Yuan B. (2024) MACHINE LEARNING-BASED IDENTIFICATION OF BIOMARKERS AND IMMUNE INFILTRATION IN PEDIATRIC TELOMERE-ASSOCIATED ALLERGIC RHINITIS AND ASTHMA: IMPLICATIONS FOR RESPIRATORY FUNCTION AND PHYSICAL PERFORMANCE. Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte vol. 24 (96) pp. 585- 604.
DOI: https://doi.org/10.15366/rimcafd2024.96.036

ORIGINAL

MACHINE LEARNING-BASED IDENTIFICATION OF BIOMARKERS AND IMMUNE INFILTRATION IN PEDIATRIC TELOMERE-ASSOCIATED ALLERGIC RHINITIS AND ASTHMA: IMPLICATIONS FOR RESPIRATORY FUNCTION AND PHYSICAL PERFORMANCE

Xinyue Zhang, Yunjiao Jing, Demei Wu, Kang Lian, Zijian Shao, Bin Yuan*

Department of Pediatrics, Affiliated Hospital of Nanjing University of Chinese Medicine, Nanjing, CN 210028, Jiangsu Province, China.
E-mail: yuanbin68358@163.com
E-mail: cicelyxyz@163.com

Recibido 23 de octubre de 2023 **Received** October 23, 2023
Aceptado 22 de puede de 2024 **Accepted** May 22, 2024

ABSTRACT

Background: Childhood Allergic Rhinitis and Asthma Syndrome (CARAS) is a chronic inflammatory condition affecting both the upper and lower airways, significantly impairing respiratory function, exercise capacity, and physical performance in children. The syndrome is closely linked to immune dysregulation, yet the biomarkers associated with its pathogenesis remain inadequately explored. Identifying key genetic and immune-related markers is crucial for improving early diagnosis, targeted interventions, and the management of physical activity limitations in affected children. **Objective:** This study aimed to identify telomere-associated biomarkers of CARAS in children using machine learning approaches and to evaluate the role of immune cell infiltration in disease progression, with potential implications for optimizing respiratory function, physical fitness, and sports participation. **Methods:** Gene expression data from the GSE19187 dataset was retrieved from the Gene Expression Omnibus (GEO) database for pooled analysis. Differentially expressed genes (DEGs) were identified using the 'limma' package in R software. Least Absolute Shrinkage and Selection Operator (LASSO) regression and Support Vector Machine Recursive Feature Elimination (SVM-

RFE) algorithms were applied to detect key candidate biomarkers. Diagnostic accuracy was assessed using Receiver Operating Characteristic (ROC) curve analysis, while CIBERSORT was utilized to quantify the immune cell infiltration in CARAS samples. Correlations between immune cells and identified biomarkers were analyzed using Pearson's correlation test. **Results:** A total of five telomere-associated genes (ABCC3, NT5DC2, RRAGD, ST6GAL1, and YBX3) were identified as potential biomarkers of CARAS using machine learning algorithms. Significant correlations were observed between these genes, with NT5DC2 emerging as the most prominent key gene. ROC analysis confirmed strong diagnostic efficacy for these biomarkers. Immune cell infiltration analysis revealed distinct immunological profiles associated with CARAS, highlighting their role in disease pathogenesis and potential implications for exercise-induced respiratory adaptation in children with asthma. **Conclusions:** The study identifies ABCC3, NT5DC2, RRAGD, ST6GAL1, and YBX3 as potential biomarkers of CARAS, with NT5DC2 being the most critical. The ferroptosis pathway emerged as the most significantly enriched signaling pathway, suggesting a novel target for therapeutic intervention. These findings have direct implications for the development of exercise-based rehabilitation strategies, as optimizing immune response and respiratory function in children with CARAS may enhance aerobic capacity, endurance, and participation in physical activities. Future research should explore the role of these biomarkers in exercise-induced airway modulation and personalized rehabilitation programs for children with respiratory conditions.

KEYWORDS: Childhood Allergic Rhinitis Asthma Syndrome (CARAS); Allergic Rhinitis; Asthma; Immune Factors Biomarkers; Machine Learning; GSE19187 Dataset; Gene Expression Omnibus (GEO); Telomere Related Genes; ABCC3; NT5DC2; RRAGD; ST6GAL1; YBX3

1. INTRODUCTION

Childhood Allergic Rhinitis and Asthma Syndrome (CARAS) is a chronic inflammatory disorder that simultaneously affects the upper and lower airways, leading to persistent symptoms such as nasal congestion, sneezing, airway hyperresponsiveness, and episodic bronchoconstriction. It is one of the most prevalent respiratory conditions in children, with a significant impact on pulmonary function, daily activities, and physical performance. The growing prevalence of childhood asthma and allergic rhinitis has raised concerns about their long-term effects on aerobic capacity, sports participation, and overall physical well-being, necessitating more in-depth research into their pathogenesis and molecular underpinnings (Anderson, 2008; Komlósi et al., 2022; Michaeloudes et al., 2022). CARAS is characterized by a complex interplay of genetic predisposition, immune dysregulation, environmental triggers, and airway inflammation. Studies have shown that prolonged airway inflammation leads to airway remodeling, increased airway resistance, and

reduced exercise tolerance, which can negatively affect cardiorespiratory fitness and physical activity levels in children. Given that physical activity plays a crucial role in respiratory health and overall fitness, children with CARAS often experience exercise-induced bronchospasms, reduced oxygen uptake efficiency, and limited participation in physical activities or sports. Understanding the molecular basis of CARAS-associated immune dysregulation and telomere-related genetic factors may offer insights into potential therapeutic targets and rehabilitation strategies that can enhance lung function, immune resilience, and sports performance in affected children (Abu-Shaheen et al., 2016; Platts-Mills, 2015). Telomeres, the protective DNA-protein structures at the ends of chromosomes, are crucial for maintaining genomic stability, cellular lifespan, and immune cell function. Emerging evidence suggests that telomere attrition is closely linked to chronic inflammatory diseases, including asthma and allergic rhinitis. Shortened telomeres have been associated with increased oxidative stress, immune cell dysfunction, and heightened airway inflammation, all of which contribute to disease severity and impaired respiratory function in children with CARAS. Identifying telomere-associated biomarkers can provide valuable insights into disease progression, treatment response, and rehabilitation strategies aimed at improving respiratory health and physical endurance. Recent advancements in machine learning (ML) techniques have revolutionized the field of biomedical research, enabling the discovery of novel biomarkers with high accuracy and reliability. ML-based algorithms, such as Least Absolute Shrinkage and Selection Operator (LASSO) regression and Support Vector Machine Recursive Feature Elimination (SVM-RFE), have proven effective in identifying key genetic signatures associated with various diseases (Bousquet et al., 2008). These techniques allow researchers to filter high-dimensional genomic data, eliminate redundant features, and pinpoint essential biomarkers that could serve as potential therapeutic targets (Ozdoganoglu & Songu, 2012). In the context of CARAS, ML-driven biomarker discovery can help predict disease severity, immune cell infiltration patterns, and response to physical activity interventions (Brożek et al., 2017). Furthermore, integrating immune infiltration analysis using computational tools like CIBERSORT provides a deeper understanding of immune cell dynamics in airway inflammation (Y. Chen et al., 2020; Zhang et al., 2013). This knowledge can be utilized to design personalized rehabilitation programs that enhance immune resilience, optimize pulmonary function, and improve exercise performance in children with CARAS. (Nappi et al., 2022). Some investigators have proposed that CARAS refers to a syndrome characterized by allergic rhinitis (AR) and asthma (Ferreira et al., 2019). However, only limited studies have been carried out to investigate the underlying molecular immune mechanism implicated in the development of CARAS till now. Identifying DGEs could help to elucidate the mechanism of childhood CARAS pathogenesis and develop new immunotherapies (Lucena-Zurita et al., 2022). While machine learning is widely used in many studies as

a practical and efficient aid for investigating important cell types and diagnostic markers and can be employed to identify the characteristics of relevant biomarkers as well as classify and validate these biomarkers (Jin et al., 2022; Waljee et al., 2022), to the best of our knowledge, it has not yet been utilized for the identification of biomarkers associated with CARAS in children (Pérez-Vigo et al., 2022). In the present study, the biomarkers of allergic rhinitis combined with asthma were screened in CARAS children via LASSO regression and SVM-PRE algorithm, and the degree of immune infiltration was further evaluated in these children, providing an opportunity for more precise diagnosis in children with CARAS. First, relevant data were obtained from the GEO for GO function enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis. Then, LASSO logistic regression algorithm and SVM-RFE feature ranking were employed for dimensionality reduction and screening of DEGs, accurate biomarkers were obtained. Finally, CIBERSORT was used to measure the infiltration of immune cells, and Pearson correlation analysis was used to examine the correlation between immune cells and biomarkers. This study provides a new perspective for researches on immune molecular mechanisms of CARAS.

2. Materials and Methods

2.1 Selecting and Preprocessing Data

The GSE19187 dataset was obtained from GEO. Data were from 11 healthy controls and 6 children with allergic rhinitis and concomitant uncontrolled asthma (UA).

2.2 Data Pre-Processing

GSE19187 was obtained with the 'limma' package in R software. The probe expression matrix was converted to the gene expression matrix using a platform annotation file. The Robust Multi-Array Average (RMA) was used to normalize the expression matrix of array data. In cases where more than one probe corresponded to a single gene, a mean value was calculated to obtain a representative expression value of that gene.

2.3 Assay of Differentially Expressed Genes

A DEG analysis was conducted in 11 healthy controls and 6 children with UA. The "limma" package in R was used for screening of DEGs between 11 healthy and 6 UA children. The criteria for screening of DEGs were: P. adj (P. adjust, corrected $P<0.05$, and $|\log2FC|>1$. ".

The "pheatmap" package and "ggplot2" package were used to accomplish the visualization of the volcanic map of DEG expression with log2FC as the abscissa and - log10 (P.adj) as the ordinate.

2.4 Functional Enrichment Analysis of DEGs

"Cluster Profiler" package was used for GO functional enrichment and KEGG enrichment analysis. GO analysis incorporates three main components: biological processes, cellular components, and molecular functions. Benjamin Hochberg FDR was employed for adjusting the p-values to account for multiple hypothesis testing. FDR<0.05 indicates statistically significant difference in the enrichment analysis and visualization was done using R software.

2.5 Screening of Candidate Diagnostic Biomarkers

In this study, two machine learning models, were applied to select important prognostic variables. LASSO logistic regression algorithm was used to analyze DEG, and the feature ranking of SVM-RFE was used to further analyze the DEGs selected with LASSO algorithm, aiming to improve the accuracy of identified diagnostic biomarkers, and these biomarkers were finally obtained. The "pROC" package was used to plot the ROC based on the data of 11 healthy and 6 allergic rhinitis and concomitant uncontrolled asthma children. The AUC was calculated for the evaluation of diagnostic value of these biomarkers.

2.6 Immune Cell Infiltration Analysis by CIBERSORT

The network tool CIBERSORT was used to evaluate immune cell infiltration in normal and UA samples. The reference set was 22 immune cell genes(LM22). The "corrplot" R package was used to visualize 22 infiltrated immune cells and analyze the correlation. The number of permutations was 1000. The "ggplot2" package in the R software was used to visualize the results. Pearson correlation analysis was utilized to determine the correlation coefficient between immune cells.

2.6 Potential Therapeutic Drugs for CARAS in Children

DGIdb database (a database related to gene - drug interaction) (https://www.dgidb.org/) was employed to identify potential therapeutic drugs based on the core genes. The KnockTF's online website was used to predict the transcription factors of hub gene. The predicted results were displayed in tables.

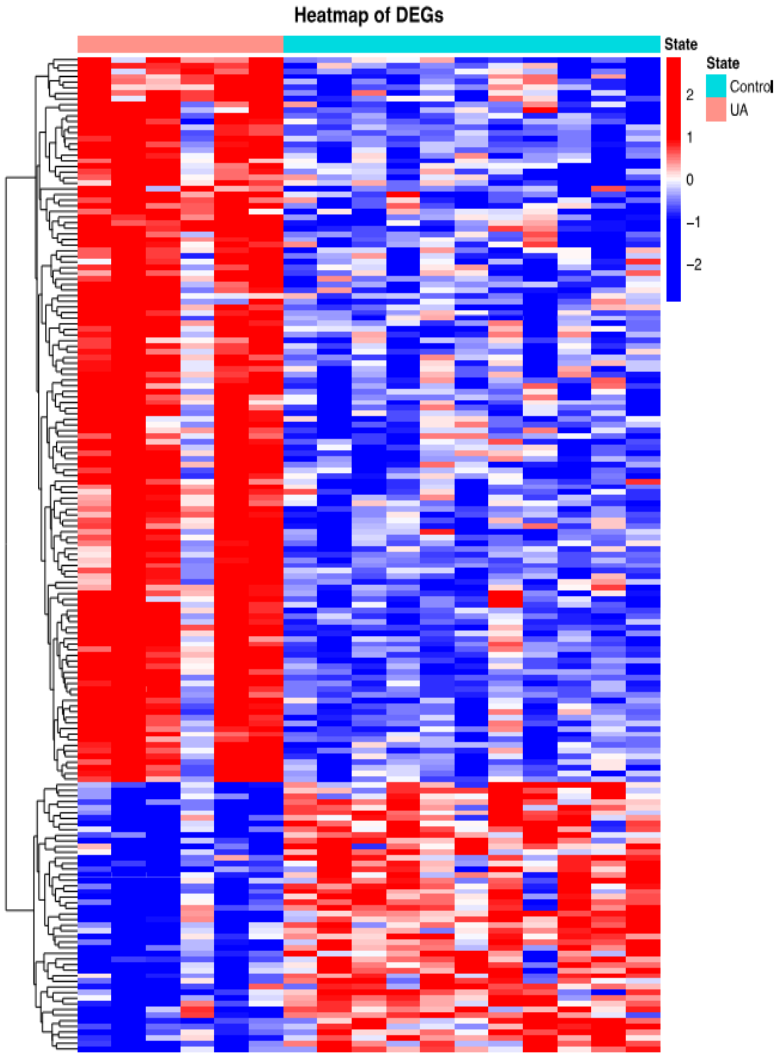
3. Results

3.1 Processing Data and Analyzing Enrichment

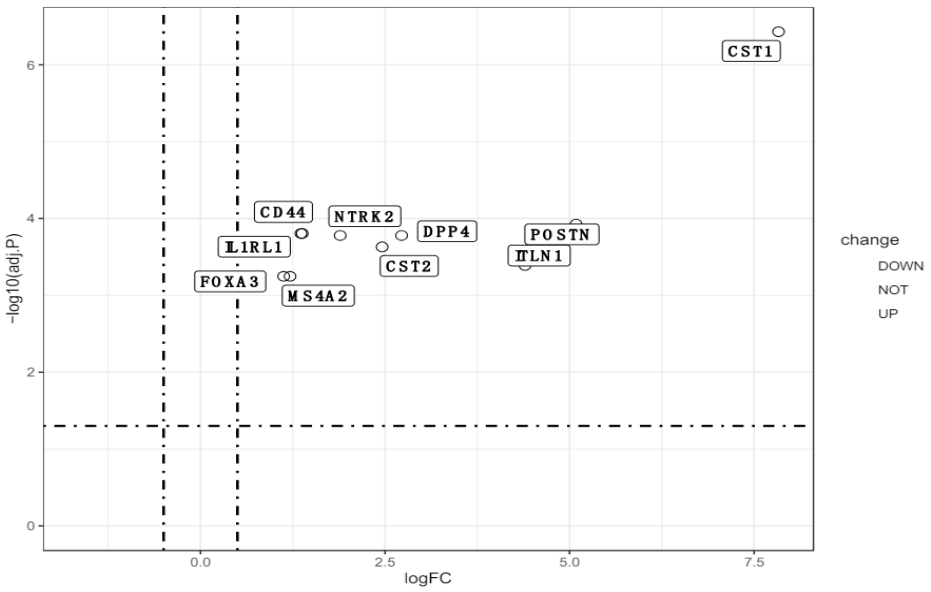
A total of 177 DEGs were identified from the combined gene expression matrix between healthy control group and UA group. Figures 1A and 1B displayed the expression of DEG. The metscape database was used for GO and pathway analysis, aiming to determine the biological function of DEG.

Figure 2A and Figure 2B displayed the functional correlation analysis results of GO and KEGG. GO analysis showed DEG was mainly involved in three cellular functions: BP, CC, and MF. The main changes in BP were related to the regulation of endopeptidase activity, regulation of peptidase activity, wound healing and glycoprotein metabolic process.

The main changes in CC were the college-containing extracellar matrix, apical part of cells, basal part of cells, basal plasma membrane, and basolateral plasma membrane. The most evident changes in MF were the sialytransfer activity and anion: cation symporter activity. Figure 2B displayed the results of KEGG pathway analysis. Results showed that DEGs were enriched in ferritosis, mucin type O-glycan biosynthatis, thyroid hormone synthesis, ECM-receptor interaction, cysteine and methionine metabolism, salivary secretion, histidine metabolism, renin-angiotensin system, pancreatic secretion, protein digestion and absorption, arachidonic acid metabolism, among which ferritosis was the most significantly enriched signal pathway.

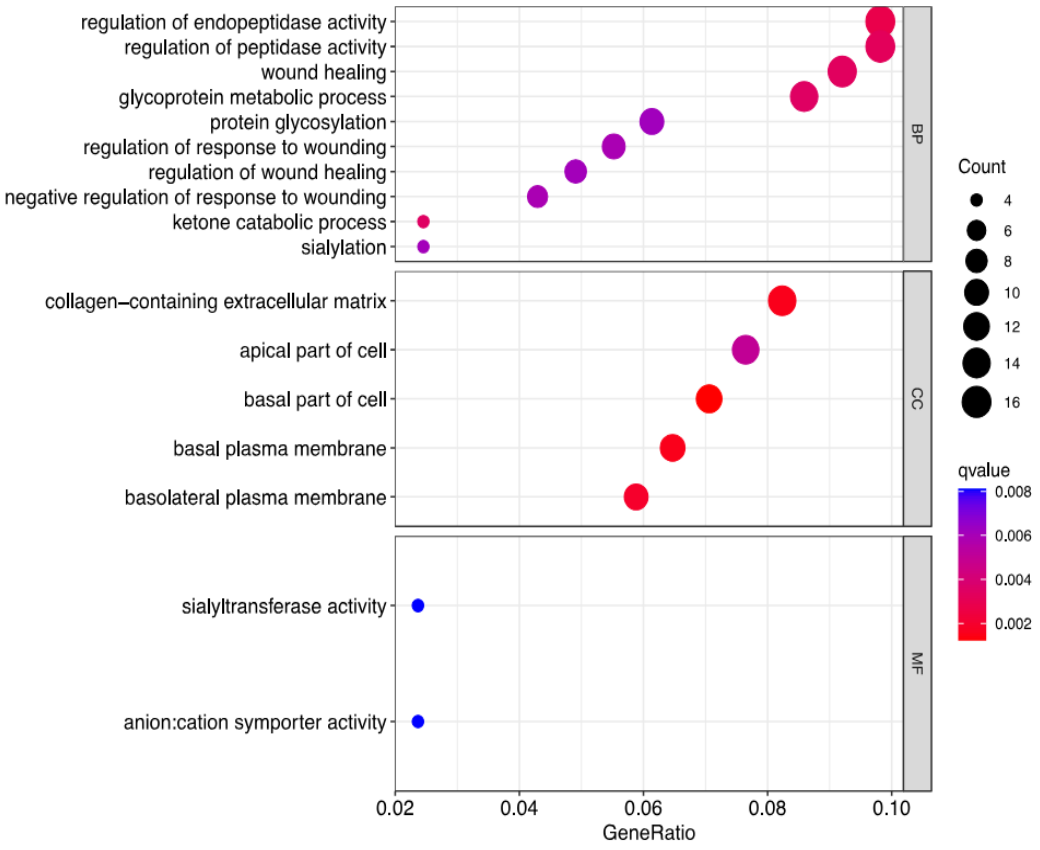


A

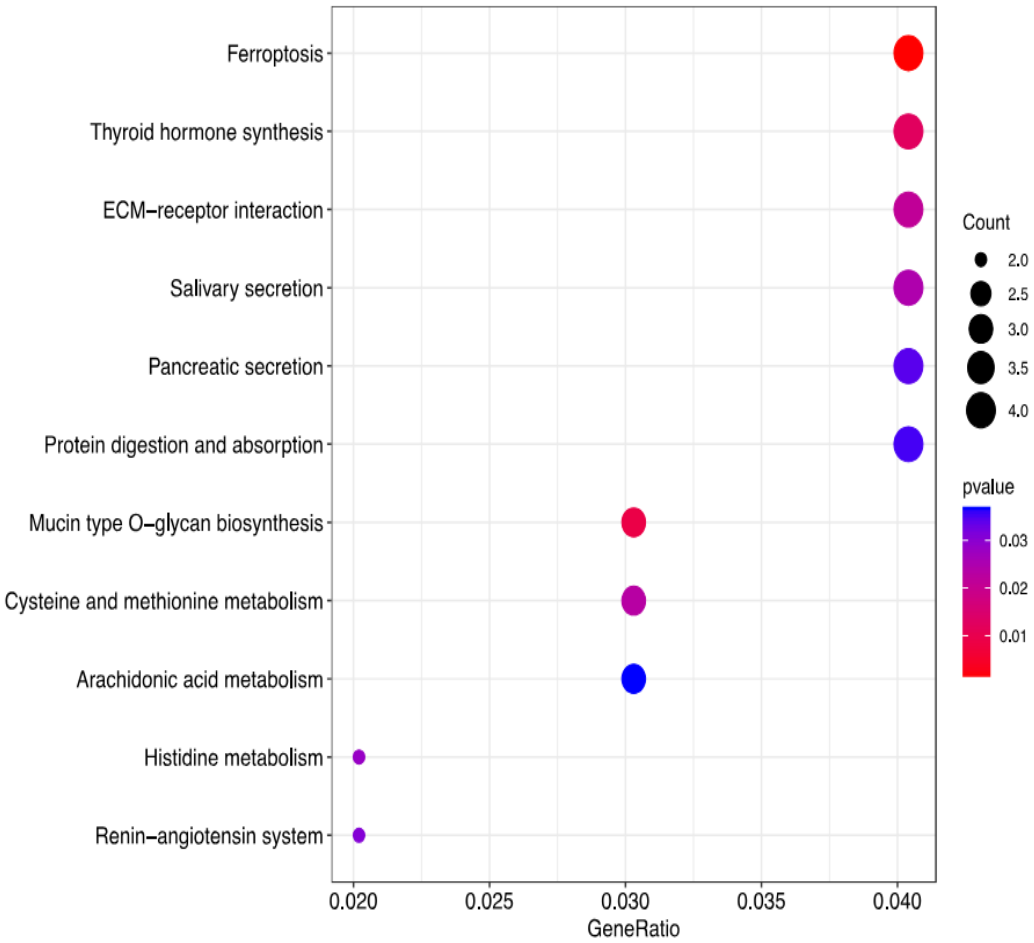


B

Figure 1: (A) Heat map of DEGs. The higher the mRNA expression, the darker the color is (red, upregulation; blue, downregulation). Left trees shows the clustering results for RNA with significant difference in different samples. Right grids represent groups (blue, control group; red, UA group). (B) Volcano map of DEG. Red, upregulated DGEs; gray, DGEs without significant difference; blue, downregulated DGEs.



A

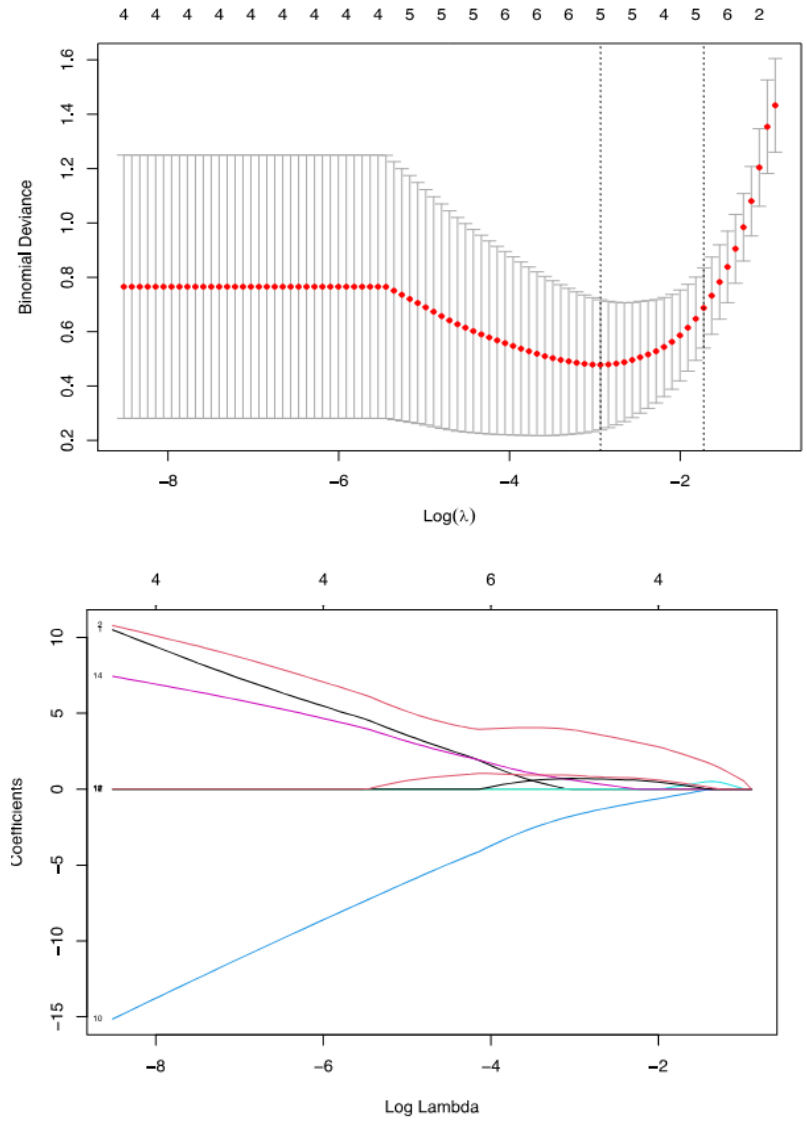


B

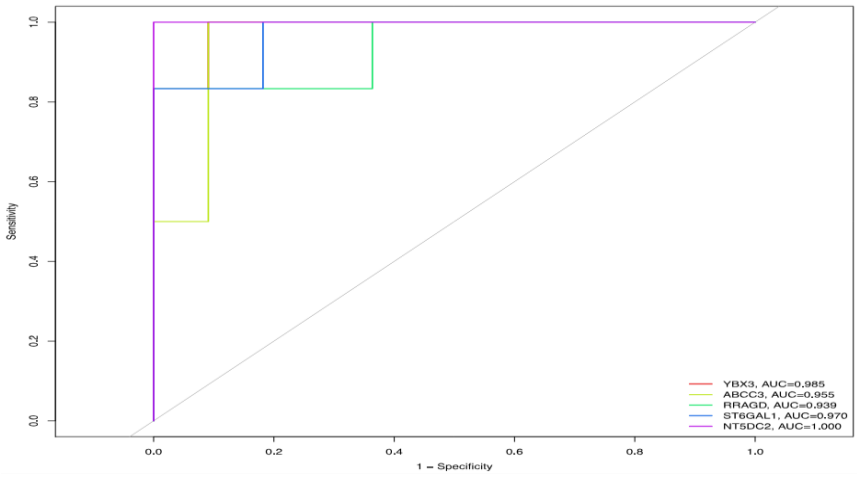
Figure 2: (A) GO enrichment analysis of the DEGs. The size of circles represents the number of DEGs; the color represents the p-values. GO analysis involves biological processes (BP), cellular components (CC), and molecular functions (MF). (B) KEGG enrichment analysis of the DEPs. The size of circles represents the number of DEG, and the color represents the p-values.

3.2 Identification of Biomarkers

In the identification of key immune biomarkers related to CARAS in children, LASSO regression was used to inference the effects of different DEGs, and the characteristic variables related to CARAS were further determined in children (Figure 3A). LASSO regression and SVM-RFE algorithms were used for the screening for DEGs, and ultimately five genes (YBX3, ABCC3, RRAGD, ST6GAL1, and NT5DC2) related to the diagnosis of CARAS were overlapped in two algorithms. ROC analysis (Figure 3B) was used to evaluate the accuracy of five diagnostic biomarkers in differentiating healthy controls from UA children. The AUC of YBX3, ABCC3, RRAGD, ST6GAL1 and NT5DC2 was 0.985, 0.955, 0.939, 0.970 and 1, respectively, indicating that these five biomarkers have high diagnostic value. However, due to the small sample size in this study, the possibility of over fitting could not be excluded in the analysis with AUC.



A



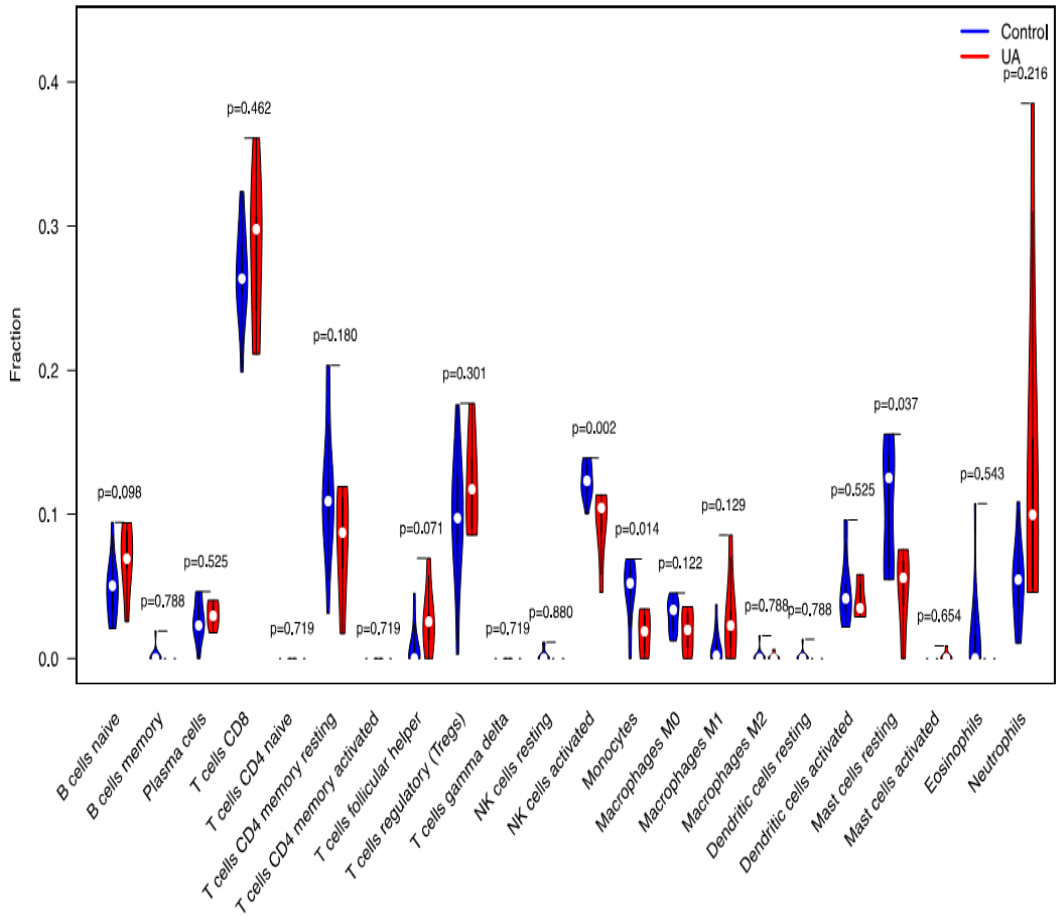
B

Figure 3: (A) LASSO regression analysis of infiltration of different immune cells. (B) Diagnostic performance of different biomarkers in the children with CARAS (ROC analysis)

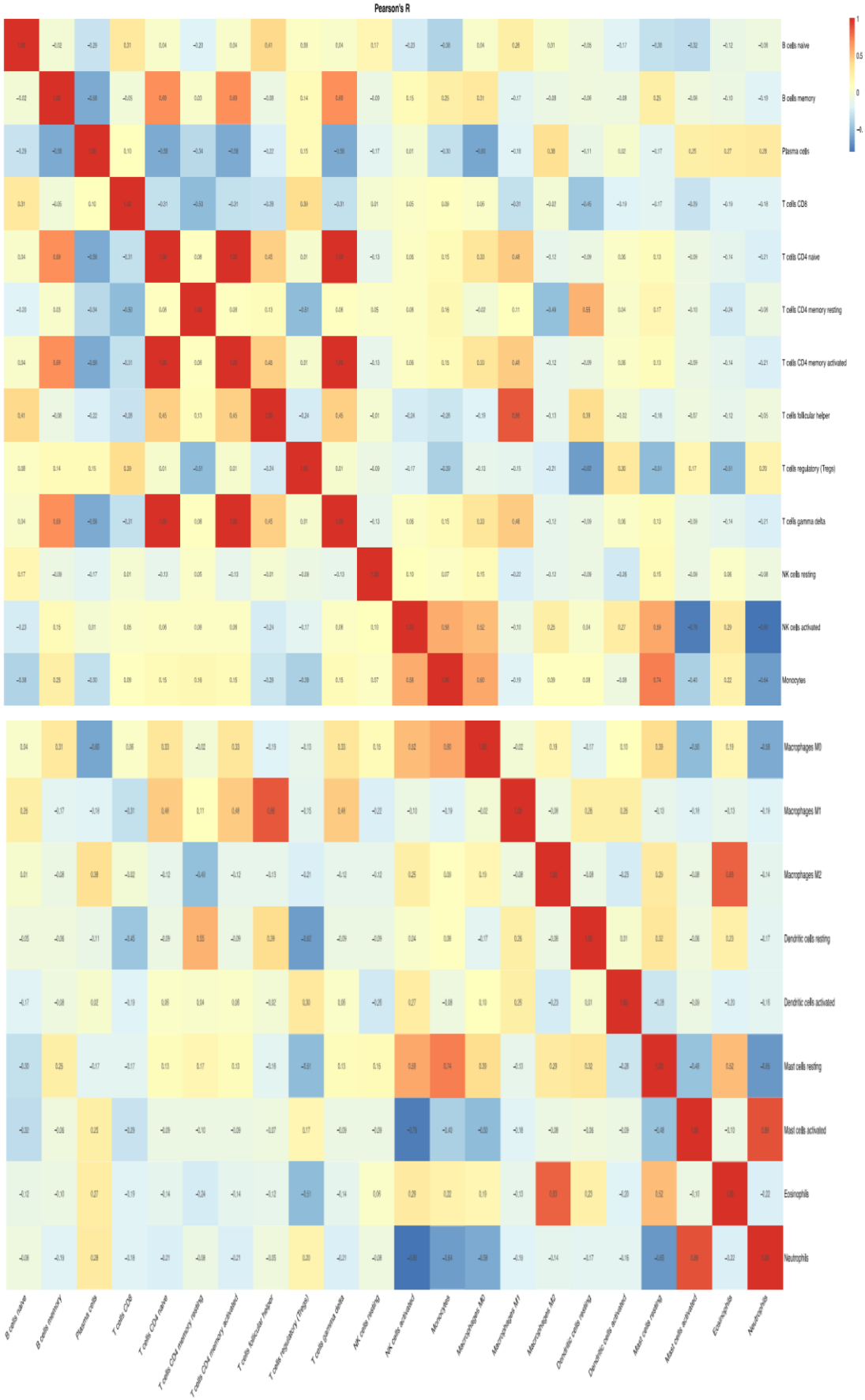
3.3 Immune Cell Infiltration Landscape

First, the infiltration ratio of immune cells in the tissues of UA children and normal tissues was calculated using the CIBERSORT algorithm. Our results demonstrated a significant increase in the proportion of NK cells activated (P=0.002) and monocytes infiltrated (P=0.014) in the tissues of UA children (Figure 4A) as compared to the control group. In addition, the correlation among 22 types of infiltrating immune cells was evaluated (Figure 4B).

Results showed T cells CD4 naive cells had positive relationships with T cells CD4 memory activated and T cells gamma delta, but NK cells activated had a negative correlation with neutrophils. Moreover, the correlations among five groups of genes were also evaluated (Figure 4C). Results showed NT5DC2 had a close correlation with other genes. The correlations of 5 groups of genes with 22 types of infiltrating immune cells were further assessed (Figure 4D). Results showed 5 groups of genes were related to immune cells infiltrated to different degrees, and mast cells resting was the most relevant infiltrated immune cells for all genes.



A


Pearson's R
1
0.5
0
-0.
B cells naive
B cells memory
Plasma cells
T cells CD8
T cells CD4 naive
T cells CD4 memory resting
T cells CD4 memory activated
T cells follicular helper
T cells regulatory (Tregs)
T cells gamma delta
NK cells resting
NK cells activated
Monocytes
Macrophages M0
Macrophages M1
Macrophages M2
Dendritic cells resting
Dendritic cells activated
Mast cells resting
Mast cells activated
Eosinophils
Neutrophils

B

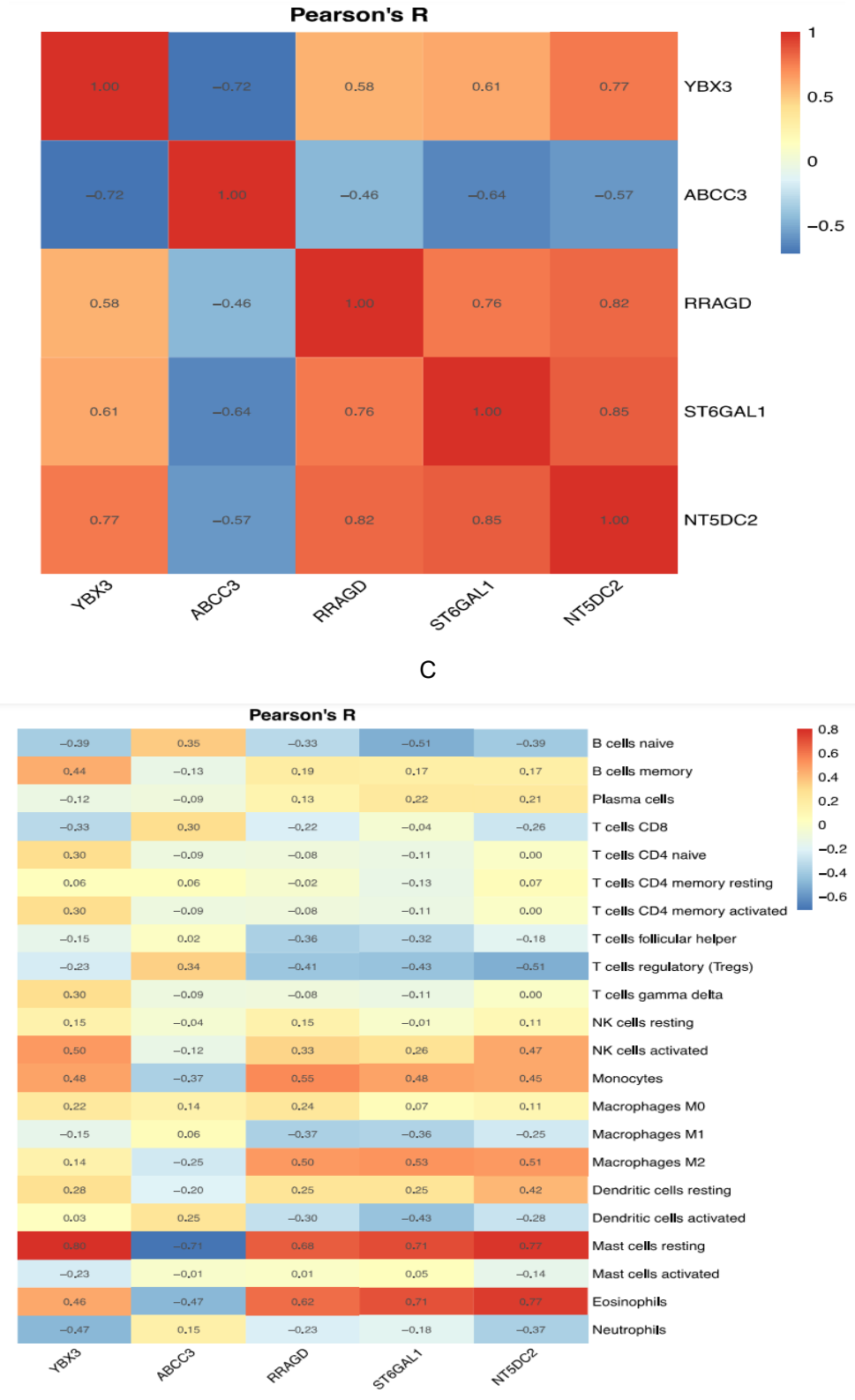


Figure 4: (A) Boxplot of percentage of 22 immune cells infiltrated. Red in the box represents the children with UA, and blue represents healthy controls. (B) The heat map of immune cell infiltration. (C) Heat map of correlations among five genes. (D) Heat map of correlations between 5 genes and 22 types of immune cells. In figures B, C and D, red represents positive correlation and blue represents negative correlation.

3.4 Potential therapeutic Drugs for CARAS in Children

Based on the key DEGs, the gene - drug interaction data from the DGIdb database were used to identify potential therapeutic drugs and confirm the interactions among drugs, genes, and immune cells. We identified six potential drugs (METHOTREXATE, VINCRISTINE, CISPLATIN, DOXORUBICIN, CYCLOPHOSPHAMIDE and PACLITAXEL), that interacted with ABCC3. The Knock TF online website was used to predict the transcription factors of hub gene, and the prediction results were presented in a table.

Table 1: Predicted Transcription Factors

TF	INTER GENES	THE NUMBER OF INTER GENES	P-VALUE
YBX1	ST6GAL1;ABCC3;NT5DC2;RRAGD	4	0.00477
MSX1	RRAGD;ABCC3;ST6GAL1;NT5DC2	4	0.0142
SAFB2	NT5DC2;ABCC3;RRAGD;ST6GAL1	4	0.0225
SNAI2	RRAGD;ABCC3;ST6GAL1;NT5DC2	4	0.0252
PCBP2	ABCC3;ST6GAL1;NT5DC2;RRAGD	4	0.0263
TAL1	RRAGD;NT5DC2;ST6GAL1;ABCC3	4	0.0442
ZNF254	ST6GAL1;NT5DC2;RRAGD	3	0.0026
TCF4	NT5DC2;RRAGD;ABCC3	3	0.00849
FOXO1	ST6GAL1;RRAGD;NT5DC2	3	0.0204
SYNCRIP	RRAGD;ST6GAL1;ABCC3	3	0.0414

4. Discussion

CARAS is a disease that manifests as both AR and asthma, and the development of the disease is believed to involve adaptive immune responses (Leite-de-Moraes et al., 2012). At present, the therapeutic effect of CARAS in children is often unsatisfactory, leading to the progression of the disease, which damages the physical and mental health of the sick children. Therefore, it is crucial to investigate the molecular mechanism of biomarkers related to CARAS and identify therapeutic targets. The aim of this study was to identify important diagnostic biomarkers and analyze the pattern of immune cell infiltration in the CARAS children. Five key genes related to CARAS were identified, including YBX3, ABCC3, RRAGD, ST6GAL1, and NT5DC2. These genes were associated with immune cell infiltration to varying degrees, suggesting that they hold potential as diagnostic biomarkers for CARAS. The identification of new genes and immune infiltration pattern have expanded the understanding of mechanisms underlying the CARAS. A total of 177 DEGs were identified and bioinformatics analysis was performed to explore the biological functions and pathways of molecules related to CARAS, which may be helpful for the identification of biomarkers related to the occurrence and development of CARAS. GO enrichment analysis indicated that most of the DEGs were primarily related to the regulation of endopeptidase activity, regulation of

peptidase activity. Therefore, it is speculated that these genes may play vital roles in the pathogenesis of CARAS through these processes. Studies have shown that tissue dependent peptide hydrolysis induced by proline endopeptidase is involved in water and electrolyte homeostasis (Irazusta et al., 2001). The key role of neutral endopeptidases in limiting and regulating neurogenic inflammation has been confirmed. In addition, viral infection, allergen exposure, cigarette smoke, and other respiratory irritants can reduce the activity of neutral endopeptidase, thereby enhancing the tachykinin in the airway. The decreased neutral endopeptidase activity may increase and perpetuate the harmful effects of airway inflammation (NADEL, 1992). On the other hand, the majority of these 177 target genes were enriched in ferroptosis, mucin type O-glycan biosynthesis, thyroid hormone synthesis, ECM receptor interaction, cysteine and methionine metabolism, salivary secret, histidine metabolism, renin-angiotensin system, pancreatic secretion, protein digestion and absorption, and acidic acid metabolism. Among them, ferroptosis was the most significantly enriched one. Ferroptosis, which regulates cell death, is characterized by abnormal peroxidation of lipids and accumulation of reactive oxygen species in cells. Ferroptosis is a form of programmed cell death that is dependent upon iron and involved in the progression of various diseases. In recent years, studies have shown a strong link between ferroptosis and the immune system. The regulatory factors of ferroptosis could be the secretions and cells of the immune system. Lv et al.(Lv et al., 2022) found that non-immune cells related to asthma were also closely related to ferroptosis, and regulating the signaling pathway related to ferroptosis may affect the target cells and/or molecules to promote the recovery of asthma. Gu (Gu et al., 2023) confirmed that ferroptosis plays an important role in the nasal epithelial injury caused by pm2.5, especially in AMPK-mediated autophagy. A case-control study of AR and healthy people in mainland China evaluated the relationship between AR risk and ferroptosis-related genes. The results of this study showed that ferroptosis-related genes increased AR risk (Yang et al., 2023). In addition, multiple studies (Gunawardhana et al., 2014; Xing et al., 2021) have indicated that the ECM receptor interaction pathway is related to the phenotype of asthma, and ECM-receptor interact in a highly enriched way and generally shared among downregulated DEPs in asthma. In a clinical observational study, salivary gland function was investigated in AR patients (Fussbroich et al., 2020). Results showed, as compared to healthy controls, the saliva amount in patients with allergies significantly reduced. Wu et al (Wu et al., 2015) investigated the microRNA expression profile of extracellular vesicles in the nasal mucus of AR patients, and results showed salivary secretion was the most significant signaling pathway enriched in the differential pathways related to vesicular miRNA pattern. A previous studies investigated the transcriptional regulation of characteristic genes in the blood samples from three patients with CARAS (Elad et al., 2006), and only phagosome was identified as the significantly enriched signaling pathway. Our findings provide more possibilities and provides a new

perspective for the diagnosis, treatment, and researches on the mechanism of CARAS. LASSO regression and SVM-RFE algorithms were used for the analysis of DEGs, and finally five diagnostic related genes (YBX3, ABCC3, RRAGD, ST6GAL1, and NT5DC2) overlapped in two algorithms were identified and regarded as potential biomarkers. These genes were related to immune cell infiltration to different degrees and may serve as potential diagnostic biomarkers. Biomarkers analyzed in this study showed good performance in the diagnosis of CARAS with FeNO (Fractional Exhaled Nitric Oxide) as a reference, which is the first available noninvasive maker of asthmatic airway inflammation (Mao et al., 2018), Zhang (Zhang et al., 2013) reported that the AUC of FeNO was 0.711, indicating uncontrolled asthma in their clinical evaluation. The AUC of all proposed biomarkers are higher than 0.939, among them, NT5DC2 had the best performance. YBX3 is a DNA/RNA binding protein, which can be used to aid the diagnosis of Idiopathic Pulmonary Fibrosis (Di Mauro et al., 2020)and nasopharyngeal carcinoma (Fan et al., 2021). ABCC3 serves as a biomarker for diverse tumor types, with potential utility as an early diagnostic and prognostic biomarker for lung cancer (Liu et al., 2012). Some studies have identified the effect of ST6Gal1 on cancers, and the upregulation of ST6Gal1 expression has been reported in many malignant tumors (Gc et al., 2022). NT5DC2 is a regulator of Ferroptosis whose target gene knockout sensitises cells to known iron-shedding inducers (Y.-C. Chen et al., 2020). Existing studies have shown that high expression of NT5DC2 reduces the median overall survival of patients with lung adenocarcinoma or squamous cell carcinoma (Schulze et al., 2022).

Meanwhile, NT5DC2 plays a crucial role in the occurrence and development of non-small cell lung cancer and glioma stem cell-like cells. It may be reasonable to consider NT5DC2 as an immunoinduction therapy site. To date, there is no conclusive evidence that confirm the promotive effect of above genes on the occurrence of CARAS. However, these genes are associated with the development of CARAS, and these genes may serve as targets for the future treatments of CARAS. In the present study, functional enrichment analysis was employed to identify the mechanism, and then LASSO logistic regression algorithm and SVM-RFE were used for cross validation aiming to identify the genes with significant difference. CIBERSORT analysis was employed to explore the patterns of immune cell infiltration. Among them, mast cells resting is a type of immune infiltration with the highest correlation with all genes. A previous study showed that mast cells play an important role in the exacerbation of airway inflammation in asthmatic patients, which was consistent with our findings. However, the role of diagnostic markers in children CARAS still have limitations. First, there are no clinical trials that confirm the accuracy of these biomarkers in the diagnosis of CARAS. Thus, silencing or over-expression of these genes may be performed in animal models and cells to evaluate the functional impact of these genes. Secondly, the sample size of

this study was still small, which may bias our findings.

5. Conclusion

This study utilized machine learning-based approaches to identify telomere-associated biomarkers and immune infiltration patterns in childhood allergic rhinitis and asthma syndrome (CARAS), providing new insights into the molecular mechanisms underlying airway inflammation, immune dysregulation, and their impact on respiratory function. By integrating gene expression analysis, computational immune profiling, and statistical modeling, we identified five key telomere-associated genes (ABCC3, NT5DC2, RRAGD, ST6GAL1, and YBX3) that play a crucial role in the pathogenesis of CARAS. Among them, NT5DC2 emerged as the most significant gene, showing strong correlations with immune cell infiltration and disease severity. The immune infiltration analysis revealed distinct alterations in immune cell populations, emphasizing the role of dysregulated inflammatory responses in airway obstruction and disease progression. These findings provide a foundation for developing targeted therapies that modulate immune function, reduce inflammation, and ultimately improve lung function and exercise capacity in children with CARAS.

5.1 Clinical and Sports Science Implications

CARAS significantly affects physical activity, exercise tolerance, and sports participation in children. Exercise-induced bronchoconstriction, reduced aerobic capacity, and impaired respiratory efficiency are common issues in affected individuals. By identifying biomarkers associated with immune infiltration and telomere dysfunction, this study paves the way for personalized interventions that enhance respiratory health and optimize physical performance. The implications for sports and exercise science are substantial:

5.2 Precision Medicine Approaches for Pediatric Athletes

Understanding telomere-related genetic variations can help personalize exercise prescriptions for children with CARAS. Targeted interventions may include graded aerobic training, respiratory muscle strengthening, and pulmonary rehabilitation tailored to individual genetic and immune profiles.

5.3 Optimizing Physical Activity Programs for Children with CARAS

5.3.1 Exercise plays a Critical Role in Modulating Immune Responses and Improving Lung Function.

Low-to-moderate intensity endurance training may help reduce airway inflammation and enhance cardiorespiratory fitness, while high-intensity interval training (HIIT) requires careful monitoring to avoid exacerbating

symptoms.

5.3.2 Enhancing Rehabilitation Strategies in Sports Medicine

Sports physiotherapists and clinicians can incorporate breathing techniques, airway clearance exercises, and adaptive training programs to help young athletes manage symptoms and improve performance. Telomere-targeted therapies could complement existing respiratory training programs, improving long-term outcomes.

5.4 Future Directions

While this study provides valuable genetic and immunological insights into CARAS, several areas require further exploration:

5.4.1 Longitudinal Studies on Telomere Dynamics and Physical Performance:

Future research should track telomere length changes over time in children with CARAS undergoing different exercise interventions. This could help determine how physical activity influences telomere-related immune responses and respiratory health over time.

5.4.2 Intervention-Based Studies on Exercise and Immune Regulation:

Investigating how different training modalities (e.g., endurance vs. resistance training) affect immune cell infiltration and inflammatory responses in CARAS patients. Assessing whether specific exercise regimens can help maintain telomere integrity and reduce disease severity. Validation of Biomarkers in Clinical and Sports Medicine Settings:

Larger multicenter studies should validate the identified biomarkers for early disease detection and risk stratification. Exploring whether genetic profiling could be integrated into sports medicine assessments for children with exercise-induced respiratory conditions.

5.5 Final Thoughts

This study highlights the importance of integrating machine learning, genomics, and sports science to improve diagnostic accuracy, treatment strategies, and physical activity recommendations for children with CARAS. The identification of telomere-associated biomarkers provides a new avenue for personalized interventions, ensuring that children with chronic airway inflammation can safely participate in physical activities while minimizing respiratory distress. As sports medicine continues to evolve, leveraging biomedical research and exercise physiology will be crucial in developing tailored exercise programs that enhance pulmonary function, optimize immune

responses, and improve overall quality of life for children with CARAS.

REFERENCES

Abu-Shaheen, A. K., Nofal, A., & Heena, H. (2016). Parental perceptions and practices toward childhood asthma. *BioMed research international*, *2016*.

Anderson, G. P. (2008). Endotyping asthma: new insights into key pathogenic mechanisms in a complex, heterogeneous disease. *The Lancet*, *372*(9643), 1107-1119.

Bousquet, J., Khaltaev, N., Cruz, A. A., Denburg, J., Fokkens, W., Togias, A., Zuberbier, T., Baena-Cagnani, C. E., Canonica, G., & Van Weel, C. (2008). Allergic rhinitis and its impact on asthma (ARIA) 2008. *Allergy*, *63*, 8-160.

Brożek, J. L., Bousquet, J., Agache, I., Agarwal, A., Bachert, C., Bosnic-Anticevich, S., Brignardello-Petersen, R., Canonica, G. W., Casale, T., & Chavannes, N. H. (2017). Allergic Rhinitis and its Impact on Asthma (ARIA) guidelines—2016 revision. *Journal of Allergy and Clinical Immunology*, *140*(4), 950-958.

Chen, Y.-C., Oses-Prieto, J. A., Pope, L. E., Burlingame, A. L., Dixon, S. J., & Renslo, A. R. (2020). Reactivity-based probe of the iron (II)-dependent interactome identifies new cellular modulators of ferroptosis. *Journal of the American Chemical Society*, *142*(45), 19085-19093.

Chen, Y., Zhu, J., Lyu, J., Xia, Y., Ying, Y., Hu, Y., Qu, J., Tong, S., & Li, S. (2020). Association of maternal prepregnancy weight and gestational weight gain with children's allergic diseases. *JAMA network open*, *3*(9), e2015643-e2015643.

Di Mauro, S., Scamporrino, A., Fruciano, M., Filippello, A., Fagone, E., Gili, E., Scionti, F., Purrazzo, G., Di Pino, A., & Scicali, R. (2020). Circulating coding and long non-coding RNAs as potential biomarkers of idiopathic pulmonary fibrosis. *International Journal of Molecular Sciences*, *21*(22), 8812.

Elad, S., Heisler, S., & Shalit, M. (2006). Saliva secretion in patients with allergic rhinitis. *International archives of allergy and immunology*, *141*(3), 276-280.

Fan, X., Xie, X., Yang, M., Wang, Y., Wu, H., Deng, T., Weng, X., Wen, W., & Nie, G. (2021). YBX3 mediates the metastasis of nasopharyngeal carcinoma via PI3K/AKT signaling. *Frontiers in Oncology*, *11*, 617621.

Ferreira, L. K. P., Ferreira, L. A. P., Monteiro, T. M., Bezerra, G. C., Bernardo, L. R., & Piuvezam, M. R. (2019). Combined allergic rhinitis and asthma syndrome (CARAS). *International Immunopharmacology*, *74*, 105718.

Fussbroich, D., Kohnle, C., Schwenger, T., Driessler, C., Dücker, R., Eickmeier, O., Gottwald, G., Jerkic, S., Zielen, S., & Kreyenberg, H. (2020). A combination of LCPUFAs regulates the expression of miRNA-146a-5p in a murine asthma model and human alveolar cells. *Prostaglandins & Other Lipid Mediators*, *147*, 106378.

Gc, S., Bellis, S. L., & Hjelmeland, A. B. (2022). ST6Gal1: Oncogenic signaling pathways and targets. *Frontiers in Molecular Biosciences*, *9*, 962908.

Gu, W., Hou, T., Zhou, H., Zhu, L., Zhu, W., & Wang, Y. (2023). Ferroptosis is involved in PM2. 5-induced acute nasal epithelial injury via AMPK-mediated autophagy. *International Immunopharmacology*, *115*, 109658.

Gunawardhana, L. P., Gibson, P. G., Simpson, J. L., Benton, M. C., Lea, R. A., & Baines, K. J. (2014). Characteristic DNA methylation profiles in peripheral blood monocytes are associated with inflammatory phenotypes of asthma. *Epigenetics*, *9*(9), 1302-1316.

Irazusta, J., Silveira, P. F., Gil, J., Varona, A., & Casis, L. (2001). Effects of hydrosaline treatments on prolyl endopeptidase activity in rat tissues. *Regulatory peptides*, *101*(1-3), 141-147.

Jin, J., Schorpp, K., Samaga, D., Unger, K., Hadian, K., & Stockwell, B. R. (2022). Machine learning classifies ferroptosis and apoptosis cell death modalities with TfR1 immunostaining. *ACS Chemical Biology*, *17*(3), 654-660.

Komlósi, Z. I., van de Veen, W., Kovács, N., Szűcs, G., Sokolowska, M., O'Mahony, L., Akdis, M., & Akdis, C. A. (2022). Cellular and molecular mechanisms of allergic asthma. *Molecular aspects of medicine*, *85*, 100995.

Leite-de-Moraes, M., Hammad, H., & Dy, M. (2012). Crosstalk between innate and adaptive cells on allergic process. *Journal of Allergy*, *2012*.

Liu, L., Liu, N., Liu, B., Yang, Y., Zhang, Q., Zhang, W., Yu, P., Jin, Y., Guo, J., & Guan, S. (2012). Are circulating autoantibodies to ABCC3 transporter a potential biomarker for lung cancer? *Journal of cancer research and clinical oncology*, *138*, 1737-1742.

Lucena-Zurita, M., Latorre-Román, P., Párraga-Montilla, J., & Manjón-Pozas, D. (2022). Validation of risk perception instrument in physical education teachers. *International Journal of Medicine & Science of Physical Activity & Sport/Revista Internacional de Medicina y Ciencias de la Actividad Física y del Deporte*, *22*(88).

Lv, X., Dong, M., Tang, W., Qin, J., Wang, W., Li, M., Teng, F., Yi, L., Dong, J., & Wei, Y. (2022). Ferroptosis, novel therapeutics in asthma. *Biomedicine & Pharmacotherapy*, *153*, 113516.

Mao, Z., Shi, Y., Cao, Q., Chen, Y., Sun, Y., Liu, Z., Zhang, Q., & Huang, M. (2018). Transcriptional regulation on the gene expression signature in combined allergic rhinitis and asthma syndrome. *Epigenomics*, *10*(2), 119-131.

Michaeloudes, C., Abubakar-Waziri, H., Lakhdar, R., Raby, K., Dixey, P., Adcock, I. M., Mumby, S., Bhavsar, P. K., & Chung, K. F. (2022). Molecular mechanisms of oxidative stress in asthma. *Molecular aspects of medicine*, *85*, 101026.

NADEL, J. (1992). Regulation of Neurogenic Inflammation by Neutral Endopeptidase1, 2.

Nappi, E., Paoletti, G., Malvezzi, L., Ferri, S., Racca, F., Messina, M., Puggioni, F., Heffler, E., & Canonica, G. (2022). Comorbid allergic rhinitis and asthma: important clinical considerations. *Expert Review of Clinical Immunology*, *18*(7), 747-758.

Ozdoganoglu, T., & Songu, M. (2012). The burden of allergic rhinitis and asthma. *Therapeutic advances in respiratory disease*, *6*(1), 11-23.

Pérez-Vigo, C., Sanchez-Lastra, M., Martínez-de-Quel, O., & Ayan, C. (2022). Reliability and validity of the V-Sit-and-Reach and Toe-Touch tests in preschoolers. *International Journal of Medicine & Science of Physical Activity & Sport/Revista Internacional de Medicina y Ciencias de la Actividad Física y del Deporte*, *22*(88).

Platts-Mills, T. A. (2015). The allergy epidemics: 1870-2010. *Journal of Allergy and Clinical Immunology*, *136*(1), 3-13.

Schulze, A. B., Kuntze, A., Schmidt, L. H., Mohr, M., Marra, A., Hillejan, L., Schulz, C., Görlich, D., Hartmann, W., & Bleckmann, A. (2022). High expression of NT5DC2 is a negative prognostic marker in pulmonary adenocarcinoma. *Cancers*, *14*(6), 1395.

Waljee, A. K., Weinheimer-Haus, E. M., Abubakar, A., Ngugi, A. K., Siwo, G. H., Kwakye, G., Singal, A. G., Rao, A., Saini, S. D., & Read, A. J. (2022). Artificial intelligence and machine learning for early detection and diagnosis of colorectal cancer in sub-Saharan Africa. *Gut*, *71*(7), 1259-1265.

Wu, G., Yang, G., Zhang, R., Xu, G., Zhang, L., Wen, W., Lu, J., Liu, J., & Yu, Y. (2015). Altered microRNA expression profiles of extracellular vesicles in nasal mucus from patients with allergic rhinitis. *Allergy, asthma & immunology research*, *7*(5), 449.

Xing, Q., You, Y., Zhao, X., Ji, J., Yan, H., Dong, Y., Ren, L., Ding, Y., & Hou, S. (2021). iTRAQ-based proteomics reveals Gu-Ben-Fang-Xiao decoction alleviates airway remodeling via reducing extracellular matrix deposition in a murine model of chronic remission asthma. *Frontiers in Pharmacology*, *12*, 588588.

Yang, Y.-F., Zheng, R., Gu, X., Gao, F., Chen, M., Du, M.-L., Zhang, Z.-D., Lu, M.-P., & Cheng, L. (2023). Evaluation of genetic variants in ferroptosis-related genes and house dust mite-induced allergic rhinitis risk. *International Immunopharmacology*, *115*, 109707.

Zhang, Y. M., Zhang, J., Liu, S. L., Zhang, X., Yang, S. N., Gao, J., Zhao, J., Chen, H., Chen, X. X., & Sun, F. X. (2013). Prevalence and associated risk factors of allergic rhinitis in preschool children in Beijing. *The Laryngoscope*, *123*(1), 28-35.