Journal of Inflammation Research

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Integrated Multi-Omics Analyses Reveal Innovative Diagnostic and Therapeutic Targets Associated with Atopic Dermatitis

Xiangjie Chen[1–3,*], Bochun Cao[2,*], Zhiren Tan[2], Xiaoping Li[4], Wenrong Xu[5], Ying Liu[2], Fang Gong[1,2]

[1]Department of Laboratory Medicine, Jiangnan University Medical Center (Wuxi No. 2 People's Hospital), Wuxi, Jiangsu, People's Republic of China; [2]Wuxi School of Medicine, Jiangnan University, Wuxi, Jiangsu, People's Republic of China; [3]State Key Laboratory of Food Science and Technology, Jiangnan University, Wuxi, Jiangsu, People's Republic of China; [4]Department of Laboratory Medicine, the First Affiliated Hospital of Soochow University, Soochow University, Suzhou, People's Republic of China; [5]Department of Dermatology, The Affiliated Wuxi People' s Hospital of Nanjing Medical University, Wuxi, People's Republic of China

*These authors contributed equally to this work

Correspondence: Fang Gong; Ying Liu, Email 9862023157@jiangnan.edu.cn; 8202310001@jiangnan.edu.cn

Background: Atopic dermatitis (AD) is a chronic skin disorder that impacts patients' physical and mental health. Diagnosing AD mainly depends on evaluating medical history and symptoms, as there are no universally accepted biomarkers for it. Identifying novel, reliable biomarkers is crucial to enhance diagnostic accuracy, reduce healthcare costs, and aid in developing new treatments.

Methods: Data from the Gene Expression Omnibus database were used to identify potential AD biomarkers through Weighted Gene Co-expression Network Analysis and machine-learning. External datasets confirmed these biomarkers' diagnostic utility and their effectiveness in assessing clinical treatment. We also gathered peripheral blood mononuclear cells from healthy individuals and AD patients to validate these biomarkers' diagnostic capability for AD. Correlation analyses linked these biomarkers to AD severity indicators. Euclidean distance clustering was employed to assess the ability of these biomarkers to distinguish between healthy individuals and AD patients. The study also examined their relationships with major inflammatory pathways in AD to understand their mechanisms.

Results: The study identified Ribonucleotide Reductase Regulatory Subunit M2 (RRM2), Late Cornified Envelope 3D (LCE3D), Cornifelin (CNFN), and Small Proline Rich Protein 2G (SPRR2G) as biomarkers with greater diagnostic value for AD than traditional biomarkers like eosinophil count and IgE levels. Treatment led to decreased expression of RRM2, LCE3D, and CNFN in AD patients' skin, indicating their potential as markers for evaluating treatment efficacy. LCE3D, CNFN, and SPRR2G correlated with AD severity indicators such as the SCORAD score and serum IgE levels. Additionally, the overexpression of these biomarkers was linked to the activation of inflammatory pathways, suggesting their role in AD pathogenesis and progression.

Conclusion: Our study identifies RRM2, LCE3D, CNFN, and SPRR2G as novel biomarkers for diagnosing AD in peripheral blood and lesional tissues, with potential for assessing disease severity, evaluating treatment efficacy, and serving as targets for diagnosis and treatment.

Keywords: atopic dermatitis, AD, biomarker, transcriptome, bioinformatics

Introduction

Atopic Dermatitis (AD) is characterized as an inflammatory skin disorder resulting from a complex interplay of genetic, immune, and environmental factors, manifesting as chronic recurrent eczema accompanied by pruritus.[1–3] Its clinical symptoms are similar to those of skin diseases such as Allergic Contact Dermatitis (ACD), resulting in difficult correct diagnosis in some cases.[4] In addition, the complex clinical manifestations of AD significantly diminish the quality of life for both patients and their families.[5,6] Therefore, enhancing the diagnostic capabilities for AD and implementing effective clinical management strategies are of paramount importance.

Received: 10 March 2025
Accepted: 4 June 2025
Published: 17 June 2025

The diagnosis of disease typically necessitates the identification of reliable biomarkers. In contrast to numerous other chronic diseases, the diagnosis of AD predominantly depends on clinical assessments rather than biochemical markers.[7] Consequently, the establishment of a reliable biomarker could mitigate the subjectivity inherent in clinical evaluations. In the lesional skin or peripheral blood of AD patients, numerous genes exhibit abnormal expression, and these aberrantly expressed genes may serve as potential biomarkers for AD. Identifying and assessing Differentially Expressed Genes (DEGs) with diagnostic relevance could yield reliable biomarkers for AD diagnosis. The integration of these biomarkers into existing clinical diagnostic frameworks has the potential to enhance the diagnostic accuracy of AD. Recent studies have investigated the development of novel diagnostic biomarkers for AD. Blood eosinophils and serum Immunoglobulin E (IgE) are the most commonly utilized biomarkers for AD diagnosis, as they are elevated in the blood of AD patients.[8] Additionally, NOS2/iNOS, hBD-2, and MMP8/9 are highly expressed in the lesional skin of AD patients, suggesting their potential as candidate biomarkers for AD diagnosis.[8] Furthermore, certain biomarkers have the dual capability of diagnosing the onset of AD and monitoring its severity. EDN has been shown to correlate with AD severity, making it a potential candidate biomarker for predicting disease severity.[9,10] SCCA2 is associated with the Scoring Atopic Dermatitis (SCORAD) score, and its expression decreases as AD improves, thus establishing it as a biomarker for reflecting AD severity.[11,12] Serum TARC/CCL17 serves as a significant T cell chemoattractant and is recognized as a biomarker for monitoring AD severity.[13–15] In summary, analyzing differences in gene expression patterns in the skin or peripheral blood of AD patients compared to healthy individuals, not only facilitates the discovery of novel diagnostic markers for AD but is also crucial for elucidating the role of these molecules in AD pathogenesis and identifying potential therapeutic targets.

High-throughput sequencing represents a powerful technique for investigating alterations in gene expression profiles associated with diseases, facilitating the screening of pivotal genes and the identification of novel diagnostic biomarkers.[16] Analyzing variations in gene expression levels can provide significant insights into the potential pathogenesis of diseases. The application of machine learning algorithms to the analysis of DEGs pertinent to diseases can efficiently identify genes of substantial biological significance, ascertain innovative diagnostic biomarkers for diseases, enable earlier and more precise diagnosis of AD, and consequently promote timely intervention and enhance patient prognosis.[17–19] Importantly, these methods have been widely used to identify innovative biomarkers of skin diseases, such as ACD, psoriasis and keloids.[20–22]

In this study, transcriptome sequencing and single-cell sequencing datasets related to AD were sourced from the Gene Expression Omnibus (GEO) database. DEGs in AD were analyzed, and comprehensive bioinformatics analyses were conducted to identify Ribonucleotide Reductase Regulatory Subunit M2 (RRM2), Late Cornified Envelope 3D (LCE3D), Cornifelin (CNFN), and Small Proline Rich Protein 2G (SPRR2G) as potential diagnostic biomarkers for AD. Our experimental findings indicated that the expression levels of RRM2, LCE3D, CNFN, and SPRR2G in the peripheral blood of AD patients were significantly elevated compared to those in healthy individuals. The diagnostic efficacy of these biomarkers was found to be comparable to traditional markers for AD diagnosis, such as peripheral blood eosinophil count and serum IgE levels. Furthermore, we observed that the expression of RRM2, LCE3D, and CNFN in the lesional skin tissue of AD patients treated with crisaborole or dupilumab were reduced. Receiver Operating Characteristic (ROC) analysis demonstrated that these biomarkers could serve as valuable indicators for assessing the clinical efficacy of AD treatment. Finally, the innovative biomarkers identified in this study were significantly associated with the severity of AD. Correlation analysis demonstrated a strong relationship between these biomarkers and SCORAD scores, serum IgE levels, as well as various effector T cell subtypes indicative of AD severity. These findings underscore the substantial utility of RRM2, LCE3D, CNFN, and SPRR2G in evaluating disease severity in AD patients. Subsequent analysis revealed that the overexpression of these four biomarkers is implicated in the activation of multiple inflammatory pathways. Collectively, our research emphasizes the potential of RRM2, LCE3D, CNFN, and SPRR2G as biomarkers for AD and investigates their functional roles in the disease. These biomarkers hold promise for enhancing clinical diagnosis and treatment strategies for AD.

Methods

Human Samples

This study was conducted in accordance with the protocol approved by the Ethics Committee of Jiangnan University Medical Center (approval number: 2024-Y-205), with the informed consent of all subjects. Furthermore, we proved that this research was conducted in accordance with the ethical standards stipulated in the Declaration of Helsinki. The patients were recruited through the outpatient clinic of Jiangnan University Medical Center. Patients with AD were diagnosed according to Guideline for diagnosis and treatment of atopic dermatitis in China (2020).[23] All patients were in the acute stage of AD and had not yet used therapeutic drugs. Blood samples were collected and Peripheral Blood Mononuclear Cells (PBMCs) were isolated using lymphocyte isolation medium (Da you, DAKEWE) according to manufacturer's instructions.

Datasets

In this study, we obtained several high-throughput sequencing datasets for AD from the GEO database. These datasets included lesional skin tissue and peripheral blood from patients with AD. The GSE121212 dataset contains 21 AD lesional samples, 27 AD non-lesional samples, and 38 healthy samples. GSE197023 dataset contains 7 AD lesional samples and 6 healthy samples. GSE182740 dataset contains 10 AD lesional samples and 6 healthy samples. GSE237920 dataset contains 4 AD lesional samples and 4 healthy samples. GSE36842 dataset contains 16 AD lesional samples and 8 AD non-lesional samples. GSE168694 dataset contains PBMCs samples from 4 patients with AD or 4 healthy individuals. The GSE133477 dataset included 40 patients with AD who were treated with Crisaborole twice daily for 14 days. The gene expression array was detected on the 8th and 15th day of treatment, respectively. Twenty-two patient skin samples were collected on day 8, and 40 patient skin samples were collected on day 15. The GSE130588 dataset included 51 AD lesional samples and 20 healthy samples. Detailed information about these datasets is provided in Table S1.

Analysis of Differentially Expressed Genes

The "DESeq2" package was used to screen DEGs between AD and Healthy Control (HC) skin specimens. DEGs were screened according to "adjusted P<0.05" and "|log2FoldChange|>1.5". Volcano points or heat maps were generated to visualize DEGs.

Functional Correlation Analysis

The Kyoto Encyclopedia of Genes and Genomes (KEGG) is a widely used database that quickly targets the signaling pathways that play a dominant role in a disease. Gene Ontology (GO) analysis is a common technique used to conduct large-scale functional enrichment studies, which has great advantages in resolving the Biological Process (BP), Cellular Component (CC), and Molecular Function (MF) of DEGs in a disease. The "org.Hs.eg.db" and "clusterProfiler" packages were used for KEGG and GO enrichment analysis, and the screening criteria were P(adj)<0.05. Then the "GOplot" and "ggplot2" packages were used to visualize the analysis results. In addition, the Gene Set Enrichment Analysis (GSEA) of the gene expression matrix was performed using the "clusterProfiler" software package. Maximum gene module 1000 and Minimum gene module 20. P(adj) < 0.05 were considered statistically significant.

scRNA-Seq Data Processing

We utilized the "Seurat" R package for processing ScRNA-seq data. In brief, cells with gene expression levels below 300 genes or above 7000 genes, as well as those exhibiting mitochondrial gene expression exceeding 10%, were excluded from the analysis. Subsequently, we applied the "NormalizeData", "FindVariableFeatures", "ScaleData" function to normalize and scale the raw counts, followed by Principal Component Analysis (PCA). The "Harmony" R package was employed to eliminate of batch effects in the original scRNA-seq data. Visualization of the resulting units was achieved using the Uniform Manifold Approximation and Projection (UMAP) algorithm. Seurat's "FindMarkers" functionality was used to identify DEGs within each cell cluster. This was followed by annotation of each cell cluster using classical

cell type marker genes. Finally, we visualized specific expression patterns of identified genes at single-cell resolution using the "scRNAtoolVis" package.

Weighted Gene Co-Expression Network Analysis

Weighted Gene Co-expression Network Analysis (WGCNA) could construct disease-related modules based on gene expression profiles. In our study, we calculated the Pearson correlation coefficient and constructed a similarity matrix. Then, select the soft threshold =5, convert the similarity matrix into an adjacency matrix, and then convert the adjacency matrix into a Topological Overlap Matrix (TOM). Using TOM average linkage hierarchical clustering to classify gene modules. After merging similar gene modules, 10 different modules were obtained, among which the greenyellow module containing 4601 genes is considered the core module of AD.

High Dimensional Weighted Gene Co-Expression Network Analysis

High dimensional Weighted Gene Co-expression Network Analysis (hdWGCNA) builds scale-free networks at the single-cell level. After the scale-free topology model fitting threshold is set to >0.85, the soft threshold is selected to 7 to obtain the best connectivity. Then run the hdWGCNA analysis with the default parameters.[24,25]

Identification of AD Biomarkers by Machine Learning

Three machine-learning algorithms, Least Absolute Shrinkage and Selection Operator (LASSO), Random Forests (RF) and Boruta (Boruta), were used to screen AD biomarkers. The "glmnet" package was used for LASSO regression analysis, 10-fold cross-validation was adopted, and the minimum lambda value was used as the optimal solution. RF analysis was performed using the "randomForest" package. Boruta analysis using the "Boruta" package.

Correlations Between Potential Biomarkers and Inflammatory Pathways in AD

Using "Mantel test" and "Pearson correlation coefficient", the correlation between RRM2, LCE3D, CNFN, SPRR2G and multiple core genes of the inflammatory pathway was calculated using the R software package "linkET".

Consensus Clustering Analysis

Based on the relative expression levels of RRM2, LCE3D, CNFN, and SPRR2G, the samples contained in datasets GSE121212 and GSE130588 were grouped. The R package "ConsensusClusterPlus" was used in combination with consensus clustering methods, employing the k-nearest neighbor algorithm and Euclidean distance, repeated 1000 times to ensure the stability of the analysis results. The optimal number of clusters was determined based on the CDF curve. Finally, visualization was performed using the "pheamap" package.

RNA Isolation and Real-Time PCR

Total RNAs were extracted from cells with TRIzol reagent (Vazyme, R401-01). cDNA was produced by HiScript III reverse transcriptase (Vazyme, R312). Then quantitative PCR(Q-PCR) analysis was performed using Gata3, Foxp3, RRM2, LCE3D, CNFN, and SPRR2G specific primers and SYBR Green mix (Vazyme, Q411-02). The primer sequence is as follows:

Rrm2, 5′-ctggctcaagaaacgaggactg-3′ and 5′-ctctcctccgatggtttgtgtac-3′;
Lce3d, 5′-agtacagtgtctgcctccagct-3′ and 5′-tgtcacaggagttgggcctctg-3′;
Cnfn, 5′-tcagtgactggcacacaggtct-3′ and 5′-cggtgcggatggagtgcagg-3′;
Sprr2g, 5′-acgccaaagtgcccagag-3′ and 5′-gcacaggagggcatttatcc-3′;
Gata3, 5′-accacaaccacactctggagga-3′ and 5′-tcggtttctggtctggatgcct-3′;
Foxp3, 5′-ggcacaatgtctcctccagaga-3′ and 5′-cagatgaagccttggtcagtgc-3′;

Serum Biochemical Test

Serum IgE levels were measured by Maglumi4000 automatic chemiluminescence analyzer.

Blood Count

The number of Eosinophils in the peripheral blood of AD patients was analyzed by Mindray CAL8000 blood analyzer.

Statistical Analysis

Data analysis was performed using GraphPad Prism 9.0 software. Different groups were compared by a two-tailed Student's *T*-test, one-way analysis of variance, or two-way analysis of variance. All differences were regarded as statistically significant when *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

Results

Screening of DEGs in AD Transcriptome Dataset

The research methodology of this study is depicted in Figure S1. To identify DEGs associated with AD, an mRNA expression microarray dataset (GSE121212) was retrieved from the GEO database. The "DESeq2" package was employed to analyze the DEGs between the AD and HC groups. A total of 1371 DEGs were identified, comprising 828 upregulated and 543 downregulated genes in the AD group relative to the HC group. A volcano plot was utilized to visually represent genes with significant differential expression between the AD and HC groups (Figure 1A). To further elucidate the biological functions of these DEGs, KEGG and GO enrichment analyses were conducted. The KEGG enrichment analysis indicated that the DEGs were predominantly associated with pathways such as the Chemokine signaling pathway, IL-17 signaling pathway, JAK-STAT signaling pathway, Lipid and atherosclerosis, Neuroactive ligand-receptor interaction, NF-kappa B signaling pathway, PPAR signaling pathway, Retinol metabolism, Th1 and Th2 cell differentiation, Th17 cell differentiation, and TNF signaling pathway (Figure 1B). The results of the GO analysis indicated that these DEGs were associated with a variety of biological processes, including the regulation of cytokine production, leukocyte cell-cell adhesion, leukocyte-mediated immunity, mononuclear cell differentiation, regulation of T cell activation, leukocyte migration, lymphocyte differentiation, positive regulation of the MAPK cascade, regulation of inflammatory response, and cell killing (Figure 1C and D). We observed that several inflammatory signaling pathways were significantly linked to AD. Consequently, we conducted a further analysis of the activation of these signaling pathways in AD. GSEA analysis revealed that the NF-κB, IL-17, TNF-α, MAPK, Chemokine, and JAK-STAT signaling pathways were upregulated in the AD group compared to the HC group (Figure 1E).

Identification of Co-Expression Gene Modules in AD Transcriptome Dataset

The WGCNA analysis begins with the construction of a gene co-expression network, followed by the identification of gene modules that exhibit significant correlations with AD within this network. To assign co-expressed genes, the "average correlation degree" and "pearson correlation coefficient" are computed to cluster the samples in the dataset, with outliers being excluded (Figure S2A). Based on a scale independence criterion of greater than 0.8, a "soft threshold power β" of 5 was selected to ensure the formation of biologically meaningful scale-free networks (Figure 2A). Seventeen modules were identified through hierarchical clustering analysis and the dynamic TreeCut method (Figure S2B and Figure 2B). The results of the module-trait correlation analysis revealed a significant positive association (Correlation = 0.81) between the greenyellow module and AD, prompting its selection for further investigation (Figure 2C). In the greenyellow module, a total of 4,601 genes were identified. The scatter plot demonstrated a positive correlation between Module Membership (MM) and Gene Significance (GS), suggesting that these genes, which are highly associated with AD, play a significant role within the greenyellow module (Figure 2D). A Venn diagram was employed to assess the overlap between DEGs and the genes of the greenyellow module within the GSE121212 dataset, identifying 938 common genes (Figure 2E). These 938 genes have been designated as "GSE121212-Hub Genes".

Identification of Hub Genes in AD Single-Cell Sequencing Dataset

Subsequently, we utilized a single-cell RNA sequencing (scRNA-seq) dataset from AD patients to conduct an in-depth analysis of key genes associated with AD. The scRNA-seq dataset (GSE197023) was sourced from the GEO database, allowing us to compare lesional AD skin (LS) with HC skin. The volcano plot (Figure 3A) illustrates that, in comparison

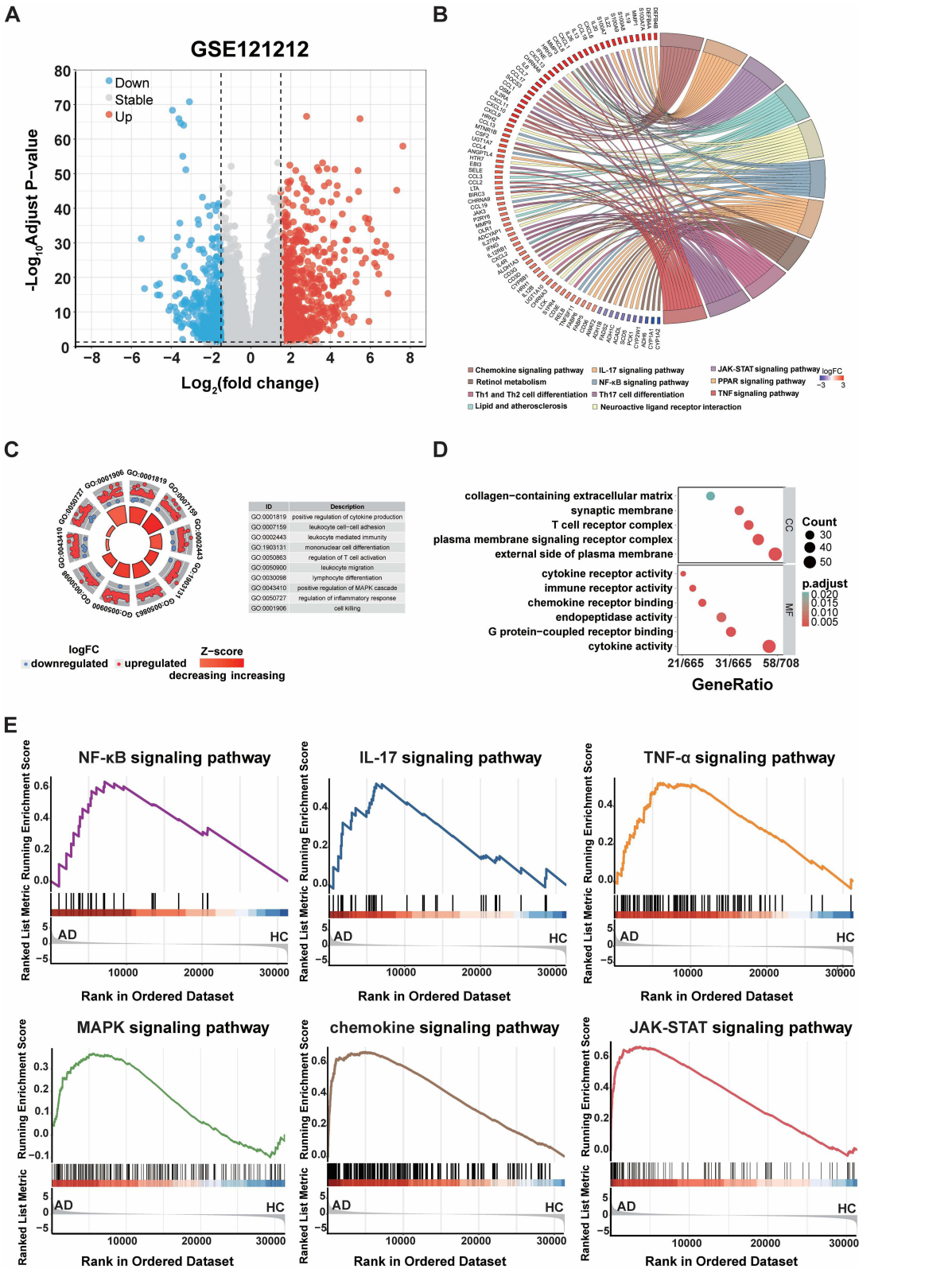

Figure 1 Identified the DEGs in AD and conducted functional enrichment analyses. (**A**) The volcano plot revealed 1371 DEGs in skin tissues from AD patients compared to healthy individuals in the GSE121212 dataset. (**B**) KEGG enrichment analysis indicated that DEGs were associated with multiple inflammatory signaling pathways. (**C** and **D**) GO analysis was performed to explore the biological processes (BP), cellular components (CC), and molecular functions (MF) of the DEGs. (**E**) Gene Set Enrichment Analysis (GSEA) was conducted to assess the enrichment of DEGs in AD and HC groups.

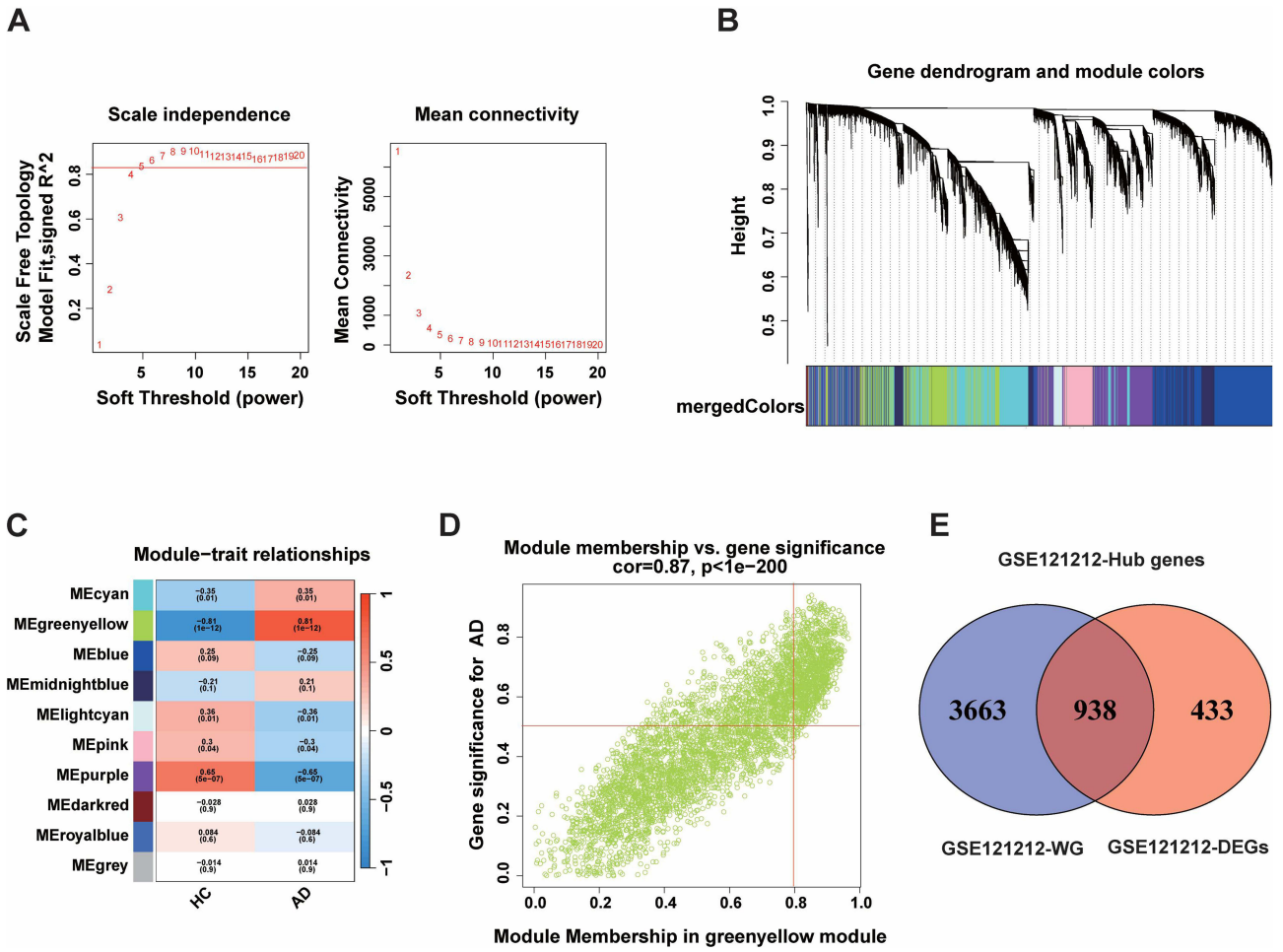

Figure 2 Identification of gene co-expression network modules associated with AD through WGCNA analysis. (**A**) Optimal soft threshold selection. (**B**) Dynamic shearing tree merging similar module genes. (**C**) Module-trait relationships. (**D**) The correlation between module membership and genetic significance in the greenyellow module. (**E**) Venn diagram of greenyellow module genes versus GSE121212-DEGs.

to the HC group, a total of 592 DEGs were identified in the AD group, comprising 331 upregulated and 261 downregulated genes. UMAP analyses revealed nine distinct cell clusters. However, upon classification using canonical marker genes, three cell populations, designated as clusters 2, 3, and 5, could not be reliably identified (Figure S3A and B). Consequently, these three cell populations were excluded, and the UMAP analysis was repeated, resulting in the identification of nine cell clusters (Figure 3B). Among all cells, six types were annotated as the main labels of the clusters. Utilizing canonical marker genes, we labeled cell clusters 0, 4, and 5 as upper epidermis cells; clusters 1 and 2 as leukocytes; cluster 6 as smooth muscle cells; cluster 3 as ECM/fibroblasts; cluster 7 as sweat glands; and cluster 8 as lower epidermis cells (Figure 3C–E). To more accurately delineate the distinct patterns of DEG signatures across these six cell subtypes, we ranked each gene based on its relative expression in individual cells. The top ten DEGs for each cell subtype were selected, and their associated biological functions were investigated (Figure 3F).

hdWGCNA is frequently employed to identify co-expressed gene modules in single-cell sequencing datasets and to further analyze the key members of these modules.[24] Subsequently, we utilized hdWGCNA to elucidate the primary molecular characteristics of AD. By applying a soft threshold of 7, we constructed a scale-free network for AD, identifying a total of five distinct gene co-expression modules (Figure 4A–C). Our analysis revealed that these five modules were activated in different cell populations: the blue module was predominantly activated in upper and lower epidermis cells; the turquoise module was primarily activated in smooth muscle cells, ECM/fibroblasts, and sweat gland cells; the brown module was chiefly activated in leukocytes; the yellow module was mainly activated in epidermis cells and leukocytes; and the green module was primarily activated in epidermis cells (Figure 4D).

https://doi.org/10.2147/JIR.S526983

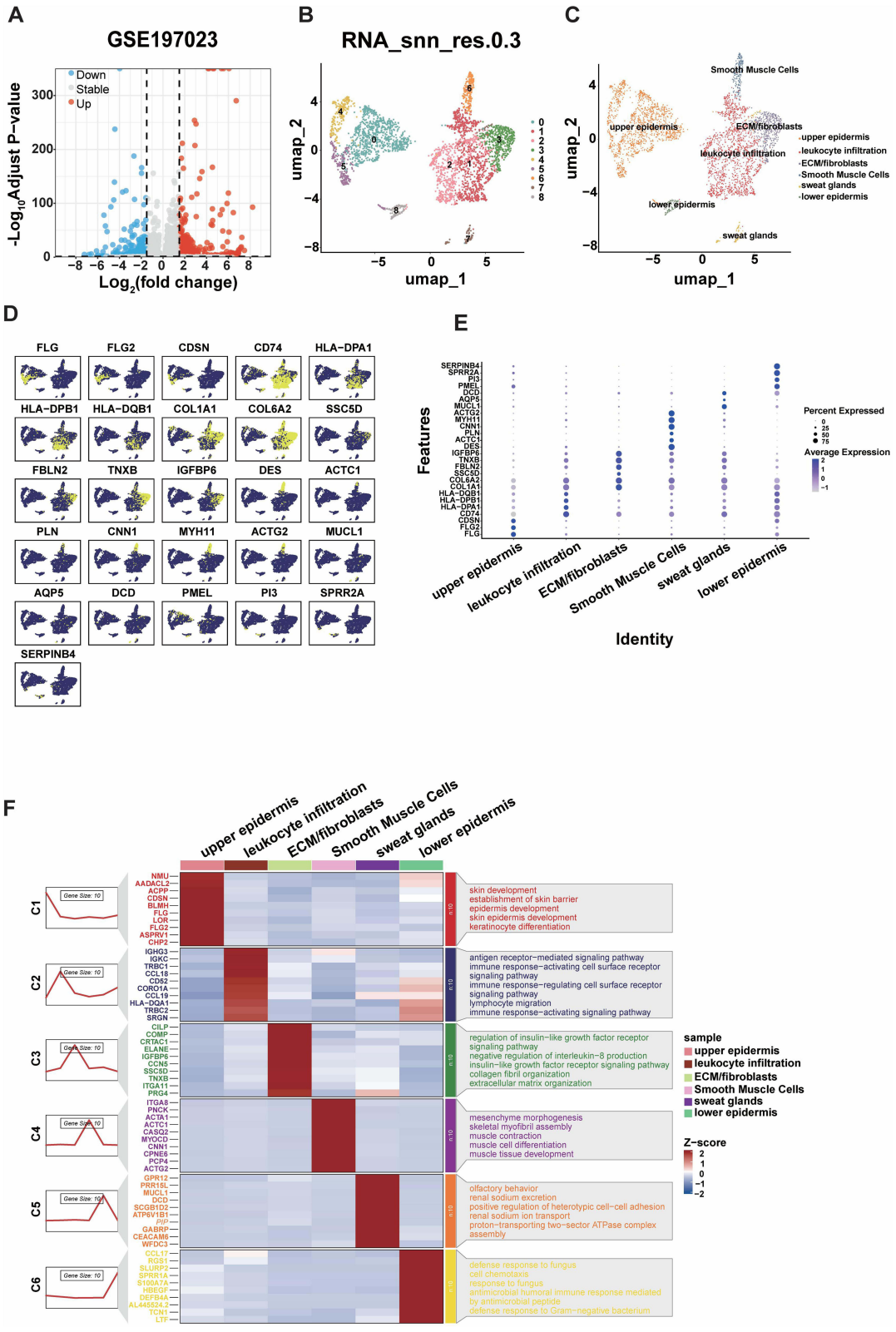

Figure 3 Comprehensive single-cell analysis of skin tissue in AD patients. (**A**) The volcano plot revealed 892 DEGs in the skin tissues of AD patients compared to healthy individuals in the single-cell sequencing dataset (GSE197023). (**B** and **C**) UMAP visualization of all 3500 cells, with cells colored according to their cell type. (**D** and **E**) Characteristic markers for each cell type. (**F**) *Left panels*: A series of diagrams depicting the dynamic expression patterns of representative DEGs in each cell population. *Middle panels*: Displaying the top 10 representative DEGs in each cell population. *Right panels*: Showing the representative enriched GO terms for each cell population.

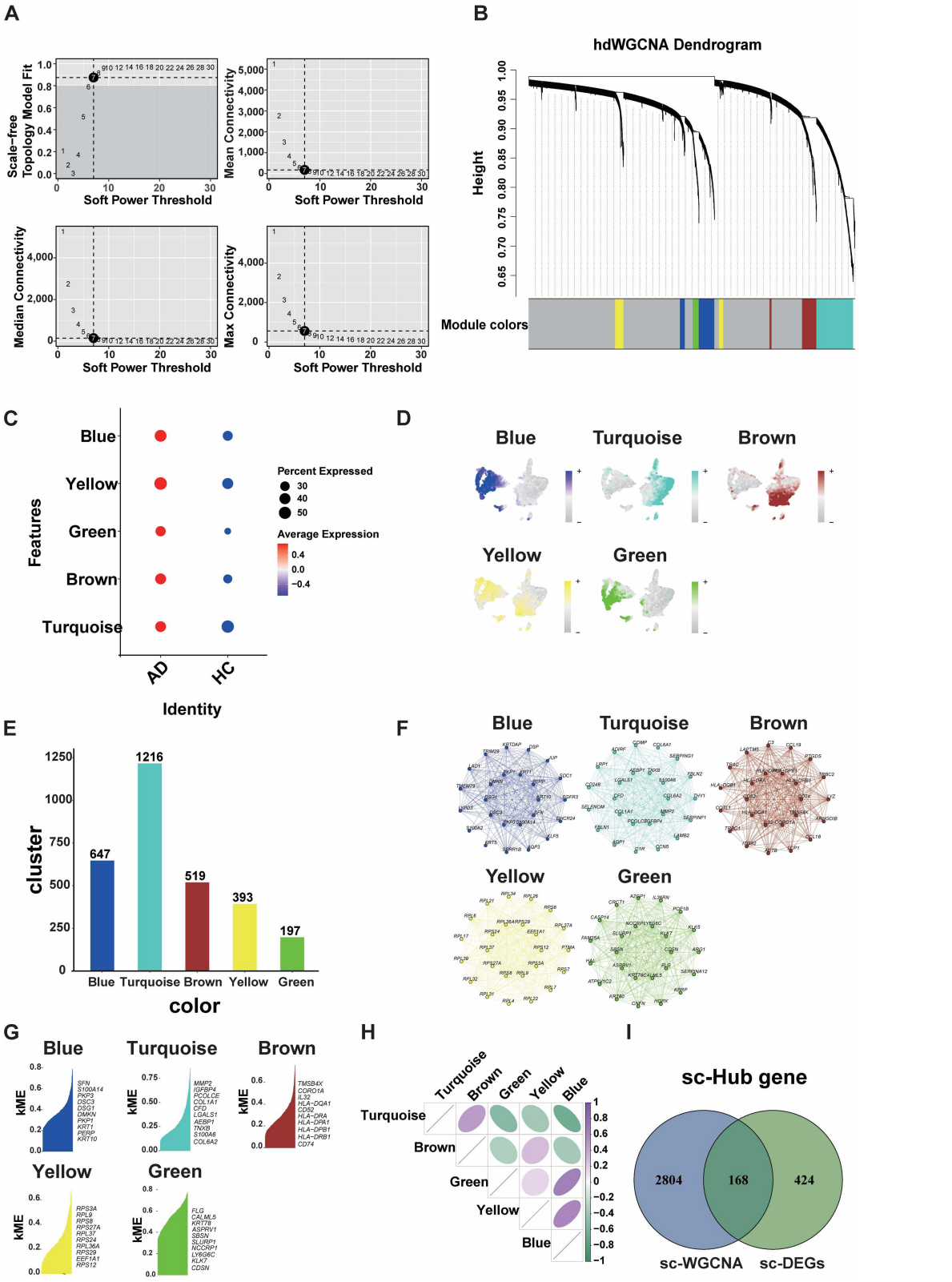

Figure 4 Identification of potential AD-related genes by hdWGCNA. (**A**) Optimal soft threshold selection. (**B**) The dendrogram illustrated the five modules (turquoise, green, blue, brown, and yellow) in the scale-free network. (**C**) The hdWGCNA algorithm was employed to assess module activity in both AD and HC groups. (**D**) t-SNE points revealed the expression distribution of each module hub genes in 6 different cell populations. (**E**) The bar chart revealed the number of genes contained in each module. (**F**) The protein-protein interaction (PPI) network of the top 20 identified hub genes in each module. (**G**) The top 10 hub genes in each module were ranked by kME. (**H**) The matrix plot visually showed the correlation between the different modules. (**I**) Venn diagram of hdWGCNA genes versus sc-DEGs.

https://doi.org/10.2147/JIR.S526983

Among the five modules analyzed, the turquoise module exhibited the highest gene count, comprising a total of 1,216 genes. In contrast, the green module contained the fewest genes, with only 197. The blue, brown, and yellow modules comprised 647, 519, and 393 genes, respectively (Figure 4E). Furthermore, Protein-Protein Interaction (PPI) networks were constructed for the top 20 feature genes within each module (Figure 4F). According to eigengene-based connectivity (KME), the top 10 characteristic genes for each module were ranked (Figure 4G). Additionally, an analysis of pairwise correlations between the modules was conducted (Figure 4H). To identify hub genes within the AD single-cell RNA sequencing dataset, a Venn diagram was employed to assess the overlap between DEGs and the five co-expression modules genes, identifying 168 shared genes (Figure 4I). These genes were designated as "sc-Hub genes".

Screening and Validation of Diagnostic Markers

By employing a venn diagram to examine the overlapping regions between GSE121212-Hub Genes and sc-Hub genes, we identified 57 common genes (Figure 5A), whose expression patterns were subsequently visualized using heatmaps (Figure S4). Subsequently, three distinct machine learning algorithms—LASSO, RF, and Boruta—were employed to identify potential diagnostic biomarkers for AD. The LASSO regression algorithm facilitated the construction of a diagnostic model, leading to the selection of seven genes (Figure 5B). Similarly, the RF algorithm was applied to develop a diagnostic model, resulting in the selection of 52 genes when the Mean Decrease Accuracy>0 (Figure 5C). Additionally, the Boruta algorithm was employed to construct a diagnostic model, culminating in the selection of 42 genes (Figure 5D and E). Ultimately, the integration of these three algorithms led to the identification of five novel biomarkers for AD: RRM2, LCE3D, CNFN, SPRR2G, and COMP (Figure 5F).

Subsequently, we employed ROC analysis to assess the diagnostic efficacy of these five genes in AD. Initially, we examined multiple datasets comprising lesional skin from AD patients and healthy skin from healthy subjects. In dataset GSE121212, the area under the curve (AUC) values for all five genes exceeded 0.8 (Figure 6A). In dataset GSE182740, the AUC values for RRM2, LCE3D, CNFN, and SPRR2G were greater than 0.8, whereas the AUC for COMP was 0.667 (Figure 6A). Similarly, in dataset GSE237920, the AUC values for RRM2, LCE3D, CNFN, and SPRR2G exceeded 0.8, while the AUC for COMP was only 0.562 (Figure 6A). Subsequently, we analyzed two datasets containing lesional and non-lesional skin from AD patients. In dataset GSE121212, the AUC values for RRM2, LCE3D, CNFN, and SPRR2G were above 0.7, but the AUC for COMP was 0.555 (Figure 6B). In dataset GSE36842, the AUC values for RRM2, LCE3D, CNFN, and SPRR2G exceeded 0.7, with the AUC for COMP at 0.586 (Figure 6B). Finally, we evaluated a dataset (GSE168694) comprising PBMCs from healthy individuals and AD patients. In this dataset, the AUC values for LCE3D, CNFN, SPRR2G, and COMP were greater than 0.9, while the AUC for RRM2 was 0.688 (Figure 6C). Given that the AUC for COMP is consistently below 0.6 across multiple datasets, we conclude that it is not a viable candidate for an innovative diagnostic marker for AD.

To further investigate the potential diagnostic and therapeutic value of the genes RRM2, LCE3D, CNFN, and SPRR2G in AD, we conducted an analysis of the GSE133477 dataset. This dataset comprises 40 AD patients who underwent treatment with Crisaborole, administered twice daily over a 14-day period. Gene expression array was performed on the 8th and 15th days of treatment. Our findings revealed that in patients with mild to moderate AD treated with Crisaborole, there was a significant reduction in the expression levels of RRM2, LCE3D, and CNFN in lesional skin (Figure 6D), indicating their diagnostic potential (Figure S5A). Similarly, in patients with severe AD treated with Dupilumab, a decrease in the expression of these genes was observed (Figure S5B). Collectively, our study demonstrates that RRM2, LCE3D, and CNFN consistently exhibit high diagnostic value across various datasets and can effectively reflect the therapeutic impact of treatments in AD patients ranging from mild to severe. These findings suggest that these genes hold promise as potential biomarkers for the diagnosis and management of AD.

RRM2, LCE3D, CNFN and SPRR2G Were up-Regulated in Peripheral Blood of AD Patients and Had Diagnostic Value

Next, we analyzed the expression patterns of RRM2, LCE3D, CNFN, and SPRR2G using GSE197023 dataset. These four genes were predominantly expressed in upper epidermis cells and leukocytes (Figure 7A). Additionally, their

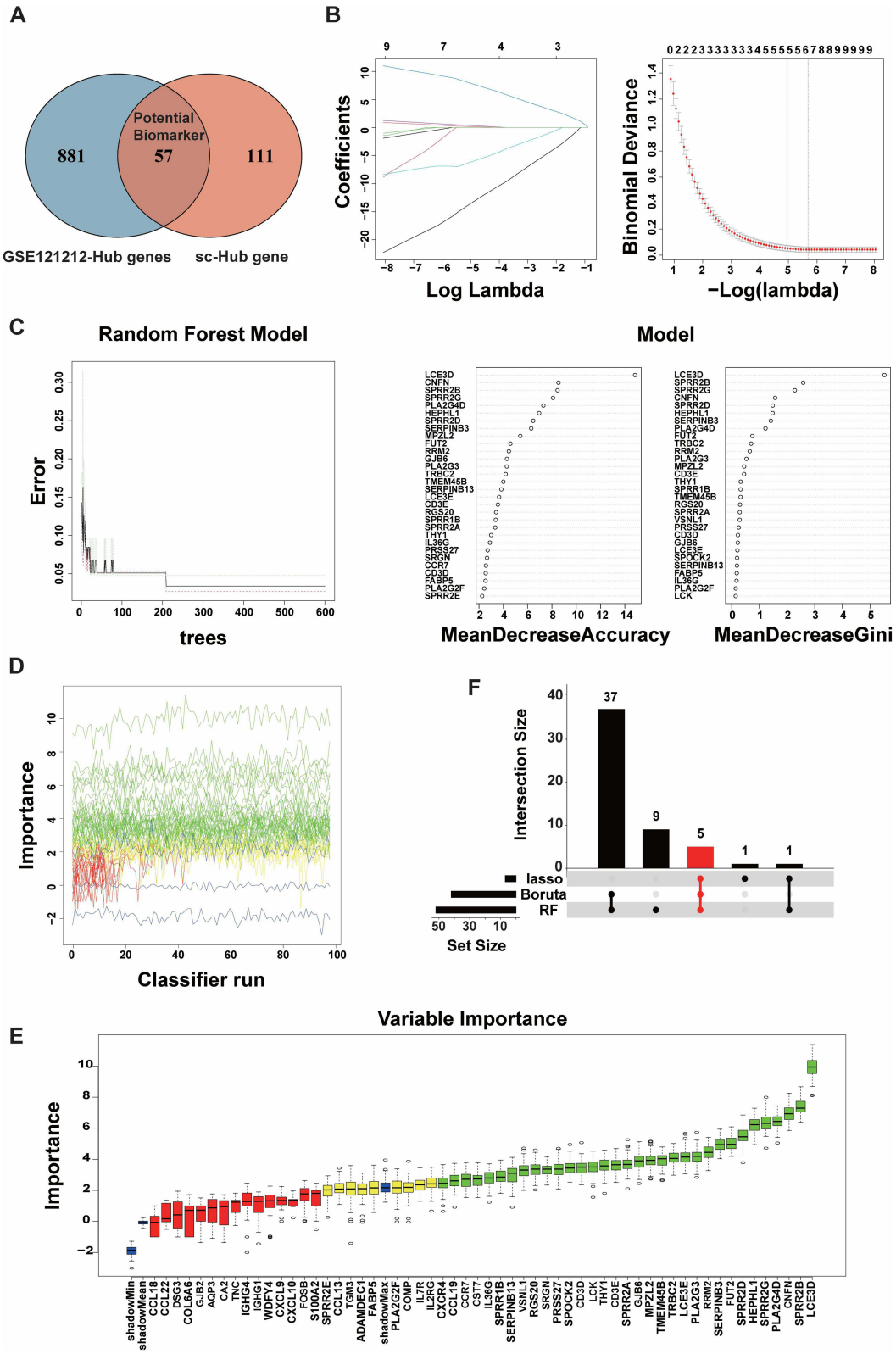

Figure 5 The identification of innovative biomarkers for AD was achieved via three machine learning algorithms. (**A**) Venn diagram of transcriptome data hub genes versus single-cell sequencing hub genes. (**B**) Seven hub genes were identified using the Lasso regression algorithm. (**C**) The Random Forest algorithm ranked all genes to determine their significance in the model. (**D** and **E**) The Boruta algorithm was used to screen for 42 hub genes. (**F**) A total of five genes were identified by the three machine learning algorithms and were considered as potential biomarkers for AD.

https://doi.org/10.2147/JIR.S526983

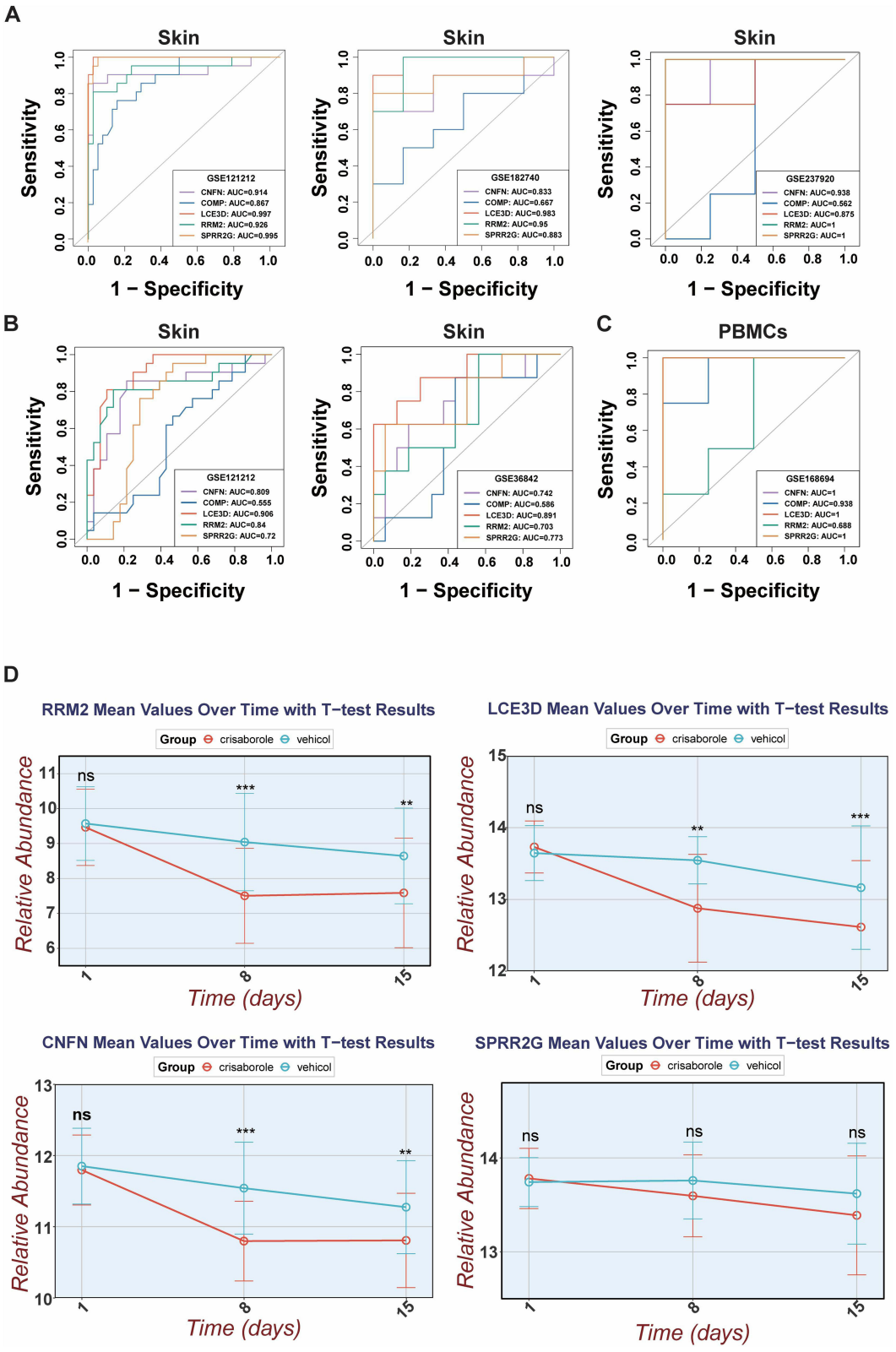

Figure 6 RRM2, LCE3D, CNFN, and SPRR2G serve as diagnostic biomarkers for both skin lesions and peripheral blood in AD. (**A**) ROC analysis was performed for RRM2, LCE3D, CNFN, SPRR2G, and COMP to evaluate their diagnostic performance in distinguishing between AD lesions and healthy individuals skin tissue in the datasets GSE121212, GSE182740, and GSE237920. (**B**) ROC analysis was performed for RRM2, LCE3D, CNFN, SPRR2G, and COMP to evaluate their diagnostic performance in distinguishing between AD lesional and non-lesional skin tissue in the datasets GSE121212 and GSE36842. (**C**) ROC analysis was performed for RRM2, LCE3D, CNFN, SPRR2G, and COMP to evaluate their diagnostic performance in distinguishing PBMCs from AD patients and healthy individuals in the GSE168694 dataset. (**D**) The expression levels of RRM2, LCE3D, CNFN, and SPRR2G in the skin lesion sites of patients with mild-to-moderate AD treated with continuous Crisaborole for 14 days were analyzed using the GSE133477 dataset (two-way analysis of variance). ns: not significant, **p < 0.01 and ***p < 0.001.

expression was notably elevated in the lesional skin tissue of patients with AD (Figure 7B). We further investigated the expression levels of RRM2, LCE3D, CNFN, and SPRR2G across various cell populations in both AD and HC groups. LCE3D, CNFN, and SPRR2G were upregulated in the upper epidermis cells, leukocytes, smooth muscle cells, ECM/fibroblasts, and sweat gland cells of AD patients. RRM2 was upregulated only in the upper epidermis cells, leukocytes, and ECM/fibroblasts of AD patients (Figure 7C). We also assessed the expression levels of RRM2, LCE3D, CNFN, and SPRR2G in PBMCs of AD patients using RT-qPCR. Notably, the mRNA levels of these genes were significantly increased in non-medicated AD patients compared to age- and sex-matched healthy controls (Figure 7D, F, H and J). ROC analysis indicated that the AUC was 0.755 for RRM2, 0.768 for LCE3D, 0.730 for CNFN, and 0.730 for SPRR2G (Figure 7E, G, I and K). These findings indicate that RRM2, LCE3D, CNFN, and SPRR2G exhibit significant diagnostic potential for AD. To further explore this potential, we collected serum samples from 112 AD patients and peripheral blood cell data from 181 patients, who treated at Jiangnan University Medical Center between February 1, 2024, and

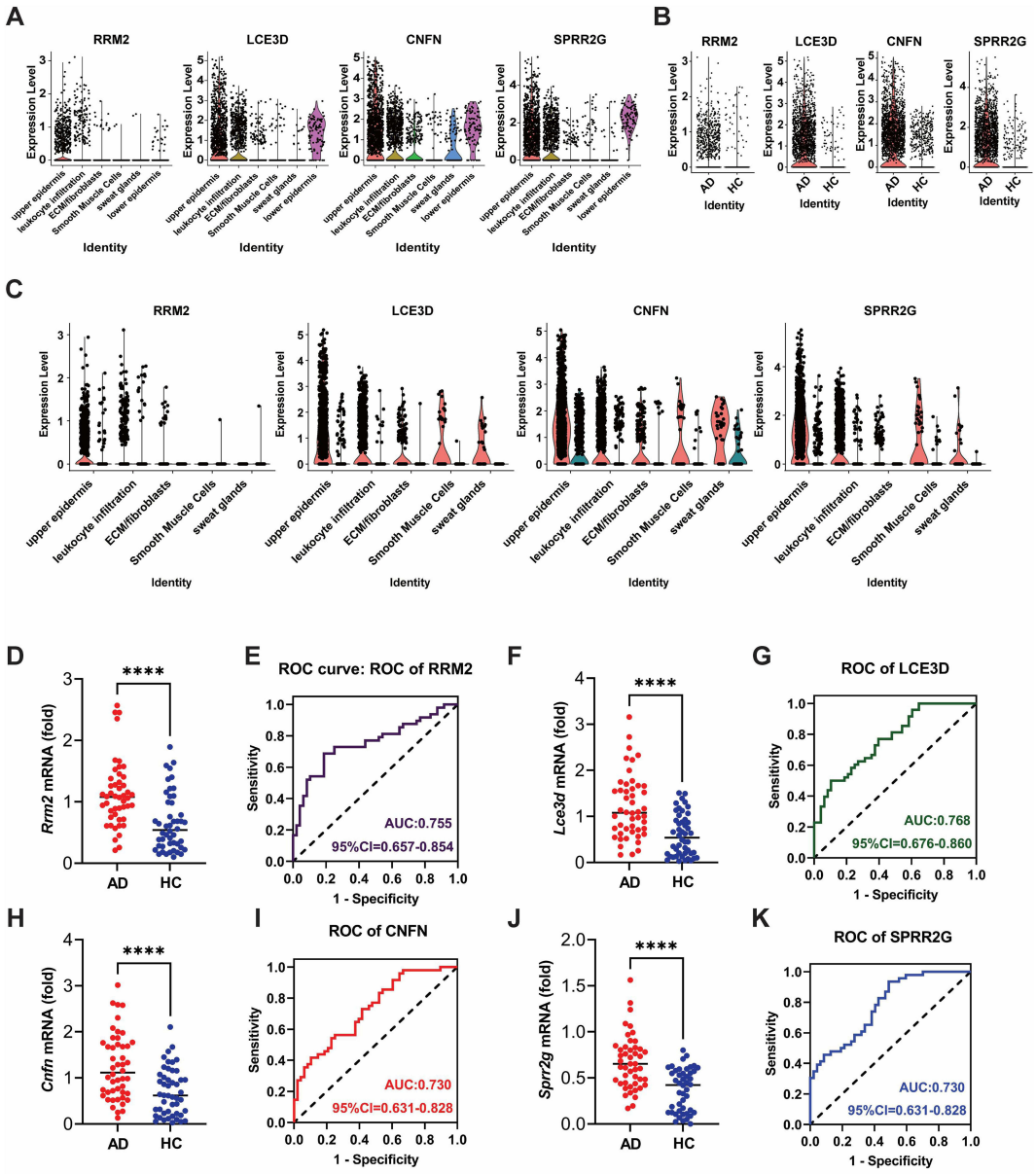

Figure 7 Continued.

https://doi.org/10.2147/JIR.S526983

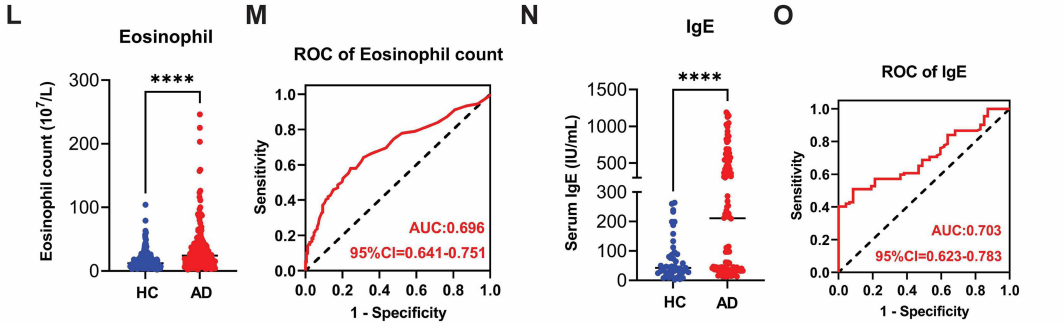

Figure 7 RRM2, LCE3D, CNFN and SPRR2G were highly expressed in multiple cell populations in AD patients. (**A**) Expression levels of RRM2, LCE3D, CNFN, and SPRR2G in different cell populations. (**B**) RRM2, LCE3D, CNFN and SPRR2G were highly expressed at the lesion site of AD patients. (**C**) High expression of RRM2, LCE3D, CNFN, and SPRR2G was observed in various cell populations at the lesion site of patients with AD. (**D**) RT-qPCR analysis was performed to measure RRM2 mRNA levels in human PBMCs from AD patients (n = 48) and healthy controls (n = 48). The relative expression of the RRM2 was normalized to GAPDH mRNA (two-tailed unpaired Student's *t*-test). (**E**) ROC analysis was conducted to evaluate the diagnostic performance of RRM2 expression in PBMCs from AD patients and the healthy control group. (**F**) RT-qPCR analysis was performed to measure LCE3D mRNA levels in human PBMCs from AD patients (n = 48) and healthy controls (n = 48). The relative expression of the LCE3D was normalized to GAPDH mRNA (two-tailed unpaired Student's *t*-test). (**G**) ROC analysis was conducted to evaluate the diagnostic performance of LCE3D expression in PBMCs from AD patients and the healthy control group. (**H**) RT-qPCR analysis was performed to measure CNFN mRNA levels in human PBMCs from AD patients (n = 48) and healthy controls (n = 48). The relative expression of the CNFN was normalized to GAPDH mRNA (two-tailed unpaired Student's *t*-test). (**I**) ROC analysis was conducted to evaluate the diagnostic performance of CNFN expression in PBMCs from AD patients and the healthy control group. (**J**) RT-qPCR analysis was performed to measure SPRR2G mRNA levels in human PBMCs from AD patients (n = 47) and healthy controls (n = 46). The relative expression of the SPRR2G was normalized to GAPDH mRNA (two-tailed unpaired Student's *t*-test). (**K**) ROC analysis was conducted to evaluate the diagnostic performance of SPRR2G expression in PBMCs from AD patients and the healthy control group. (**L**) The Mindray CAL8000 blood analyzer was used to measure the number of eosinophils in patients with AD (n = 181) and healthy controls (n = 172) (two-tailed unpaired Student's *t*-test). (**M**) ROC analysis was conducted to evaluate the diagnostic performance of the number of eosinophils from AD patients and the healthy control group. (**N**) The Maglumi4000 automatic chemiluminescence analyzer was used to measure the serum IgE levels in patients with AD (n = 112) and healthy controls (n = 47) (two-tailed unpaired Student's *t*-test). (**O**) ROC analysis was conducted to evaluate the diagnostic performance of the serum IgE levels from AD patients and the healthy control group. ****$p < 0.0001$.

May 30, 2024. We assessed the levels of serum IgE and peripheral blood eosinophils, as these biomarkers are commonly utilized in the diagnosis of AD. Our analysis demonstrated that both serum IgE and peripheral blood eosinophil levels were markedly elevated in AD patients (Figure 7L and N). ROC curve analysis yielded AUC values of 0.696 and 0.703 for serum IgE and eosinophils, respectively, which were lower than the AUC values obtained for RRM2, LCE3D, CNFN, and SPRR2G (Figure 7M and O).

Utilizing the relative expression levels of the genes RRM2, LCE3D, CNFN, and SPRR2G, we performed a euclidean distance clustering analysis on 59 samples (healthy skin from healthy individuals and diseased skin from AD patients) from the GSE121212 dataset, categorizing them into two optimally distinct groups, labeled as Group 1 and Group 2 (Figure 8A–C). Group 1 was characterized by uniformly low expression levels of RRM2, LCE3D, CNFN, and SPRR2G, whereas Group 2 was defined by high expression levels of these genes. Notably, Group 1 mainly comprised skin samples from healthy individuals, while all samples from AD lesions were allocated to Group 2 (Figure 8D). To corroborate these findings, we conducted a validation study using an independent dataset, GSE130588. Consistent with our initial results, the validation revealed that skin samples from healthy individuals were predominantly clustered in Group 1, which exhibited low gene expression, whereas AD lesion samples were mainly associated with Group 2, characterized by elevated gene expression (Figure 8E–H). In summary, the high expression of RRM2, LCE3D, CNFN, and SPRR2G in PBMCs and lesional skin tissue of AD patients underscores their potential utility as diagnostic biomarkers, effectively distinguishing between healthy individuals and AD patients.

RRM2, LCE3D, CNFN and SPRR2G Were Correlated with AD Severity

In the aforementioned study, we observed an upregulation of RRM2, LCE3D, CNFN, and SPRR2G expression in both PBMCs and lesional skin of AD patients, thereby affirming their potential diagnostic and therapeutic relevance. Subsequently, we undertook a series of investigations to explore the association between the expression levels of these four genes and the severity of AD. The SCORAD score, a widely utilized metric in clinical research and practice, was employed to assess AD severity.[26] In the GSE182740 dataset, SCORAD scores for each patient were meticulously documented. We conducted an analysis to determine the correlation between the expression levels of these four genes and

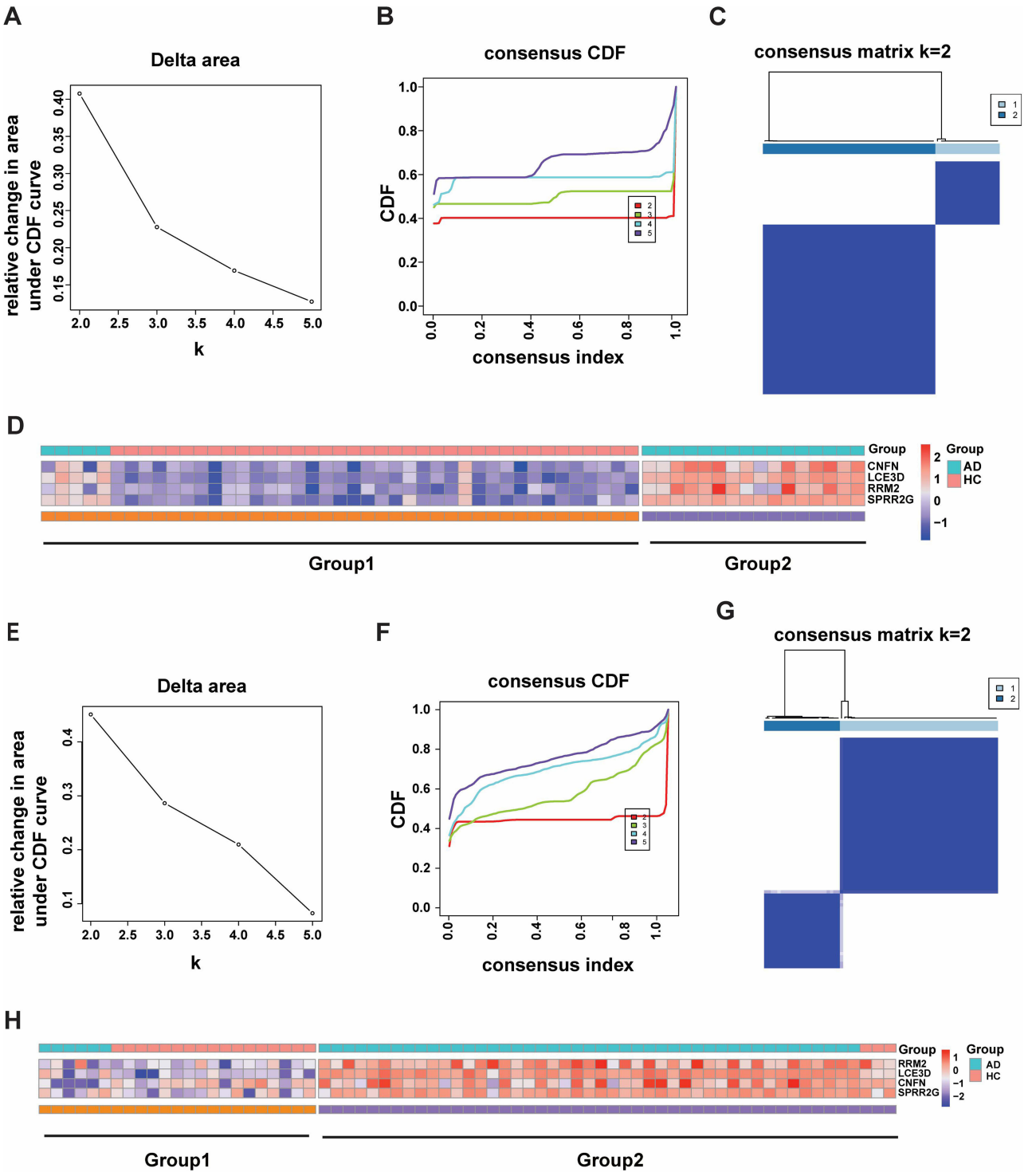

Figure 8 Distinguishing AD patients from healthy individuals based on RRM2, LCE3D, CNFN, and SPRR2G expression levels. (**A**) In the GSE121212 dataset, delta area curves for consensus clustering illustrate the relative change in the area under the cumulative distribution function (CDF) curve for each category number k compared to k-1. The horizontal axis represents the category number k, and the vertical axis represents the relative change in the area under the CDF curve. (**B**) Consensus among clusters for each category number k in the GSE121212 dataset. (**C**) In the GSE121212 dataset, all samples were categorized into two optimal groups based on the consensus clustering matrix (k=2). (**D**) Unsupervised hierarchical clustering was performed on the expression levels of RRM2, LCE3D, CNFN, and SPRR2G in the skin tissues in the GSE121212 dataset. In this heatmap, red indicates high expression levels of genes, while blue indicates low expression levels of genes. (**E**) Delta area curves for consensus clustering of validation dataset GSE130588. (**F**) Consensus among clusters for each category number k in the GSE130588 dataset. (**G**) In the GSE130588 dataset, all samples were categorized into two optimal groups based on the consensus clustering matrix (k=2). (**H**) Unsupervised hierarchical clustering was performed on the expression levels of RRM2, LCE3D, CNFN, and SPRR2G in the skin tissues in the GSE130588 dataset.

https://doi.org/10.2147/JIR.S526983

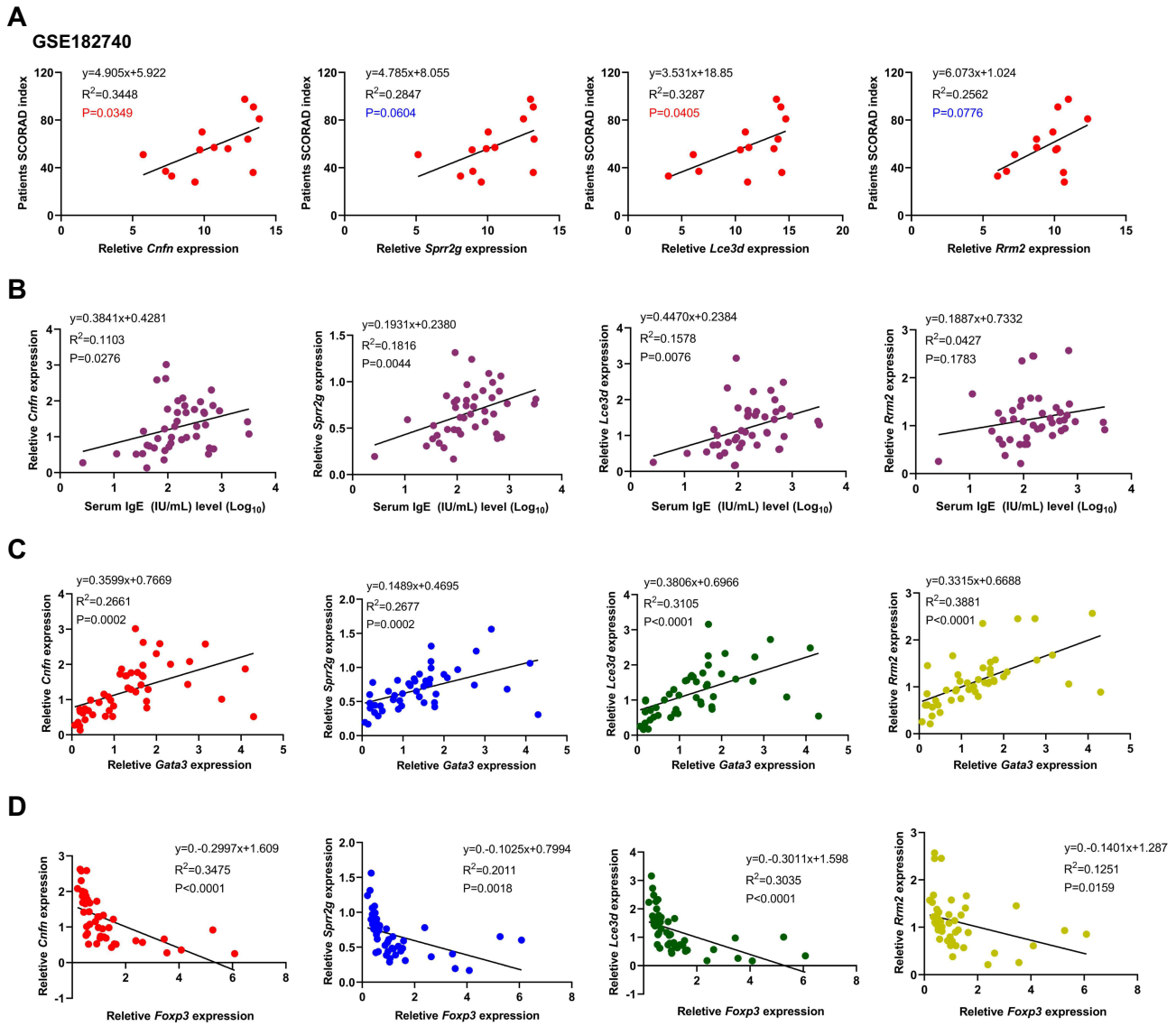

Figure 9 RRM2, LCE3D, CNFN and SPRR2G were correlated with AD severity. (**A**) The correlation between the expression levels of RRM2, LCE3D, CNFN, and SPRR2G in lesional skin tissues and SCORAD scores in patients with AD was analyzed using the GSE182740 dataset. (**B**) The correlation between the expression levels of RRM2, LCE3D, CNFN, and SPRR2G in PBMCs and serum IgE levels in patients with AD was analyzed. (**C**) An examination of the correlation between the expression levels of RRM2, LCE3D, CNFN, and SPRR2G in PBMCs of patients with AD and the expression level of Gata3 in these cells. (**D**) An examination of the correlation between the expression levels of RRM2, LCE3D, CNFN, and SPRR2G in PBMCs of patients with AD and the expression level of Foxp3 in these cells.

the patients' SCORAD scores. Our results indicated a statistically significant positive correlation between the expression of CNFN, LCE3D and the SCORAD scores of patients (P<0.05) (Figure 9A). Although the expression of RRM2 and SPRR2G also exhibited a positive correlation with SCORAD scores, this was not statistically significant. We speculate that the lack of statistical significance may be attributed to the limited sample size.

Serum IgE is a well-established serological biomarker indicative of AD severity. In this study, we investigated the correlation between the expression levels of RRM2, LCE3D, CNFN, and SPRR2G in PBMCs of AD patients and their serum IgE levels. Our findings demonstrated that LCE3D, CNFN, and SPRR2G were significantly positively correlated with serum IgE levels, whereas RRM2 exhibited a weak correlation (Figure 9B). T cells play a critical role in the pathogenesis and progression of AD. Specifically, activated Th2 cells secrete cytokines such as IL-4, IL-5, IL-13, and IL-31, which enhance B cell-mediated IgE production, thereby contributing to the pathological mechanisms underlying AD.[27] In contrast, regulatory T (Treg) cells are essential for modulating inflammation resulting from skin

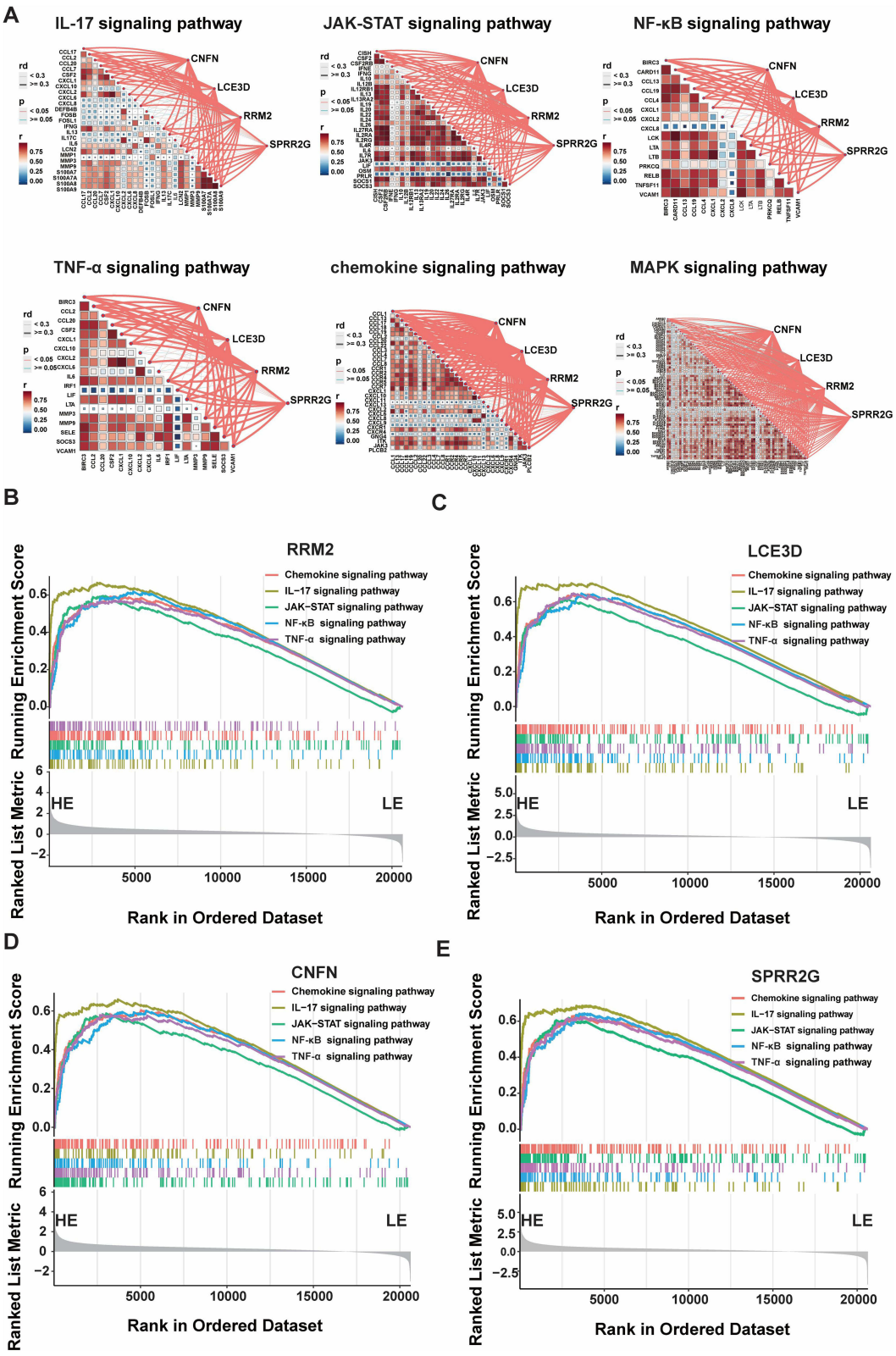

Figure 10 The expression of RRM2, LCE3D, CNFN, and SPRR2G was positively associated with the activation of multiple inflammatory pathways. (**A**) RRM2, LCE3D, CNFN and SPRR2G were positively correlated with most genes in the IL-17 signaling pathway, JAK-STAT signaling pathway, NF-κB signaling pathway, TNF-α signaling pathway, Chemokine signaling pathway and MAPK signaling pathway. (**B–E**) Overexpression of RRM2, LCE3D, CNFN or SPRR2G were involved in the activation of IL-17 signaling pathway, JAK-STAT signaling pathway, NF-κB signaling pathway, TNF-α signaling pathway, Chemokine signaling pathway and MAPK signaling pathway.
Abbreviations: HE, high expression; LE, low expression.

https://doi.org/10.2147/JIR.S526983

lesions and allergen-specific immune responses, serving as critical inhibitors of skin inflammation in AD patients.[28] The expression of GATA3, a pivotal transcription factor that facilitates Th2 cell differentiation, is increased in PBMCs of AD patients (Figure S6). Foxp3, which is integral to the differentiation of Treg cells, and its expression is decreased in PBMCs in AD patients (Figure S6). We investigated the correlation between the expression levels of RRM2, LCE3D, CNFN, and SPRR2G in PBMCs of AD patients and these key T cell transcription factors. Our analysis revealed that the expression levels of RRM2, LCE3D, CNFN, and SPRR2G were significantly positively correlated with GATA3, and significantly negatively correlated with Foxp3 (Figure 9C and D). In conclusion, our study indicates that LCE3D, CNFN, and SPRR2G are associated with multiple indicators of AD severity, suggesting their potential utility as biomarkers for assessing the severity of AD.

Four Potential AD Biomarkers are Closely Related to the Activation of Inflammatory Pathways

Inflammatory cytokines and chemokines play a crucial role in the pathogenesis and progression of AD.[29] As previously discussed, we identified the enrichment of several inflammatory pathways through KEGG and GO analyses. Subsequently, we investigated the association between four potential AD biomarkers and these enriched pathways. Our findings indicate that these biomarkers exhibit a positive correlation with the majority of genes within the IL-17, JAK-STAT, NF-κB, TNF-α, Chemokine, and MAPK signaling pathways (Figure 10A). Furthermore, we also revealed that the overexpression of RRM2, LCE3D, CNFN, and SPRR2G is implicated in the activation of these pathways (Figure 10B–E).

Discussion

AD is a chronic, recurrent inflammatory dermatosis characterized by pruritus and pleomorphic lesions with a propensity for exudation.[30] Approximately 50% of individuals with AD manifest the condition during infancy, with some patients experiencing recurrent episodes in adulthood.[31] The diagnosis of AD is primarily based on medical history, clinical manifestations, and histopathological examination through skin biopsy. In recent years, as research has advanced, numerous biomarkers have been identified as instrumental in the diagnosis of AD and the evaluation of disease activity. Serum IgE levels are typically elevated in individuals with AD; however, this marker lacks specificity due to its also elevation in other allergic conditions. Eosinophil counts are increased in both the skin and peripheral blood of patients with AD. Furthermore, certain cytokines and chemokines, including IL-4, IL-5, IL-13, TNF-α, and CXCL10, are upregulated in the skin and blood of these patients, thereby facilitating clinical diagnosis.[32] Additionally, abnormalities in the expression of skin barrier proteins such as filaggrin, involucrin, and loricrin are implicated in the pathogenesis of AD.[33–35] Despite the utility of these biomarkers in facilitating the diagnosis or assessment of disease activity in AD, there remains an absence of widely adopted AD-specific biomarkers in clinical practice. The current clinical diagnosis of AD necessitates a comprehensive evaluation of the patient's medical history and clinical manifestations.[36,37] Consequently, the development of reliable indicators for AD is essential for enhancing the accuracy of clinical diagnoses, reducing healthcare costs, and providing critical targets and evaluation metrics for novel drug development. In our study, we identified RRM2, LCE3D, CNFN, and SPRR2G as potential diagnostic biomarkers for AD. Notably, these biomarkers effectively differentiate between skin tissues and peripheral blood samples from AD patients and healthy individuals, as well as between lesional and non-lesional tissues in AD patients. Furthermore, we observed that LCE3D, CNFN, and SPRR2G exhibit strong correlations with various indicators reflecting the severity of AD, indicating their potential utility in assessing AD severity and evaluating the efficacy of therapeutic interventions.

Bioinformatics represents an interdisciplinary domain that integrates computer science, mathematics, statistics, biology, and other related fields. Its primary objective is to address biological challenges through the application of computational technology and algorithmic strategies.[38] The high-throughput data processing capabilities inherent in bioinformatic analysis facilitate the discovery of potential factors associated with disease onset and progression. Consequently, bioinformatics has become an essential tool in contemporary life science research. In this study, we employed a comprehensive array of bioinformatic analytical techniques to identify novel biomarkers for AD and to

explore their potential biological functions. Analysis of the AD transcriptome dataset, sourced from the GEO database, revealed 1371 DEGs between the skin tissues of AD patients and healthy controls, with 828 genes exhibiting upregulation in the AD group. Enrichment analyses, including GO, KEGG, and GSEA, indicated the activation of numerous inflammatory signaling pathways and aberrant T cell differentiation within the skin tissues of AD patients. We conducted an analysis of a single-cell sequencing dataset from patients with AD and identified 592 DEGs. Utilizing WGCNA and hdWGCNA, we identified co-expression gene modules associated with AD, ultimately pinpointing 57 DEGs with strong associations to the condition. The application of machine learning algorithms in the analysis of high-throughput data enhances both the efficiency and accuracy of data interpretation, serving as a robust tool for uncovering novel biological processes and protein functions. We employed three machine learning algorithms—LASSO, RF, and Boruta—to identify RRM2, LCE3D, CNFN, and SPRR2G as potential biomarkers for AD. By integrating multiple transcriptome datasets, we conclusively identified RRM2, LCE3D, CNFN, and SPRR2G as AD biomarkers with diagnostic potential.

T cells are pivotal in the pathogenesis of AD. Activated Th2 cells secrete interleukins IL-4, IL-5, IL-13, and IL-31, which facilitate IgE production by B cells, thereby exacerbating the pathological progression of AD.[27] Conversely, Treg cells are crucial in modulating inflammation resulting from skin injury and allergen-specific immune responses, serving as significant inhibitors of skin inflammation in AD patients.[28] The abundance of these cell types is often associated with AD severity; an increase in Th2 cells correlates with heightened disease severity, whereas an increase in Treg cells is typically indicative of symptom amelioration. In our study, we observed that the expression levels of RRM2, LCE3D, CNFN, and SPRR2G were positively associated with GATA3 (a Th2 cell transcription factor) and negatively associated with Foxp3 (a Treg cell transcription factor) in the PBMCs of AD patients. The SCORAD score is a clinical tool employed to assess AD severity, recognized as the most effective clinical measure for reflecting AD severity, and is extensively utilized in both clinical evaluations and scientific research pertaining to AD. Our study demonstrated a significant positive correlation between the relative expression levels of LCE3D and CNFN in the lesional skin tissue of AD patients and their SCORAD scores. The elevation of serum IgE levels serves as a crucial diagnostic marker for AD and is positively associated with the disease's severity. Furthermore, we observed that the relative expression levels of LCE3D, CNFN, and SPRR2G in PBMCs of AD patients were significantly positively correlated with serum IgE levels. These findings suggest that LCE3D, CNFN, and SPRR2G may serve as biomarkers indicative of AD severity.

Inflammatory signaling pathways are markedly activated in AD, leading to an upregulation in the expression of cytokines, including IL-4, IL-5, IL-13, and IL-22, as well as chemokines such as CCL2, CCL17, and CCL22. This upregulation facilitates the activation and migration of inflammatory cells, thereby exacerbating the inflammatory response in the skin. Correlation analysis and GSEA analysis have identified that RRM2, LCE3D, CNFN, SPRR2G were significantly positively correlated with key genes in several inflammatory signaling pathways, such as IL-17, JAK-STAT, NF-κB, TNF-α, Chemokine, and MAPK signaling pathways. The overexpression of these four genes is implicated in the activation of these signaling pathways. These findings suggest that RRM2, LCE3D, CNFN, and SPRR2G may play a critical role in the pathogenesis and progression of AD, highlighting their potential as therapeutic targets for the treatment of this condition.

RRM2 is a subunit of ribonucleotide reductase (RNR), it is involved in the generation of deoxynucleotides during DNA synthesis and repair. It is integral to cellular growth, proliferation, and DNA repair mechanisms. Notably, RRM2 is overexpressed in various cancers, facilitating tumor cell proliferation, invasion, and resistance to chemotherapy. Consequently, RRM2 is considered a potential biomarker for liver cancer, lung adenocarcinoma, breast cancer, and other malignancies.[39] Recent research has demonstrated elevated RRM2 expression in PBMCs of patients with rheumatoid arthritis, suggesting its potential diagnostic utility in this condition.[40] Furthermore, serum RRM2 levels may serve as a biomarker for diagnosing and monitoring liver fibrosis in chronic hepatitis B.[41] In this study, we report that RRM2 is significantly overexpressed in the skin tissue and PBMCs of patients with AD, indicating its diagnostic value. Additionally, GSEA revealed a positive correlation between RRM2 overexpression and the activation of multiple inflammatory pathways. Thus, RRM2 may play a crucial role in the pathogenesis and progression of AD and represents a potential therapeutic target for this condition.

LCE3D, a constituent of the Late Cornified Envelope (LCE) family, encodes structural proteins in keratinocytes and is predominantly expressed in the epithelial cells of the skin and mucous membranes. Peng et al demonstrated that

LCE3D is significantly upregulated in the lesional skin tissue of patients with AD, suggesting its utility as a potential diagnostic biomarker for AD skin tissue.[42] Building on this foundation, our research identified elevated expression levels of LCE3D in the PBMCs of AD patients, proposing its potential as a diagnostic biomarker for AD in peripheral blood. Furthermore, LCE3D expression is strongly associated with various clinical indicators reflecting AD severity, indicating its prospective value as a biomarker for assessing disease severity. Collectively, these findings underscore the potential diagnostic significance of LCE3D in AD.

CNFN and SPRR2G are predominantly expressed in human skin, yet their biological functions remain underexplored. Some studies have indicated that CNFN and SPRR2G are highly expressed in the lesional skin of psoriasis patients.[43,44] However, to date, no investigations have been conducted on the association between CNFN, SPRR2G, and AD. Our research identified an upregulation of CNFN and SPRR2G in AD skin tissues and PBMCs. This upregulation may contribute to the chronic inflammation characteristic of AD, as the overexpression of CNFN and SPRR2G is linked to the activation of multiple inflammatory pathways. Nonetheless, this hypothesis requires validation through further research.

This study has elucidated the potential biological functions of RRM2, LCE3D, CNFN, and SPRR2G in AD. Nonetheless, due to inherent limitations, our sample size of AD patients was restricted. Future research should focus on conducting larger-scale, multicenter studies to comprehensively assess the diagnostic potential of RRM2, LCE3D, CNFN, and SPRR2G within the context of AD. Furthermore, while our findings underscore the diagnostic significance and suggest potential biological mechanisms involving these biomarkers in AD, their specific roles in the pathogenesis and progression of the disease warrant further investigation.

Conclusion

In this study, we conducted an analysis of AD transcriptomes and single-cell sequencing datasets. Through an integrated bioinformatics approach, we identified RRM2, LCE3D, CNFN and SPRR2G as potential diagnostic indicators for AD patients and validated their diagnostic efficacy using multiple external datasets. In addition, we found that the expression of RRM2, LCE3D, and CNFN in the lesional skin of AD patients after treatment was significantly decreased, which is of diagnostic value. Subsequent analyses revealed correlations between LCE3D, CNFN, and SPRR2G with several indicators that directly reflect AD severity. Collectively, these findings demonstrate that LCE3D and CNFN possess not only diagnostic value for AD but also the capacity to indicate disease severity and reflect clinical therapeutic outcomes. Additionally, we discovered that the overexpression of RRM2, LCE3D, CNFN, and SPRR2G is positively associated with the activation of multiple inflammatory pathways. In conclusion, our research identifies novel diagnostic markers in peripheral blood and lesional skin tissue for AD, elucidates the potential of these biomarkers in assessing AD severity, and proposes innovative targets for the diagnosis and treatment of AD.

Data Sharing Statement

The datasets supporting the conclusions of this article are available in the GEO database, https://www.ncbi.nlm.nih.gov/geo/. Further inquiries can be directed to the corresponding author(s).

Ethics Approval and Consent to Participate

The study was carried out in strict compliance with a protocol (approval number: 2024-Y-205) that had been approved by the Ethics Committee of Jiangnan University Medical Center, and informed consent was obtained from all participants involved. All participant data were anonymized and deidentified prior to analysis. Given that this study was observational rather than a clinical trial, it was not registered in a clinical trials registry.

Acknowledgments

Thanks to the lab members for valuable discussions and suggestions.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically

reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This work was supported by the Foundation for the Natural Science Foundation of China (82270034), Jiangsu Specially-Appointed Professor (1286010241241040), Top Talent Support Program for young and middle-aged people of Wuxi Health Committee (BJ2023028), the "Taihu Light" technical research basic project (K20231058), Boxi Youth Natural Science Foundation (2023012), Suzhou Applied Basic Research (Medical and Health) (SYW2024076).

Disclosure

The authors declare no conflicts of interest in this work.

References

1. Weidinger S, Novak N. Atopic dermatitis. *Lancet*. 2016;387(10023):1109–1122. doi:10.1016/s0140-6736(15)00149-x
2. Weidinger S, Beck LA, Bieber T, Kabashima K, Irvine AD. Atopic dermatitis. *Nat Rev Dis Primers*. 2018;4(1):1. doi:10.1038/s41572-018-0001-z
3. Torres T, Mendes-Bastos P, Cruz MJ, et al. Interleukin-4 and Atopic Dermatitis: why Does it Matter? A Narrative Review. *Dermatol Ther*. 2025;15(3):579–597. doi:10.1007/s13555-025-01352-y
4. Sargen M, Sasaki A, Maskey AR, Li XM. Biomarkers to aid in diagnosis of allergic contact dermatitis. *Front Allergy*. 2025;6:1564588. doi:10.3389/falgy.2025.1564588
5. Sroka-Tomaszewska J, Trzeciak M. Molecular Mechanisms of Atopic Dermatitis Pathogenesis. *Int J Mol Sci*. 2021;22(8):4130. doi:10.3390/ijms22084130
6. Napolitano M, Fabbrocini G, Martora F, Genco L, Noto M, Patruno C. Children atopic dermatitis: diagnosis, mimics, overlaps, and therapeutic implication. *Dermatol Ther*. 2022;35(12):e15901. doi:10.1111/dth.15901
7. Fishbein AB, Silverberg JI, Wilson EJ, Ong PY. Update on Atopic Dermatitis: diagnosis, Severity Assessment, and Treatment Selection. *J Allergy Clin Immunol Pract*. 2020;8(1):91–101. doi:10.1016/j.jaip.2019.06.044
8. Yu L. Potential biomarkers of atopic dermatitis. *Front Med Lausanne*. 2022;9:1028694. doi:10.3389/fmed.2022.1028694
9. Gomułka K, Wójcik E, Szepietowski JC. Serum Levels of Eosinophil-Derived Neurotoxin, Platelet-Activating Factor and Vascular Endothelial Growth Factor in Adult Patients with Atopic Dermatitis—A Pilot Study. *Biomedicines*. 2022;10(12):3109. doi:10.3390/biomedicines10123109
10. Hs K, jh K, Seo YM, et al. Eosinophil-derived neurotoxin as a biomarker for disease severity and relapse in recalcitrant atopic dermatitis. *Ann Allergy Asthma Immunol*. 2017;119(5):441–445. doi:10.1016/j.anai.2017.06.022
11. Nakahara T, Onozuka D, Nunomura S, et al. The ability of biomarkers to assess the severity of atopic dermatitis. *J Allergy Clin Immunol Glob*. 2024;3(1):100175. doi:10.1016/j.jacig.2023.100175
12. Nagao M, Inagaki S, Kawano T, et al. SCCA2 is a reliable biomarker for evaluating pediatric atopic dermatitis. *J Allergy Clin Immunol*. 2018;141(5):1934–1936.e11. doi:10.1016/j.jaci.2018.01.021
13. Mastraftsi S, Vrioni G, Bakakis M, et al. Atopic Dermatitis: striving for Reliable Biomarkers. *J Clin Med*. 2022;11(16):4639. doi:10.3390/jcm11164639
14. Halling AS, Rinnov MR, Ruge IF, et al. Skin TARC/CCL17 increase precedes the development of childhood atopic dermatitis. *J Allergy Clin Immunol*. 2023;151(6):1550–1557.e6. doi:10.1016/j.jaci.2022.11.023
15. Asahina R, Ueda K, Oshima Y, et al. Serum canine thymus and activation-regulated chemokine (TARC/CCL17) concentrations correlate with disease severity and therapeutic responses in dogs with atopic dermatitis. *Vet Dermatol*. 2020;31(6):446–455. doi:10.1111/vde.12894
16. Liu Q, Yang S, Tan Y, Cui L. High-throughput sequencing technology facilitates the discovery of novel biomarkers for antiphospholipid syndrome. *Front Immunol*. 2023;14:1128245. doi:10.3389/fimmu.2023.1128245
17. Komorowski M, Green A, Tatham KC, Seymour C, Antcliffe D. Sepsis biomarkers and diagnostic tools with a focus on machine learning. *EBioMedicine*. 2022;86:104394. doi:10.1016/j.ebiom.2022.104394
18. Ps R, Reel S, Pearson E, Trucco E, Jefferson E. Using machine learning approaches for multi-omics data analysis: a review. *Biotechnol Adv*. 2021;49:107739. doi:10.1016/j.biotechadv.2021.107739
19. Balakrishnan V, Kherabi Y, Ramanathan G, Paul SA, Tiong CK. Machine learning approaches in diagnosing tuberculosis through biomarkers - A systematic review. *Prog Biophys Mol Biol*. 2023;179:16–25. doi:10.1016/j.pbiomolbio.2023.03.001
20. Sasaki A, Sargen M, Maskey AR, Li XM. Scratching the surface: biomarkers and neurobiomarkers for improved allergic contact dermatitis management. *Front Allergy*. 2025;6:1564528. doi:10.3389/falgy.2025.1564528
21. Zhou Y, Wang Z, Han L, et al. Machine learning-based screening for biomarkers of psoriasis and immune cell infiltration. *Eur J Dermatol*. 2023;33(2):147–156. doi:10.1684/ejd.2023.4453
22. Xiao K, Wang S, Chen W, et al. Identification of novel immune-related signatures for keloid diagnosis and treatment: insights from integrated bulk RNA-seq and scRNA-seq analysis. *Hum Genomics*. 2024;18(1):80. doi:10.1186/s40246-024-00647-z
23. Group ADW, Group I, Dermatology CSO, Dasgeb B, Leila Y, Saeidian AH, et al. Guidelines for Diagnosis and Treatment of Atopic Dermatitis in China (2020)#. *Int J Dermatol Venereol*. 2021;4(1):70–75. doi:10.1097/JD9.0000000000000170.
24. Morabito S, Reese F, Rahimzadeh N, Miyoshi E, Swarup V. hdWGCNA identifies co-expression networks in high-dimensional transcriptomics data. *Cell Rep Methods*. 2023;3(6):100498. doi:10.1016/j.crmeth.2023.100498
25. Sun Y, Wu J, Zhang Q, Wang P, Zhang J, Yuan Y. Single-cell hdWGCNA reveals metastatic protective macrophages and development of deep learning model in uveal melanoma. *J Transl Med*. 2024;22(1):695. doi:10.1186/s12967-024-05421-2

26. European Task Force on Atopic Dermatitis. Severity scoring of atopic dermatitis: the SCORAD index. Consensus Report of the European Task Force on Atopic Dermatitis. *Dermatology*. 1993;186(1):23–31. doi:10.1159/000247298
27. Akdis CA, Arkwright PD, Brüggen MC, et al. Type 2 immunity in the skin and lungs. *Allergy*. 2020;75(7):1582–1605. doi:10.1111/all.14318
28. Ma X, Deng G, Tian N, et al. Calycosin enhances Treg differentiation for alleviating skin inflammation in atopic dermatitis. *J Ethnopharmacol*. 2024;326:117883. doi:10.1016/j.jep.2024.117883
29. Nedoszytko B, Sokołowska-Wojdyło M, Ruckemann-Dziurdzińska K, Roszkiewicz J, Nowicki RJ. Chemokines and cytokines network in the pathogenesis of the inflammatory skin diseases: atopic dermatitis, psoriasis and skin mastocytosis. *Postepy Dermatol Alergol*. 2014;31(2):84–91. doi:10.5114/pdia.2014.40920
30. Kim J, Be K, Leung DYM. Pathophysiology of atopic dermatitis: clinical implications. *Allergy Asthma Proc*. 2019;40(2):84–92. doi:10.2500/aap.2019.40.4202
31. Chittock J, Kay L, Brown K, et al. Association between skin barrier development and early-onset atopic dermatitis: a longitudinal birth cohort study. *J Allergy Clin Immunol*. 2024;153(3):732–741.e8. doi:10.1016/j.jaci.2023.10.017
32. Renert-Yuval Y, Thyssen JP, Bissonnette R, et al. Biomarkers in atopic dermatitis-a review on behalf of the International Eczema Council. *J Allergy Clin Immunol*. 2021;147(4):1174–1190.e1. doi:10.1016/j.jaci.2021.01.013
33. Moosbrugger-Martinz V, Leprince C, Méchin MC, et al. Revisiting the Roles of Filaggrin in Atopic Dermatitis. *Int J Mol Sci*. 2022;23(10):5318. doi:10.3390/ijms23105318
34. Furue M. Regulation of Filaggrin, Loricrin, and Involucrin by IL-4, IL-13, IL-17A, IL-22, AHR, and NRF2: pathogenic Implications in Atopic Dermatitis. *Int J Mol Sci*. 2020;21(15):5382. doi:10.3390/ijms21155382
35. Bakker D, de Bruin-Weller M, Drylewicz J, van Wijk F, Thijs J. Biomarkers in atopic dermatitis. *J Allergy Clin Immunol*. 2023;151(5):1163–1168. doi:10.1016/j.jaci.2023.01.019
36. Maliyar K, Sibbald C, Pope E, Gary Sibbald R. Diagnosis and Management of Atopic Dermatitis: a Review. *Adv Skin Wound Care*. 2018;31(12):538–550. doi:10.1097/01.ASW.0000547414.38888.8d
37. Davis DM, Waldman A, Jacob S, et al. Diagnosis, comorbidity, and psychosocial impact of atopic dermatitis. *Semin Cutan Med Surg*. 2017;36(3):95–99. doi:10.12788/j.sder.2017.028
38. Zhang Y, Cheng L, Chen G, Alghazzawi D. Evolutionary Computation in bioinformatics: a survey. *Neurocomputing*. 2024;591:127758. doi:10.1016/j.neucom.2024.127758
39. Zuo Z, Zhou Z, Chang Y, et al. Ribonucleotide reductase M2 (RRM2): regulation, function and targeting strategy in human cancer. *Genes Dis*. 2024;11(1):218–233. doi:10.1016/j.gendis.2022.11.022
40. Wu B, Peng R, Chen Y, et al. Identification of RRM2 in peripheral blood mononuclear cells as a novel biomarker for the diagnosis of rheumatoid arthritis. *Clin Exp Rheumatol*. 2022;40(11):2109–2118. doi:10.55563/clinexprheumatol/ixdaek
41. Zhan Y, Tao Q, Lang Z, et al. Serum ribonucleotide reductase M2 is a potential biomarker for diagnosing and monitoring liver fibrosis in chronic hepatitis B patients. *J Med Virol*. 2023;95(10):e29157. doi:10.1002/jmv.29157
42. Peng S, Chen M, Yin M, Feng H. Identifying the Potential Therapeutic Targets for Atopic Dermatitis Through the Immune Infiltration Analysis and Construction of a ceRNA Network. *Clin Cosmet Invest Dermatol*. 2021;14:437–453. doi:10.2147/ccid.S310426
43. Michibata H, Chiba H, Wakimoto K, et al. Identification and characterization of a novel component of the cornified envelope, cornifelin. *Biochem Biophys Res Commun*. 2004;318(4):803–813. doi:10.1016/j.bbrc.2004.04.109
44. Zhen Y, Li X, Huang S, et al. LncRNA lnc-SPRR2G-2 contributes to keratinocyte hyperproliferation and inflammation in psoriasis by activating the STAT3 pathway and downregulating KHSRP. *Mol Cell Probes*. 2024;76:101967. doi:10.1016/j.mcp.2024.101967