ImmunoTargets and Therapy

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Genetic and Plasma Proteomic Approaches to Identify Therapeutic Targets for Graves' Disease and Graves' Ophthalmopathy

Chenxin Ke[1,*], Yuefeng Yu[1,*], Jiang Li[1], Yuetian Yu[1], Ying Sun[1], Yuying Wang[1], Bin Wang[1], Yingli Lu[1], Mengjun Tang[2], Ningjian Wang[1], Yi Chen[1]

[1]Institute and Department of Endocrinology and Metabolism, Shanghai Ninth People's Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, People's Republic of China; [2]Orthopedic Department, Taizhou Hospital of Zhejiang Province, Zhejiang University, Taizhou, People's Republic of China

*These authors contributed equally to this work

Correspondence: Yi Chen, Email chenyi9h@126.com; Ningjian Wang, Email wnj486@126.com

Background: The blood proteome is a major source of biomarkers and therapeutic targets. We aimed to identify the causal proteins and potential targets for Graves' disease (GD) and Graves' ophthalmopathy (GO) via systematic genetic analyses.

Methods: Genome-wide association studies (GWASs) on the UK Biobank- Pharma Proteomics Project (UKB-PPP) collected 2923 Olink proteins from 54,219 participants. We conducted a proteome-wide Mendelian randomization (MR) study with cis-pQTLs to identify candidate proteins for GD and GO risk. Colocalization analysis and the Heidi test were used to examine whether the identified proteins and diseases shared the same variant. More proteins with potential causal associations were identified in Summary-data-based MR (SMR) analyses using trans-pQTLs. Then, downstream analyses were performed to detect protein interactions, gene function, cell type-specific expression and druggable information.

Results: This study genetically predicted levels of 62 plasma proteins were associated with GD risk. Four proteins (CD40, TINAGL1, GMPR and CXCL10) were prioritized with the evidence of sharing the same variants with GD. Specifically, some proteins had potential associations with GD with trans-pQTLs mapping in CD40. The four prioritized protein-coding genes were mainly enriched in the regulation of apoptotic and death processes. In addition, GMPR was associated with both GO and GD in a consistent direction. BTN1A1 and FCRL1 were prioritized as the causal proteins for GO onset and were not associated with GD.

Conclusion: By synthesizing proteomic and genetic data, we identified several protein biomarkers for GD, with one linked to both GD and GO and two other protein biomarkers specific to GO onset, which provides valuable insights into the etiology and potential therapeutic targets for the two diseases.

Keywords: plasma proteomics, Graves' disease, Olink, Mendelian randomization

Introduction

Graves' disease (GD) is an autoimmune disorder specifically targeting the thyroid gland and is the most common cause of hyperthyroidism. It is driven by circulating autoantibodies that stimulate the thyroid-stimulating hormone receptor (TSH-R). The disease arises from the breakdown of immune tolerance to thyroid antigens, primarily involving the TSH receptor.[1,2] Epidemiological data show that the incidence of GD is 20–40 cases per 100,000 population per year,[3] which significantly impacts the quality of life and is correlated with an increased risk of death.[4] Research has been conducted to better prevent and control GD. In uncomplicated cases of GD, antithyroid medication remains the first line of treatment in Europe.[5,6] Moreover, radioactive iodine is increasingly being replaced by antithyroid drugs. Methimazole and propylthiouracil are widely used in GD treatment because they block thyroid hormone synthesis through the inhibition

Received: 3 October 2024
Accepted: 22 January 2025
Published: 7 February 2025

of thyroid peroxidase. However, both methimazole and propylthiouracil have a high risk of relapse once the drug is discontinued. Therefore, the development of new target drugs is especially important.

In an effort to increase the understanding of the essence of the disease, several studies involving whole-genome sequencing (GWAS) of GD have now been reported, combining a series of downstream analyses to provide a detailed resolution of the genetic architecture of GD.[7–9] However, few research has been done on the relationship between plasma proteins and GD risk, which indicates that proteomic etiology of GD remains to be fully explored. It is well known that proteome-wide Mendelian randomization studies, combining proteomics with genetic data, have identified new risk factors and potential therapeutic targets for other autoimmune diseases, such as rheumatoid arthritis and psoriasis.[10] Assessment of the proteomic basis of GD is also important to better understand the underlying pathophysiology and possibly manage disease progression.

Furthermore, Graves' ophthalmopathy (GO), an inflammatory disease affecting the extraocular muscles and the orbital tissues, is the most common extra-thyroidal manifestation of GD.[11] Histopathologic changes, including immune cell infiltration, hyaluronan accumulation, and lipogenesis, lead to orbital tissue expansion and muscle hypertrophy. These structural alterations manifest as eyelid retraction, proptosis, and even vision loss, leading to a heavy burden for GD patients.[12] Assessment of the shared and differential proteomic basis between GD and GO is important to better understand the underlying pathophysiology and possibly manage disease progression. Recently, some studies in small-population cohorts have utilized proteomic analysis to identify potential biomarkers of GO activity and progression.[13,14] For example, the proinflammatory cytokine, IL-17 was identified as a significant component that promotes the progression of GO.[15] However, existing studies have yet to fully explore these dimensions, and there is an insufficient understanding of the associated risk proteins.

In order to systematically identify circulating protein biomarkers associated with GD and GO risk, a proteome-wide MR analysis was employed in our study. Subsequently, colocalization analyses and heterogeneity in dependent instruments (HEIDI) tests were conducted to verify the associations of these proteins with GD. Furthermore, more detailed information was performed about single cell-type gene expression, protein interactions, functional enrichment and druggability evaluation of the identified protein-coding genes.

Methods

Study Population

There were a total of 54,219 individuals, including 46,595 randomly selected UKB participants at the baseline visit to provide base group representation, as well as a targeted complement of 6,376 participants with specific phenotypes or diseases based on the research objectives of the UKB-PPP consortium members, and 1,268 individuals who participated in COVID-19-related studies at multiple visits.[16]

Proteomic Data Sources

GWAS from a large-scaled proteomic study with available full summary-level data, derived from a platform (antibody-based Olink[17]), was used in our study. The UK Biobank Pharma Proteomics Project (UKB-PPP) is a precompetitive consortium of 13 biopharmaceutical companies that funded the generation of multiplexed, population-scale proteomic data, which utilized the Olink platform to perform proteomic analysis on plasma samples from 54,219 participants, collecting data on 2,923 proteins.[16] The detailed plan for plasma sample collection and processing, as well as proteomic analysis techniques, has been presented in the Supplementary Material (Supplementary Methods).

GWAS Data Sources

We used the latest release data on GD from the FinnGen study R10, which included 3,176 cases and 409,005 controls.[18] The definition of GD in FinnGen study R10 is according to the ICD (E4_Graves' disease). The FinnGen GWAS is a public–private partnership research project, integrating imputed genotype data generated from newly collected and legacy samples from the Finnish Biobank and digital health record data from the Finnish Health Registry. No sample overlaps were found between it and proteomic datasets. Likewise, data on GO (E4_Graves' ophthalmopathy) also

obtained from the FinnGen GWAS R10, including 691 cases and 411,490 controls. In sensitivity analysis, GWAS for GD based on UK Biobank was used, which was conducted from 1,678 cases and 456,942 controls.[19]

MR Analysis

We performed two-sample MR analyses using the proteins with cis-pQTLs (protein Quantitative Trait Loci) identified as exposures and GWAS of GD or GO as outcomes. SNPs were selected within 1 Mb from the gene encoding the protein and associated with the levels of plasma proteins at the genome-wide significance level ($p < 5*10^{-8}$) (Supplementary Information). SNPs without linkage disequilibrium, with thresholds of r2 >0.001 on the basis of the 1000 Genomes European panel, were included as instrumental variables. The statistical power satisfied statistical requirements, for the F statistics of all instrumental variables were higher than 30. Single nucleotide polymorphisms (SNPs) for pQTLs that were not available in outcome data were substituted with SNP proxies exhibiting high linkage disequilibrium ($R^2 \geq 0.8$) based on the 1000 Genomes European panel.[20] In total, 2080 plasma proteins with at least one pQTL were included in the present study. Missing SNPs without suitable SNP proxies were removed from the analysis. In the case of a single independent instrumental variable (IV), Wald Ratio was applied, otherwise, inverse variance weighted (IVW) estimates were reported. Radial MR and MR-presso were used to remove IVs with pleiotropy and heterogeneity as much as possible. We used the Bonferroni method in order to correct for multiple testing, with a p-value threshold of 2.4E-5 (0.05/2080 proteins). For the proteins significantly associated with GD or GO risk, sensitivity analyses were carried out to display the possibility that causality is confounded by the pleiotropy and heterogeneity. In addition, we performed a reverse MR analysis to estimate the associations of GD liability with the levels of identified proteins to explore if the results could be distorted by reverse causality.

Colocalization Analysis and HEIDI Test

To distinguish the associations of proteins with GD driven by a shared causal variant from linkage disequilibrium, we conducted HEIDI test and colocalization analysis.

HEIDI test is under a frequentist framework, which considers the entire cis-pQTL region as a whole and treats the null hypothesis as colocalization.[21] Hypothesis testing for this method is based on the theory that there is consistency in the estimates of the association between exposure and outcome based on any SNP within LD. The LD matrix of the enrolled SNPs is taken into account in the algorithm. A p-value >0.05 distinguishes the association caused by the same SNP from confounded by two SNPs with linkage disequilibrium.

The colocalization analysis identified the specific SNP sharing between the protein level and the diseases, based on a Bayesian model.[22] We included SNPs within ±200 kb of the SNP in each pQTL, and performed the highest colocalization probability among all pQTLs of a protein. We set the prior probability that a SNP is associated with a protein or a disease as 1×10^{-4}, and set the prior probability that a SNP is associated both a protein and a disease as 1×10^{-5}. For the case where multiple pQTLs exist for one protein, we performed a colocalization analysis for each pQTL and showed the maximum probability. Moreover, we included all the pQTLs of a protein as a whole and repeated the colocalization analysis under the sum of single effects (SuSiE) framework to isolate the statistical support of colocalization for protein and disease on each SNP, which could provide more precise inferences when there are multiple SNPs simultaneously and significantly associated with both exposure and outcome.[23] The posterior probability for H4 (PPH4) greater than 0.5 was considered evidence of colocalization between protein and disease, and greater than 0.8 was strong evidence.

Identification of Prioritizing Probable Trans-pQTLs

The trans-pQTLs are defined as signals beyond 1 Mb from the gene encoding the protein, meeting the threshold of genome-wide significance level. Summary-data-based Mendelian randomization (SMR)[21] was adapted to assess the causal associations of each independent trans-pQTL with Graves' disease. For each trans-pQTL significantly associated with Graves' disease, we queried the trans-pQTL sentinel variant in cis-pQTLs from UKB-PPP to identify cis-pQTLs that could mediate the trans-pQTL. Referring to ProGeM tool,[24] a prioritize probable trans-pQTL satisfies the following

conditions: 1) the sentinel variant mapped in a protein identified in cis-pQTL MR analyses; 2) the sentinel variant was the same as the cis-pQTL sentinel variant for the protein, or the two were in high LD ($r2 > 0.8$).

Differential Gene Expression Analysis

The cell type-specific expression of genes of identified plasma protein was further evaluated by using scRNA-seq data of cross-tissue immune cells across the entire body.[25] A total of 45 immune cells from 16 different tissues (five lymphoid tissues and 11 non-lymphoid tissues) were taken into consideration. The differential gene expression between each cell type and other cell types was performed based on Wilcoxon's Rank Sum test to examine whether the encoded genes of identified proteins were highly expressed in a particular cell type.[26]

Functional Enrichment and Protein–Protein Interaction (PPI)

Furthermore, Gene Ontology annotation enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis were conducted to characterize the biological properties of the identified proteins, including their molecular function, cellular component, biological process, signalling pathways and metabolic pathways.[27] These enrichment analyses were carried out using tools such as ClusterProfiler and DAVID, which enable comprehensive and high-throughput functional analysis of large proteomics datasets. A PPI network was constructed to explore the potential interactions between the identified proteins using the STRING database, which integrates known and predicted interactions based on experimental data, computational predictions, and text mining.[28]

Druggability Evaluation

To assess whether the identified proteins could be potential therapeutic targets, we searched the Open Target platform (https://genetics.opentargets.org), an open-access resource that helps scientists translate GWAS and Biobank data into target genes across thousands of traits using customizable workflows. Drug names and the development process of drugs that target identified proteins were recorded on this platform.[29]

Results

MR Analysis to Identify the Causal Proteins for GD

An overview of the study design is shown in Figure 1. After Bonferroni correction, 62 proteins were identified to be significantly associated with GD (Figure 2 and Supplementary Table 1). No causality was confounded by heterogeneity and pleiotropy in sensitivity analyses for proteins with more than two pQTLs (Supplementary Table 2). Eighteen of the 62 causal proteins had consistent significance in replication analysis with GD GWAS from the UKB dataset (Supplementary Table 3).

Reverse MR Analysis to Test the Causal Direction

Reverse MR was completed on the identified 62 proteins to estimate the genetic liability of GD with level of proteins and revealed one protein (TBC1D17) with the reverse causality for genetic predisposition to GD on protein level (Supplementary Table 4), which was excluded from following analyses. The results indicated that 61/62 of the associations had the correct causal directions from the proteins to GD.

Colocalization Analysis and Heidi Test

Forty-eight out of 62 proteins did not reject the null hypothesis of Heidi tests, suggesting that there was a potential possibility of sharing genetic variants between the protein and GD (Supplementary Table 5). On the other hand, six (CD40, TINAGL1, GMPR, CXCL10, BTN3A2, and HLA-DRA) proteins were supported by evidence of genetic colocalization (PP4 > 0.5), indicating a specific shared causal variant between the protein level and GD. In the colocalization analysis based on a SuSiE framework, we identified three proteins (CD40, BTN3A2, and BCL2L15) sharing multiple variants with GD (Supplementary Table 5). Ultimately, with the selection of proteins that pass the Heidi test and colocalization, four prioritized proteins (CD40, TINAGL1, GMPR and CXCL10) were identified (Figure 3).

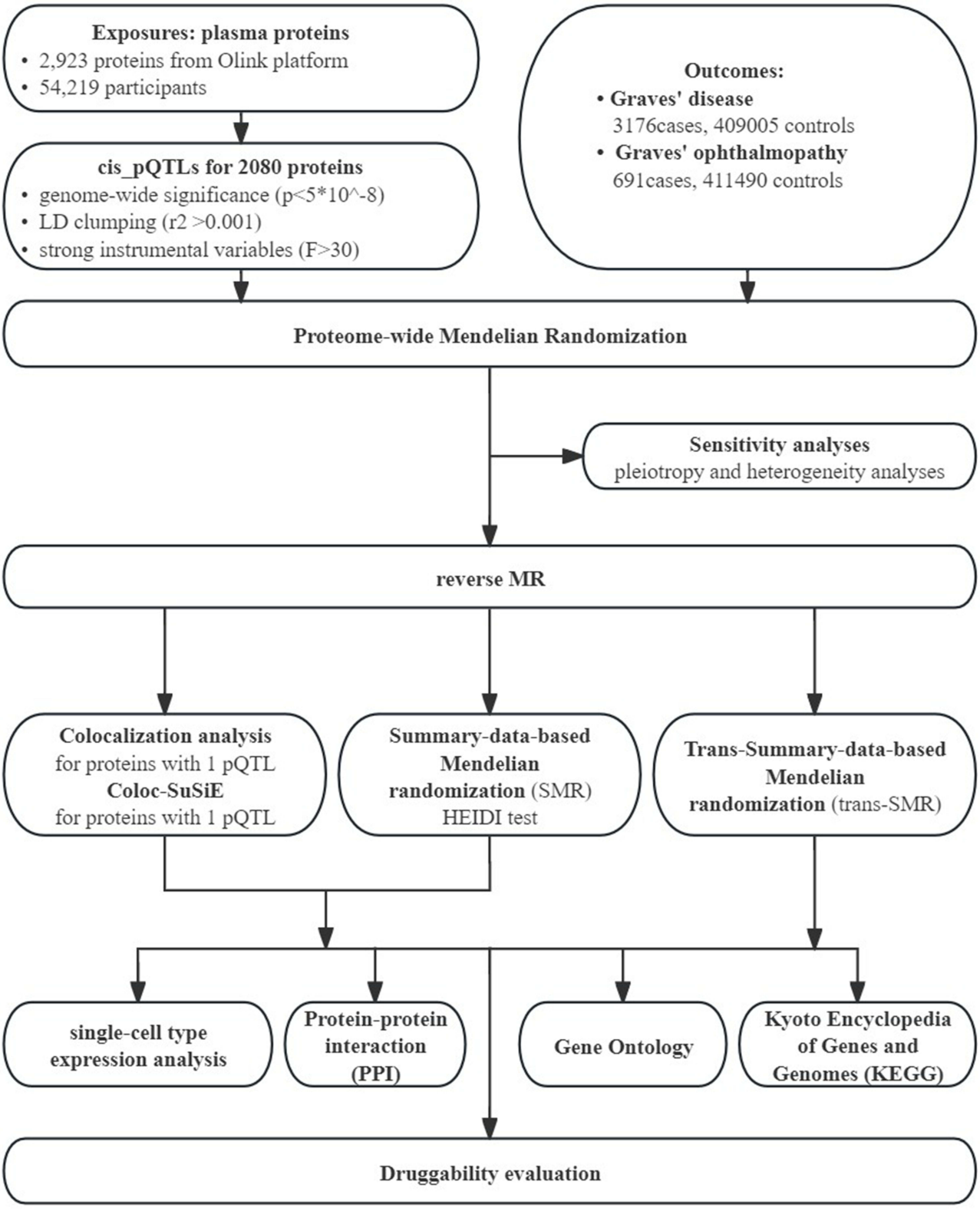

Figure 1 Flowchart of the study.

https://doi.org/10.2147/ITT.S494692

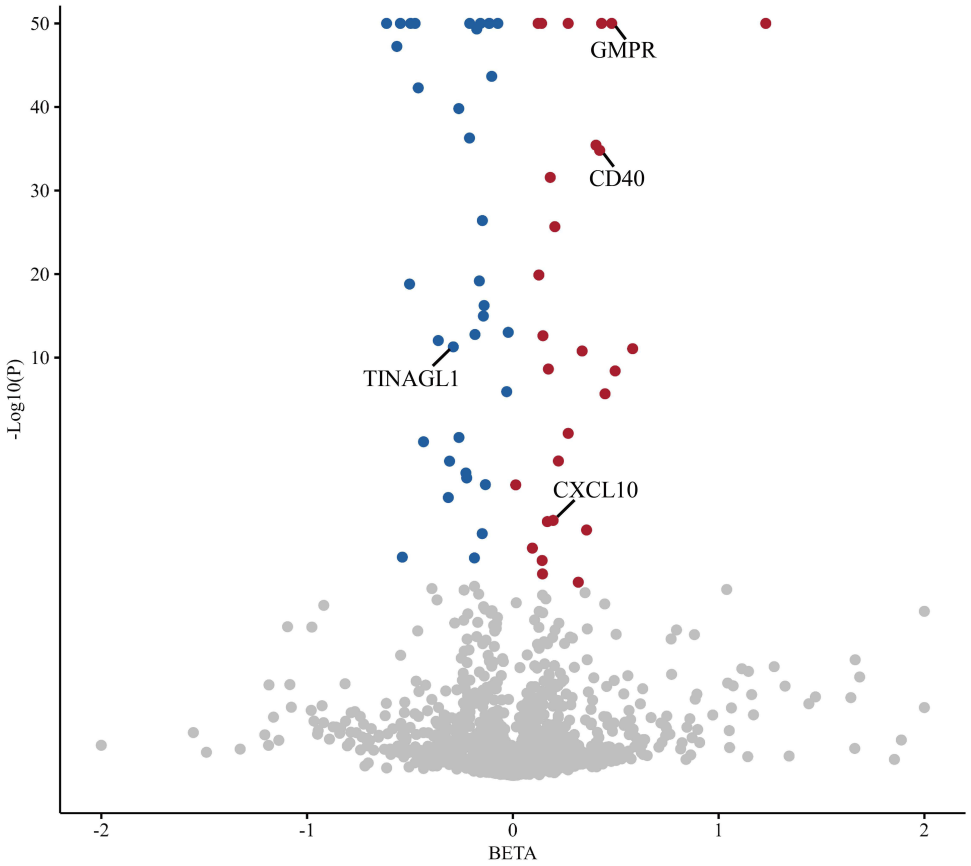

Figure 2 Causal effects of the plasma proteins on GD. Volcano plot for the results of 2080 proteins on GD. Blue plots represented negative association, red plots represented positive association, while gray plots represented associations that did not reach the threshold of significance.

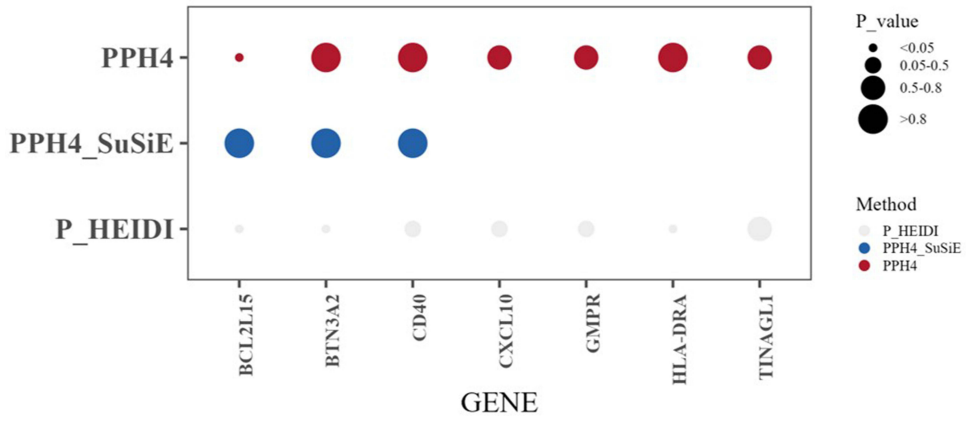

Figure 3 Bubble chat of colocalization analysis of 7 proteins with evidence of sharing variants with GD. The red bubbles represented the results of the traditional colocalization analysis. The blue bubbles represented the results of colocalization analysis in the SuSiE framework. Grey bubbles represented the probability of passing HEIDI test. The larger size represents the higher probability that the protein and the disease share the risk genetic loci.

Cell-Type Specificity mRNA Expression of Causal Proteins

Since the onset of GD is closely related to abnormalities in the immune system, we performed single cell-type expression analyses covering 45 immune cells throughout the body. Among four prioritized proteins, CD40 had the highest expression, who was enriched in both mononuclear phagocytes and B-cell subsets, led by dendritic cells (Figure 4D and E, H) and memory B cells (Figure 4A–C), respectively. In addition, GMPR and CXCL10 were only enriched in one cell-type (alveolar macrophages and

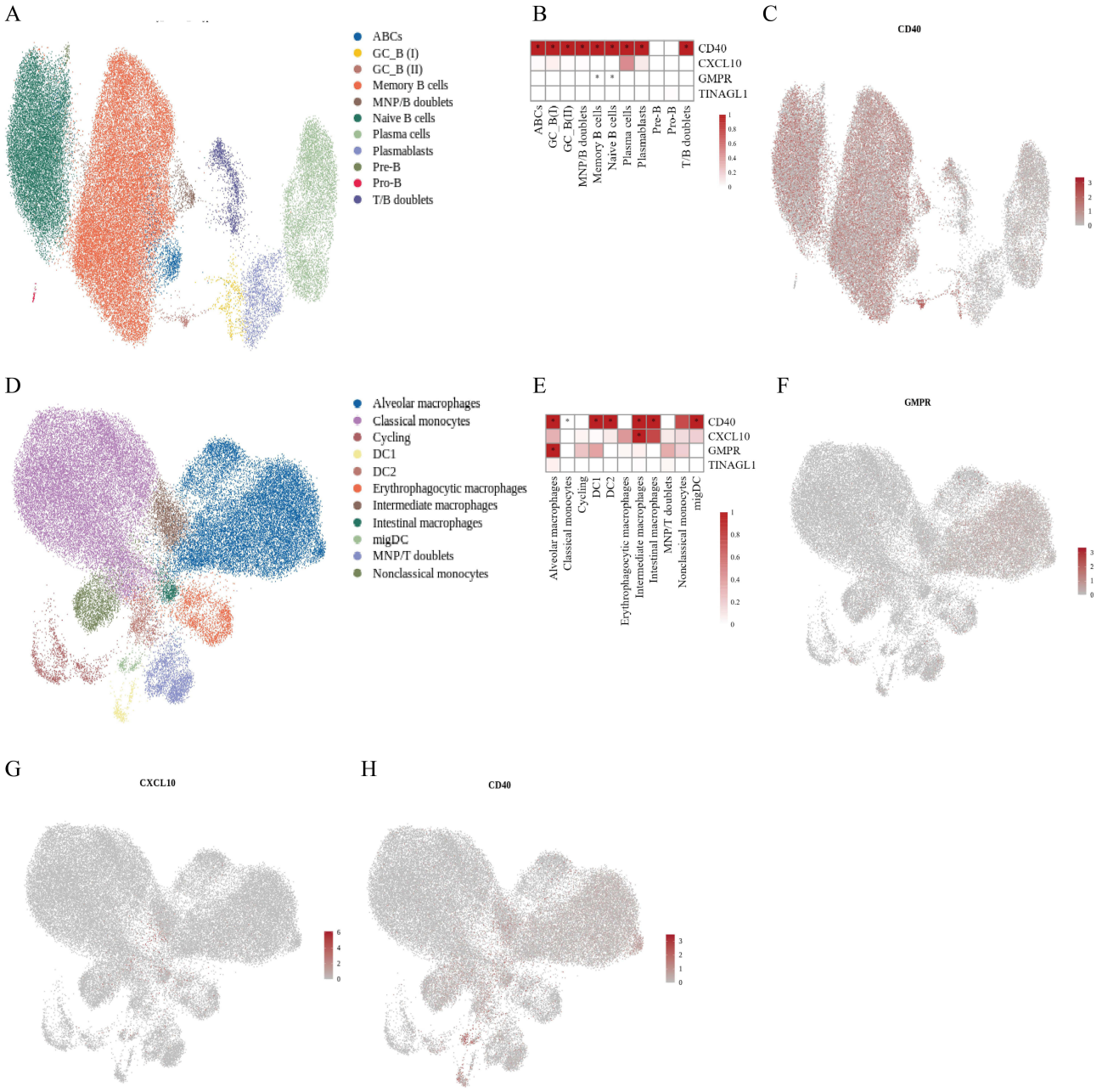

Figure 4 Cell-type-specific mRNA expression of 4 prioritized causal proteins. (**A**) UMAP plot of 11 B cells. (**B**) Heatmap of significant enrichments of protein-encoding genes in B cells. (**C**) Plots of expression patterns of CD40 in 11 B cells. (**D**) UMAP plot of 11 monocyte macrophages. (**E**) Heatmap of significant enrichments of protein-encoding genes in monocyte macrophages. (**F–H**) Plots of expression patterns of GMPR, CXCL10 and CD40 in 11 monocyte macrophages. *represent p-values less than the Bonferroni-corrected threshold of 0.0045 (0.05/11).

intermediate macrophages, respectively) (Figure 4D–G). However, no differential enrichment was found for TINAGL1 in immune cells. In addition, the immune cell enrichment of 61 causal proteins was displayed in Supplementary Table 6 and Supplementary Figure 1. Most were concentrated in mononuclear phagocytes led by macrophages, and about half genes were enriched in the B-cell subsets and T-cell subsets, respectively.

https://doi.org/10.2147/ITT.S494692

Function and Network Prediction of the Causal Proteins

The protein–protein interaction analysis found 56 interactions between the 61 potential causal proteins (Supplementary Figure 2). For example, the interaction among CD40, IRAK4 and CASP8 with each other were experimentally determined. In addition, there were potential interactions of CXCL10 with CD40, IRAK4 and CASP8.

Functional annotation analysis results indicated that in the biological process category, CD40 and CXCL10 are associated with regulation of apoptotic and death process. In the cellular component category, TINAGL1 mainly enriched extracellular matrix. In the molecular function category, CD40 was correlated with enzyme binding, ubiquitin protein ligase binding and so on. For the KEGG pathway analysis, it displayed some enriched pathways, such as "Toll-like receptor signaling pathway" and "NF-kappa B signaling pathway" (Supplementary Table 7).

Druggability Evaluation on the Potentials of Therapeutic Targets

Among 4 prioritized causal proteins, CD40 and CXCL10 were investigated as potential targets, which were involved in indications of autoimmune diseases (Supplementary Table 8). For other 57 causal proteins, 17 proteins had also been reported as targets for drug development, among which drugs for 6 targets have been approved. Information on the drug name and the process of drug development were displayed in Supplementary Table 8.

Identification of Prioritizing Probable Trans-pQTLs Associated With GD

SMR was conducted to identify trans-pQTLs, with SNPs mapped in the regions of 61 causal proteins, significantly associated with GD (Supplementary Table 9). For example, FCRL1, FCER2, CD22 and CD40LG that had trans-pQTLs in the gene region of CD40 were significantly associated with GD risk in trans-SMR analysis. Furthermore, FCRL1, FCER2, CD22, CD40LG, together with CD40, have close protein interactions with each other, and most of them are enriched in the immune cell activation pathway and are associated with cell adhesion functions (Supplementary Figure 3, Supplementary Table 10). These results indicates that variant mutations within the CD40 gene region not only affect CD40 expression levels but also a group of CD40-related proteins, which in turn have a significant impact on GD onset. In addition, four more groups of proteins with trans-pQTLs were identified that shared variant mutations with BTN3A2, LTB, HLA-DRA and TRIM40, respectively, and influenced GD onset (Supplementary Tables 9, 10 and Supplementary Figure 3).

Differential Associations of Plasma Proteins With GD and GO

Among the 61 causal proteins for GD, genetically predicted levels of 8 proteins were associated with GO, in which 6 proteins had consistent directions of the associations with both GD and GO, including a prioritized protein (GMPR) for GD (Supplementary Table 11). On the other hand, 55 novel proteins were identified for GO risk that were not associated with GD (Supplementary Table 12 and Supplementary Figure 4). The causal association of BTN1A1 and FCRL1 with GO was supported by colocalization analysis with PPH4 >0.5 (Supplementary Table 13). BTN1A1 was enriched in plasmablasts, FCRL1 was enriched in memory B cells (Supplementary Table 14, Supplementary Figure 5). Furthermore, 12 proteins were identified as targets with known drugs (Supplementary Table 15).

Discussion

In this study, we used proteomic-wide MR to identify 61 genetically predicted plasma proteins associated with GD risk. CD40, TINAGL1, GMPR and CXCL10 were highlighted for sharing causal variants with GD. Further downstream analysis revealed interactions, functional enrichment and potential drug targets of these GD risk proteins. Several of these proteins were associated with GO in a direction consistent with that for GD. In addition, several differential risk proteins associated with GO onset rather than GD onset were identified.

Our study confirmed that among the 61 risk proteins identified, some previously reported proteins, such as CD40, are associated with GD. CD40 belongs to the tumor necrosis factor (TNF) receptor superfamily and plays a pivotal role in autoimmune diseases such as psoriasis and multiple sclerosis, by inducing cell activation and cytokine production.[30,31] Our study found that elevated levels of CD40 in the blood were associated with the risk of GD. This finding was

supported by the association between the CD40 polymorphism and the risk of sporadic GD genetic susceptibility.[32] In addition, we provided potential evidence that CD40 might exert its function through B cells. It is proven that CD40 expression is increased in B cells among people with GD.[33] Research suggests that GD is a B cell-mediated autoimmune disease that occurs by a mechanism that requires the destruction of tolerance so that self-reactive B cells are activated and participate in the disease.[34] These findings indicate that the results of the current analysis have a high degree of reliability. There are research teams that have conducted clinical trials of drugs for CD40 that are not yet on the market, such as lucatumumab (CHIR 12.12 or HCD122), a human monoclonal antibody against CD40.[35] In multiple myeloma and chronic lymphocytic leukemia, two diseases that are closely linked to the immune system and involve immune dysfunction, the companies worked to determine optimal dosing in three Phase 1 trials, but the results are unclear.[35]

In addition, CD40 combined with FCRL1, FCER2, CD22 and CD40LG had synergistic effects on the GD risk. CD40LG encodes CD40L, which can combine with CD40 on antigen presenting cells as a receptor. FCRL1, a gene encoding one of the several Fc receptor-like glycoproteins, is downregulated in GD patients compared with normal individuals.[36] The relevance of other proteins has not yet been demonstrated.

CXCL10 is a chemokine that binds to its specific receptor CXCR3, regulating the immune response by recruiting and activating T cells, monocytes, and NK cells.[37] Under typical circumstances, some endocrine cells stimulated by IFN-g secrete CXCL10, which in turn recruits Th1 lymphocytes and secretes IFN-g, therefore, a significant loop of sustained inflammatory response occurs.[38] A previous study revealed that CXCL10 was associated with the risk of GD; however, compared with those in euthyroid thyroid autoantibody-negative patients, CXCL10 levels in patients with newly diagnosed GD were significantly greater.[39] In the presence of IFNγ+ TNFα, fenofibrate dose-dependently inhibited CXCL10 release, which in turn improves the condition of patients with GD.[40] Changes in CXCL10 partially explain the mechanism by which immunosuppression treats GD. Furthermore, drugs specifically targeting CXCL10 are already in the development stage, including Eldelumab.[41]

Although the two remaining prioritized proteins have no corresponding drug targets to develop, the association with GD is well founded. GMPR is a key enzyme in the nucleotide anabolic pathway that converts guanosine monophosphate (GMP) to inosine monophosphate (IMP). GMPR plays an important role in immune-related diseases, such as providing nucleotides to support the proliferation of immune cells, synthesize new DNA and RNA during the regulation of inflammatory responses and immune responses.[42] Previous studies support GMPR as a potential biomarker of response to immune cell function[43]. This is consistent with our results regarding the risk of excessive GMPR levels for GD pathogenesis. In addition, a previous study demonstrated the feasibility of using GMPR as a drug target.[44] TINAGL1 is an important pro-angiogenic matricellular protein. TINAGL1 plays an important role as an angiogenic protein in the inflammatory response, which emphasizes that vascular response is the heart of the inflammatory response.[45] The onset of GD involves an inflammatory response, which provides evidence for the reliability of our results.

Furthermore, FCRL1 and BTN1A1 were found to be related to GO but not GD. FCRL1, a member of the Ig superfamily, is expressed predominantly by B cells.[46] Studies have suggested that B-cell activation and function may be regulated by FCRL1.[47] GO is a B-cell-mediated disease and B cells participate in antigen presentation and autoantibody production. This finding is consistent with our findings that FCRL1 is associated with the onset of GO and is enriched mainly in the B-cell subset. Despite the lack of information on drugs targeting FCRL1, study has revealed possibilities for immunotherapeutic means targeting FCRL1 as well as signaling blocking drug development.[47] Another protein, Butyrophilin 1a1 (BTN1A1), is a member of the Ig superfamily.[48] It has been shown that the autoimmune response in some multiple sclerosis patients may be regulated by molecular mimicry between the IgI structural domain of BTN1A1 and similar Ig-folds in myelin oligodendrocyte glycoproteins.[48] This finding identified BTN1A1 as being associated with certain autoimmune diseases, providing possible evidence that BTN1A1 levels may contribute to the risk of GO development.

The strengths of this study are as follows. First, the two-stage proteome-wide MR design allows for a more systematic and rigorous study of the relationship between plasma protein biomarkers and GD risk, which could provide a better aetiologically important role of associated proteins in disease. Second, the single-cell type expression analysis, PPI and druggability assessment made the analysis results more informative and the potential pathogenic effects of the candidate proteins on GD clearer, laying the foundation for the subsequent identification of druggable targets. However, several

limitations should also be considered. First, the small sample size of the study may have prevented the detection of additional metabolic proteins associated with GD and GO. Second, the analysis was limited to European populations, without validation of the findings in other ancestries. Furthermore, differences in sequencing methods between the two platforms led to incomplete overlap in detected proteins, representing a limitation rather than a false-negative result. In addition, due to financial and technical constraints, our study did not fully elucidate the biological mechanisms behind the identified genetic and proteomic associations. Future studies should aim to address these limitations by increasing sample sizes, incorporating multi-ethnic populations, and conducting functional experiments to explore the roles of key genes and proteins in greater depth.

Conclusions

In conclusion, our proteome-wide MR and colocalization analyses identified four prioritized causal proteins (CD40, TINAGL1, GMPR and CXCL10) for GD onset and two (BTN1A1 and FCRL1) for GO onset. By integrating gene function and drug development information, promising therapeutic targets for GD and GO were identified, which has provided valuable insights for precision medicine and new drug development. Future work could focus on validating and functionally investigating these targets in large-scale clinical studies and accelerating their translation into effective clinical treatment.

Abbreviations

GD, Graves' disease; GO, Graves' ophthalmopathy; GWAS, genome sequencing; HEIDI, heterogeneity in dependent instruments; IV, instrumental variable; IVW, inverse variance weighted; KEGG, Kyoto Encyclopedia of Genes and Genomes; MR, Mendelian randomization; SMR, Summary-data-based MR; PPH4, posterior probability for H4; PPI, Protein–protein interaction; SNPs, Single nucleotide polymorphisms; SuSiE, sum of single effects; TSH, thyroid-stimulating hormone; UKB-PPP, UK Biobank Pharma Proteomics Project.

Data Sharing Statement

Original data generated and analyzed during this study are included in this published article.

Ethics Approval and Informed Consent

The GWAS summary data used for our analysis are fully publicly available. The authors of the published GWAS have provided the corresponding ethical certificates for the cohorts they used. Therefore, we were not required to submit ethical certificates. For example, the data we used for OLINK is based on the UK Biobank. The North West Multi-Center Research Ethics Committee Study approved the UKB study (No. REC reference 11/NW/0382), and all participants provided written informed consent. All relevant ethical regulations were followed the original study of the datasets, and the authors of the source studies had also obtained informed consent from participants. Ethic approval was exempted for this study by Shanghai Jiaotong University because the data were obtained from public resource.

Acknowledgments

We thank the authors of studies, that conducted GWAS of Olink proteins, Graves' disease and Graves' ophthalmopathy, for providing summary data used in this study.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This study was supported by the National Natural Science Foundation of China (81800694, 82170870, and 82370862). The funders had no role in the study design, data collection and analysis, publication decision, or manuscript preparation.

Disclosure

The authors report no conflicts of interest in this work.

References

1. McLachlan SM, Rapoport B. Breaking tolerance to thyroid antigens: changing concepts in thyroid autoimmunity. *Endocr Rev*. 2014;35(1):59–105. doi:10.1210/er.2013-1055
2. Romagnani S. The Th1/Th2 paradigm and allergic disorders. *Allergy*. 1998;53(46 Suppl):12–15. doi:10.1111/j.1398-9995.1998.tb04951.x
3. Davies TF, Andersen S, Latif R, et al. Graves' disease. *Nat Rev Dis Primers*. 2020;6(1):52. doi:10.1038/s41572-020-0184-y
4. Smith TJ, Longo DL, Hegedüs L. Graves' Disease. *N Engl J Med*. 2016;375(16):1552–1565. doi:10.1056/NEJMra1510030
5. Bartalena L, Burch HB, Burman KD, Kahaly GJ. A 2013 European survey of clinical practice patterns in the management of Graves' disease. *Clin. Endocrinol*. 2015;84(1):115–120. doi:10.1111/cen.12688
6. Burch HB, Burman KD, Cooper DS. Survey of Clinical Practice Patterns in the Management of Graves' Disease. *J Clin Endocrinol Metab*. 2012;97 (12):4549–4558. doi:10.1210/jc.2012-2802
7. ResearchGate. A genome-wide association study identifies two new risk loci for Graves' disease. *Nature Genet*. 2011;43(9):897–901. doi:10.1038/ng.898
8. Zhao S-X, Liu W, Liang J, et al. Assessment of Molecular Subtypes in Thyrotoxic Periodic Paralysis and Graves Disease Among Chinese Han Adults. *JAMA Network Open*. 2019;2(5):3348. doi:10.1001/jamanetworkopen.2019.3348
9. Liu W, Wang H-N, Gu Z-H, et al. Identification of BACH2 as a susceptibility gene for Graves' disease in the Chinese Han population based on a three-stage genome-wide association study. *Hum Genet*. 2013;133(5):661–671. doi:10.1007/s00439-013-1404-2
10. Zheng J, Haberland V, Baird D, et al. Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. *Nature Genet*. 2020;52(10):1122–1131. doi:10.1038/s41588-020-0682-6
11. Bartalena L, Fatourechi V. Extrathyroidal manifestations of Graves' disease: a 2014 update. *J Endocrinol Invest*. 2014;37(8):691–700. doi:10.1007/s40618-014-0097-2
12. Ponto KA, Merkesdal S, Hommel G, Pitz S, Pfeiffer N, Kahaly GJ. Public health relevance of Graves' orbitopathy. *J Clin Endocrinol Metab*. 2013;98(1):145–152. doi:10.1210/jc.2012-3119
13. Jiang L, Rong A, Wei R, Diao J, Ding H, Wang W. Tear proteomics of orbital decompression for disfiguring exophthalmos in inactive thyroid-associated ophthalmopathy. *Exp Ther Med*. 2020;20(6):1. doi:10.3892/etm.2020.9383
14. Yang M, Chung Y, Lang S, Yawata N, Seah LL, Looi A. The tear cytokine profile in patients with active Graves' orbitopathy. *Endocrine*. 2017;59 (2):402–409. doi:10.1007/s12020-017-1467-2
15. Chen Q. The expression of interleukin-15 and interleukin-17 in tears and orbital tissues of Graves ophthalmopathy patients. *J Cell Biochem*. 2018;120(4):6299–6303. doi:10.1002/jcb.27916
16. Sun BB, Chiou J, Traylor M, et al. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature*. 2023;622(7982):329–338. doi:10.1038/s41586-023-06592-6
17. Lundberg M, Eriksson A, Tran B, Assarsson E, Fredriksson S. Homogeneous antibody-based proximity extension assays provide sensitive and specific detection of low-abundant proteins in human blood. *Nucleic Acids Res*. 2011;39(15):e102. doi:10.1093/nar/gkr424
18. Kurki MI, Karjalainen J, Palta P, et al. FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature*. 2023;613 (7944):508–518. doi:10.1038/s41586-022-05473-8
19. Sakaue S, Kanai M, Tanigawa Y, et al. A cross-population atlas of genetic associations for 220 human phenotypes. *Nat Genet*. 2021;53 (10):1415–1424. doi:10.1038/s41588-021-00931-x
20. Yuan S, Xu F, Li X, et al. Plasma proteins and onset of type 2 diabetes and diabetic complications: proteome-wide Mendelian randomization and colocalization analyses. *Cell Rep Med*. 2023;4(9):101174. doi:10.1016/j.xcrm.2023.101174
21. Zhu Z, Zhang F, Hu H, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nature Genet*. 2016;48 (5):481–487. doi:10.1038/ng.3538
22. Williams SM, Giambartolomei C, Vukcevic D, et al. Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. *PLoS Genetics*. 2014;10(5). doi:10.1371/journal.pgen.1004383
23. Wallace C. A more accurate method for colocalisation analysis allowing for multiple causal variants. *PLOS Genetics*. 2021;17(9):e1009440. doi:10.1371/journal.pgen.1009440
24. Stacey D, Fauman EB, Ziemek D, et al. ProGeM: a framework for the prioritization of candidate causal genes at molecular quantitative trait loci. *Nucleic Acids Res*. 2019;47(1):e3. doi:10.1093/nar/gky837
25. Domínguez Conde C, Xu C, Jarvis LB, et al. Cross-tissue immune cell analysis reveals tissue-specific features in humans. *Science*. 2022;376(6594). doi:10.1126/science.abl5197
26. Jagadeesh KA, Dey KK, Montoro DT, et al. Identifying disease-critical cell types and cellular processes by integrating single-cell RNA-sequencing and human genetics. *Nature Genet*. 2022;54(10):1479–1492. doi:10.1038/s41588-022-01187-9
27. Bao C, Wang S, Jiang L, et al. OpenXGR: a web-server update for genomic summary data interpretation. *Nucleic Acids Res*. 2023;51(W1):W387–w396. doi:10.1093/nar/gkad357
28. Szklarczyk D, Kirsch R, Koutrouli M, 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(D1):D638–d646. doi:10.1093/nar/gkac1000
29. Ghoussaini M, Mountjoy E, Carmona M, et al. Open Targets Genetics: systematic identification of trait-associated genes using large-scale genetics and functional genomics. *Nucleic Acids Res*. 2021;49(D1):D1311–d1320. doi:10.1093/nar/gkaa840
30. Liu S, Xu J, Wu J. The Role of Co-Signaling Molecules in Psoriasis and Their Implications for Targeted Treatment. *Front Pharmacol*. 2021;12:717042. doi:10.3389/fphar.2021.717042
31. Vermersch P, Granziera C, Mao-Draayer Y, et al. Inhibition of CD40L with Frexalimab in Multiple Sclerosis. *N Engl J Med*. 2024;390(7):589–600. doi:10.1056/NEJMoa2309439

32. Kurylowicz A, Kula D, Ploski R, et al. Association of CD40 Gene Polymorphism (C-1T) with Susceptibility and Phenotype of Graves' Disease. *Thyroid*. 2005;15(10):1119–1124. doi:10.1089/thy.2005.15.1119
33. Hwang CJ, Afifiyan N, Sand D, et al. Orbital Fibroblasts from Patients with Thyroid-Associated Ophthalmopathy Overexpress CD40: CD154 hyperinduces IL-6, IL-8, and MCP-1. *Investigat Opthalmol Vis Sci*. 2009;50(5):2262–2268. doi:10.1167/iovs.08-2328
34. Smith MJ, Rihanek M, Coleman BM, Gottlieb PA, Sarapura VD, Cambier JC. Activation of thyroid antigen-reactive B cells in recent onset autoimmune thyroid disease patients. *J Autoimmun*. 2018;89:82–89. doi:10.1016/j.jaut.2017.12.001
35. Hassan SB, Sørensen JF, Olsen BN, Pedersen AE. Anti-CD40-mediated cancer immunotherapy: an update of recent and ongoing clinical trials. *Immunopharmacol Immunotoxicol*. 2014;36(2):96–104. doi:10.3109/08923973.2014.890626
36. Rostamzadeh D, Dabbaghmanesh M, Shabani M, Hosseini A, Amirghofran Z. Expression Profile of Human Fc Receptor-Like 1, 2, and 4 Molecules in Peripheral Blood Mononuclear Cells of Patients with Hashimoto's Thyroiditis and Graves' Disease. *Hormone Metab Res*. 2015;47(09):693–698. doi:10.1055/s-0035-1545280
37. Ferrari SM, Fallahi P, Santaguida G, et al. Circulating CXCL10 is increased in non-segmental vitiligo, in presence or absence of autoimmune thyroiditis. *Autoimmunity Rev.* 2017;16(9):946–950. doi:10.1016/j.autrev.2017.07.006
38. Romagnani P, Serio M, Romagnani S, Chiovato L, Rotondi M. Role of Chemokines in Endocrine Autoimmune Diseases. *Endocrine Reviews*. 2007;28(5):492–520. doi:10.1210/er.2006-0044
39. Cakir M, Levendoglu F, Kiyici A, Coskun Y. Serum CXCL10 levels and neuromuscular manifestations in patients with autoimmune thyroid diseases. *Autoimmunity.* 2011;44(6):496–503. doi:10.3109/08916934.2011.561818
40. Ferrari SM, Ragusa F, Paparo SR, et al. Differential modulation of CXCL8 versus CXCL10, by cytokines, PPAR-gamma, or PPAR-alpha agonists, in primary cells from Graves' disease and ophthalmopathy. *Autoimmunity Rev.* 2019;18(7):673–678. doi:10.1016/j.autrev.2019.05.004
41. Sandborn WJ, Colombel J-F, Ghosh S, et al. Eldelumab [Anti-IP-10] Induction Therapy for Ulcerative Colitis: a Randomised, Placebo-Controlled, Phase 2b Study. *J Crohns Colitis*. 2016;10(4):418–428. doi:10.1093/ecco-jcc/jjv224
42. Ma EH, Bantug G, Griss T, et al. Serine Is an Essential Metabolite for Effector T Cell Expansion. *Cell Metab*. 2017;25(2):345–357. doi:10.1016/j.cmet.2016.12.011
43. Lee SY, Choi B-S, Yoon C-H, Kang C, Kim K, Kim K-C. Selection of biomarkers for HIV-1 latency by integrated analysis. *Genomics*. 2019;111(3):327–333. doi:10.1016/j.ygeno.2018.02.007
44. Gaiya DD, Muhammad A, Aimola IA, et al. Potential of Onchocerca ochengi inosine-5′-monophosphate dehydrogenase (IMPDH) and guanosine-5′-monophosphate oxidoreductase (GMPR) as druggable and vaccine candidates: immunoinformatics screening. *J Biomol Struct Dyn*. 2023;41(24):14832–14848. doi:10.1080/07391102.2023.2184171
45. Cotran RS, Pober JS. Cytokine-endothelial interactions in inflammation, immunity, and vascular injury. *J Am Soc Nephrol*. 1990;1(3):225–235. doi:10.1681/asn.V13225
46. Davis RS. Fc Receptor-Like Molecules. *Ann Rev Immunol*. 2007;25(1):525–560. doi:10.1146/annurev.immunol.25.022106.141541
47. Zhao X, Xie H, Zhao M, et al. Fc receptor–like 1 intrinsically recruits c-Abl to enhance B cell activation and function. *Sci Adv.* 2019;5(7). doi:10.1126/sciadv.aaw0315
48. Ogg SL, Weldon AK, Dobbie L, Smith AJH, Mather IH. Expression of butyrophilin (Btn1a1) in lactating mammary gland is essential for the regulated secretion of milk–lipid droplets. *Proc Natl Acad Sci*. 2004;101(27):10084–10089. doi:10.1073/pnas.0402930101