

RESEARCH

Open Access



# Identification of plasma protein markers of allergic disease risk: a mendelian randomization approach to proteomic analysis

Ziqin Cao<sup>1,2</sup>, Qiangxiang Li<sup>2</sup>, Yajia Li<sup>2,3\*</sup> and Jianhuang Wu<sup>1,2\*</sup>

## Abstract

**Background** While numerous allergy-related biomarkers and targeted treatment strategies have been developed and employed, there are still significant limitations and challenges in the early diagnosis and targeted treatment for allergic diseases. Our study aims to identify circulating proteins causally associated with allergic disease-related traits through Mendelian randomization (MR)-based analytical framework.

**Methods** Large-scale cis-MR was employed to estimate the effects of thousands of plasma proteins on five main allergic diseases. Additional analyses including MR Steiger analyzing and Bayesian colocalisation, were performed to test the robustness of the associations; These findings were further validated utilizing meta-analytical methods in the replication analysis. Both proteome- and transcriptome-wide association studies approach was applied, and then, a protein-protein interaction was conducted to examine the interplay between the identified proteins and the targets of existing medications.

**Results** Eleven plasma proteins were identified with links to atopic asthma (AA), atopic dermatitis (AD), and allergic rhinitis (AR). Subsequently, these proteins were classified into four distinct target groups, with a focus on tier 1 and 2 targets due to their higher potential to become drug targets. MR analysis and extra validation revealed STAT6 and TNFRSF6B to be Tier 1 and IL1RL2 and IL6R to be Tier 2 proteins with the potential for AA treatment. Two Tier 1 proteins, CRAT and TNFRSF6B, and five Tier 2 proteins, ERBB3, IL6R, MMP12, ICAM1, and IL1RL2, were linked to AD, and three Tier 2 proteins, MANF, STAT6, and TNFSF8, to AR.

**Conclusion** Eleven Tier 1 and 2 protein targets that are promising drug target candidates were identified for AA, AD, and AR, which influence the development of allergic diseases and expose new diagnostic and therapeutic targets.

**Keywords** Proteome, Plasma protein, Mendelian randomization, Allergy, Atopic dermatitis, Allergic asthma, Allergic rhinitis

\*Correspondence:

Yajia Li  
zndxlyj1996@csu.edu.cn  
Jianhuang Wu  
wujianhuang11@csu.edu.cn

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Introduction

Allergic diseases arise from inappropriate initiation of type 2 immune responses against innocuous environmental antigens and include disorders such as atopic dermatitis (AD), allergic asthma (AA), allergic rhinitis (AR), allergic conjunctivitis (AC) and allergic urticaria (AU) [1, 2]. The development of AD in early life, followed by other allergies, such as asthma, has been described as the atopic march [3, 4]. Allergic disease incidence has increased over the past 3 decades to affect an estimated 433 million people worldwide and exert a considerable economic and social burden [5, 6]. A combination of genetics, exposure to environmental allergens and irritants, microbial interactions, and abnormal immune responses contribute to inflammation and the atopic march [7, 8]. However, allergic disease pathogenesis is complex and poorly understood with limited drug options to tackle the recurrent and potentially life-long nature. These observations illustrate the need for further mechanistic studies and identification of drug targets.

The human plasma proteome participates in inter-tissue communications via metabolic, signaling, and physiological pathways and includes potential drug targets [9–13]. The link between levels of specific plasma proteins and allergic disease risk has been reported previously but residual confounders and reverse causation mean that observational studies do not demonstrate causality [14, 15]. Mendelian randomization (MR) uses genetic variation as an instrumental variable (IV) to explore the causal effect of exposure on outcomes. The impact of genes or proteins is difficult to interpret from isolated GWAS outcomes due to the occurrence of single nucleotide polymorphisms (SNPs). However, proteome-wide association studies (PWAS) give a clearer picture of the influence of proteins on disease [16]. In addition, functional summary-based imputation software allows transcriptome-wide association studies (TWAS) to assess the impact of whole blood protein-coding gene expression on allergic disease risk [17].

The current study used a PWAS approach, combining GWAS data with protein quantitative trait locus (pQTL) and TWAS with expression quantitative trait loci (eQTL), to investigate potential proteinaceous diagnostic and therapeutic targets in allergic disease. Bayesian colocalization analysis and phenotype scanning were conducted to verify causality between candidate proteins and disease pathogenesis. A protein-protein interaction (PPI) network was constructed and pathway enrichment was performed to illuminate potential mechanisms. The detailed research flowchart is presented in Fig. 1.

## Materials and methods

### Data sources and selection of IVs

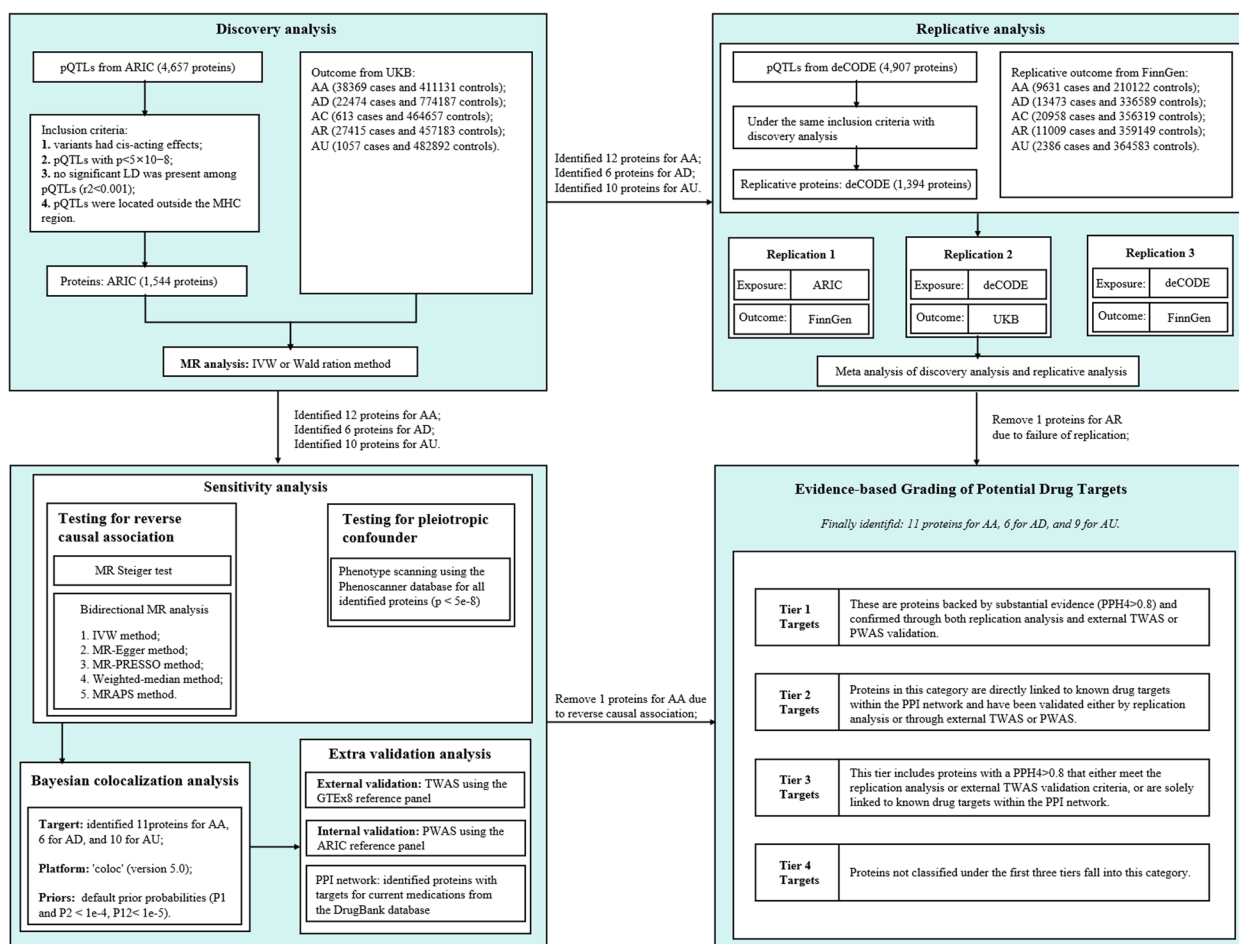
Whole blood pQTL data concerning 4,657 proteins derived from 7,213 participants of European descent in the ARIC study was analyzed [18]. SomaScan technology on the v.4.1 platform (SomaLogic) was used for proteomic profiling. pQTLs were filtered according to the following criteria: (1) variants had cis-acting effects, defined as being situated within a 1 Mb region up- or downstream of the gene encoding the plasma protein; (2) pQTLs met the genome-wide significance threshold of  $p < 5 \times 10^{-8}$ ; (3) no significant linkage disequilibrium (LD) was present among pQTLs ( $r^2 < 0.001$ ); (4) pQTLs were located outside the major histocompatibility complex (MHC) region (chr6, 26–34 Mb). The analysis was repeated using pQTL data from 35,559 Icelandic individuals, including 4,907 plasma proteins (Icelandic Cancer Project and deCODE genetics, Reykjavík, Iceland) on SomaScan v 4.0 platform [19].

Data for individual allergic diseases, including 38,369 cases and 411,131 controls for AA, 613 cases and 464,657 controls for AC, and 1,057 cases and 482,892 controls for AU were obtained from a UKB cohort GWAS study [20]. A UKB cohort from a trans-biobank meta-analysis study [21] yielded data for 22,474 cases of AD and 774,187 controls. Data on 27,415 cases of AR and 457,183 controls came from a separate GWAS study [22]. Summary statistics for the allergic diseases under scrutiny were obtained from the FinnGen Biobank Analysis Consortium database (<https://finngen.gitbook.io/documentation/>). Allergic diseases were diagnosed according to ICD-10 (International Classification of Diseases) criteria. Data sources are presented in detail in S. Table 1.

### Statistical analysis

#### Mendelian randomization analysis

MR analysis was performed using R software (version 4.1.3) and the TwoSampleMR package (version 0.5.6) [23]. Effect estimates of plasma proteins with a single SNP were generated using the Wald ratio [24] and of those with two or more SNPs by the Inverse Variance Weighted (IVW) method [25]. A false discovery rate (FDR,  $p < 0.05$ ) correction was applied to account for multiple comparisons in validating MR results [26]. Results are expressed as odds ratios per standard deviation increase in genetically determined plasma protein levels. Considering the genetic representativeness of the results, the ARIC/UKB combination, which includes a mixed European genetic background, was chosen as the discovery analysis.



**Fig. 1** Flowchart of evidence-based grading of potential drug targets from plasma protein on allergic diseases. AA: Allergic asthma; AC: Allergic conjunctivitis; AD: Atopic dermatitis; AR: Allergic rhinitis; AU: Allergic urticaria; PPI: Protein–protein interaction; pQTL: Protein quantitative trait loci; PPH4: posterior probability of hypothesis 4; PWAS: proteome-wide association studies; TWAS: transcriptome-wide association studies

**Sensitivity analysis**

Reverse causality between the proteins identified by discovery analysis and allergic diseases was evaluated by bidirectional MR analysis and MR Steiger test [26]. Four additional methods were employed in bidirectional MR analysis to assess the IV validity focusing on the three IV principles: strong association with exposure, direct influence on the outcome through exposure, and no association with outcome confounders. Violations of these principles could compromise the accuracy of the results. MR-Egger regression slope and intercept were used to estimate pleiotropy across IVs to give an adjusted estimate independent of IV validity [27, 28]. MR-PRESSO was used to identify outliers causing significant pleiotropy and heterogeneity to give a corrected causal effect assessment [29]. The weighted-median method enabled consistent inference even when over 50% of IVs were valid [30]. MR-Robust Adjusted Profile Score (MRAPS) increased statistical

power to give robust estimates in the presence of weak instrumental bias and horizontal pleiotropy [31]. By default, IVW results are preferred [32], but we turn to MR-Egger when significant pleiotropy is detected by the MR-Egger pleiotropy test. If the MR-PRESSO global test identifies significant outliers, we prioritize results corrected by MR-PRESSO. A Bonferroni correction was applied to address multiple comparison errors and a corrected value of  $p < 0.01$  was considered to indicate significance in reverse MR. Plasma proteins that appeared significant during both MR Steiger and reverse MR analysis, suggestive of reverse causality, were excluded from further analysis.

Potential links between all identified proteins and confounders were investigated via phenotype scanning using the Phenoscanner database [33] with a genome-wide significance threshold of  $p < 5 \times 10^{-8}$ . pQTLs linked to known allergic disease factors, indicative of pleiotropic effects, were interpreted with caution.

Then, proteins identified by discovery analysis were selected for replication analysis in further multi-center MR studies. Validation alternated between the two sets of pQTL and outcome data from FinnGen and UKB (including three validation sets: ARIC/FinnGen, deCODE/UKB, and deCODE/FinnGen), using genome-wide significant SNPs as genetic instruments. The stability of causal associations was evaluated through meta-analysis with a value of  $I^2 > 50\%$  indicating significant heterogeneity, necessitating a random-effects model [34].

#### **Bayesian colocalization analysis**

Colocalization analysis [35] was employed to determine whether a specific genetic variant influenced both an exposure factor and an outcome by modulating gene expression at common loci. Bayesian analysis to calculate the posterior probability of a shared causal variant influencing two traits was performed using the R package ‘coloc’ (version 5.0, available at <https://github.com/chr1s wallace/coloc>) with a default prior probabilities: a prior probability of  $1e-4$  for any single SNP being associated with each trait (P1 and P2) and a prior probability of  $1e-5$  for a SNP being associated with both traits (P12) [36]. Assuming a single causal variant, four hypotheses were considered: H0: no causal variants for either trait; H1: a causal variant for the first trait only; H2: a causal variant for the second trait only; H3: separate causal variants for both traits and H4: a shared causal variant for both traits [37]. Significant colocalization was inferred when the posterior probability of H4 was  $> 0.8$ , implying strong evidence of a shared causal influence [38].

#### **Extra validation analysis**

Considering that gene expression and protein synthesis are influenced by numerous factors beyond simple genetic processes, we conducted extra validation analyses to verify the results of our discovery analysis at both the tissue and protein levels. This validation was performed using both TWAS and PWAS methods predictive of gene influence on phenotype generated by Functional Summary-Based Imputation (FUSION) software (available at <http://gusevlab.org/projects/fusion>), based on the utility of GWAS summary statistics to indicate associations between GWAS phenotypes and functional phenotypes. TWAS indicated the association of protein-coding genes with allergic disease risk at the tissue level and was used as external validation analysis which utilized the pre-computed eQTL reference panel for target proteins derived from the GTEx8 (Genotype-Tissue Expression version 8) database. Likewise, the PWAS served as an internal validation analysis that integrated the GWAS summary statistics and the pre-computed plasma proteome reference weight also from the ARIC study [17]

to calculate the genetic impact on allergic disease. Thus, the impact of significant SNPs from the GWAS on protein abundance could be evaluated and candidate genes linked to allergic disease that regulate plasma protein levels identified. An similar FDR corrected p value  $< 0.05$  was the threshold of significance in the extra validation analysis.

A PPI network was constructed using the Search Tool for the Retrieval of Interacting Genes (STRING) database (version 11.5) [39] with a minimum required interaction score (IAS) threshold of 0.4 to indicate interactions among identified proteins and pre-existing anti-allergy drug targets [40]. Information on anti-allergy drug targets was sourced from the DrugBank database.

#### **Evidence-based grading of potential drug targets**

Proteins were graded according to the criteria of Feihong Ren [41].

Tier 1 Targets: substantial evidence ( $PPH4 > 0.8$ ) for drug targeting, confirmed by replication analysis and extra TWAS or PWAS validation.

Tier 2 Targets: direct linkage to known drug targets within the PPI network, validated either by replication analysis or extra TWAS or PWAS.

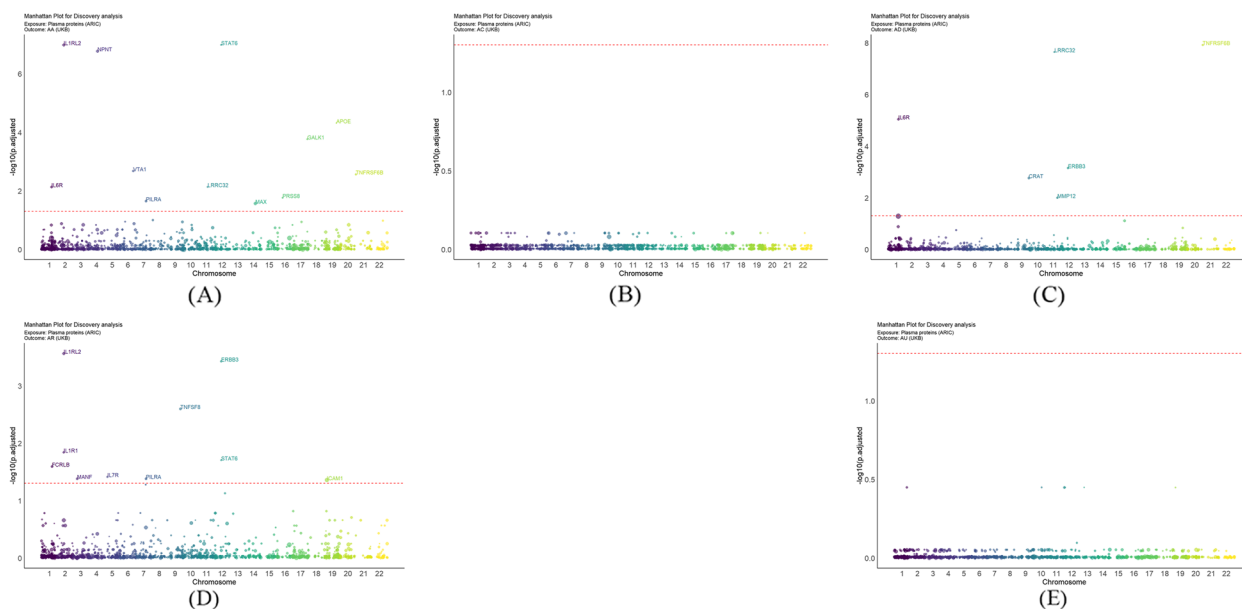
Tier 3 Targets: proteins with a  $PPH4 > 0.8$ , validated by either replication analysis or extra TWAS criteria or linked to known drug targets within the PPI network.

Tier 4 Targets: Proteins not classified under the first three tiers.

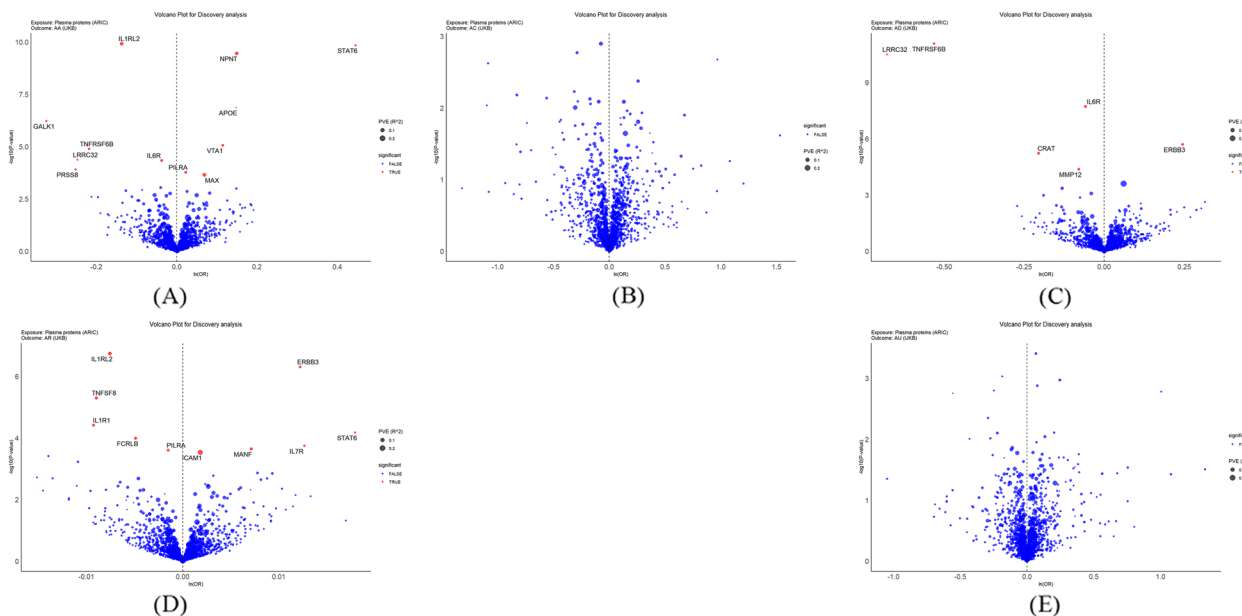
## **Results**

### **Discovery MR analysis**

A total of 1,544 proteins associated with 2,810 SNPs (S. Table 2) were identified during initial discovery analysis. Chromosomal locations are shown by Manhattan plots (Fig. 2). After FDR adjustment, 12 proteins were linked to AA, 6 to AD and 10 to AR but no proteins associated with AC or AU were found (Fig. 3). APOE (OR: 1.1591, 95% CI: 1.0972, 1.2245), MAX (OR: 1.0710, CI: 1.0326, 1.1109), NPNT (OR: 1.1607, CI: 1.1079, 1.2161), PILRA (OR: 1.0231, CI: 1.0110, 1.0354), STAT6 (OR: 1.5608, CI: 1.3622, 1.7884) and VTA1 (OR: 1.1213, CI: 1.0661, 1.1794) were associated with increased AA risk and GALK1 (OR: 0.7228, CI: 0.6363, 0.8211), IL1RL2 (OR: 0.8717, CI: 0.8360, 0.9089), IL6R (OR: 0.9632, CI: 0.9459, 0.9807), LRRC32 (OR: 0.7813, CI: 0.6941, 0.8795), PRSS8 (OR: 0.7769; CI: 0.6830, 0.8837) and TNFRSF6B (OR: 0.8038, CI: 0.7285, 0.8868) with decreased risk. ERBB3 (OR: 1.0123, CI: 1.0075, 1.0171), ICAM1 (OR: 1.0018, CI: 1.0008, 1.0028), IL7R (OR: 1.0127, CI: 1.0061, 1.0195), MANF (OR: 1.0071,



**Fig. 2** Manhattan plot illustrating the the chromosomal distribution of identified plasma proteins for allergic diseases. The standard line in the plot represents the threshold of FDR  $P=0.05$ . **A** Allergic asthma; **B** Allergic conjunctivitis; **C** Atopic dermatitis; **D** Allergic rhinitis; **E** Allergic urticaria

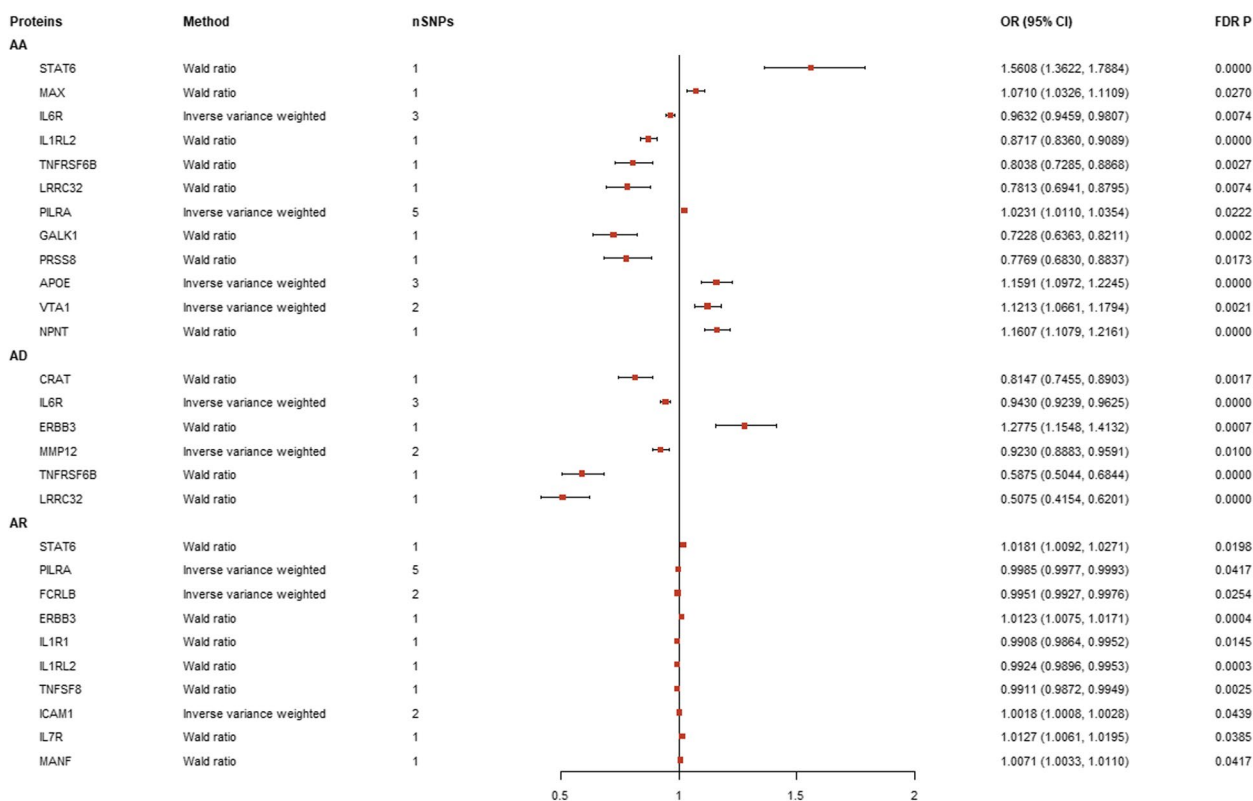


**Fig. 3** Volcano plots of the MR results from the discovery analysis, displaying the associations between 1,544 proteins from ARIC and the risk of allergic diseases. The increased OR for allergic diseases risk is represented as increments in SD of plasma protein levels. Red dots indicate significant proteins. 'ln' refers to the natural logarithm; 'PVE' stands for the proportion of variance explained. **A** Allergic asthma; **B** Allergic conjunctivitis; **C** Atopic dermatitis; **D** Allergic rhinitis; **E** Allergic urticaria

CI: 1.0033, 1.0110) and STAT6 (OR: 1.0181; CI: 1.0092, 1.0271) were associated with increased risk of AR and FCRLB (OR: 0.9951, CI: 0.9927 to 0.9976), IL1R1 (OR: 0.9908, CI: 0.9864, 0.9952), IL1RL2 (OR: 0.9924, CI: 0.9896, 0.9953), PILRA (OR: 0.9985, CI: 0.9977,

0.9993) and TNFSF8 (OR: 0.9911, CI: 0.9872, 0.9949) with decreased risk. All proteins, with the exception of ERBB3 (OR: 1.2775, CI: 1.1548, 1.4132), were linked to lower AD risk. Results are visualized in forest plots (Fig. 4).





**Fig. 4** Forest plot of MR results from the discovery analysis. AA: Allergic asthma; AD: Atopic dermatitis; AR: Allergic rhinitis

**Replicative MR and meta-analysis**

A total of 1,394 novel proteins with 4,144 SNPs were identified from deCode data by a similar IV selection process (S. Table 3) and used to validate the findings of discovery analysis from the FinnGen datasets (Table 1). Replication analysis using deCode data showed that GALK1, IL1RL2, and TNFRSF6B failed to replicate for AA, as did VTA1 and TNFRSF6B for AD and FCRLB, IL1RL2 and MANF for AR. IL7R could not be replicated during any iteration and was excluded from further analysis. Meta-analysis showed robust associations for other proteins with AA except APOE ( $p=0.8196$ ; OR: 1.0152, 95% CI: 0.8918, 1.1557) and LRRC32 ( $p=0.1864$ ; OR: 0.8048, CI: 0.5832, 1.1107, S. Figure 1). Significant associations also remained for proteins with AD, apart from LRRC32 ( $p=0.0990$ ; OR: 0.6015, 95% CI: 0.3288, 1.1002, S. Figure 2). Only STAT6 and PILRA retained a significant association with AR following post-replication analysis (S. Figure 3). IL7R in AR was excluded from further analysis due to the failure of replication.

**Sensitivity analysis**

Steiger filtering confirmed causal relationship directionality with only the relationship of APOE to AA failing to pass the Steiger test ( $R^2_{xy}=3.36 \times 10^{-6}$ ;  $p=0.878$ ).

Further bidirectional MR analysis confirmed the significant reverse causal association between AA and APOE (OR: 2.5883, CI: 1.8731, 3.5766, S. Table 4). Consequently, AA was excluded from further analysis.

Phenotype scanning indicated potential pleiotropic effects with APOE, GALK1, ICAM1, MAX, PRSS8, and VTA1 being associated with body mass and blood lipid levels. APOE and PILRA have been previously linked to diabetes, a condition that may be comorbid with allergic diseases. Furthermore, APOE has been associated with the phenotype of maternal diabetes, a condition that correlates with a higher risk of allergic disease in offspring. In addition, ERBB3, IL1R1, IL1RL2, IL6R, LRRC32, STAT6, and TNFRSF6B have been strongly linked to allergic diseases, such as asthma and rhinitis (S. Table 5).

**Colocalization analysis and extra validation**

Colocalization analysis was performed to assess any shared genetic signals between the proteins and AA or AD with a predefined threshold of  $PPH4 \geq 0.8$ . The majority of proteins were found to colocalize with AA, except for APOE ( $PPH4=1.54\%$ ), IL1RL2 ( $PPH4=0.00\%$ ), MAX ( $PPH4=75.70\%$ ), and PILRA ( $PPH4=0.19\%$ ). Only LRRC32 ( $PPH4=27.4\%$ ) failed to show colocalization with AD. However, substantial colocalization with

**Table 1** Details of the discovery analysis and replicative validations

Proteins	ARIC and UKB		FDRP		OR (95%CI)		FDRP		OR (95%CI)		Decode and UKB		FDRP		OR (95%CI)		Models	I2	Meta-analysis	P	Outcome
	Uniprot ID	OR (95%CI)	FDRP	OR (95%CI)	FDRP	OR (95%CI)	FDRP	OR (95%CI)	Decode and UKB	FDRP	OR (95%CI)	Decode and FinnGen	FDRP	OR (95%CI)							
APOE	P02649	1.1591 (1.0972, 1.2245)	4.83E-05	1.1039 (0.9955, 1.2241)	6.25E-01	0.9144 (0.8578, 0.9748)	2.84E-01	0.8934 (0.7844, 1.0176)	6.68E-01	91.90%	Random-effect	1.0152 (0.8918, 1.1557)	8.20E-01	AA							
GALK1	P51570	0.7228 (0.6363, 0.8211)	1.71E-04	0.5958 (0.4337, 0.8184)	1.18E-01	NA	NA	NA	NA	18.43%	Fixed-effect	0.7036 (0.6251, 0.7920)	5.80E-09								
IL1RL2	Q9HB29	0.8717 (0.8360, 0.9089)	1.02E-07	0.7943 (0.7252, 0.8700)	4.75E-04	NA	NA	NA	NA	69.78%	Random-effect	0.8397 (0.7680, 0.9182)	1.27E-04								
IL6R	P08887	0.9632 (0.9459, 0.9807)	7.38E-03	0.9417 (0.9132, 0.9711)	2.84E-02	0.9683 (0.9243, 1.0145)	7.83E-01	1.0119 (0.9466, 1.0818)	9.51E-01	27.47%	Fixed-effect	0.9611 (0.9473, 0.9751)	6.93E-08								
LRRC32	Q14392	0.7813 (0.6941, 0.8795)	7.38E-03	0.5497 (0.4057, 0.7448)	2.84E-02	1.1925 (0.9403, 1.5123)	7.64E-01	0.7919 (0.4868, 1.2882)	9.35E-01	82.15%	Random-effect	0.8048 (0.5832, 1.1107)	1.86E-01								
MAX	P61244	1.0710 (1.0326, 1.1109)	2.70E-02	1.1617 (1.0769, 1.2531)	2.84E-02	1.1605 (1.0715, 1.2568)	3.98E-02	1.3949 (1.1820, 1.6461)	2.71E-02	77.83%	Random-effect	1.1637 (1.0666, 1.2696)	6.49E-04								
NPNT	Q6UX19	1.1607 (1.1079, 1.2161)	1.69E-07	1.1261 (1.0326, 1.2279)	2.45E-01	1.3415 (1.2238, 1.4705)	2.23E-07	1.2637 (1.0654, 1.4989)	3.11E-01	68.91%	Random-effect	1.2096 (1.1139, 1.3135)	6.04E-06								
PILRA	Q9UKJ1	1.0231 (1.0110, 1.0354)	2.22E-02	1.0055 (0.9797, 1.0320)	9.83E-01	1.0110 (0.9906, 1.0319)	8.76E-01	1.0164 (0.9711, 1.0639)	9.39E-01	0.00%	Fixed-effect	1.0180 (1.0085, 1.0275)	1.94E-04								
PRSS8	Q16651	0.7769 (0.6830, 0.8837)	1.73E-02	0.7719 (0.6011, 0.9911)	5.53E-01	0.5990 (0.4561, 0.7866)	3.98E-02	0.5975 (0.3538, 1.0089)	6.05E-01	16.93%	Fixed-effect	0.7398 (0.6671, 0.8205)	1.15E-08								
STAT6	P42226	1.5608 (1.3622, 1.7884)	1.02E-07	1.7634 (1.3338, 2.3314)	2.84E-02	1.0315 (0.8989, 1.1838)	9.79E-01	1.1330 (0.8091, 1.5866)	9.39E-01	86.95%	Random-effect	1.3364 (1.0367, 1.7228)	2.52E-02								
TNFRSF6B	O95407	0.8038 (0.7285, 0.8868)	2.69E-03	0.6561 (0.5222, 0.8243)	4.97E-02	NA	NA	NA	NA	61.00%	Random-effect	0.7462 (0.6161, 0.9039)	2.76E-03								
VTA1	Q06263	1.1213 (1.0661, 1.1794)	2.09E-03	1.1228 (0.9921, 1.2708)	6.27E-01	1.9728 (1.4726, 2.6428)	2.19E-03	1.6285 (0.8949, 2.9632)	7.23E-01	80.45%	Random-effect	1.3524 (1.0139, 1.8039)	4.00E-02								
CRAT	P43155	0.8147 (0.7455, 0.8903)	1.70E-03	0.8067 (0.7184, 0.9058)	3.40E-02	NA	NA	NA	NA	0.00%	Fixed-effect	0.8104 (0.7630, 0.8606)	7.54E-12								
ERBB3	P21860	1.2775 (1.1548, 1.4132)	7.03E-04	1.2871 (1.1333, 1.4617)	1.94E-02	1.5768 (1.3069, 1.9024)	2.43E-03	1.5989 (1.2620, 2.0256)	3.63E-02	39.45%	Fixed-effect	1.3306 (1.2517, 1.4144)	5.06E-20								
IL6R	P08887	0.9430 (0.9239, 0.9625)	9.14E-06	0.9363 (0.8546, 1.0258)	8.69E-01	0.9494 (0.9261, 0.9732)	1.62E-02	0.9423 (0.9136, 0.9720)	3.63E-02	0.00%	Fixed-effect	0.9445 (0.9316, 0.9576)	4.00E-16								
LRRC32	Q14392	0.5075 (0.4154, 0.6201)	2.28E-08	0.3207 (0.2462, 0.4176)	2.23E-14	1.1788 (0.8348, 1.6645)	9.90E-01	1.3638 (0.9030, 2.0597)	7.65E-01	94.14%	Random-effect	0.6015 (0.3288, 1.1002)	9.90E-02								
MMP12	P39900	0.9230 (0.8883, 0.9591)	9.98E-03	0.8911 (0.8539, 0.9300)	5.38E-05	0.9255 (0.8890, 0.9634)	4.86E-02	0.9051 (0.8636, 0.9485)	3.48E-02	0.00%	Fixed-effect	0.9083 (0.8914, 0.9254)	7.42E-24								
TNFRSF6B	O95407	0.5875 (0.5044, 0.6844)	1.20E-08	0.4312 (0.3547, 0.5243)	2.23E-14	NA	NA	NA	NA	76.92%	Random-effect	0.4819 (0.3909, 0.5942)	8.49E-12								

**Table 1** (continued)

Proteins	Uniprot ID	OR (95%CI)		FDR P		OR (95%CI)		FDR P		OR (95%CI)		FDR P		Models	I2	Meta-analysis	OR (95%CI)	P	Outcome
		ARIC and UKB	ARIC and UKB	ARIC and FinnGen	ARIC and FinnGen	Decode and UKB	Decode and UKB	Decode and FinnGen	Decode and FinnGen										
ERBB3	P21860	1.0123 (1.0075, 1.0171)	3.72E-04	0.9825 (0.8542, 1.1300)	9.99E-01	0.9776 (0.9689, 0.9862)	3.13E-04	0.9677 (0.7461, 1.2551)	9.90E-01	93.55%	Random-effect	0.9941 (0.9640, 1.0251)	7.05E-01	AR					
FCRLB	Q6BAA4	0.9951 (0.9927, 0.9976)	2.54E-02	0.9206 (0.8547, 0.9915)	6.16E-01	NA	NA	NA	NA	76.34%	Random-effect	0.9659 (0.8969, 1.0403)	3.60E-01						
ICAM1	Q5NKV9	1.0018 (1.0008, 1.0028)	4.39E-02	1.0017 (0.9734, 1.0309)	9.99E-01	0.9982 (0.9964, 1.0000)	5.95E-01	1.0930 (1.0448, 1.1435)	2.27E-02	88.78%	Random-effect	1.0182 (0.9813, 1.0565)	3.39E-01						
IL1R1	P14778	0.9908 (0.9864, 0.9952)	1.45E-02	0.8672 (0.7585, 0.9916)	6.41E-01	1.0117 (1.0063, 1.0171)	5.65E-03	0.8750 (0.7436, 1.0298)	7.87E-01	92.85%	Random-effect	0.9667 (0.9068, 1.0305)	2.99E-01						
IL1RL2	Q9HB29	0.9924 (0.9896, 0.9953)	2.71E-04	0.8438 (0.7757, 0.9179)	5.15E-02	NA	NA	NA	NA	92.99%	Random-effect	0.9203 (0.7853, 1.0785)	3.05E-01						
IL7R	P16871	1.0127 (1.0061, 1.0195)	3.85E-02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA						
MANF	P55145	1.0071 (1.0033, 1.0110)	4.17E-02	1.1749 (1.0514, 1.3128)	3.82E-01	NA	NA	NA	NA	86.46%	Random-effect	1.0765 (0.9270, 1.2502)	3.34E-01						
PILRA	Q9UKJ1	0.9985 (0.9977, 0.9993)	4.17E-02	1.0053 (0.9799, 1.0315)	9.99E-01	0.9983 (0.9970, 0.9996)	3.81E-01	1.0092 (0.9673, 1.0530)	9.78E-01	0.00%	Fixed-effect	0.9984 (0.9977, 0.9991)	1.10E-05						
STAT6	P42226	1.0181 (1.0092, 1.0271)	1.98E-02	1.5063 (1.1623, 1.9523)	2.51E-01	1.0152 (1.0035, 1.0269)	3.81E-01	1.2486 (0.9088, 1.7155)	8.43E-01	71.58%	Random-effect	1.0174 (1.0103, 1.0245)	1.36E-06						
TNFSF8	P32971	0.9911 (0.9872, 0.9949)	2.51E-03	1.0534 (0.9425, 1.1773)	9.73E-01	1.0120 (0.9984, 1.0259)	6.60E-01	1.1328 (0.7257, 1.7683)	9.71E-01	69.58%	Random-effect	1.0023 (0.9822, 1.0228)	8.23E-01						

OR Odds ratio, FDR False discovery rate, AA Allergic asthma, AD Atopic dermatitis



AR was only observed for ERBB3 (PPH4=98.40%) and TNFSF8 (PPH4=88.00%, S. Table 6 and S. Figures 4–6).

Two different sets of reference panels from the GTEx database were used for TWAS of AD skin tissue and AA or AR respiratory tract tissue. IL1RL2 and NPNT were mismatched and only STAT6 (Ptwas=0.0011), TNFRSF6B (Ptwas=0.0058), and VTA1 (Ptwas=0.0052) showed significant association with AA among the remaining ten proteins. Similarly, ICAM1, IL1RL2, TNFSF8, and FCRLB were absent from the database but significant associations with AR were found for IL7R (Ptwas=0.0023), ERBB3 (Ptwas=0.0444), MANF (Ptwas=0.0122) and PILRA (Ptwas=0.0486). CRAT (Ptwas=0.0000), TNFRSF6B (Ptwas=0.0002), and LRRC32 (Ptwas=0.0000) levels correlated with AD. PWAS was performed on the reference panel from the ARIC study and 4 proteins, STAT6, TNFRSF6B, GALK1, and CRAT, were excluded due to mismatches. Other proteins maintained stable associations with the exception of APOE (ppwas=0.4915), NPNT (ppwas=0.1310), and LRRC32 (ppwas=0.0831) for AA and IL1R1 (ppwas=0.1330) for AR (Table 2 and S. Figure 7).

A PPI network was constructed from DrugBank data to illustrate interactions among anti-allergic drug targets and proteins of interest (S. Table 7). Interactions were found between STAT6, TNFRSF6B, IL1RL2, IL6R, and established drug targets, as evidenced by an IAS greater than 0.4, in the PPI network for AA (S. Figure 8). Similarly, interactions were found for ERBB3, IL6R, and MMP12 in the AD-specific PPI network (S. Figure 9) and for ERBB3, ICAM1, IL1RL2, MANF, STAT6 and TNFSF8 in the AR-specific PPI network (S. Figure 10). Further gene-disease enrichment analysis revealed STAT6 was enriched in various allergic diseases, most notably in AA, with a strength of 1.86 and an FDR P-value of 1.8e-13. IL1RL2, IL6R, ERBB3, ICAM1, IL1R1, and GALK1 were also enriched in categories like immune system disease, autoimmune disease, skin disease, and disease of anatomical entity (S. Table 8). We also conducted KEGG and GO enrichment analyses on genes within the PPI network. With the exception of GALK1, NPNT, PRSS8, VTA1, LRRC32, PILRA, MMP12, MANF, TNFSF8, and FCRLB, genes of other proteins were successfully enriched in multiple pathways, particularly STAT6, IL6R, and IL1R1 proteins, which were involved in several immune and inflammation-related pathways. Except for FCRLB, MANF, and TNFSF8, other proteins were all successfully enriched in specific biological processes (see S. Table 9, and 10 and S. Figure 11).

#### **Potential drug targets**

Ultimately, 11 proteins for allergic asthma (AA), 6 for atopic dermatitis (AD), and 9 for allergic rhinitis (AR)

were evaluated for their potential as drug targets. Among these, STAT6 was identified as excellent Tier 1 potential drug targets for AA as were CRAT and TNFRSF6B for AD. No Tier 1 proteins were identified for AR. Tier 2 proteins included TNFRSF6B, and IL1RL2 for AA; ERBB3, IL6R, MMP12, ICAM1 and IL1RL2 for AD, and ICAM1, IL1RL2, MANF, STAT6 and TNFSF8 for AR. Other proteins were assigned to Tier 3 or below (Table 2).

#### **Discussion**

To the best of our knowledge, this study is the first to scrutinize causal associations between plasma proteins and allergic diseases through the integrated approach of MR, colocalization, Steiger filtering analysis, PAV assessment, eQTLs overlap determination, PPI analysis, pathway enrichment, and drug target evaluation. Eleven plasma proteins were identified with links to AA, AD, and AR. MR analysis and extra validation revealed STAT6 and TNFRSF6B to be Tier 1 and IL1RL2 and IL6R to be Tier 2 proteins with the potential for AA treatment. Two Tier 1 proteins, CRAT and TNFRSF6B, and five Tier 2 proteins, ERBB3, IL6R, MMP12, ICAM1, and IL1RL2, were linked to AD and three Tier 2 proteins, MANF, STAT6, and TNFSF8, to AR.

Many novel biomarkers have been identified by proteomic and metabolomic analyses, although studies on allergic diseases have generally used low throughput methods [42]. Niet-Fontarigo et al. [43] identified 18 potential biomarkers of asthma phenotype and disease severity, including HSPG2 and IGFALS for AA, through a bottom-up/non-targeted proteomics approach. The sample size was modest with 32 healthy controls, 43 AR patients, and 192 asthmatics and failed to distinguish protein biomarkers from pathogenic factors for allergic disease due to the reverse causal characteristics of an observational study.

Several proteins identified by the current study have previously been linked to allergic disease by epidemiological or laboratory studies. Indeed, STAT6 is known to participate in IL-4 signaling and its role in asthma has been extensively studied since both doctor-diagnosed asthma and blood eosinophil counts are known to be linked to STAT6 signaling and the IL-1 receptor family [44]. Baris S et al. [45] have identified a novel inborn error of immunity arising from a STAT6 gain-of-function mutation causing severe allergic dysregulation which is treated by Janus kinase inhibitor therapy. The TNFRSF6 (also called Fas) receptor binds TNFSF6 (FasL) ligands expressed on CD8+ T cells and oligodendrocytes [46–48]. Th1 cells secrete IFN- $\gamma$  to activate the Fas/FasL system and induce keratinocyte apoptosis in the spongiosis area which may influence the progression of AD. INF- $\gamma$  and Fas ligand are secreted by activated CD4+ T

**Table 2** Assessment of druggability and evidence grading

Proteins	Discovery analysis		Meta-analysis of Replication analysis		Colocalization analysis		External validation		Steigjer direction test		PPI network	Levels of Evidence
	OR (95%)	FDR P	OR (95%)	FDR P	PPH4	Pass/Fail	TWAS Z, FDR P	PWAS Z, FDR P	P	Pass/Fail		
AA												
APOE	1.1591 (1.0972, 1.2245)	4.83E-05	1.0152 (0.8918, 1.1557)	8.20E-01	1.5%	Fail	-1.0730, 0.7651	-0.6880, 0.9665	8.78E-01	Fail	Pass	Excluded
STAT6	1.5608 (1.3622, 1.7884)	1.02E-07	1.3364 (1.0367, 1.7228)	2.52E-02	100.0%	Pass	3.2652, 0.0432	NA	4.67E-10	Pass	Pass	Tier 1
TNFRSF6B	0.8038 (0.7285, 0.8868)	2.69E-03	0.7462 (0.6161, 0.9039)	2.76E-03	99.4%	Pass	-2.7590, 0.1572	NA	9.80E-16	Pass	Pass	Tier 2
IL1RL2	0.8717 (0.8360, 0.9089)	1.02E-07	0.8397 (0.7680, 0.9182)	1.27E-04	0.0%	Fail	NA	-4.0486, 0.0013	6.00E-88	Pass	Pass	Tier 2
IL6R	0.9632 (0.9459, 0.9807)	7.38E-03	0.9611 (0.9473, 0.9751)	6.93E-08	100.0%	Pass	-1.4100, 0.6611	5.6215, 0.0000	4.28E-38	Pass	Pass	Tier 3
GALK1	0.7228 (0.6363, 0.8211)	1.71E-04	0.7036 (0.6251, 0.7920)	5.80E-09	99.9%	Pass	0.7682, 0.8242	NA	4.62E-11	Pass	Fail	Tier 3
NPNT	1.1607 (1.1079, 1.2161)	1.69E-07	1.2096 (1.1139, 1.3135)	6.04E-06	100.0%	Pass	NA	1.5119, 0.7020	2.48E-90	Pass	Fail	Tier 3
PRSS8	0.7769 (0.6830, 0.8837)	1.73E-02	0.7398 (0.6671, 0.8205)	1.15E-08	84.6%	Pass	-1.4135, 0.5748	3.8410, 0.0055	1.78E-11	Pass	Fail	Tier 3
VTA1	1.1213 (1.0661, 1.1794)	2.09E-03	1.3524 (1.0139, 1.8039)	4.00E-02	97.8%	Pass	2.7957, 0.1446	-2.6109, 0.1967	1.28E-27	Pass	Fail	Tier 3
LRRC32	0.7813 (0.6941, 0.8795)	7.38E-03	0.8048 (0.5832, 1.1107)	1.86E-01	95.4%	Pass	-1.9494, 0.5000	1.7330, 0.4283	6.28E-08	Pass	Fail	Tier 4
MAX	1.0710 (1.0326, 1.1109)	2.70E-02	1.1637 (1.0666, 1.2696)	6.49E-04	75.7%	Fail	-1.6300, 0.6860	-3.6018, 0.0164	1.22E-115	Pass	Fail	Tier 4
PILRA	1.0231 (1.0110, 1.0354)	2.22E-02	1.0180 (1.0085, 1.0275)	1.94E-04	0.2%	Fail	-0.7261, 0.9126	2.6438, 0.1449	3.26E-37	Pass	Fail	Tier 4
AD												
CRAT	0.8147 (0.7455, 0.8903)	1.70E-03	0.8104 (0.7630, 0.8606)	7.54E-12	99.3%	Pass	4.2501, 0.0046	NA	1.16E-43	Pass	Fail	Tier 1
TNFRSF6B	0.5875 (0.5044, 0.6844)	1.20E-08	0.4819 (0.3909, 0.5942)	8.49E-12	98.0%	Pass	-3.6822, 0.0215	NA	9.77E-16	Pass	Fail	Tier 1
ERBB3	1.2775 (1.1548, 1.4132)	7.03E-04	1.3306 (1.2517, 1.4144)	5.06E-20	99.7%	Pass	-0.5643, 0.9467	-4.7549, 0.0001	6.15E-32	Pass	Pass	Tier 2
IL6R	0.9430 (0.9239, 0.9625)	9.14E-06	0.9445 (0.9316, 0.9576)	4.00E-16	100.0%	Pass	0.9614, 0.8304	5.7955, 0.0000	4.26E-38	Pass	Pass	Tier 2
MMP12	0.9230 (0.8883, 0.9591)	9.98E-03	0.9083 (0.8914, 0.9254)	7.42E-24	97.0%	Pass	0.9690, 0.8193	3.8311, 0.0102	5.37E-34	Pass	Pass	Tier 2

**Table 2** (continued)

Proteins	Discovery analysis		Meta-analysis of Replication analysis		Colocalization analysis		External validation		Steigjer direction test		PPI network	Levels of Evidence
	OR (95%)	FDR P	OR (95%)	FDR P	PPH4	Pass/Fail	TWAS Z, FDR P	PWAS Z, FDR P	P	Pass/Fail		
LRRC32	0.5075 (0.4154, 0.6201)	2.28E-08	0.6015 (0.3288, 1.1002)	9.90E-02	27.4%	Fail	4.3107, 0.0031	3.1004, 0.0515	6.27E-08	Pass	Fail	Tier 4
AR												
IL7R	1.0127 (1.0061, 1.0195)	3.85E-02	NA	NA	0.0%	Fail	-3.0526, 0.1818	4.3853, 0.0008	5.07E-18	Pass	Pass	Excluded
ICAM1	1.0018 (1.0008, 1.0028)	4.39E-02	1.0182 (0.9813, 1.0565)	3.39E-01	8.1%	Fail	NA	-3.6235, 0.0106	8.65E-244	Pass	Pass	Tie 2
IL1RL2	0.9924 (0.9896, 0.9953)	2.71E-04	0.9203 (0.7853, 1.0785)	3.05E-01	0.0%	Fail	NA	-3.1346, 0.0340	6.05E-88	Pass	Pass	Tie 2
MANF	1.0071 (1.0033, 1.0110)	4.17E-02	1.0765 (0.9270, 1.2502)	3.34E-01	24.1%	Fail	-2.5070, 0.6411	-3.4132, 0.0482	1.06E-48	Pass	Pass	Tie 2
STAT6	1.0181 (1.0092, 1.0271)	1.98E-02	1.0174 (1.0103, 1.0245)	1.36E-06	47.5%	Fail	0.2935, 0.9346	NA	4.68E-10	Pass	Pass	Tie 2
TNFSF8	0.9911 (0.9872, 0.9949)	2.51E-03	1.0023 (0.9822, 1.0228)	8.23E-01	88.0%	Pass	NA	4.3603, 0.0007	1.91E-47	Pass	Pass	Tie 2
ERBB3	1.0123 (1.0075, 1.0171)	3.72E-04	0.9941 (0.9640, 1.0251)	7.05E-01	98.4%	Pass	-2.0102, 0.5548	-5.0264, 0.0000	6.21E-32	Pass	Pass	Tie 3
IL1R1	0.9908 (0.9864, 0.9952)	1.45E-02	0.9667 (0.9068, 1.0305)	2.99E-01	0.0%	Fail	1.3200, 0.8961	1.5022, 0.6188	1.57E-36	Pass	Pass	Tie 3
FCRLB	0.9951 (0.9927, 0.9976)	2.54E-02	0.9659 (0.8969, 1.0403)	3.60E-01	12.0%	Fail	NA	-2.5711, 0.1947	8.03E-49	Pass	Fail	Tier 4
PILRA	0.9985 (0.9977, 0.9993)	4.17E-02	0.9984 (0.9977, 0.9991)	1.10E-05	0.0%	Fail	1.9719, 0.4586	-2.2443, 0.3289	3.28E-37	Pass	Fail	Tier 4

Tier 1 Targets: These are proteins backed by substantial evidence (PPH4>0.8) and confirmed through both replication analysis and external TWAS or PWAS validation

Tier 2 Targets: Proteins in this category are directly linked to known drug targets within the PPI network and have been validated either by replication analysis or through external TWAS or PWAS

Tier 3 Targets: This tier includes proteins with a PPH4>0.8 that either meet the replication analysis or external TWAS validation criteria, or are solely linked to known drug targets within the PPI network

Tier 4 Targets: Proteins not classified under the first three tiers fall into this category

OR Odds ratio, FDR False discovery rate, MM Malignant melanoma, SCC Squamous-cell carcinoma, BCC Basal cell carcinoma

cells, TNFRSF6 expressed on keratinocytes, and tumor necrosis factor (TNF)- $\alpha$  secreted by both the activated CD4+ T cells and keratinocytes, with cell-mediated cytotoxicity induced by perforin and granzyme B released by CD8+ T cells [46–48].

The IL-36 receptor (IL-36R, IL-1Rrp2, IL1RL2, or IL-1R6) binds all  $\alpha$ ,  $\beta$ , and  $\gamma$  members of the IL-36 family. IL-36R is expressed in skin, mammary, and mucosal epithelial cell lines and IL-36 mediates intracellular signaling through the IL-36R and IL-1 receptor accessory protein (IL-1RAcP) [49]. IL-36 $\alpha$ , IL-36 $\beta$ , and IL-36 $\gamma$  bind IL-36R, form a signal transduction complex with IL-1RAcP, and recruit myeloid differentiation factor 88 to activate mitogen-activated protein kinases mediated by c-Jun N-terminal kinase, extracellular regulated protein kinases 1/2 and the nuclear factor kappa B pathway. The resulting inflammatory mediators have roles in adaptive immunity. IL-36 cytokines released from keratinocytes increase the immunoglobulin (Ig)E production mediated by IL-4 in B cells from AD patients and treatment with anti-IL-36R antibodies decreases IgE and alleviates the disease phenotype [50, 51]. The RNA helicase, DDX5, which regulates the alternative splicing of IL-36R pre-mRNA was found to be down-regulated in keratinocytes from AD patients which promoted the inflammatory response [52]. TNFSF8 (CD30L) is a ligand for the cell surface antigen and marker for Hodgkin lymphoma and related hematologic malignancies, TNFRSF8/CD30. It is considered that inhibition of CD30L/CD30 signaling may constitute a novel biological therapy for AR, since CD30L was shown to amplify Th2 cell effector response in animal models of AR. In vivo treatment with anti-CD30 antibody suppressed AR development and this may be a sufficient target for the treatment of allergic inflammation [53].

Some novel proteins were suggested to have potential causal effects on allergic diseases by the current work. For example, CRAT is a mitochondrial enzyme that transfers acetyl groups between CoA and carnitine during lipid metabolism and links with dermatitis have not been previously reported. CRAT is a key regulator of mitochondrial dysfunction-induced cellular senescence in dermal fibroblasts [54, 55]. Silencing of CRAT is known to cause mitochondrial dysfunction, inflammation and senescence via activation of the cGAS-STING and NF- $\kappa$ B pathways [54]. In addition, functional variants of the IL6R have been linked to increased risk of AA and AD but mechanisms remain unclear, although IL-6/soluble IL-6R trans-signaling may affect AD and AA development [56–58]. Genetic variants of ERBB3 have been identified as AD susceptibility factors and serum MMP12 may be an indicator of AD and AR disease pathways [57, 59–61]. Adhesion molecules are known to be involved in

T cell homing to skin lesions in AD patients, one example being ICAM-1 which is highly expressed and may have a pathogenic role [62, 63]. Lastly, little attention has been paid to MANF, although the protein is measurable in serum and reflective of extracellular biomarkers in AD [64]. However, these putative mechanisms are speculative and experimental mechanistic studies are required to extend the findings of the present study.

We acknowledge several limitations to the current study. First, the current focus was on serum proteins which may differ from those within cells and tissues and should also be explored for disease associations. Second, European participants accounted for the vast majority of the current cohort and findings may not be generalizable to populations with different ethnicities. Thirdly, a cis-pQTL coding variant might change a protein's amino acid sequence without necessarily impacting its function or level. Equating sequence alterations with functional changes could lead to incorrect conclusions. Caution should be exercised in the interpretation of these results. Last, publicly available datasets were used and represent data resources for target identification which are not new, although novel insights and perspectives may be drawn from them.

## Conclusion

A MR analysis was conducted to explore the proteomic pathogenesis of allergic disease. Examples of Tier 1 and 2 protein targets that are promising drug target candidates were STAT6, TNFRSF6B, IL1RL2, and IL6R for AA; CRAT, TNFRSF6B, ERBB3, IL6R, MMP12, ICAM1 and IL1RL2 for AD, and ICAM1, IL1RL2, MANF, STAT6 and TNFSF8 for AR. These proteins may influence the development of allergic diseases and expose new diagnostic and therapeutic targets. Further experiments are required to validate the current findings regarding proteinaceous allergic disease markers.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-024-10412-0>.

Supplementary Material 1.

Supplementary Material 2.

## Acknowledgements

The authors would like to express their gratitude to EditSprings (<https://www.editsprings.com/>) for the expert linguistic services provided.

## Authors' contributions

Dr. Yajia Li and Pro. Jianhuang Wu had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Concept and design: Jianhuang Wu and Yajia Li; Acquisition, analysis, or interpretation of data: Ziqin Cao and Yajia Li; Drafting of the manuscript: All of the authors; Critical revision of the manuscript for important

intellectual content: Ziqin Cao and Jianhuang Wu; Statistical analysis: Ziqin Cao, Yajia Li and Qiangxiang Li; Administrative, technical, or material support: Ziqin Cao and Qiangxiang Li. Dr. Yajia Li and Pro. Jianhuang Wu contributed equally to the work.

### Funding

This work was supported by the National Key R&D Program of China [No. 2019YFA0112100]. The study funders/sponsors had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

### Availability of data and materials

All analyses were conducted using publicly available data. The data that support this study are openly available in UK Biobank at <https://www.ukbiobank.ac.uk/>, and FinnGen, at <https://www.finnngen.fi/en>. Code Availability: the analysis code in R is available on request.

### Declarations

#### Ethics approval and consent to participate

The manuscript does not contain clinical studies or patient data. Our study is based on the large-scale GWAS datasets, and not the individual-level data. All participants gave informed consent in all the corresponding original studies, and no additional ethical approval was applicable. No animal subjects were used in this work.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare no competing interests.

#### Author details

<sup>1</sup> Department of Spine Surgery and Orthopaedics, Xiangya Hospital, Central South University, Xiangya Road 87, Changsha 410000, China. <sup>2</sup> National Clinical Research Center for Geriatric Disorders, Central South University, Changsha, China. <sup>3</sup> Department of Dermatology, Xiangya Hospital, Central South University, Changsha, Hunan 410011, China.

Received: 25 February 2024 Accepted: 15 May 2024

Published online: 22 May 2024

### References

- Julia V, Macia L, Dombrowicz D. The impact of diet on asthma and allergic diseases. *Nat Rev Immunol*. 2015;15:308–22.
- Yamamoto T, Endo Y, Onodera A, Hirahara K, Asou HK, Nakajima T, Kanno T, Ouchi Y, Uematsu S, Nishimasu H, et al. DUSP10 constrains innate IL-33-mediated cytokine production in ST2(hi) memory-type pathogenic Th2 cells. *Nat Commun*. 2018;9:4231.
- Aw M, Penn J, Gauvreau GM, Lima H, Sehmi R. Atopic march: collegium internationale allergologicum update 2020. *Int Arch Allergy Immunol*. 2020;181:1–10.
- Toledo MF, Saraiva-Romanholo BM, Oliveira RC, Saldiva PH, Silva LF, Nascimento LF, Solé D. Changes over time in the prevalence of asthma, rhinitis and atopic eczema in adolescents from Taubaté, São Paulo, Brazil (2005–2012): relationship with living near a heavily travelled highway. *Allergol Immunopathol (Madr)*. 2016;44:439–44.
- GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet*. 2020;396:1204–22. [https://doi.org/10.1016/S0140-6736\(20\)30925-9](https://doi.org/10.1016/S0140-6736(20)30925-9). Erratum in: *Lancet*. 2020 Nov 14;396(10262):1562.
- GBD 2019 IMID Collaborators. Global, regional, and national incidence of six major immune-mediated inflammatory diseases: findings from the global burden of disease study 2019. *EclinicalMedicine*. 2023;64:102193. <https://doi.org/10.1016/j.eclinm.2023.102193>.
- Fuchs O, Bahmer T, Rabe KF, von Mutius E. Asthma transition from childhood into adulthood. *Lancet Respir Med*. 2017;5:224–34.
- Langan SM, Irvine AD, Weidinger S. Atopic dermatitis. *Lancet*. 2020;396:345–60.
- Gudjonsson A, Gudmundsdottir V, Axelsson GT, Gudmundsson EF, Jonsson BG, Launer LJ, Lamb JR, Jennings LL, Aspelund T, Emilsson V, Gudnason V. A genome-wide association study of serum proteins reveals shared loci with common diseases. *Nat Commun*. 2022;13:480.
- Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, Burgess S, Jiang T, Paige E, Surendran P, et al. Genomic atlas of the human plasma proteome. *Nature*. 2018;558:73–9.
- Imming P, Sinning C, Meyer A. Drugs, their targets and the nature and number of drug targets. *Nat Rev Drug Discov*. 2006;5:821–34.
- Lam MP, Ping P, Murphy E. Proteomics research in cardiovascular medicine and biomarker discovery. *J Am Coll Cardiol*. 2016;68:2819–30.
- Gashaw I, Ellinghaus P, Sommer A, Asadullah K. What makes a good drug target? *Drug Discov Today*. 2011;16:1037–43.
- Schnabel CL, Jentsch MC, Lübke S, Kaiser-Thom S, Gerber V, Vrtala S, Huang HJ, Rhyner C, Wagner B, Hoffmann R, Volke D. Immunoproteomics reveal increased serum IgG3/5 binding to Dermatophagoides and yeast protein antigens in severe equine asthma in a preliminary study. *Front Immunol*. 2023;14:1293684.
- Asamoah K, Chung KF, Zounemat Kermani N, Bodinier B, Dahlen SE, Djukanovic R, Bhavsar PK, Adcock IM, Vuckovic D, Chadeau-Hyam M. Proteomic signatures of eosinophilic and neutrophilic asthma from serum and sputum. *EBioMedicine*. 2024;99:104936.
- Yao C, Chen G, Song C, Keefe J, Mendelson M, Huan T, Sun BB, Laser A, Maranville JC, Wu H, et al. Genome-wide mapping of plasma protein QTLs identifies putatively causal genes and pathways for cardiovascular disease. *Nat Commun*. 2018;9:3268.
- Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BW, Jansen R, de Geus EJ, Boomsma DI, Wright FA, et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet*. 2016;48:245–52.
- Zhang J, Dutta D, Köttgen A, Tin A, Schlosser P, Grams ME, Harvey B, Yu B, Boerwinkle E, Coresh J, Chatterjee N. Plasma proteome analyses in individuals of European and African ancestry identify cis-pQTLs and models for proteome-wide association studies. *Nat Genet*. 2022;54:593–602.
- Ferkingstad E, Sulem P, Atlason BA, Sveinbjornsson G, Magnusson MI, Styrismisdottir EL, Gunnarsdottir K, Helgason A, Oddsson A, Halldorsson BV, et al. Large-scale integration of the plasma proteome with genetics and disease. *Nat Genet*. 2021;53:1712–21.
- Sakaue S, Kanai M, Tanigawa Y, Karjalainen J, Kurki M, Koshiba S, Narita A, Konuma T, Yamamoto K, Akiyama M, et al. A cross-population atlas of genetic associations for 220 human phenotypes. *Nat Genet*. 2021;53:1415–24.
- Sliz E, Huilaja L, Pasanen A, Laisk T, Reimann E, Mägi R, Hannula-Jouppi K, Peltonen S, Salmi T, Koulou L, et al. Uniting biobank resources reveals novel genetic pathways modulating susceptibility for atopic dermatitis. *J Allergy Clin Immunol*. 2022;149:1105–1112.e1109.
- Dönertaş HM, Fabian DK, Valenzuela MF, Partridge L, Thornton JM. Common genetic associations between age-related diseases. *Nat Aging*. 2021;1:400–12.
- Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, Laurin C, Burgess S, Bowden J, Langdon R, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife*. 2018;7:7.
- Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for mendelian randomization. *Stat Methods Med Res*. 2017;26:2333–55.
- Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol*. 2013;37:658–65.
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol*. 1995;57:289–300.
- Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*. 2015;44:512–25.
- Burgess S, Thompson SG. Interpreting findings from mendelian randomization using the MR-Egger method. *Eur J Epidemiol*. 2017;32:377–89.

29. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from mendelian randomization between complex traits and diseases. *Nat Genet.* 2018;50:693–8.
30. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol.* 2016;40:304–14.
31. Zhao Q, Wang J, Hemani G, Bowden J, Small DS. Statistical inference in two-sample summary-data Mendelian randomization using robust adjusted profile score. *Ann Stat.* 2020;48:1742–69. <https://doi.org/10.1214/19-AOS1866>.
32. Slob EAW, Burgess S. A comparison of robust mendelian randomization methods using summary data. *Genet Epidemiol.* 2020;44:313–29.
33. Kamat MA, Blackshaw JA, Young R, Surendran P, Burgess S, Danesh J, Butterworth AS, Staley JR. PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. *Bioinformatics.* 2019;35:4851–3.
34. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ.* 2003;327:557–60.
35. Folkersen L, Fauman E, Sabater-Lleal M, Strawbridge RJ, Frånberg M, Sennblad B, Baldassarre D, Veglia F, Humphries SE, Rauramaa R, et al. Mapping of 79 loci for 83 plasma protein biomarkers in cardiovascular disease. *PLoS Genet.* 2017;13:e1006706.
36. Wang G, Sarkar A, Carbonetto P, Stephens M. A simple new approach to variable selection in regression, with application to genetic fine mapping. *J R Stat Soc Ser B Stat Methodol.* 2020;82:1273–300.
37. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, Plagnol V. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet.* 2014;10:e1004383.
38. Chen J, Xu F, Ruan X, Sun J, Zhang Y, Zhang H, Zhao J, Zheng J, Larsson SC, Wang X, et al. Therapeutic targets for inflammatory bowel disease: proteome-wide mendelian randomization and colocalization analyses. *EBioMedicine.* 2023;89:104494.
39. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, Gable AL, Fang T, Doncheva NT, Pyysalo S, et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* 2023;51:D638–646.
40. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 2019;47:D607–613.
41. Ren F, Jin Q, Liu T, Ren X, Zhan Y. Proteome-wide mendelian randomization study implicates therapeutic targets in common cancers. *J Transl Med.* 2023;21:646.
42. Breiteneder H, Peng YQ, Agache I, Diamant Z, Eiwegger T, Fokkens WJ, Traidl-Hoffmann C, Nadeau K, O'Hehir RE, O'Mahony L, et al. Biomarkers for diagnosis and prediction of therapy responses in allergic diseases and asthma. *Allergy.* 2020;75:3039–68.
43. Nieto-Fontarigo JJ, González-Barcala FJ, Andrade-Bulos LJ, San-José ME, Cruz MJ, Valdés-Cuadrado L, Crujeiras RM, Arias P, Salgado FJ. iTRAQ-based proteomic analysis reveals potential serum biomarkers of allergic and nonallergic asthma. *Allergy.* 2020;75:3171–83.
44. Foster PS, Maltby S, Rosenberg HF, Tay HL, Hogan SP, Collison AM, Yang M, Kaiko GE, Hansbro PM, Kumar RK, Mattes J. Modeling T(H) 2 responses and airway inflammation to understand fundamental mechanisms regulating the pathogenesis of asthma. *Immunol Rev.* 2017;278:20–40.
45. Baris S, Benamar M, Chen Q, Catak MC, Martínez-Blanco M, Wang M, Fong J, Massaad MJ, Sefer AP, Kara A, et al. Severe allergic dysregulation due to a gain of function mutation in the transcription factor STAT6. *J Allergy Clin Immunol.* 2023;152:182–194.e187.
46. Kerstan A, Bröcker EB, Trautmann A. Decisive role of tumor necrosis factor- $\alpha$  for spongiosis formation in acute eczematous dermatitis. *Arch Dermatol Res.* 2011;303:651–8.
47. Trautmann A, Akdis M, Kleemann D, Altnauer F, Simon HU, Graeve T, Noll M, Bröcker EB, Blaser K, Akdis CA. T cell-mediated Fas-induced keratinocyte apoptosis plays a key pathogenetic role in eczematous dermatitis. *J Clin Invest.* 2000;106:25–35.
48. Yawalkar N, Schmid S, Braathen LR, Pichler WJ. Perforin and granzyme B may contribute to skin inflammation in atopic dermatitis and psoriasis. *Br J Dermatol.* 2001;144:1133–9.
49. Peñaloza HF, van der Geest R, Ybe JA, Standiford TJ, Lee JS. Interleukin-36 cytokines in infectious and non-infectious lung diseases. *Front Immunol.* 2021;12:754702.
50. Ahmad F, Alam MA, Ansari AW, Jochebeth A, Leo R, Al-Abdulla MN, Al-Khawaga S, AlHammadi A, Al-Malki A, Al Naama K, et al. Emerging role of the IL-36/IL-36R Axis in multiple inflammatory skin diseases. *J Invest Dermatol.* 2024;144:206–24.
51. Fukaura R, Akiyama M. Targeting IL-36 in inflammatory skin diseases. *BioDrugs.* 2023;37:279–93.
52. Ni X, Xu Y, Wang W, Kong B, Ouyang J, Chen J, Yan M, Wu Y, Chen Q, Wang X, et al. IL-17D-induced inhibition of DDX5 expression in keratinocytes amplifies IL-36R-mediated skin inflammation. *Nat Immunol.* 2022;23:1577–87.
53. Fuchiwaki T, Sun X, Fujimura K, Yamada H, Shibata K, Muta H, Podack ER, Kawauchi H, Yoshikai Y. The central role of CD30L/CD30 interactions in allergic rhinitis pathogenesis in mice. *Eur J Immunol.* 2011;41:2947–54.
54. Song MJ, Park CH, Kim H, Han S, Lee SH, Lee DH, Chung JH. Carnitine acetyltransferase deficiency mediates mitochondrial dysfunction-induced cellular senescence in dermal fibroblasts. *Aging Cell.* 2023;22:e14000.
55. Davies MN, Kjalarsdóttir L, Thompson JW, Dubois LG, Stevens RD, Ilkayeva OR, Brosnan MJ, Rolph TP, Grimsrud PA, Muoio DM. The acetyl group buffering action of carnitine acetyltransferase offsets macronutrient-induced lysine acetylation of mitochondrial proteins. *Cell Rep.* 2016;14:243–54.
56. Zhang M, Bai Y, Wang Y, Cui H, Tang M, Wang L, Wang X, Gu D. Cumulative evidence for associations between genetic variants in interleukin 6 receptor gene and human diseases and phenotypes. *Front Immunol.* 2022;13:860703.
57. Esparza-Gordillo J, Schaarschmidt H, Liang L, Cookson W, Bauerfeind A, Lee-Kirsch MA, Nemat K, Henderson J, Paternoster L, Harper JJ, et al. A functional IL-6 receptor (IL6R) variant is a risk factor for persistent atopic dermatitis. *J Allergy Clin Immunol.* 2013;132:371–7.
58. Wang Y, Hu H, Wu J, Zhao X, Zhen Y, Wang S, Li W, Liang M, Wu B, Ma G. The IL6R gene polymorphisms are associated with sIL-6R, IgE and lung function in Chinese patients with asthma. *Gene.* 2016;585:51–7.
59. Martel BC, Litman T, Hald A, Norsgaard H, Lovato P, Dyring-Andersen B, Skov L, Thestrup-Pedersen K, Skov S, Skak K, Poulsen LK. Distinct molecular signatures of mild extrinsic and intrinsic atopic dermatitis. *Exp Dermatol.* 2016;25:453–9.
60. Brunner PM, Suárez-Fariñas M, He H, Malik K, Wen HC, Gonzalez J, Chan TC, Estrada Y, Zheng X, Khattri S, et al. The atopic dermatitis blood signature is characterized by increases in inflammatory and cardiovascular risk proteins. *Sci Rep.* 2017;7:8707.
61. Zhou Y, Xu M, Gong W, Kang X, Guo R, Wen J, Zhou D, Wang M, Shi D, Jing Q. Circulating MMP-12 as potential biomarker in evaluating disease severity and efficacy of sublingual immunotherapy in allergic rhinitis. *Mediators Inflamm.* 2022;2022:3378035.
62. Marinović Kulišić S, Takahashi M, Himelreich Perić M, Mužić Radović V, Jurakić Tončić R. Immunohistochemical analysis of adhesion molecules e-selectin, intercellular adhesion molecule-1, and vascular cell adhesion molecule-1 in inflammatory lesions of atopic dermatitis. *Life (Basel).* 2023;13:933.
63. Lugović L, Cupić H, Lipozencić J, Jakić-Razumović J. The role of adhesion molecules in atopic dermatitis. *Acta Dermatovenerol Croat.* 2006;14:2–7.
64. Bang H, Kim JE, Lee HS, Park SM, Park DJ, Lee EJ. Integrated bioinformatic analysis of gene expression profiling data to identify combinatorial biomarkers in inflammatory skin disease. *Sci Rep.* 2022;12:5889.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.