levels in hyperthyroidism and hypothyroidism groups were increased, while L-glutamate and L-citrulline and taurine levels were reduced [13]. Pathway analysis showed that thyroid dysfunction might be mainly related to steroid hormone, arginine, butanoate, glutamine and glutamate metabolism, and citrate cycle. Endogenous arginine is mainly derived from the conversion of citrulline in the proximal convoluted tubules of the kidney, while citrulline was synthesized from glutamate and glutamine in the intestinal tract. And the biosynthetic pathway of steroid hormones was related to sugar metabolism.

5. Conclusions

To identify the novel biomarker panels for THE and THO diagnosis, integrative proteomics and metabolomics analyses of plasma were performed. A total of 757 proteins and 2487 metabolites were identified and quantified. ROC analysis of proteins showed 16 proteins were the potential markers with an AUC value more than 0.7. Among them, four proteins were significantly different between the thyroid dysfunction and healthy groups. Similarly, four metabolites including L-arginine, L-proline, cortisol, and cortisone were recognized as the potential biomarkers. Furthermore, integration of proteomic and metabolomic correlation and LASSO analyses indicated complement C4-A, C3/C4 convertase, L-arginine, L-proline and APOL1, ITIH4, cortisol and cortisone could be the combined biomarkers for THO and THE, respectively. And the independent samples were used for the validation of these biomarkers using ELISA assays. In future, studies of the larger sample groups and strongly validated experiments will be benefit for the deeper understanding of thyroid dysfunction. However, the present results provide a potential way for the early diagnosis of THE and THO.

Data Access

The proteomics raw data, peak lists and result files have been deposited in the ProteomeXchange Consortium via the PRIDE partner repository under dataset identifier PXD029880. The metabolomics raw data files and the identified metabolites and proteins tables are freely available via scientific data repository Zenodo.org <https://doi.org/10.5281/zenodo.5722878>.

Author contributions

JT, XT, YW, WZ, HX contributed to the idea and design. HX, YL, SL contributed to methodology. HX, YL, HH contributed to the validation. HX, SL, HH contributed to the data Curation. JT, XT, YW, WZ contributed to the project administration, funding acquisition, resources. CZ, LW QJ contributed to the investigation. HX, WZ contributed to the manuscript writing and revision. HX contributed to the data analysis by software. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

All plasma samples were obtained from patients with hyperthyroid and hypothyroid in Zhejiang Provincial People's Hospital, and written informed consent was obtained from each patient. This study was approved by the ethics committee for clinical studies of Zhejiang Provincial People's Hospital, code: No.2021QT355.

Acknowledgment

Thanks to Amin Liu from Zhejiang University for helping to revise the article.

Funding

This work was supported by National Natural Science Foundation of China (81971172), Distinguished Young Scientists, Natural Science Foundation of Zhejiang Province, China (LR21H090001), The Key Research and Development Program of Zhejiang Province (WKJ-ZJ-1914).

Conflict of interest

The authors declare no conflict of interest.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at <https://www.imrpress.com/journal/FBL/27/3/10.31083/j.fbl2703090>.

References

- [1] Sinha RA, Singh BK, Yen PM. Direct effects of thyroid hormones on hepatic lipid metabolism. *Nature Reviews Endocrinology*. 2018; 14: 259–269.
- [2] Taylor PN, Albrecht D, Scholz A, Gutierrez-Buey G, Lazarus JH, Dayan CM, *et al.* Global epidemiology of hyperthyroidism and hypothyroidism. *Nature Reviews Endocrinology*. 2018; 14: 301–316.
- [3] Klein I, Danzi S. Thyroid Disease and the Heart. *Current Problems in Cardiology*. 2016; 41: 65–92.
- [4] Alemu A, Terefe B, Abebe M, Biadgo B. Thyroid hormone dysfunction during pregnancy: A review. *International Journal of Reproductive Biomedicine*. 2016; 14: 677–686.
- [5] Chaker L, Bianco AC, Jonklaas J, Peeters RP. Hypothyroidism. *Lancet*. 2017; 390: 1550–1562.
- [6] Zhu WZ, Olson A, Portman M, Ledee D. Sex impacts cardiac function and the proteome response to thyroid hormone in aged mice. *Proteome Science*. 2020; 18: 11.
- [7] Ucal Y, Ozpinar A. Proteomics in thyroid cancer and other thyroid-related diseases: A review of the literature. *Biochimica Et Biophysica Acta Proteins and Proteomics*. 2020; 1868: 140510.
- [8] Coelho M, Raposo L, Goodfellow BJ, Atzori L, Jones J, Mandadas B. The Potential of Metabolomics in the Diagnosis of Thyroid Cancer. *International Journal of Molecular Sciences*. 2020; 21: 5272.
- [9] Masood A, Benabdelkamel H, Ekhzaimy AA, Alfadda AA. Plasma-Based Proteomics Profiling of Patients with Hyperthyroidism after Antithyroid Treatment. *Molecules*. 2020; 25: 2831.
- [10] Alfadda AA, Benabdelkamel H, Masood A, Jammah AA, Ekhzaimy AA. Differences in the plasma proteome of patients with hypothyroidism before and after thyroid hormone replacement: A proteomic analysis. *International Journal of Molecular Sciences*. 2018; 19: 88.
- [11] Piras C, Pibiri M, Leoni VP, Balsamo A, Tronci L, Arisci N, *et al.* Analysis of metabolomics profile in hypothyroid patients before and after thyroid hormone replacement. *Journal of Endocrinological Investigation*. 2021; 44: 1309–1319.
- [12] Piras C, Arisci N, Poddighe S, Liggi S, Mariotti S, Atzori L. Metabolomic profile in hyperthyroid patients before and after antithyroid drug treatment: Correlation with thyroid hormone

and TSH concentration. *The International Journal of Biochemistry & Cell Biology*. 2017; 93: 119–128.

- [13] Liu J, Fu J, Jia Y, Yang N, Li J, Wang G. Serum Metabolomic Patterns in Patients with Autoimmune Thyroid Disease. *Endocrine Practice*. 2020; 26: 82–96.
- [14] Sun Z, Liu Y, Ouyang Q, Liu Z, Liu Y. Research progress of omics technology in the field of tumor resistance: From single-omics to multi-omics combination application. *Journal of Central South University Medical Sciences*, 2021; 46: 620–627.
- [15] Saraswat M, Joenväärä S, Tohmola T, Sutinen E, Vartiainen V, Koli K, *et al.* Label-free plasma proteomics identifies haptoglobin-related protein as candidate marker of idiopathic pulmonary fibrosis and dysregulation of complement and oxidative pathways. *Scientific Reports*. 2020; 10: 7787.
- [16] Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology*. 2008; 26: 1367–1372.
- [17] Choi M, Chang C, Clough T, Broudy D, Killeen T, MacLean B, *et al.* MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *Bioinformatics*. 2014; 30: 2524–2526.
- [18] von Mering C, Jensen LJ, Snel B, Hooper SD, Krupp M, Foglierini M, *et al.* STRING: known and predicted protein-protein associations, integrated and transferred across organisms. *Nucleic Acids Research*. 2005; 33: D433–D437.
- [19] Wozniak JM, Mills RH, Olson J, Caldera JR, Sepich-Poore GD, Carrillo-Terrazas M, *et al.* Mortality Risk Profiling of *Staphylococcus aureus* Bacteremia by Multi-omic Serum Analysis Reveals Early Predictive and Pathogenic Signatures. *Cell*. 2020; 182: 1311–1327.e14.
- [20] Neumann U, Genze N, Heider D. EFS: an ensemble feature selection tool implemented as R-package and web-application. *BioData Mining*. 2017; 10: 21.
- [21] Kravets I. Hyperthyroidism: Diagnosis and Treatment. *American Family Physician*. 2016; 93: 363–370.
- [22] Rizzo LFL, Mana DL, Serra HA. Drug-induced hypothyroidism. *Medicina*. 2017; 77: 394–404.
- [23] Wang H, Liu M. Complement C4, infections, and autoimmune diseases. *Frontiers in Immunology*, 2021; 12: 2682.
- [24] Blanchin S, Estienne V, Durand-Gorde J, Carayon P, Ruf J. Complement activation by direct C4 binding to thyroperoxidase in Hashimoto's thyroiditis. *Endocrinology*. 2003; 144: 5422–5429.
- [25] Yu H, Yang Y, Zhang M, Lu H, Zhang J, Wang H, *et al.* Thyroid status influence on adiponectin, acylation stimulating protein (ASP) and complement C3 in hyperthyroid and hypothyroid subjects. *Nutrition & Metabolism*. 2006; 3: 13.
- [26] Jafarzadeh A, Poorgholami M, Izadi N, Nemati M, Rezayati M. Immunological and hematological changes in patients with hyperthyroidism or hypothyroidism. *Clinical and Investigative Medicine Medecine Clinique Et Experimentale*. 2010; 33: E271–E279.
- [27] Albaugh VL, Pinzon-Guzman C, Barbul A. Arginine-Dual roles as an onconutrient and immunonutrient. *Journal of Surgical Oncology*. 2017; 115: 273–280.
- [28] Rodríguez-Gómez I, Moliz JN, Quesada A, Montoro-Molina S, Vargas-Tendero P, Osuna A, *et al.* L-Arginine metabolism in cardiovascular and renal tissue from hyper- and hypothyroid rats. *Experimental Biology and Medicine*. 2016; 241: 550–556.
- [29] Wan G, Zhaorigetu S, Liu Z, Kaini R, Jiang Z, Hu CA. Apolipoprotein L1, a novel Bcl-2 homology domain 3-only lipid-binding protein, induces autophagic cell death. *The Journal of Biological Chemistry*. 2008; 283: 21540–21549.
- [30] Cubedo J, Padró T, Alonso R, Mata P, Badimon L. ApoL1 levels in high density lipoprotein and cardiovascular event presentation in patients with familial hypercholesterolemia. *Journal of Lipid Research*. 2016; 57: 1059–1073.
- [31] Benvenega S, Alesci S, Trimarchi F. High-density lipoprotein-facilitated entry of thyroid hormones into cells: a mechanism different from the low-density lipoprotein-facilitated entry. *Thyroid*. 2002; 12: 547–556.
- [32] Krisp C, Randall SA, McKay MJ, Molloy MP. Towards clinical applications of selected reaction monitoring for plasma protein biomarker studies. *Proteomics Clinical Applications*. 2012; 6: 42–59.
- [33] Ferrari E, Wittig A, Basilico F, Rossi R, De Palma A, Di Silvestre D, *et al.* Urinary Proteomics Profiles Are Useful for Detection of Cancer Biomarkers and Changes Induced by Therapeutic Procedures. *Molecules*. 2019; 24: 794.
- [34] Ohara N, Kobayashi M, Tuchida M, Koda R, Yoneoka Y, Iino N. Isolated Adrenocorticotrophic Hormone Deficiency and Primary Hypothyroidism in a Patient Undergoing Long-Term Hemodialysis: A Case Report and Literature Review. *The American Journal of Case Reports*, 2020; 21: e922376.
- [35] Mousa SS, Davis FB, Davis PJ, Mousa SA. Human Platelet Aggregation and Degranulation is Induced in Vitro by L-Thyroxine, but not by 3,5,3'-Triiodo-L-Thyronine or Diiodothyropropionic Acid (DITPA). *Clinical and Applied Thrombosis/Hemostasis*. 2010; 16: 288–293.