scientific reports

Check for updates

OPEN

Dissecting immune-mediated pathways in rheumatoid arthritis: A multivariate mediation analysis of antibodies and circulating proteins

Jingmin Niu[1,4], Pingxin Zhang[1,4], Wei Liu[1], Song Sun[1], Yingkai Zhang[1], Jinghao Sang[1], Jialin Yang[2], Qin Zhang[3]✉ & Limin Chai[1]✉

Rheumatoid arthritis (RA) is a chronic inflammatory disorder with complex etiologies involving immune responses and circulating proteins. This study investigates the causal relationships between antibody immune responses, plasma circulating proteins, and the development of RA using Mendelian Randomization (MR) analysis; A two-sample and multivariate MR analysis was conducted to explore the mediating causal relationship between 46 antibody immune responses and RA through 4,907 plasma circulating proteins. Genetic variations were utilized as instrumental variables (IVs) to infer causality, ensuring that they met the assumptions of relevance, independence, and exclusion restriction. Data were sourced from the FinnGen R10 dataset, UK Biobank, and the SomaScan platform, providing a robust foundation for the analysis. Statistical methods including IVW, weighted median, and mode-based approaches were employed, complemented by sensitivity analyses to ensure the robustness of the findings; The study identified significant causal relationships between six antibody immune responses and RA, with three specific responses—Epstein-Barr virus EBNA-1, Epstein-Barr virus ZEBRA, and Anti-polyomavirus 2 IgG seropositivity—showing strong associations. However, reverse causality was detected for EBNA-1 and ZEBRA, leading to their exclusion from further analysis. Additionally, 12 plasma circulating proteins were found to have significant causal relationships with RA, with KCNIP3 emerging as a key protective factor. Multivariate MR analysis revealed that KCNIP3 mediates the relationship between Anti-polyomavirus 2 IgG seropositivity and RA, suggesting a potential protective mechanism. This study highlights the intricate relationships between specific antibody responses, circulating proteins, and RA risk. The findings suggest that certain proteins, particularly KCNIP3, may mediate the effects of immune responses on RA development, offering potential targets for therapeutic intervention.

Keywords Rheumatoid arthritis, Mendelian randomization, Antibody immune response, Plasma Circulating proteins, Epstein-Barr virus, Polyomavirus.

Rheumatoid arthritis (RA) is a chronic, progressive autoimmune disorder that primarily affects synovial joints, leading to inflammation, pain, and eventual joint destruction[1–3]. It affects approximately 1% of the global population, with a higher prevalence in women and older adults[4]. The disease not only causes significant morbidity and reduced quality of life for affected individuals but also imposes a substantial economic burden on healthcare systems worldwide[5]. Despite advances in treatment options, including disease-modifying antirheumatic drugs (DMARDs) and biologic therapies, RA remains a challenging condition to manage, with many patients experiencing inadequate disease control or treatment-related side effects[6].

The etiology of RA is complex and multifactorial, involving an intricate interplay of genetic predisposition, environmental factors, and dysregulated immune responses[7]. While the exact mechanisms underlying RA pathogenesis are not fully elucidated, it is well-established that both adaptive and innate immune systems play crucial roles[8]. The hallmark of RA is the presence of autoantibodies, such as rheumatoid factor (RF) and anti-

[1]Key Laboratory of Chinese Internal Medicine of Ministry of Education and Beijing, Dongzhimen Hospital, Beiing University of Chinese Medicine, Beijing, China. [2]Zhongshan College of Dalian Medical University, Dalian, Liaoning Province, China. [3]Beijing Hospital of Traditional Chinese Medicine, Capital Medical University, Beijing, China. [4]Jingmin Niu and Pingxin Zhang contributed equally to this work. ✉email: qinzhang1974@126.com; liminchai@hotmail.com

citrullinated protein antibodies (ACPAs), which can be detected years before the onset of clinical symptoms[9]. These autoantibodies contribute to the initiation and perpetuation of the inflammatory cascade that characterizes RA.

Recent research has highlighted the potential role of infectious agents in triggering or exacerbating RA[10]. Several pathogens, including Epstein-Barr virus (EBV), Porphyromonas gingivalis, and Mycoplasma, have been implicated in RA pathogenesis[11]. The proposed mechanisms include molecular mimicry, where pathogen-derived antigens share structural similarities with host proteins, leading to cross-reactive immune responses[12]. Additionally, infections may alter the host immune environment, promoting the breakdown of self-tolerance and the development of autoimmunity[13].

The plasma proteome, comprising thousands of circulating proteins, represents a rich source of potential biomarkers and therapeutic targets in RA. These proteins, derived from various tissues and cell types, reflect the systemic nature of RA and may provide insights into disease activity, progression, and treatment response. Previous studies have identified several plasma proteins associated with RA, including cytokines (e.g., TNF-α, IL-6), acute-phase reactants (e.g., C-reactive protein), and matrix metalloproteinases[14]. However, the causal relationships between these proteins, immune responses, and RA development remain poorly understood.

Understanding the complex interplay between antibody immune responses, plasma circulating proteins, and RA is crucial for several reasons. First, it may lead to the identification of novel biomarkers for early diagnosis, disease stratification, and monitoring of treatment response[15]. Early intervention in RA is associated with better long-term outcomes, and reliable biomarkers could facilitate timely and targeted treatment initiation. Second, elucidating the causal pathways involved in RA pathogenesis may reveal new therapeutic targets, potentially leading to the development of more effective and personalized treatment strategies. Finally, a deeper understanding of the role of specific immune responses and circulating proteins in RA may shed light on the mechanisms of action of existing therapies and inform the rational design of combination treatments.

Traditional observational studies have provided valuable insights into RA pathogenesis, but they are limited by confounding factors and reverse causation, making it challenging to establish causal relationships[16]. Mendelian Randomization (MR) analysis offers an approach to overcome these limitations by using genetic variants as instrumental variables to infer causality. This method leverages the random assortment of genetic variants during meiosis, which is analogous to a randomized controlled trial, to provide more robust evidence for causal associations[17].

In this study, we employ a comprehensive two-sample and multivariate MR analysis to investigate the causal relationships between antibody immune responses, plasma circulating proteins, and the development of RA. By utilizing large-scale genetic data from well-characterized cohorts, including the FinnGen R10 dataset and the UK Biobank, we aim to provide a systems-level understanding of the complex biological processes underlying RA pathogenesis. Our approach examines 46 antibody immune responses and 4,907 plasma circulating proteins, offering a unprecedented breadth of analysis in the context of RA. The primary objectives of this study are to: (1) Identify specific antibody immune responses causally associated with RA risk. (2) Determine plasma circulating proteins that have causal relationships with RA development. (3) Investigate potential mediating effects of plasma proteins on the relationship between antibody immune responses and RA. (4) Explore the possibility of reverse causality and pleiotropy in the observed associations.

By addressing these objectives, we aim to contribute to the growing body of knowledge on RA pathogenesis and potentially identify novel targets for therapeutic intervention. The findings from this study may have significant implications for clinical practice, including improved risk stratification, early diagnosis, and personalized treatment approaches for individuals with RA or at risk of developing the disease.

Materials and methods

Study design

In this study, we employed two-sample and multivariate MR analysis to evaluate the mediating causal relationship between antibody immune responses and rheumatoid arthritis through the influence of plasma circulating proteins (Fig. 1). MR analysis leverages genetic variations as instrumental variables (IVs) to infer causality, requiring these IVs to meet three critical assumptions: (1) Relevance: The genetic variation must be directly related to the exposure of interest[18]. (2) Independence: The genetic variation should not be associated with any potential confounders that could bias the relationship between the exposure and the outcome[19]. (3)

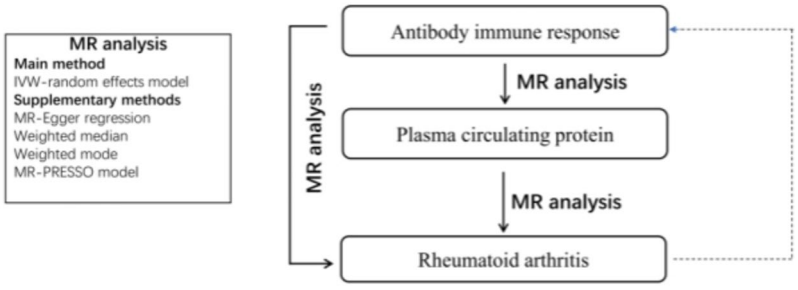

Fig. 1. The research design of this article uses antibody immune response as exposure, plasma circulating protein as mediator, and rheumatoid arthritis as outcome mediator for MR analysis.

Exclusion Restriction: The genetic variation should influence the outcome only through the exposure, without alternative pathways[20].

The data for rheumatoid arthritis were sourced from the FinnGen R10 dataset, which included 14,818 cases and 287,796 controls[21]. This large and well-characterized dataset provided a robust foundation for the analysis, enabling the identification of potential causal pathways between immune responses, circulating proteins, and rheumatoid arthritis.

GWAS data for all antibody immune response and plasma Circulating proteins

The study utilized data on 46 antibody immune responses, which were sourced from the GWAS catalog[22]. This dataset included up to 10,000 infectious disease serological measurements and whole-genome genotyping data from the UK Biobank cohort. The data represented immune responses to 13 different pathogens, categorized into 46 phenotypes: 15 serum-positive case-control phenotypes and 31 quantitative antibody measurement phenotypes. To analyze this data, the study employed a GWAS using the fastGWA linear mixed model software package, known for its efficiency in handling large-scale genetic data[23]. Additionally, human leukocyte antigen (HLA) classical allele and amino acid residue association analyses were performed using Lasso regression, providing detailed insights into the genetic underpinnings of these immune responses (PMID: 33204752).

For plasma circulating proteins, data was obtained from the SomaScan platform, identifying 28,191 genetic associations (with $P < 1.8e\text{-}9$) related to 4,907 aptamers among 35,559 Icelandic individuals[24]. The data was derived from two major projects: the Icelandic Cancer Project (ICP), which included 52% of the participants, and various genetic projects at deCODE Genetics in Reykjavik, Iceland, which accounted for the remaining 48%. The summary statistics for these associations were pre-calculated and analyzed using recursive conditional analysis. This method was used to identify the most significant variation in each region (± 1 Mb), categorized as sentinel protein quantitative trait loci (pQTL, $n = 18{,}084$), with other variations classified as secondary pQTLs ($n = 10{,}107$)[25]. This GWAS successfully replicated 83% of the pQTLs reported in the INTERVAL study (based on SomaScan) and 64% of those from the SCALLOP consortium (based on Olink) (PMID: 34857953).

Selection of instrumental variables (IVs)

Given the direct correlation between genetic variation and exposure, a stringent significance threshold was set for selecting IVs in the study. For each antibody immune response, the significance level was set at 1×10^{-5}, while for circulating proteins, a more stringent significance level of 5×10^{-8} was applied. These thresholds were chosen to ensure that the selected IVs were robust and highly relevant to the exposures being studied[26].

To ensure the independence of the selected IVs, we applied a linkage disequilibrium (LD) threshold of $R^2 < 0.001$. This threshold was used in conjunction with a 10,000 kb aggregation distance to group the "TwoSampleMR" data, helping to identify independent SNPs that could be used as reliable instruments in the analysis[27].

For the analysis of rheumatoid arthritis, we maintained the significance level at 5×10^{-8}, which is the standard for indicating genome-wide significance in GWAS studies[17]. The same LD threshold and aggregation distance were applied to ensure that only independent and statistically significant SNPs were included.

Additionally, we calculated the F-value for each SNP to assess its strength as an instrumental variable. SNPs with F-values less than 10 were excluded to avoid weak instrument bias, which could compromise the validity of the causal inference[28]. Palindromic SNPs, which can cause strand ambiguity, were also excluded to prevent potential errors in the analysis[29].

Statistical analysis

All statistical analyses were performed using R software version 4.4.1, a widely adopted environment for statistical computing and graphical representation (available at http://www.Rproject.org). The analysis was facilitated by the "TwoSampleMR" package (version 0.5.7), which is specifically designed for conducting MR studies[30]. This package provides a comprehensive suite of tools for estimating causal effects, performing hypothesis testing, and conducting sensitivity analyses.

The primary method used in this analysis was the IVW method, a standard approach in MR that combines Wald estimates from multiple genetic variants. These estimates are calculated as the ratio of SNP-outcome associations to SNP-exposure associations, and they are weighted by the inverse variance of each SNP-outcome association[31]. This method provides a powerful and statistically efficient means of estimating causal effects.

In addition to the IVW method, supplementary methods such as the weighted median and mode-based approaches were employed[32]. These methods are particularly useful when some instrumental variables may not fully meet the assumptions required for valid causal inference. They allow for robust causal estimates even in the presence of some invalid instruments, provided that a majority of the instruments are valid.

To further ensure the robustness of the findings, rigorous sensitivity analyses were conducted. This included Cochran's Q-test, which was used to examine heterogeneity among the instrumental variables[33]. The presence of heterogeneity could indicate that the instruments are not all estimating the same causal effect, so this test is crucial for assessing the consistency of the results.

Finally, multivariate MR analysis was performed to evaluate the stability and reliability of the mediation results[34]. This approach allows for the simultaneous consideration of multiple exposures or mediators, providing a more nuanced understanding of the causal relationships and ensuring that the final results are robust and reliable.

Mendelian mediation analysis

The analysis began by exploring the causal relationship between antibody immune responses and rheumatoid arthritis, with an emphasis on eliminating any potential reverse causality that could confound the results. This

step was crucial to ensure that the observed associations accurately reflect the direction of causality, confirming that immune responses are driving factors rather than consequences of rheumatoid arthritis.

Subsequently, the study focused on analyzing the causal effects of plasma circulating proteins on rheumatoid arthritis. By identifying and selecting positive associations for both antibody immune responses and plasma circulating proteins, we could conduct a more targeted MR analysis. This approach allowed us to explore how these circulating proteins might mediate the effects of immune responses on the risk of developing rheumatoid arthritis.

The overall impact of antibody immune responses on rheumatoid arthritis was then decomposed into direct and indirect effects. The indirect effect represents the pathway through which antibody immune responses influence rheumatoid arthritis via plasma circulating proteins. This effect was quantified by multiplying the causal relationship (β value) between antibody immune responses and plasma circulating proteins by the causal relationship (β value) between plasma circulating proteins and rheumatoid arthritis[35].

The direct effect, on the other hand, represents the portion of the immune response's impact on rheumatoid arthritis that is not mediated by circulating proteins. It was calculated by subtracting the indirect effect from the total effect[36]. This decomposition provides a nuanced understanding of how antibody immune responses contribute to the development of rheumatoid arthritis, both through direct mechanisms and via the mediation of circulating proteins.

When calculating OR (odds ratio) in the study, logistic regression models were mainly used to estimate the association between SNPs and disease outcomes. Used the generate_olds_ratios function in R software for conversion[37].

Results

The causal relationship between antibody immune response and rheumatoid arthritis

Using the IVW method, we identified a causal relationship between six antibody immune responses and rheumatoid arthritis at a significance level of 0.05 (Fig. 2) (Supplementary Material 1). The MR Egger, weighted median, and IVW methods[38,39] of these six antibodies all point towards the same direction of effect, indicating the robustness of the observed associations.

After applying the Bonferroni correction to the IVW results, adjusting the significance threshold to $P < 0.0011$ (0.05/46), it was found that three specific antibody immune responses showed a positive causal relationship with rheumatoid arthritis (Fig. 3) (Supplementary Material 2). These antibodies are: (1) Epstein-Barr virus EBNA-1 antibody levels: The IVW method showed a significant association with rheumatoid arthritis, with a p-value of < 0.001 and an odds ratio (OR) of 0.701 (95% CI: 0.598–0.821), indicating a protective effect. However, a reverse causal relationship was noted, and thus, it was excluded from further analysis. (2) Epstein-Barr virus ZEBRA antibody levels: Similarly, a significant association was found with an OR of 1.314 (95% CI: 1.121–1.541) under the IVW method. The reverse causal relationship was also observed here, which led to its exclusion from further analysis. (3) Anti-polyomavirus 2 IgG seropositivity: This antibody showed a significant positive association with rheumatoid arthritis, with an OR of 1.325 (95% CI: 1.140–1.540) under the IVW method, suggesting a strong causal effect.

The forest plot in Fig. 3 highlights these significant associations, with clear consistency across different MR methods. Importantly, no pleiotropy was detected in the remaining positive antibody immune responses, as indicated by the results of the sensitivity analyses (Supplementary Material 3). This strengthens the validity of the causal inferences drawn from these findings.

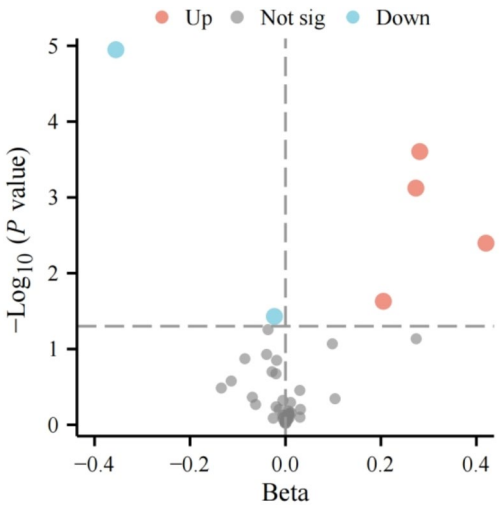

Fig. 2. IVW method displays the number of antibody immune responses and positive cases of rheumatoid arthritis.

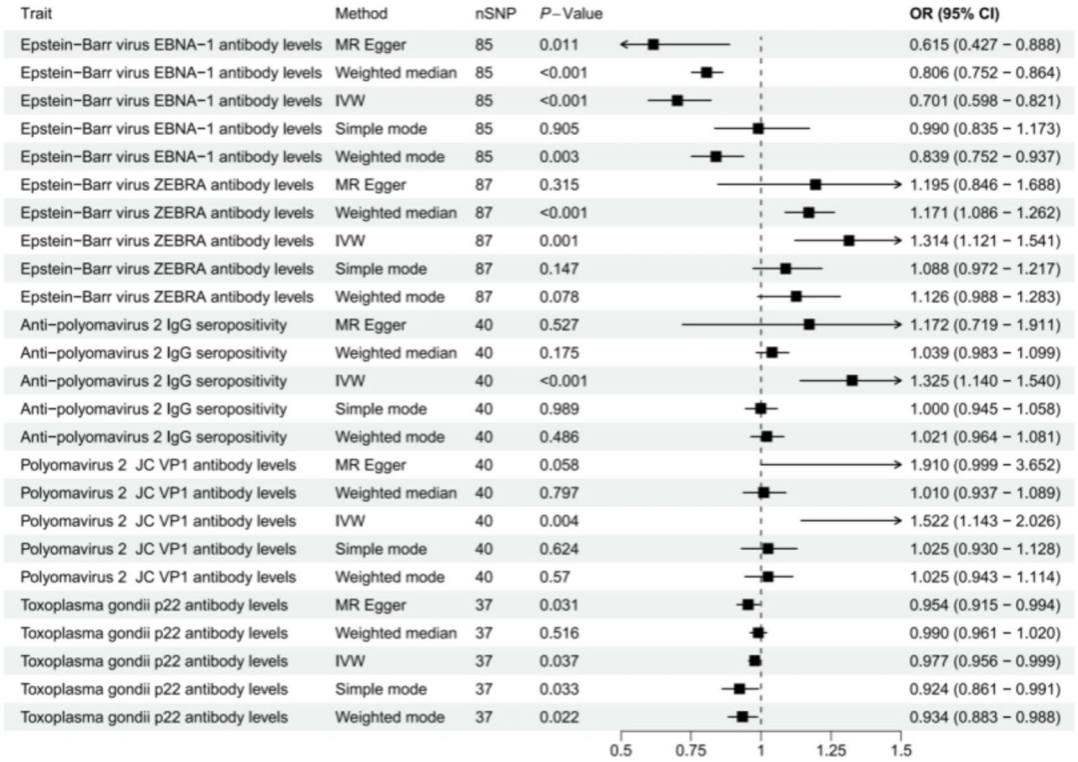

Fig. 3. Analysis of the causal relationship between antibody immune response and rheumatoid arthritis.

These results suggest that specific immune responses, particularly those related to Epstein-Barr virus and polyomavirus, may play a significant role in the development of rheumatoid arthritis, although reverse causality must be carefully considered in future analyses.

The causal relationship between plasma Circulating proteins and rheumatoid arthritis

Using the IVW method (for analyses with more than one SNP) and the Wald ratio method (for analyses with only one SNP) at a significance level of 0.05, a causal relationship was identified between 159 plasma circulating proteins and rheumatoid arthritis (Fig. 4). To adjust for multiple comparisons, the Bonferroni correction was applied, setting the threshold for significance at $P < 0.000014$ (0.05/3598).

After applying this correction, we identified 12 plasma circulating proteins with significant causal relationships with rheumatoid arthritis (Fig. 5) (Supplementary Material 4). Among these proteins: (1) KCNIP3: Identified using the Wald ratio method with a p-value of < 0.001 and an OR of 0.037 (95% CI: 0.022–0.065), suggesting a protective effect against rheumatoid arthritis. (2) SPAG11B: This protein was identified as having a significant causal relationship through multiple MR methods, including IVW (OR = 0.535, 95% CI: 0.416–0.688, $p < 0.001$), suggesting a protective effect as well. However, it was excluded from further analysis due to evidence of reverse causality and pleiotropy. (3) FAM177 A1: Despite showing significance in various methods, it was also excluded from further analysis due to the presence of pleiotropy and inconsistent results across different MR methods.

Some proteins, like COL8 A1 and SPAG11B, displayed reverse causal relationships with rheumatoid arthritis (Supplementary Material 5). Others, such as SPAG11B and FAM177 A1, showed evidence of pleiotropy (Supplementary Material 6), making them unsuitable for subsequent analyses.

These findings highlight the complexity of the relationships between plasma proteins and rheumatoid arthritis, indicating that while several proteins show significant associations, the presence of pleiotropy and reverse causality must be carefully considered in interpreting these results. The identification of proteins like KCNIP3 provides potential targets for further research and therapeutic intervention.

Mediation analysis using multivariate Mendelian randomization

In this study, we performed MR analysis focusing on positive antibody immunoreactivity, specifically Anti-poliovirus 2 IgG serosity, along with identified positive plasma circulating proteins. The objective was to determine potential mediating pathways through which these proteins might influence the relationship between antibody responses and rheumatoid arthritis.

After conducting the initial MR analysis, we further assessed the stability of the results using multivariate MR, which allows for the simultaneous analysis of multiple variables to account for potential confounders and ensure the robustness of the findings.

The multivariate MR analysis revealed that Anti-polyomavirus 2 IgG serosity played a significant role, with KCNIP3 identified as a key mediating pathway. The mediating effect of KCNIP3 was quantified at −0.086, with

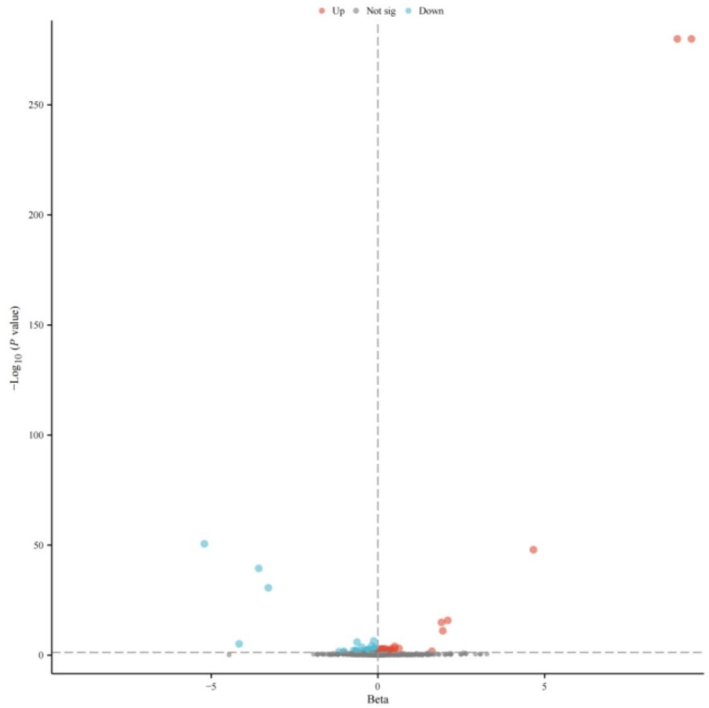

Fig. 4. IVW method displays the positive numbers of plasma circulating proteins and rheumatoid arthritis.

Trait	Method	nSNP	*P*−Value	OR (95% CI)
KCNIP3	Wald ratio	1	<0.001	0.037 (0.022 − 0.065)
ATAD2	Wald ratio	1	<0.001	0.028 (0.017 − 0.048)
HIF1A	Wald ratio	1	<0.001	0.005 (0.003 − 0.011)
CCDC25	Wald ratio	1	<0.001	7.002 (4.008 − 12.233)
COL8A1	Wald ratio	1	<0.001	106.220 (56.917 − 198.230)
NOTCH2	MR Egger	3	0.597	0.033 (0.000 − 297.429)
NOTCH2	Weighted median	3	<0.001	0.015 (0.004 − 0.067)
NOTCH2	IVW	3	<0.001	0.016 (0.003 − 0.096)
NOTCH2	Simple mode	3	0.055	0.008 (0.001 − 0.080)
NOTCH2	Weighted mode	3	0.037	0.014 (0.003 − 0.075)
SPAG11B	MR Egger	5	0.006	0.332 (0.246 − 0.449)
SPAG11B	Weighted median	5	<0.001	0.504 (0.431 − 0.590)
SPAG11B	IVW	5	<0.001	0.535 (0.416 − 0.688)
SPAG11B	Simple mode	5	0.059	0.517 (0.316 − 0.846)
SPAG11B	Weighted mode	5	0.001	0.497 (0.422 − 0.585)
ADPGK	Wald ratio	1	<0.001	6.734 (4.220 − 10.744)
FAM177A1	MR Egger	16	<0.001	0.807 (0.736 − 0.885)
FAM177A1	Weighted median	16	<0.001	0.887 (0.829 − 0.948)
FAM177A1	IVW	16	<0.001	0.883 (0.842 − 0.926)
FAM177A1	Simple mode	16	0.676	0.973 (0.859 − 1.102)
FAM177A1	Weighted mode	16	0.002	0.883 (0.826 − 0.944)
SPATA31D4	Wald ratio	1	<0.001	8.141 (4.948 − 13.393)
RNF150	Wald ratio	1	<0.001	7964.070 (4870.272 − 13023.177)
UBE2B	Wald ratio	1	<0.001	12151.726 (7261.219 − 20336.039)

Fig. 5. Analysis of the causal relationship between plasma circulating proteins and rheumatoid arthritis.

a confidence interval ranging from −0.172 to −0.023. This negative mediating effect suggests that the presence of KCNIP3 may reduce the risk associated with Anti-polyomavirus 2 IgG serosity, highlighting a potential protective mechanism against rheumatoid arthritis.

These findings underscore the importance of KCNIP3 in the complex interplay between immune responses and plasma proteins in the development of rheumatoid arthritis. The use of multivariate MR adds a layer of confidence to these results, indicating that KCNIP3 could be a valuable target for further research and therapeutic strategies aimed at mitigating the impact of specific antibody responses on rheumatoid arthritis.

Discussion

This study discovered a potential relationship between antibody immune response, plasma circulating proteins, and rheumatoid arthritis through MR. The genetic variations we use are fixed at conception, thus representing lifelong exposure and independent of the measurement time of biomarkers such as antibody levels[40]. And MR utilizes lifelong exposure inferred from genetic susceptibility, overcoming the limitations of cross-sectional measurements, which allows us to draw conclusions about potential causal relationships[41].

The causal relationship between serum positivity for anti-SARS-CoV-2 IgG and RA risk suggests that we have made progress in our research on environmental triggers of autoimmunity. This discovery is consistent with many pieces of evidence indicating that viral infections play an important role in the occurrence and transmission of autoimmune diseases[42]. The association between multi tumor virus exposure and RA risk adds a new dimension to the "multiple strikes" hypothesis of RA pathogenesis, which suggests that a combination of genetic susceptibility and environmental factors is a necessary condition for disease onset[43]. Our results suggest that polyomavirus infection could be one such environmental factor, potentially interacting with genetic predisposition to breach immunological tolerance.

The implications of this finding extend beyond mere association. It raises intriguing questions about the mechanisms through which polyomavirus infection might contribute to RA development. Possible pathways include molecular mimicry, where viral antigens share structural similarities with host proteins, leading to cross-reactive immune responses[44]. Alternatively, polyomavirus infection might induce epigenetic changes in immune cells, altering their reactivity and promoting autoimmunity[45]. These hypotheses warrant further investigation and could lead to novel approaches in RA prevention, such as the development of polyomavirus-targeted vaccines or antiviral therapies as preventive measures in high-risk individuals.

Our study also sheds light on the complex relationship between EBV and RA. While we identified significant associations between EBV antibodies (EBNA-1 and ZEBRA) and RA, the detection of reverse causality necessitated their exclusion from further analysis. This finding underscores the intricate and possibly bidirectional relationship between EBV infection and RA. It challenges the simplistic view of EBV as a straightforward trigger for RA and suggests a more nuanced interaction where RA-related immune dysregulation might influence EBV reactivation or antibody production. This observation calls for a reevaluation of the role of EBV in RA pathogenesis and highlights the need for longitudinal studies to disentangle the temporal relationship between EBV infection, immune responses, and RA onset[46].

Perhaps the most striking finding of our study is the identification of KCNIP3 as a plasma circulating protein with a strong protective effect against RA. The magnitude of this effect (OR = 0.037, 95% CI: 0.022–0.065) is remarkable and suggests that KCNIP3 could be a major player in RA pathophysiology. This discovery refines, rather than contradicts, the current understanding of RA as a disease driven by immune dysregulation and excessive inflammation[47]. While inflammation remains central to RA pathogenesis, our findings suggest that loss of intrinsic protective factors, such as KCNIP3, may facilitate or amplify inflammatory cascades, acting as a permissive condition rather than a direct trigger. In this view, KCNIP3 insufficiency may not oppose inflammation conceptually, but rather remove an endogenous brake that normally helps maintain immune homeostasis. Thus, both the presence of pro-inflammatory stimuli and the absence of regulatory counterbalances may jointly contribute to RA onset.

The multifunctional nature of KCNIP3, also known as calsenilin or DREAM (Downstream Regulatory Element Antagonist Modulator), opens up avenues for exploration[48]. Its role as a calcium-binding protein and transcriptional regulator suggests that it might influence RA pathogenesis through multiple mechanisms[49]. For instance, KCNIP3 could modulate the expression of inflammatory genes in synovial tissues, regulate calcium signaling in immune cells, or influence synovial fibroblast behavior[50]. The protective effect of KCNIP3 might also extend beyond its direct impact on inflammation. Its involvement in pain perception through regulation of potassium channels raises the intriguing possibility that KCNIP3 could modulate the neurogenic component of RA, influencing both pain perception and neurogenic inflammation in joints[51].

Our multivariate MR analysis revealed that KCNIP3 mediates the relationship between Anti-polyomavirus 2 IgG seropositivity and RA, with a significant negative mediating effect. This finding represents a paradigm shift in our understanding of how environmental exposures interact with host regulatory mechanisms to influence RA risk. Traditionally, research on RA pathogenesis has focused on the presence of harmful factors—such as autoantibodies, pro-inflammatory cytokines, and infections—as the main drivers of disease onset[52]. However, our results suggest that the absence or insufficiency of protective proteins like KCNIP3 may also play a critical role, not merely by failing to suppress inflammation, but by actively modulating the host's ability to resist pathogenic immune perturbations. In this context, KCNIP3 may act as a molecular buffer that mitigates the downstream inflammatory consequences of polyomavirus-induced immune activation. The implications for personalized medicine are notable. Beyond risk stratification, individuals with genetically or functionally reduced KCNIP3 expression may benefit from targeted surveillance or early interventions, such as lifestyle modifications, prophylactic antivirals, or future agents aimed at restoring KCNIP3 function. Thus, therapeutic strategies may be tailored not only to suppress disease-promoting pathways but also to augment innate protective mechanisms, representing a new frontier in RA management.

The identification of KCNIP3 as a key protective factor and mediator in RA pathogenesis opens up new horizons for therapeutic interventions. Traditional approaches to RA treatment have focused primarily on suppressing inflammation and modulating immune responses[53]. Our findings suggest that enhancing protective mechanisms, specifically by upregulating KCNIP3 expression or activity, could be an equally valid and potentially more physiological approach to RA prevention and treatment. This represents a shift from the current paradigm of immunosuppression towards a strategy of enhancing endogenous protective factors.

Developing pharmacological agents that modulate KCNIP3 activity or expression could offer a novel approach to managing RA, potentially with a more favorable side effect profile than current immunosuppressive therapies[54]. Moreover, the pleiotropic effects of KCNIP3 suggest that such interventions might have benefits beyond reducing inflammation, potentially addressing multiple aspects of RA pathology, including pain modulation and synovial hyperplasia[55].

The potential of KCNIP3 as a biomarker for RA susceptibility or disease progression is another avenue for translational research[56]. If validated, KCNIP3 levels could become part of a multi-biomarker panel for RA risk assessment, enabling earlier and more accurate identification of at-risk individuals. This could facilitate pre-emptive interventions, potentially altering the course of the disease before clinical symptoms manifest.

Our study also highlights the power and limitations of Mendelian Randomization in unraveling complex disease mechanisms. The exclusion of several proteins due to pleiotropy or inconsistent results across different MR methods underscores the complexity of genetic associations in autoimmune diseases[32]. It serves as a cautionary tale against over-interpretation of genetic data and emphasizes the need for rigorous statistical approaches and validation in MR studies[57].

While our findings provide valuable insights, they also raise numerous questions for future research. The molecular mechanisms through which KCNIP3 exerts its protective effects need to be elucidated. The interaction between KCNIP3 and polyomavirus-induced immune responses warrants detailed investigation, potentially through in vitro and animal model studies. Moreover, the generalizability of these findings to diverse populations needs to be established, as genetic and environmental factors may vary across different ethnic groups.

In summary, this study provides evidence for the causal relationship between antibody immune response, plasma circulating proteins, and RA. The identification of KCNIP3 as a protective factor and mediator of multi tumor virus associated RA risk represents a paradigm shift in our understanding of the pathogenesis of RA. These findings not only contribute to the scientific knowledge base, but also provide potential avenues for fundamentally changing RA diagnosis, risk stratification, and treatment. However, there are still some shortcomings in this article: firstly, the conclusions of this study are mainly based on genetic analysis, without evidence from functional experiments and longitudinal cohort studies. In future research directions, we will strive to find evidence from clinical studies and functional experiments.

Data availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Received: 1 November 2024; Accepted: 5 May 2025
Published online: 14 May 2025

References

1. Alivernini, S., Firestein, G. S. & McInnes, I. B. The pathogenesis of rheumatoid arthritis. *Immunity* **55** (12), 2255–2270 (2022).
2. Smolen, J. S. & Aletaha, D. Rheumatoid arthritis therapy reappraisal: strategies, opportunities and challenges. *Nat. Rev. Rheumatol.* **11** (5), 276–289 (2015).
3. Firestein, G. S. & McInnes, I. B. Immunopathogenesis of rheumatoid arthritis. *Immunity* **46** (2), 183–196 (2017).
4. Cross, M. et al. The global burden of rheumatoid arthritis: estimates from the global burden of disease 2010 study. *Ann. Rheum. Dis.* **73** (7), 1316–1322 (2014).
5. Birnbaum, H. et al. Societal cost of rheumatoid arthritis patients in the U.S. *Curr. Med. Res. Opin.* **26** (1), 77–90 (2010).
6. Nam, J. L. et al. Efficacy of biological disease-modifying antirheumatic drugs: A systematic literature review informing the 2013 update of the EULAR recommendations for the management of rheumatoid arthritis. *Ann. Rheum. Dis.* **73** (3), 516–528 (2014).
7. Raychaudhuri, S. Recent advances in the genetics of rheumatoid arthritis. *Curr. Opin. Rheumatol.* **22** (2), 109–118 (2010).
8. Weyand, C. M. & Goronzy, J. J. Immunology of rheumatoid arthritis. *Nat. Rev. Immunol.* **13** (1), 39–51 (2013).
9. van der Woude, D. & van der Helm-van Mil, A. H. Update on the epidemiology, risk factors, and disease outcomes of rheumatoid arthritis. *Best Pract. Res. Clin. Rheumatol.* **32** (2), 131–150 (2018).
10. Smolik, I., Robinson, D. & Bernstein, C. N. The role of infectious agents in the development of rheumatoid arthritis. *J. Intern. Med.* **268** (5), 526–538 (2010).
11. van de Sande, M. G. et al. Mycoplasma and rheumatoid arthritis: do they trigger the disease? *Rheumatol. (Oxford).* **49** (6), 1057–1064 (2010).
12. Kokkonen, H. et al. Antibodies of IgG, IgA and IgM isotypes against Cyclic citrullinated peptide precede the development of rheumatoid arthritis. *Arthritis Res. Therapy*. **13**, 1–10 (2011).
13. Ravelli, A. & Martini, A. Juvenile idiopathic arthritis. *Lancet* **369** (9563), 767–778 (2007).
14. Miyoshi, M. et al. The impact of matrix metalloproteinases and their inhibitors on rheumatoid arthritis. *Front. Biosci.* **21**, 919–927 (2016).
15. Van Baarsen, L. G. et al. Gene expression profiling in rheumatoid arthritis: Understanding disease mechanisms and predicting treatment responses. *BMC Genom.* **11**, 214 (2010).
16. Morabia, A. History of the modern epidemiological concept of confounding. *J. Epidemiol. Community Health*. **65** (4), 297–300 (2011).
17. Burgess, S., Butterworth, A. & Thompson, S. G. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet. Epidemiol.* **37** (7), 658–665 (2013).
18. Burgess, S. & Thompson, S. G. *Mendelian Randomization: Methods for Using Genetic Variants in Causal Estimation* (Chapman & Hall/CRC, 2015).

19. Yarmolinsky, J. et al. Causal inference in cancer epidemiology: what is the role of Mendelian randomization? *Cancer Epidemiol. Biomarkers Prev.* **27** (9), 995–1010 (2018).
20. Lawlor, D. A. et al. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat. Med.* **27** (8), 1133–1163 (2008).
21. Kurki, M. I. et al. FinnGen: Unique genetic insights from combining isolated population and national health register data. medRxiv. ; preprint. (2022).
22. Bycroft, C. et al. The UK biobank resource with deep phenotyping and genomic data. *Nature* **562** (7726), 203–209 (2018).
23. Loh, P. R. et al. Efficient bayesian mixed-model analysis increases association power in large cohorts. *Nat. Genet.* **47** (3), 284–290 (2015).
24. Ho, J. E. et al. Protein biomarkers of cardiovascular disease and mortality in the community. *J. Am. Heart Association*. **7** (14), e008108 (2018).
25. Ferkingstad, E. et al. Large-scale integration of the plasma proteome with genetics and disease. *Nat. Genet.* **53** (12), 1712–1721 (2021).
26. Visscher, P. M., Brown, M. A., McCarthy, M. I. & Yang, J. Five years of GWAS discovery. *Am. J. Hum. Genet.* **90** (1), 7–24 (2012).
27. Hemani, G. et al. The MR-Base platform supports systematic causal inference across the human phenome. *eLife* **7**, e34408 (2018).
28. Andrews, I., Stock, J. H. & Sun, L. Weak instruments in instrumental variables regression: theory and practice. *Annual Rev. Econ.* **11** (1), 727–753 (2019).
29. Palmer, T. M. et al. Using multiple genetic variants as instrumental variables for modifiable risk factors. *Stat. Methods Med. Res.* **21** (3), 223–242 (2012).
30. Hemani, G., Bowden, J. & Davey Smith, G. Evaluating the potential role of Pleiotropy in Mendelian randomization studies. *Hum. Mol. Genet.* **27** (R2), R195–R208 (2018).
31. Burgess, S. & Thompson, S. G. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur. J. Epidemiol.* **32** (5), 377–389 (2017).
32. Bowden, J., Davey Smith, G., Haycock, P. C. & Burgess, S. Consistent Estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet. Epidemiol.* **40** (4), 304–314 (2016).
33. Ruppar, T. Meta-analysis: how to quantify and explain heterogeneity? *Eur. J. Cardiovasc. Nurs.* **19** (7), 646–652 (2020).
34. Sanderson, E., Davey Smith, G., Windmeijer, F. & Bowden, J. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. *Int. J. Epidemiol.* **48** (3), 713–727 (2019).
35. Böhnke, J. R. Explanation in causal inference: methods for mediation and interaction. *Q. J. Exp. Psychol. (Hove).* **69** (6), 1243–1244 (2016).
36. Namazi, M. & Namazi, N. R. Conceptual analysis of moderator and mediator variables in business research. *Procedia Econ. Finance*. **36**, 540–554 (2016).
37. Spiller, W., Davies, N. M. & Palmer, T. M. Software application profile: mrrobust—a tool for performing two-sample summary Mendelian randomization analyses. *Int. J. Epidemiol.* **48** (3), 684–690 (2019).
38. Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid instruments: effect Estimation and bias detection through Egger regression. *Int. J. Epidemiol.* **44** (2), 512–525 (2015).
39. Slob, E. A. W. & Burgess, S. A comparison of robust Mendelian randomization methods using summary data. *Genet. Epidemiol.* **44** (4), 313–329 (2020).
40. Davey Smith, G. & Hemani, G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum. Mol. Genet.* **23** (R1), R89–R98 (2014).
41. Lawlor, D. A., Harbord, R. M., Sterne, J. A., Timpson, N. & Davey Smith, G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat. Med.* **27** (8), 1133–1163 (2008).
42. Langer-Gould, A. et al. Epstein-Barr virus, cytomegalovirus, and multiple sclerosis susceptibility: A multiethnic study. *Neurology* **89** (13), 1330–1337 (2017).
43. Holers, V. M. et al. Rheumatoid arthritis and the mucosal origins hypothesis: protection turns to destruction. *Nat. Rev. Rheumatol.* **14** (9), 542–557 (2018).
44. Cusick, M. F., Libbey, J. E. & Fujinami, R. S. Molecular mimicry as a mechanism of autoimmune disease. *Clin. Rev. Allergy Immunol.* **42** (1), 102–111 (2012).
45. Mei, X. et al. An update on epigenetic regulation in autoimmune diseases. *J. Translational Autoimmun.* **5**, 100176 (2022).
46. James, J. A. & Robertson, J. M. Lupus and Epstein-Barr. *Curr. Opin. Rheumatol.* **24** (4), 383–388 (2012).
47. McInnes, I. B. & Schett, G. The pathogenesis of rheumatoid arthritis. *N. Engl. J. Med.* **365** (23), 2205–2219 (2011).
48. Formisano, L. et al. Transcriptional and epigenetic regulation of ncx1 and ncx3 in the brain. *Cell. Calcium*. **87**, 102194 (2020).
49. Buxbaum, J. D. A role for calsenilin and related proteins in multiple aspects of neuronal function. *Biochem. Biophys. Res. Commun.* **322** (4), 1140–1144 (2004).
50. Kim, H. A. et al. Regulation of calcium signaling in rheumatoid arthritis: implications for pathogenesis and treatment. *Rheumatol. Int.* **36** (11), 1485–1493 (2016).
51. Carrillo-Salinas, F. J. et al. DREAM regulates immune and glial responses in experimental autoimmune encephalomyelitis. *Glia* **62** (1), 115–127 (2014).
52. Jang, S., Kwon, E. J. & Lee, J. J. Rheumatoid arthritis: pathogenic roles of diverse immune cells. *Int. J. Mol. Sci.* **23** (2), 905 (2022).
53. Burgess, S. et al. Sensitivity analyses for robust causal inference from Mendelian randomization analyses with multiple genetic variants. *Genet. Epidemiol.* **41** (6), 508–524 (2017).
54. Sanders, K. M., McDougall, J. A., Evans, J. P., Dunstan, C. R. & Cicuttini, F. M. Treating osteoarthritis: what have we learned from targeting subchondral bone? *Nat. Rev. Rheumatol.* **14** (10), 635–647 (2018).
55. Edilova, M. I., Akram, A. & Abdul-Sater, A. A. Innate immunity drives pathogenesis of rheumatoid arthritis. *Biomedical J.* **44** (2), 172–182 (2021).
56. de Vries, R. et al. Biomarkers for rheumatoid arthritis: where are we now? *Autoimmun. Rev.* **14** (12), 1116–1120 (2015).
57. Stogiannis, D., Siannis, F. & Androulakis, E. Heterogeneity in meta-analysis: a comprehensive overview. *Int. J. Biostatistics*. **20** (1), 169–199 (2024).

Acknowledgements

We would like to express our gratitude to all the participants involved in the studies from which data was obtained, including the FinnGen R10 dataset and the UK Biobank.

Author contributions

J.N. was responsible for assisting P.Z. in conducting research in this field, processingexperimental data, and writing this article. J.N. and P.Z. contributed to the writing and reviewing of the manuscript.S.S., Y.Z., J.S., J.Y. and W.L. were involved in data analysis. All authors contributed to the article and approved the submitted version. All authors have read and agreed to the published version of the manuscript.

Funding information
This study was supported by the National Natural Science Foundation of China (82174338).

Declarations

Competing interests
The authors declare no competing interests.

Conflict of interest
The authors have no conflict of interest.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-01216-7.

Correspondence and requests for materials should be addressed to Q.Z. or L.C.

Reprints and permissions information is available at www.nature.com/reprints.

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