Journal of Translational Medicine

RESEARCH **Open Access**

Identifying prioritization of therapeutic targets for ankylosing spondylitis: a multi-omics Mendelian randomization study

Lingyu Dai[1†], Lan Xia[1†], Guannan Su[1], Yu Gao[1], Qingyan Jiang[1] and Peizeng Yang[1*]

Abstract

Background To investigate the associations of methylation, expression, and protein quantitative trait loci (mQTL, eQTL, and pQTL) with ankylosing spondylitis (AS) and find out genetically supported drug targets for AS.

Methods The summary-data-based Mendelian randomization (SMR) and Bayesian co-localization analysis were used to assess the potential causality between AS and relevant genes. The GWAS data obtained from the International Genetics of Ankylosing Spondylitis Consortium (IGAS) were set as the discovery stage, and the FinnGen and UK Biobank databases were used to replicate the analysis as an external validation. We further integrated the multi-omics results to screen overlapped genes at different levels. The protein–protein interaction (PPI) network and enrichment analyses were used to explore the biological effect of SMR-identified genes on AS. Drug prediction and molecular docking were used to validate the medicinal value of candidate drug targets.

Results Based on the results of multi-omics evidence screening, we identified potential associations of TNFRSF1A, B3GNT2, ERAP1, and FCGR2A with AS at different regulatory levels. At the protein level, AIF1, TNXB, APOM, and B3GNT2 were found to be negatively associated with AS risk, whereas higher levels of FCGR2A, FCGR2B, IL12B, TNFRSF1A, and ERAP1 were associated with an increased risk of AS. The bioinformatics analyses showed that the SMR-identified genes were mainly involved in immune response. Molecular docking results displayed stable binding between predicted candidate drugs and these aforementioned proteins.

Conclusion Our study found four AS-associated genes with multi-omics evidence and nine promising drug targets for AS, which may contribute to the understanding of the genetic mechanisms of AS and provide innovative perspectives into targeted therapy for AS.

Keywords Ankylosing spondylitis, Mendelian randomization, Multi-omics, Methylation, Gene expression, Protein, Drug target

Background

Ankylosing Spondylitis (AS), a common type of spondyloarthropathy, is a highly genetic chronic autoimmune disease. [1, 2] It mainly affects spine joints, sacroiliac joints, and adjacent soft tissues, such as tendons and ligaments. The global prevalence of ankylosing spondylitis is about 0.1% to 1.4%, with a higher prevalence in men [3]. Previous GWAS studies demonstrated ANTXR2, CARD9, ERAP1, IL12B, and IL23R, etc. were associated with AS [4]. However, these correlated variants may

[†]Lingyu Dai and Lan Xia These are co-first authors.

*Correspondence:
Peizeng Yang
peizengycmu@126.com
[1] The First Affiliated Hospital of Chongqing Medical University, Chongqing Branch (Municipality Division) of National Clinical Research Center for Ocular Diseases, Youyi Road 1, Chongqing 400016, People's Republic of China

© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/.

not be causally relevant due to the intricate linkage disequilibrium structures in genomes. Currently, despite the beneficial effectiveness of treating AS with anti-rheumatic drugs, including glucocorticoids, biological agents, and targeted synthetic drugs, there remains a significant gap between patient expectations and actual outcomes [5]. Therefore, exploring the genetic underpinnings of AS and identifying novel drug targets is paramount for improving patient prognosis.

Novel bioinformatics methodology supplies formidable facilities for elucidating molecular processes and pinpointing potential therapeutic targets for disease [6]. Integration of multi-omics is an emerging post-GWAS approach to identify pivotal regulators at diverse levels of gene regulation for providing comprehensive biological perspectives [7–9]. Although multi-omics studies have been efficient in linking genes with downstream expression associated with disease risk, they could not consistently identify clear causal relationships. The summary data-based Mendelian randomization (SMR) analysis is an effective method for inferring causality utilizing genetic variants as instrumental variables, which could support high-throughput screening and reduce the impact of confounding biases and reverse causation. Compared to the conventional MR analysis, SMR could achieve a much higher statistical power based on the top cis-quantitative trait loci (cis-QTL) [10]. In this study, we combined the SMR and multi-omics approaches to estimate the potential effects of genes on AS at perspectives of methylation, expression, and protein abundance and identify promising therapeutic targets for AS.

Methods

Study design and data sources

The overall details of our study design are displayed in Fig. 1. The exposure data of quantitative trait loci (QTL) included methylation quantitative trait loci (mQTL), expression quantitative trait loci (eQTL), and protein quantitative trait loci (pQTL). The mQTLs data included 52,916 cis and 2025 trans mQTLs measured in 1980 European ancestry individuals [11]. The eQTLs data were retrieved from the eQTLGen Consortium including 31,684 individuals with 10,317 trait-associated SNPs. The blood pQTLs data contained 4719 proteins from 35,559 Icelanders [12].

The GWAS data sources for AS were obtained from the International Genetics of Ankylosing Spondylitis Consortium (IGAS), FinnGen, and UK biobank. The IGAS study as the discovery stage included 9069 cases and 13,578 controls, and all of them were European. The FinnGen cohort and UK biobank cohorts were set as replication stages. The FinnGen cohort consists of 3162 cases and 294,770 controls, whereas the UKB cohort included 1296 cases and 461,637 controls. (Supplementary Table S1) These studies were approved by the corresponding ethical review boards and informed written consent was obtained from all subjects.

Summary-data-based MR analysis

Summary-data-based Mendelian randomization (SMR) is a novel statistical method based on the principles of MR. It could use top single nucleotide polymorphism (SNP) to identify potential causal effects of genes on diverse traits. We conducted it to assess the causal association between AS and genes at methylation, expression, and protein levels. There are three core assumptions of Mendelian randomization, including assumption 1: the genetic instruments are robustly associated with the exposure; assumption 2: the genetic instruments are not associated with any major confounders; assumption 3: the genetic instruments affect the outcome only through exposure. To meet the MR hypothesis in the study, we controlled the quality of the instrumental variables based on the following criteria: (1) SNPs reached the threshold of genome-wide significance ($P < 5 \times 10^{-8}$) were included; (2) SNPs with linkage disequilibrium $r2 > 0.9$ and < 0.05 around the top SNPs were excluded; (3) The plink clumping with default conditions was conducted based on the 1000 Genomes European panel [13]. Furthermore, the P-values were corrected using the Benjamini–Hochberg method to restrict the false discovery rate (FDR) at $\alpha = 0.05$, which could reduce the false positive results.

Heterogeneity in the dependent instrument (HEIDI) test, a sensitivity analysis for SMR, was performed by using multiple SNPs around the top SNP in a region to evaluate whether the casual associations between genes and AS risk was caused by linkage disequilibrium (LD) with the causal variant. The P-value of the HEIDI test < 0.05 considered to be caused by LD was discarded from the analysis. The SMR software tool (SMR v1.3.1, https://yanglab.westlake.edu.cn/software/smr/) was utilized to conduct SMR and HEIDI tests. The colocalization analysis was further performed in association with P_{SMR}-FDR value < 0.05 and P-HEIDI > 0.05.

Colocalization analysis

The colocalization analysis could assess whether two traits share the same casual variant in a certain region. The posterior probability is used to quantify support for all hypotheses. The basic hypothesis for colocalization in the same genomic location includes (1) PPH0: gene variants have no causality to either trait 1 or trait 2; (2) PPH1: gene variants only have causality to trait 1; (3) PPH2: gene variants only have causality to trait 2; (4) PPH3: gene variants have different causality to the two traits; and (5) PPH4: gene variants have the same

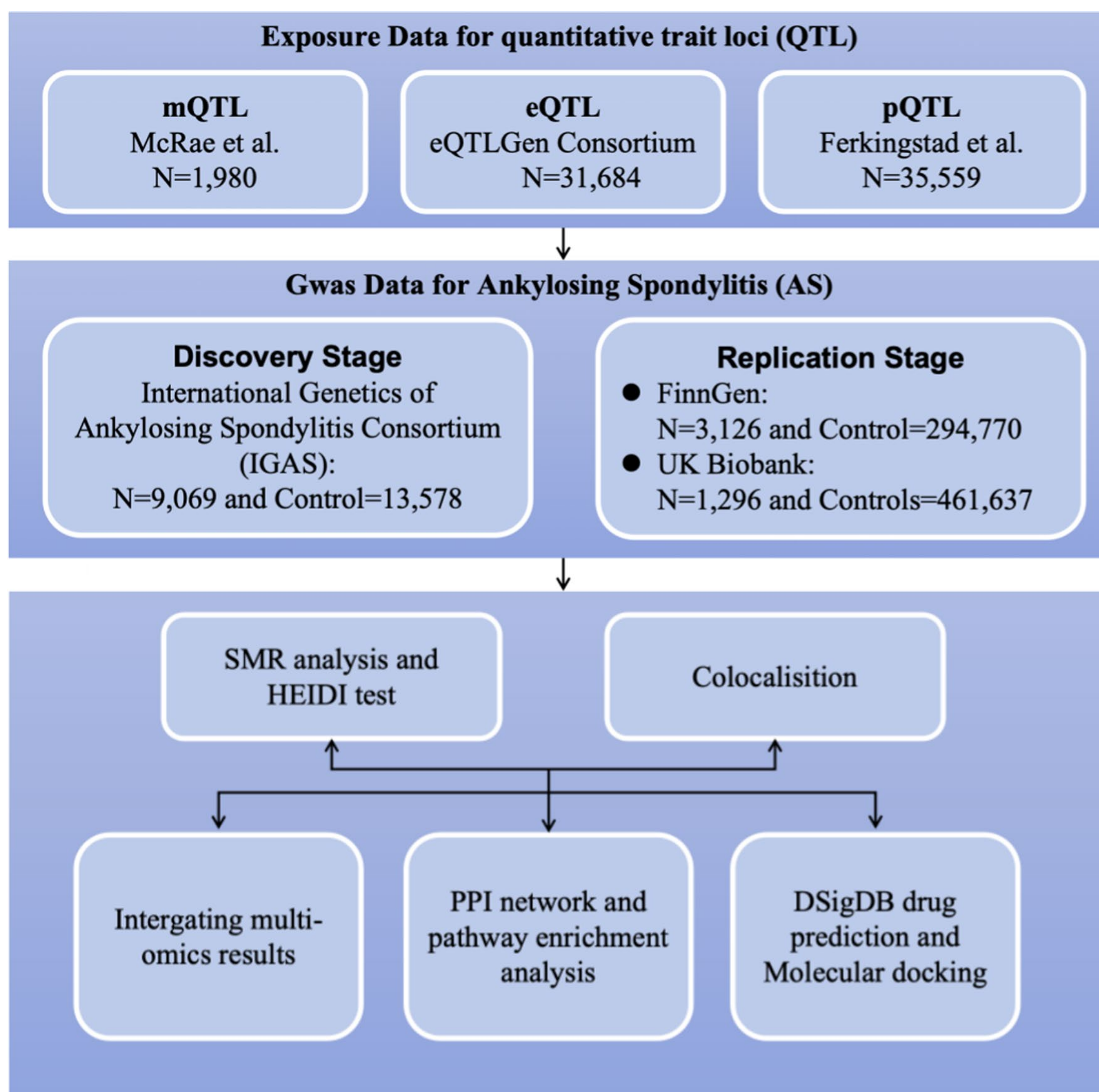

Fig. 1 The overall design of this study. *mQTL*, methylation quantitative trait loci; *eQTL*, expression quantitative trait loci; *pQTL*, protein quantitative trait loci; *SMR*, summary-data-based Mendelian randomization; *HEIDI*, heterogeneity in the dependent instrument; *FDR*, false discovery rate; *PPI*, protein–protein interaction

shared causality to the two traits. The colocalization region windows of pQTL, eQTL, and mQTL were set to ± 1000 kb, ± 1000 kb, and ± 500 kb, respectively [14]. Due to limited power in the colocalization analysis, the combined posterior probability of genes was restricted to PPH3 + PPH4 ≥ 80%. [15, 16].

The integration of mQTL, eQTL and pQTL

We integrated results from pQTLs, eQTLs, and mQTLs analyses to achieve a multidimensional landscape of connections between gene regulation and AS. AS proteins were the terminal expression products of genes, the correlation of potential candidate genes with AS needed to be identified at the protein level, and then this association was examined at other regulatory levels. Therefore, the causal genes associated with AS at the protein level needed to meet P_{SMR}-FDR < 0.05, P-HEIDI > 0.05, and PPH3 + PPH4 of colocalization ≥ 80%, and then were further categorized into three tiers based on the following standards. (1) Tier 1: genes were associated with AS at both methylation and expression levels (P_{SMR}-FDR < 0.05); (2) Tier 2: genes were associated with AS at methylation or expression levels (P_{SMR}-FDR < 0.05); (3) Tier3: genes were

associated with AS at methylation and expression levels ($P_{SMR^{-}Original}$ value < 0.05).

Protein–protein interaction (PPI) and functional enrichment analysis

PPI networks were used to estimate the associations between AS-associated proteins (P < 0.05) by using the Search Tool for the Retrieval of Interacting Genes (STRING, https://string-db.org/). The Gene Ontology (GO) and Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathway analyses were performed to assess the relationship of potentially enriched pathways with proteins.

The SMR-prioritized proteins with drug targets and Molecular docking

DSigDB is a large database of 22,527 genomes and 17,389 different compounds covering 19,531 genes connecting drugs and compounds to their target genes. We used this database to assess the pharmacological activity of SMR-prioritized proteins and determine whether these proteins had potential as practical drug targets (DSigDB, http://dsigdb.tanlab.org/DSigDBv1.0/) [17]. We further employed molecular docking to evaluate the predicted interactions between drug candidates and their targets. It could help prioritize drug targets by recognizing the characteristics of proteins and the interaction modes between proteins and drug molecules, laying the foundation for future experimental validation. The protein and drug structural data were obtained from the Protein Data Bank (https://www.rcsb.org/) and the PubChem Compound Database (https://pubchem.ncbi.nlm.nih.gov/), respectively. We removed the water molecules of proteins and ligands and added the polar hydrogen atoms. Molecular docking of the top five significant predicted drugs and the corresponding target genes encoded proteins was performed using the computerized protein–ligand docking software AutoDock 4.2.6 (http://autodock.scripps.edu/), and the outcomes were visualized using PyMol (https://www.pymol.org/).

Result

Assessing the effect of gene methylation on AS

We identified 932 CpG sites in 455 unique genes associated with the risk of AS (P < 0.05 and P-HEIDI > 0.05). With the FDR correction of P-value and colocalization screening, a total of 194 CpG sites in 87 unique genes were found to be significantly correlated with AS. (Fig. 2 and Supplementary Table S2) It is worth noting that distinct CpG loci of the same gene do not always contribute consistently to the risk of AS. For example, the genetically predicted HLA-DQB1 methylation at cg03796735, cg06255955, and cg21588215 was

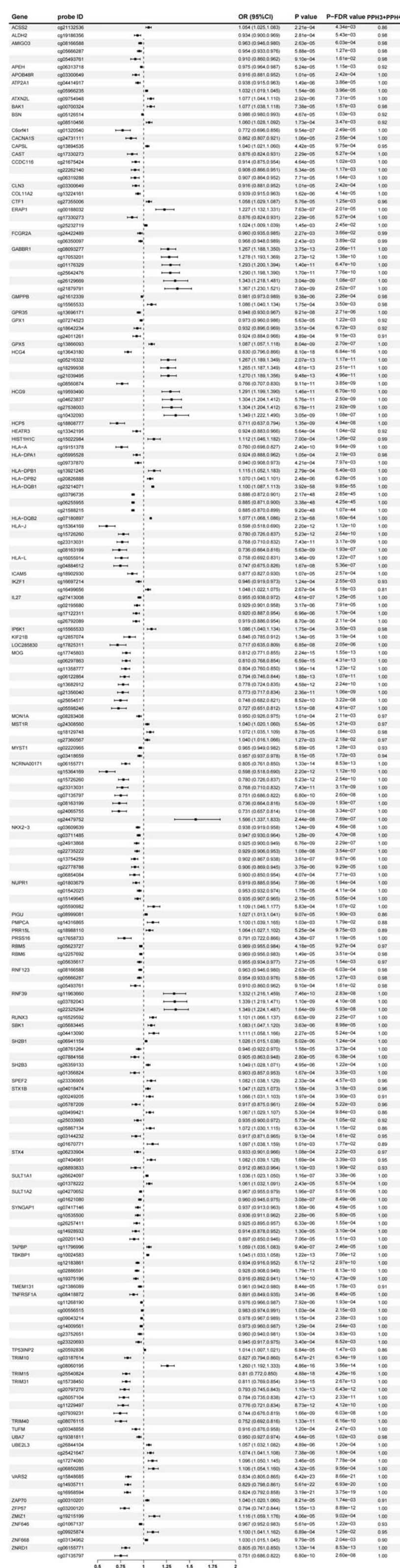

Fig. 2 The associations between gene methylation and ankylosing spondylitis. *OR* odds ratio, *CI* confidence interval, *FDR* false dicovery rate, *PPH3* + *PPH4* posterior probability of H3 + H4

negatively associated with AS risk (OR: 0.886, 95% CI 0.872, 0.901; OR: 0.885, 95% CI 0.871, 0.900; and OR: 0.885, 95% CI 0.870,0.899, respectively). However, the methylation of cg08060195 in HLA-DQB1 was associated with a higher risk of AS (OR: 1.100, 95% CI 1.087, 1.113). Furthermore, 34 CpG sites near 17 genes were simultaneously replicated in the FinnGen and UK biobank validation cohorts, such as TBKBP1 (cg10024583, cg12183861, cg02886591, cg19375196), IL27 (cg27413008, cg02195680, cg17122311, cg26792089), and TUFM (cg00348858). (Supplementary Table S3 and Supplementary Table S4).

Assessing the effect of gene expression on AS

The causal effects of gene expression on AS are shown in Supplementary Table S5. A total of 329 genes were significantly associated with AS ($P < 0.05$ and P-HEIDI > 0.05). There were 37 genes potentially associated with AS after FDR correction of P value and colocalization screening. (Table 1) The genetically

Table 1 The associations of eQTLs with ankylosing spondylitis

Gene	Symbol	OR (95%CI)	P value	P-FDR value	PPH3	PPH4	PPH3 + PPH4
ENSG00000170340	B3GNT2	0.914 (0.894,0.933)	5.28E-17	1.10E-14	100.0%	0.0%	100.0%
ENSG00000263293	EFCAB13-DT	1.178 (1.116,1.243)	2.33E-09	2.80E-07	0.0%	100.0%	100.0%
ENSG00000221988	PPT2	0.556 (0.456,0.678)	6.11E-09	6.96E-07	97.2%	2.7%	99.9%
ENSG00000259207	ITGB3	1.393 (1.240,1.564)	2.10E-08	2.20E-06	0.1%	99.9%	100.0%
ENSG00000240065	PSMB9	0.473 (0.363,0.617)	3.45E-08	3.56E-06	70.6%	29.3%	99.9%
ENSG00000176476	SGF29	0.953 (0.937,0.970)	4.79E-08	4.83E-06	2.5%	97.5%	100.0%
ENSG00000178952	TUFM	1.030 (1.019,1.041)	8.54E-08	8.14E-06	4.6%	95.3%	99.9%
ENSG00000197165	SULT1A2	1.055 (1.034,1.076)	1.12E-07	1.05E-05	4.6%	95.3%	99.9%
ENSG00000185651	UBE2L3	1.036 (1.023,1.050)	1.34E-07	1.24E-05	3.0%	97.0%	100.0%
ENSG00000188322	SBK1	0.864 (0.818,0.914)	2.90E-07	2.62E-05	3.8%	96.1%	99.9%
ENSG00000111252	SH2B3	0.896 (0.858,0.936)	7.57E-07	6.64E-05	2.0%	97.9%	99.9%
ENSG00000205609	EIF3CL	1.045 (1.026,1.063)	1.26E-06	0.000	37.8%	62.0%	99.8%
ENSG00000263766	KPNB1-DT	0.766 (0.682,0.859)	5.49E-06	0.000	3.9%	96.1%	100.0%
ENSG00000067182	TNFRSF1A	1.077 (1.043,1.112)	7.33E-06	0.001	96.9%	3.1%	100.0%
ENSG00000182179	UBA7	1.033 (1.018,1.047)	7.74E-06	0.001	3.0%	95.1%	98.1%
ENSG00000246465	LOC124903670	0.867 (0.814,0.923)	8.88E-06	0.001	3.4%	96.6%	100.0%
ENSG00000107796	ACTA2	1.021 (1.012,1.031)	9.84E-06	0.001	0.0%	97.3%	97.3%
ENSG00000169180	XPO6	0.684 (0.578,0.810)	1.04E-05	0.001	2.7%	97.3%	100.0%
ENSG00000196296	ATP2A1	1.273 (1.140,1.423)	1.91E-05	0.001	46.6%	53.2%	99.8%
ENSG00000101000	PROCR	0.855 (0.796,0.920)	2.31E-05	0.001	0.0%	97.3%	97.3%
ENSG00000185614	INKA1	1.143 (1.074,1.217)	2.48E-05	0.002	4.6%	92.6%	97.2%
ENSG00000004534	RBM6	1.027 (1.014,1.040)	3.92E-05	0.002	12.5%	79.7%	92.2%
ENSG00000099381	SETD1A	1.169 (1.085,1.259)	3.90E-05	0.002	1.6%	94.7%	96.3%
ENSG00000115073	ACTR1B	1.105 (1.053,1.160)	5.30E-05	0.003	0.0%	90.9%	90.9%
ENSG00000103496	STX4	1.050 (1.025,1.075)	6.76E-05	0.004	5.4%	82.2%	87.6%
ENSG00000161180	CCDC116	0.664 (0.543,0.813)	7.28E-05	0.004	3.1%	96.9%	100.0%
ENSG00000078747	ITCH	1.125 (1.060,1.194)	9.86E-05	0.005	0.5%	85.2%	85.7%
ENSG00000233276	GPX1	0.842 (0.772,0.919)	1.10E-04	0.006	13.1%	79.0%	92.1%
ENSG00000138134	STAMBPL1	0.740 (0.633,0.863)	1.35E-04	0.006	0.0%	97.3%	97.3%
ENSG00000178188	SH2B1	0.675 (0.551,0.828)	1.59E-04	0.007	6.0%	94.0%	100.0%
ENSG00000173402	DAG1	1.251 (1.113,1.406)	1.69E-04	0.008	13.3%	78.7%	92.0%
ENSG00000196821	ILRUN	0.949 (0.923,0.976)	2.02E-04	0.009	99.9%	0.1%	100.0%
ENSG00000224557	HLA-DPB2	0.915 (0.873,0.960)	2.59E-04	0.011	100.0%	0.0%	100.0%
ENSG00000114270	COL7A1	1.358 (1.146,1.609)	4.03E-04	0.017	0.7%	89.5%	90.2%
ENSG00000112182	BACH2	0.920 (0.877,0.966)	0.001	0.026	84.5%	7.9%	92.4%
ENSG00000170035	UBE2E3	0.902 (0.846,0.962)	0.002	0.049	99.9%	0.0%	99.9%
ENSG00000111319	SCNN1A	1.292 (1.102,1.516)	0.002	0.050	20.5%	79.5%	100.0%

OR odds ratio, *CI* confidence interval, *FDR* false dicovery rate

predicted genes such as B3GNT2(OR: 0.914, 95%CI 0.894,0.933), PPT2 (OR: 0.556 95%CI 0.456,0.678), and PSMB9 (OR: 0.473, 95%CI 0.363,0.617) were negatively correlated with AS. However, genetically predicted higher expression levels of several genes were positively associated with AS, such as EFCAB13-DT (OR: 1.178, 95%CI: 1.116, 1.243), TUFM (OR: 1.030, 95%CI 1.019,1.041), ACTA2 (OR: 1.021, 95%CI 1.012,1.031), and TNFRSF1A (OR: 1.077, 95% CI 1.043,1.112). Of these, EFCAB13-DT, SGF29, TUFM, SULT1A2, KPNB1-DT, ACTA2, ATP2A1, SH2B1, KRT18P39, and INO80E were replicated in replication cohorts at the nominally significant level. (Supplementary Table S6 and Supplementary Table S7).

Assessing the effect of plasma proteins on AS

A total of 20 SMR-identified proteins were associated with AS risk ($P<0.05$ and P-HEIDI >0.05). (Supplementary Table S8) Based on the corrections of P value and colocalization, the volcano plot shows that nine SMR-prioritized proteins were identified as potential drug targets, of which five were positively associated with AS risk (Fig. 3). Table 2 shows that genetically predicted AIF1 (OR: 0.290, 95%CI 0.244–0.343), TNXB (OR: 0.570,

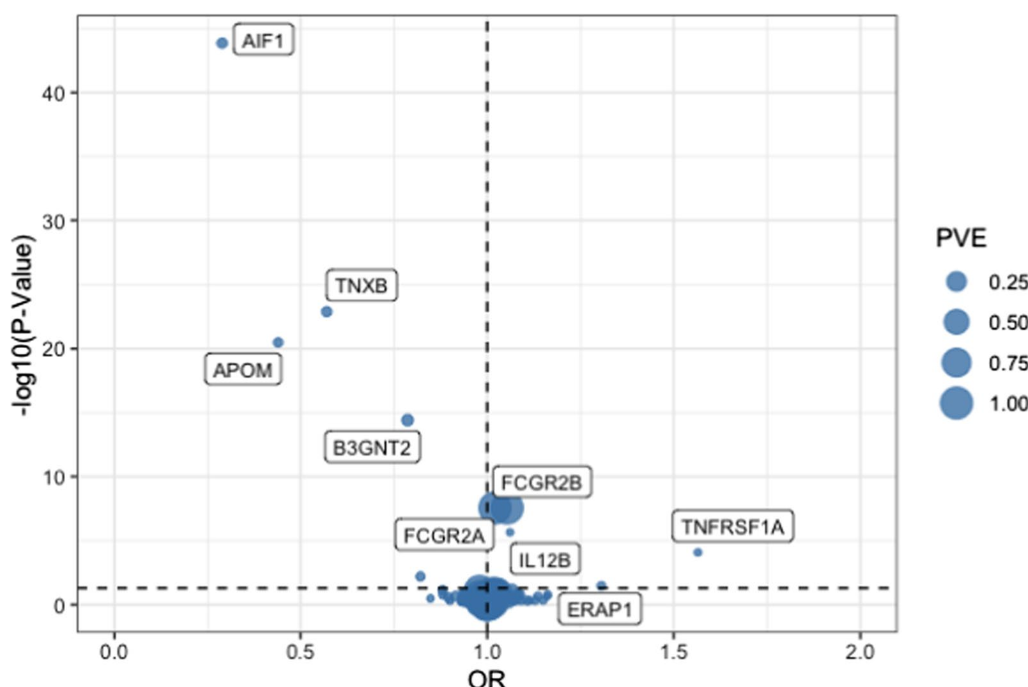

Fig. 3 Volcano plot of associations between plasma proteins and ankylosing spondylitis. A total of 9 SMR-prioritized proteins show significant associations with ankylosing spondylitis (AS), in which FCGR2A, FCGR2B, IL12B, TNFRSF1A, and ERAP1 are positively associated with AS, whereas AIF1, APOM, and B3GNT2 are negatively associated with AS. *OR* odds ratio, *PVE* proportion of variance explained

Table 2 The associations of pQTLs with ankylosing spondylitis

Protein	OR (95% CI)	P value	P-FDR value	PPH3	PPH4	PPH3 + PPH4
AIF1	0.290 (0.244–0.343)	9.68E-47	1.37E-44	10.9%	89.1%	100%
TNXB	0.570 (0.513–0.633)	1.85E-25	1.32E-23	100.0%	0.0%	100%
APOM	0.439 (0.373–0.517)	6.91E-23	3.27E-21	100.0%	0.0%	100%
B3GNT2	0.786 (0.743–0.832)	1.07E-16	3.79E-15	100.0%	0.0%	100%
FCGR2A	1.021 (1.014–1.027)	1.35E-09	2.74E-08	0.0%	100.0%	100.0%
FCGR2B	1.053 (1.036–1.071)	1.35E-09	2.74E-08	99.5%	0.0%	99.5%
IL12B	1.061 (1.038–1.085)	1.66E-07	2.14E-06	1.9%	98.1%	100%
TNFRSF1A	1.565 (1.289–1.900)	6.15E-06	7.94E-05	0.4%	99.6%	100%
ERAP1	1.306 (1.094–1.558)	3.07E-03	0.034	100.0%	0.0%	100%

OR odds ratio, *CI* confidence interval, *FDR* false dicovery rate

Table 3 Tier of genetically predicted methylation, expression, and protein of candidate genes with ankylosing spondylitis

Gene	Tier	mQTL				eQTL			pQTL		
		Probe	OR (95%CI)	P value	P-FDR value	OR (95%CI)	P value	P-FDR value	OR (95%CI)	P value	P-FDR value
TNFRSF1A	1	cg08418872	0.891 (0.849,0.935)	3.41E–06	8.46E–05	1.077 (1.043, 1.112)	7.33E–06	5.31E–04	1.565 (1.289–1.900)	6.15E–06	7.94E–05
TNFRSF1A	1	cg11268190	0.976 (0.966,0.987)	7.92E–06	1.93E–04						
TNFRSF1A	1	cg00556515	0.983 (0.974,0.991)	1.03E–04	0.002						
TNFRSF1A	1	cg09043214	0.978 (0.967,0.989)	1.15E–04	0.002						
TNFRSF1A	1	cg14009561	0.973 (0.960,0.987)	1.29E–04	0.003						
TNFRSF1A	1	cg23752651	0.960 (0.940,0.981)	1.93E–04	0.004						
TNFRSF1A	1	cg23320693	0.945 (0.917,0.975)	3.40E–04	0.007						
B3GNT2	2	–	–	–	–	0.914 (0.894,.0933)	5.28E-17	1.10E–14	0.786 (0.743–0.832)	1.07E–16	3.79E–15
ERAP1	2	cg00188032	1.227 (1.132,1.331)	7.63E–07	2.01E–05				1.306 (1.094–1.558)	3.07E–03	0.034
ERAP1	2	cg17330273	0.876 (0.824,0.931)	2.29E-05	5.27E-04	1.007 (0.961,1.055)	0.779	0.936			
ERAP1	2	cg25232719	1.024 (1.009,1.039)	0.001	0.025						
FCGR2A	2	cg24422489	0.960 (0.935,0.985)	0.002	0.037	1.060 (0.979,1.147)	0.152	0.542	1.021 (1.014–1.027)	1.35E–09	2.74E–08
FCGR2A	2	cg06350097	0.968 (0.948,0.989)	0.002	0.039						

OR odds ratio, *CI* confidence interval, *FDR* false dicovery rate

95%CI 0.513–0.633), APOM (OR: 0.439, 95%CI 0.373–0.517), and B3GNT2 (OR: 0.786, 95% CI 0.743–0.832) were negatively associated with AS. However, FCGR2A (OR: 1.021, 95% CI 1.014–1.027), FCGR2B (OR: 1.053, 95% CI 1.036–1.071), IL12B (OR: 1.061, 95%CI: 1.038–1.085), TNFRSF1A (OR: 1.565, 95% CI 1.289–1.900) and ERAP1 (OR: 1.306, 95% CI 1.094–1.558) were positively associated with AS. These SMR-prioritized proteins all had supported evidence of colocalization in AS (PPH3 + PPH4 ≥ 80%). Similarly, AIF1, TNXB, APOM, B3GNT2, IL12B, TNFRSF1A, GCKR, and SLAMF7 were identified to be associated with AS in the FinnGen cohort (Supplementary Table S9), whereas AIF1, TNXB, APOM, and FCGR3B were found to be associated with this disease in the UK biobank cohorts. (Supplementary Table S10).

Integration of multi-omics evidence

Based on the integration of multi-omics results, we found that TNFRSF1A was categorized into tier 1, whereas B3GNT2, FCGR2A, and ERAP1 into tier 2. (Table 3) In the replication cohort, not all associations between these four genes and AS reached statistical significance. However, mostly their connections had a consistent direction as those disclosed in the discovery stage. (Supplementary Table S11) The associations between gene methylation, expression, and protein abundance are shown in Supplementary Table S12. We found that the methylation sites (cg08418872, cg11268190, cg00556515, cg09043214, cg14009561, cg23752651, and cg23320693) in TNFRSF1A were negatively associated with its gene expression (all $P < 0.05$). Inversely, the expression of TNFRSF1A was positively associated with its protein level ($P = 2.31E\text{-}13$). A positive connection was observed between B3GNT2 expression and its protein. ($P = 2.67E\text{-}65$). Moreover, the methylations of ERAP1 (cg17330273) and FCGR2A (cg06350097 and cg24422489) were positively associated with their gene expressions, but negatively correlated with their corresponding proteins. (all $P < 0.05$) The colocalization analysis further supported our results, in which PPH3 + PPH4 of mQTL with eQTL, eQTL with pQTL in these genes were all ≥ 80%.

PPI and enriched pathways of the SMR-identified proteins

The PPI network plot shows the interactions between these SMR-identified proteins, which contained 20 nodes and 18 edges with a significant PPI enrichment P-value of $5.75E^{-}10$. (Fig. 4). In the GO enrichment analysis, we found that biological pathways relevant to AS mainly participate in the immune response, such as leukocyte activation, positive regulation of lymphocyte migration, natural killer cell activation, T-helper 1 cell cytokine production, interleukin-23-mediated signaling pathway, and cell adhesion. (Supplementary Table S13, Fig. 5A). The KEGG analysis identified that osteoclast differentiation, cytokine-cytokine receptor interaction, NF-kappa B signaling pathway, TNF signaling pathway, inflammatory bowel disease (IBD), leishmaniasis, and tuberculosis, etc. were associated with these proteins identified in this study (Supplementary Table S14, Fig. 5B).

Drug target of the SMR-prioritized proteins and molecular docking

The DSigDB database predicted 149 potential chemical compounds based on these SMR-prioritized proteins. (Supplementary Table S15). The top 10 potential chemical compounds with P-value < 0.05 are listed in Table 4. IL12B and TNFRSF1A could be the most important target proteins for AS, which were significantly connected to more drug candidates, such as isoproterenol (CTD 00006175), dinoprostone (CTD 00007049), and thalidomide (CTD 00006858), etc. We further used AutoDock to estimate the affinity of the top 5 predicted drugs for these proteins. A total of 9 effective docking results were found between 5 proteins and 4 candidate drugs (Fig. 6). Of these, the lowest binding energy was exhibited between IL12B and thalidomide (−7.2 kcal/mol), which was regarded as the most stable binding. (Table 5).

Discussion

In this study, a total of 9 genes were identified as promising therapeutic targets for AS, including AIF1, TNXB, APOM, B3GNT2, FCGR2A, FCGR2B, IL12B, TNFRSF1A, and ERAP1. Based on multi-omics evidence, TNFRSF1A, B3GNT2, FCGR2A, and ERAP1 genes were found to be associated with AS risk at different regulatory levels. PPI network and enrichment analysis revealed the regulatory connections among these proteins and their functional traits, primarily related to immune responses. All these results show that these genetically predicted drug targets may have potentially druggable values.

TNFRSF1A was categorized into tier 1 in our study. We found that its protein level was positively associated with AS risk. This gene-encoded protein is one of the major transmembrane receptors for TNF-α, which plays a critical role in inducing the activation of NF-κB. The activation of the NF-κB signaling pathway functions in recognizing inflammation, regulating inflammatory mediators, and inducing the release of various pro-inflammatory cytokines, such as IL-1β, TNF-α, and IL-6. [18] Previous GWAS studies have reported that TNFRSF1A is one of the susceptibility genes for AS. [19, 20] However, no study reported the regulation among TNFRSF1A methylation, expression, and protein abundance for AS. Our results expand the knowledge that up-regulating the methylation level of cg08418872,

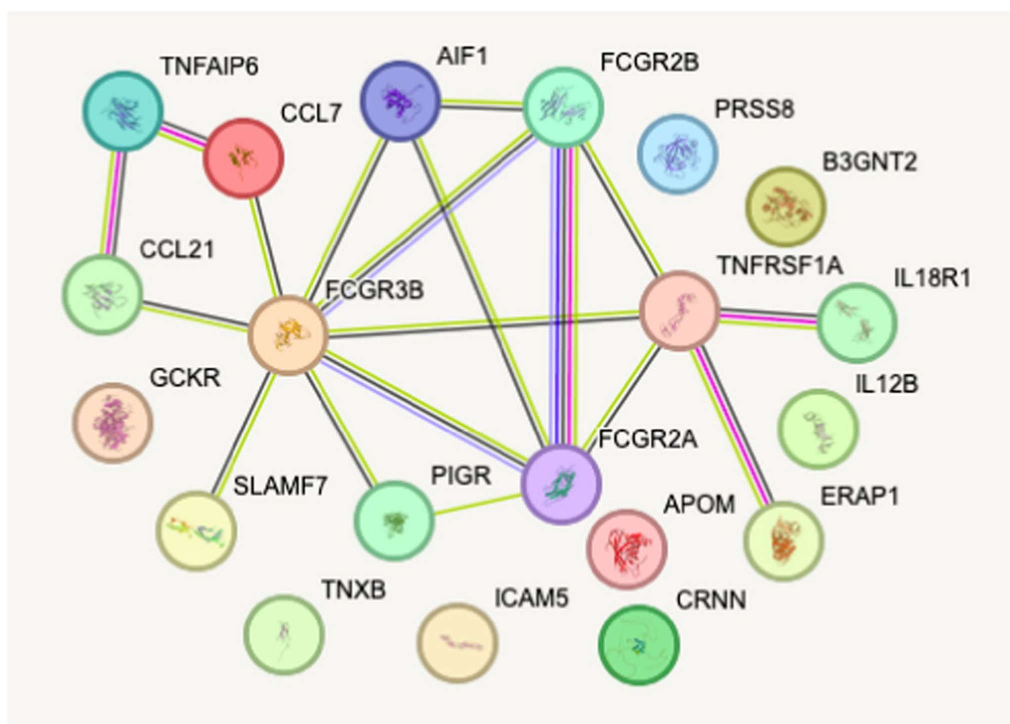

Fig. 4 Protein–protein interaction (PPI) of 20 SMR-identified proteins constructed by STRING

cg11268190, cg00556515, cg09043214, cg14009561, cg23752651, and cg23320693 in this gene could reduce the risk of AS by down-regulating its expression and protein levels.

B3GNT2 was categorized into tier 2 based on multi-omics results. Our findings showed that its decreased expression could significantly increase AS risk. It may be due to the fact that knocking out B3GNT2 could reduce the cell surface poly-N-acetyl-lactosamine and lead to hypersensitivity and hyperresponsivity of immunocytes [21]. A recent study has reported that increased mRNA expression of B3GNT2 could reduce C-reactive protein, erythrocyte sedimentation rate, and syndesmophyte formation in AS patients, suggesting that B3GNT2 was a protective factor for this disease [22]. ERAP1 and FCGR2A were also categorized into tier 2. Our findings are consistent with previous studies, in which reduced endopeptidase activity of ERAP1 could be a protective effect for this disease [20, 23]. FCGR2A encodes a member of the Fc-gamma receptors family for recognizing immunoglobulin G. It could activate immune cell functions, function in phagocytosis, and release inflammatory mediators [24]. Some studies have reported that FCGR2A is genetically associated with a lot of diseases triggered by auto-antibodies or immune complexes, such as rheumatoid arthritis (RA) [25], SLE [26], Kawasaki disease [27], and inflammatory bowel disease [28]. However, the causal relationship between FCGR2A and AS is still unclear. Our SMR results showed that this gene was positively associated with AS at the protein level. Furthermore, we revealed a negative relationship of methylation in ERAP1 (cg17330273) and FCGR2A (cg06350097 and cg24422489) with AS.

Our study demonstrated that IL12B, AIF1, TXNB, and APOM were associated with AS at the protein level. However, we failed to find significant associations between these genes and AS at the methylation and gene expression levels. A few studies have demonstrated that the polymorphisms of IL12B are associated with AS susceptibility [29, 30]. In this study, we found that IL12B encoded protein was positively associated with AS risk. This gene encodes the p40 component of IL-12 and IL-23, which is involved in both the IL12/Th1 and IL23/Th17 pathways [31]. Notably, the IL-23/IL-17 signaling pathway plays a critical role in the pathogenesis of AS [32]. Inversely, our findings showed that the increased protein levels of AIF1, TNXB, and APOM could reduce the risk of AS. Meanwhile, these significantly negative correlations were also found in FinnGen and UK biobank cohorts. The AIF1 gene encodes a calcium-responsive cytoplasmic scaffold protein, which mediates innate and adaptive immune responses within dendritic cells (DC) and macrophages [33, 34]. It has been reported that AIF1 is correlated with various autoimmune diseases, such as

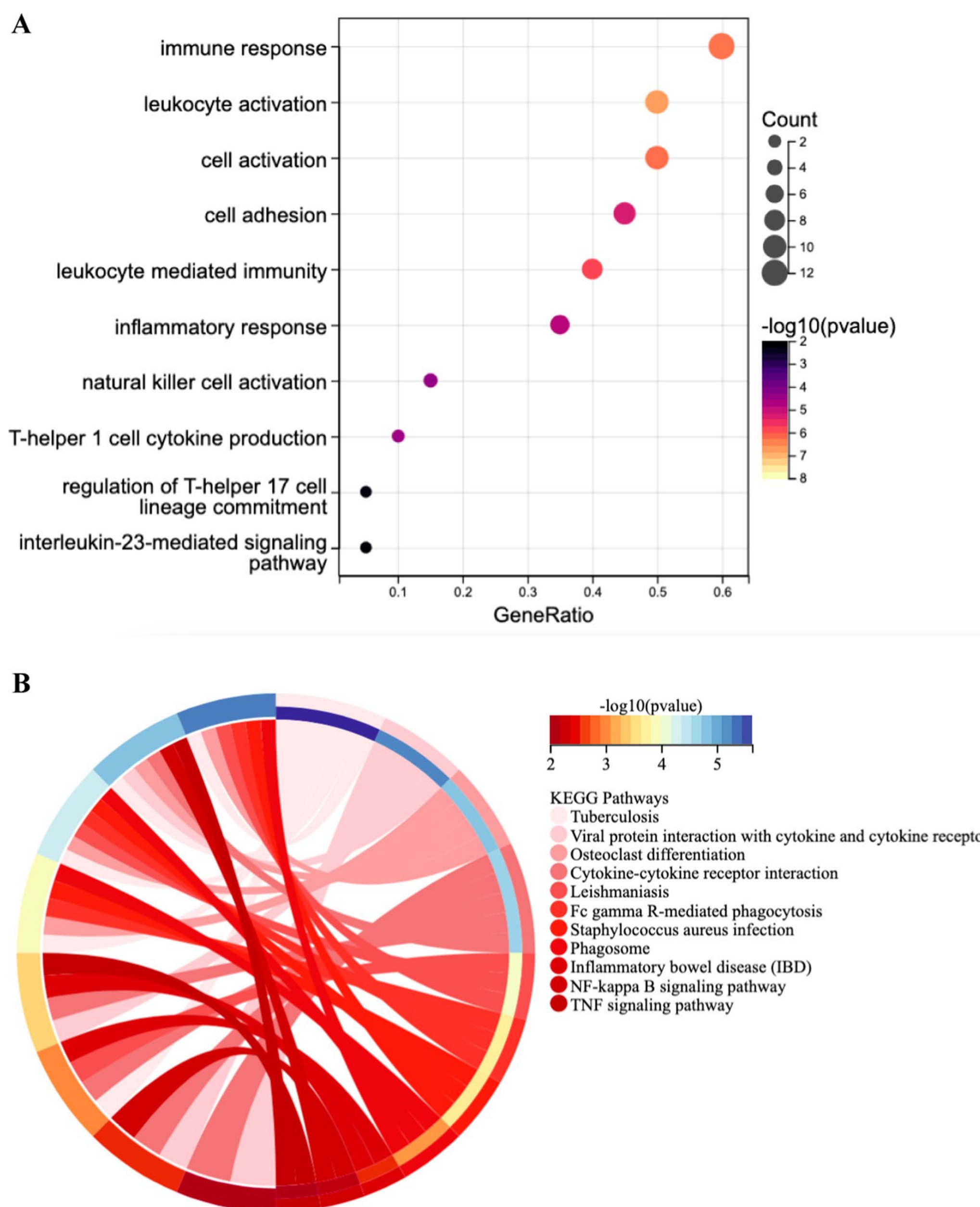

Fig. 5 Functional enrichment analysis of 20 SMR-identified proteins. **A** The Gene Ontology (GO) analysis; **B** Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathway analyses

RA [35, 36] and systemic sclerosis [37]. The TNXB gene is responsible for the production of tenascin-X (TNX), an extracellular matrix glycoprotein abundantly expressed in skin, muscle, tendon sheath, peripheral nerve, and blood vessels [38, 39]. The APOM protein is mainly associated with high-density lipoprotein in human plasma and functions with anti-inflammatory effects [40]. However, the evidence on the causal effect of AIF1, TNXB, and APOM on AS from epidemiological and experimental studies was relatively limited. Therefore, experimental validation is warranted to further explore the biological mechanisms of these genes on AS in the future.

Our enrichment analysis demonstrated that the proteins were associated with the TNF pathway, IL23/Th17 pathway, and NF-kappa B signaling pathway. Of note, these pathways are the focal point of drug development for AS [41]. Currently, a few biological agents have been approved for AS by targeting key regulators

Table 4 Candidate drug predicted using DSigDB

Drugs	Genes	P value
Isoproterenol CTD 00006175	IL12B; TNFRSF1A	5.45E-04
Cimetidine PC3 UP	TNXB; FCGR2A;ERAP1	7.13E-04
Dinoprostone CTD 00007049	IL12B; TNFRSF1A	1.40E-03
Thalidomide CTD 00006858	IL12B; TNFRSF1A	1.72E-03
NICKEL SULFATE CTD 00001417	IL12B; FCGR2B;TNFRSF1A	1.75E-03
R-atenolol PC3 DOWN	ERAP1; B3GNT2	1.77E-03
AGN-PC-0JHFVD BOSS	IL12B; FCGR2B	3.13E-03
Capsaicin CTD 00005570	IL12B; TNFRSF1A	3.62E-03
Dimethyl sulfoxide CTD 00005842	APOM; TNFRSF1A	3.66E-03
PHENCYCLIDINE CTD 00005881	IL12B; TNFRSF1A	4.34E-03

in these signaling cascades, such as anti-TNF-α agents (e.g. infliximab, adalimumab, etanercept) and an IL-17A inhibitor (secukinumab) [42, 43]. However, some patients still respond poorly to these therapies or develop resistance over time, suggesting that the signaling pathways involved in the pathogenesis of AS are diverse and complex. In this study, drug prediction and molecular docking results further demonstrated the medicinal potential of these SMR-prioritized proteins, especially TNFRSF1A and IL12B. Compared to the TNF inhibitor, the TNFRSF1A inhibitor could block the interaction of TNF with TNFRSF1A directly at the receptor level without affecting the signaling of TNF bound to the cell membrane and other receptors. Therefore, it may reduce the adverse effects (e.g. infections, malignancies) caused by non-selective TNF inhibitors. Furthermore, the IL12B inhibitors may reduce pro-inflammatory cytokines like IFN-γ and IL-17 by simultaneously modulating multiple immune pathways, such as IL12/Th1 and IL23/Th17 pathways. Therefore, targeting these targets for the development of new alternative strategies may be promising for AS patients unresponsive to the current therapies.

This study has some advantages. The SMR method could minimize reverse causation and confounding bias. Meanwhile, the HEIDI test and colocalization approach could attenuate the potential bias caused by linkage disequilibrium and strengthen our conclusion. Moreover, a large sample of GWAS data contributes to increasing the statistical power of this study. However, there were also some inevitable limitations. The limitations of this study and corresponding recommendations for future research were listed as follows. Firstly, this study predominantly focused on European descent.

Table 5 Molecular docking between target genes and predicted drugs

Target Genes	PDB ID	Drugs	PUBCHEM ID	Binding energy (kcal/mol)
ERAP1	2YD0	Cimetidine	2756	−5.6
FCGR2A	1FCG	Cimetidine	2756	−4.0
IL12B	1F42	Isoproterenol	3779	−5.9
IL12B	1F42	Dinoprostone	5,280,360	−5.8
IL12B	1F42	Thalidomide	5426	−7.2
TNFRSF1A	1EXT	Isoproterenol	3779	−6.4
TNFRSF1A	1EXT	Dinoprostone	5,280,360	−4.7
TNFRSF1A	1EXT	Thalidomide	5426	−5.9
TXNB	2CUH	Cimetidine	2756	−4.3

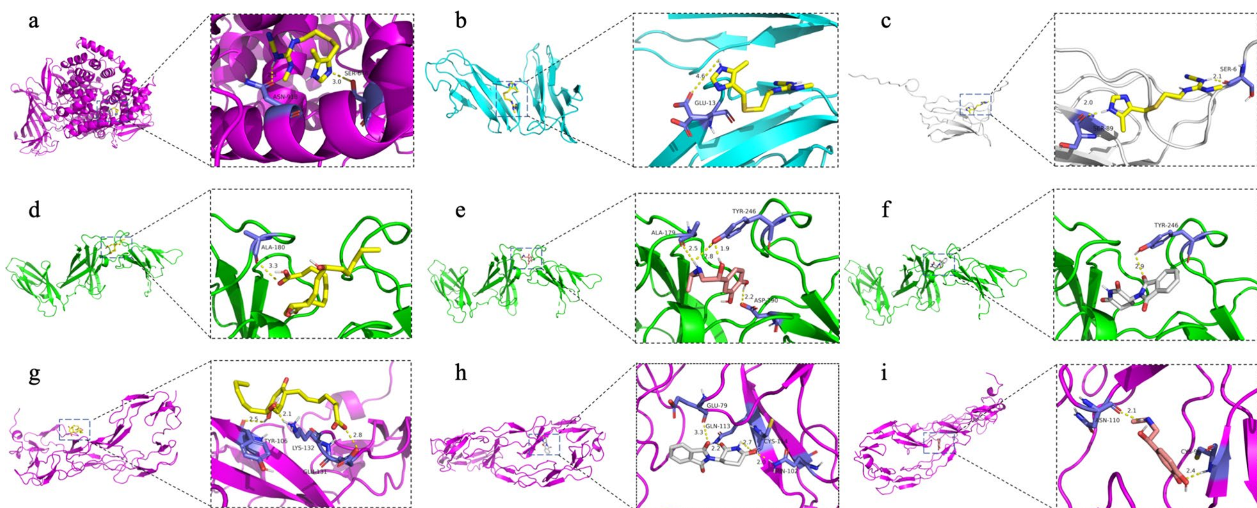

Fig. 6 Molecular docking. **a** ERAP1-cimetidine; **b** FCGR2A-cimetidine; **c** TNXB-cimetidine: **d** IL12B-dinoprostone; **e** IL12B-isoproterenol; **f** IL12B-thalidomide; **g** TNFRSF1A-dinoprostone; **h** TNFRSF1A-isoproterenol; **i** TNFRSF1A-thalidomide

Due to differences in allele frequencies and linkage disequilibrium patterns in diverse populations, genetic variants often exhibit population-specificity. For example, HLA-B27, a well-known genetic risk factor for AS, exhibits different prevalence rates across ethnicities, being highly prevalent in European and certain Asian populations but much less common in African populations [44]. Therefore, future research should focus on validating these findings in populations with diverse genetic backgrounds to identify both shared and differential genetic risk for AS, which may contribute to developing more personalized and effective therapeutic strategies.

Secondly, the QTL data were just derived from peripheral blood, which may result in a lack of understanding of tissue-specific regulatory mechanisms. Previous studies have reported that methylation patterns, expression profiles, and protein interactions of certain genes may differ in various tissues [45]. Therefore, incorporating QTL data from AS mainly involved tissues (e.g. joint and spine tissues) in future studies will be essential to improve the comprehensive understanding of AS pathogenesis.

Thirdly, although molecular docking predicts the interactions of potential drugs and targets, their accuracy largely depends on the quality of the ligand and protein structure, which may fail to account for the complexity of biological systems. However, experimental studies could find off-target effects and toxicity and reduce the risk of failure in drug development. The in vitro experiments could confirm whether the predicted interactions translate into real-world biological activity. The animal models and clinical trials could comprehensively assess the pharmacokinetics, efficacy, and safety of potential therapies. Therefore, future studies should focus on experimental validations to confirm the biological relevance and therapeutic potential of the identified drug-target interactions presented in this study.

Conclusion

This study conducted SMR analysis to assess the causal associations of gene methylation, expression, and protein levels with AS. We demonstrated 9 proteins as potential drug targets for AS, which provided a novel insight into the therapeutic strategy for this disease in the future. However, more functional experiments and genetically engineered animal models are needed to further address the role of these genes in the pathogenesis of AS.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12967-024-05925-x.

Supplementary material 1

Acknowledgements

All authors are grateful to all those who participated in this study and thank the International Genetics Consortium for Ankylosing Spondylitis (IGAS), FinnGen database, UK Biobank database, and quantitative trait loci sources for providing publicly available statistics.

Author contributions

Lingyu Dai: Conceptualization, Data curation, Investigation, Methodology, Software, Project administration, Validation, Writing – original draft. Lan Xia: Data curation, Investigation, Software, Validation. Guannan Su: Data curation, Investigation. Yu Gao: Data curation, Investigation. Peizeng Yang: Conceptualization, Supervision, Writing – review & editing, Funding acquisition.

Funding

The work was supported by National Natural Science Foundation Key Program (82230032), National Natural Science Foundation Key Program (81930023), Key Project of Chongqing Science and Technology Bureau (CSTC2021jscx-gksb-N0010), Chongqing Outstanding Scientists Project (2019), and Chongqing Science and Technology Bureau Mountaineering Project (cyyy-xkdfjh-jcyj-202301, cyyy-xkdfjh-lcyj-202303, cyyy-xkdfjh-cgzh202302).

Availability of data and materials

All data related to our study are presented in the article or uploaded as supplementary information. Reasonable requests for further data queries could be addressed to the corresponding author.

Declarations

Ethics approval and consent to participate

All data in this study were obtained from public databases approved by the corresponding ethical review boards. Informed written consent was obtained from all subjects.

Consent for publication

Not applicable.

Competing interests

The authors declare no conflict of interests.

Received: 24 August 2024 Accepted: 27 November 2024
Published online: 20 December 2024

References

1. Taurog JD, Chhabra A, Colbert RA. Ankylosing spondylitis and axial spondyloarthritis. N Engl J Med. 2016;374(26):2563–74.
2. Navarro-Compán V, Sepriano A, El-Zorkany B, van der Heijde D. Axial spondyloarthritis. Ann Rheum Dis. 2021;80(12):1511–21.
3. Exarchou S, Lindström U, Askling J, et al. The prevalence of clinically diagnosed ankylosing spondylitis and its clinical manifestations: a nationwide register study. Arthritis Res Ther. 2015;17(1):118.
4. Farh KK, Marson A, Zhu J, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature. 2015;518(7539):337–43.
5. Danve A, Deodhar A. Treatment of axial spondyloarthritis: an update. Nat Rev Rheumatol. 2022;18(4):205–16.
6. Yang HC, Liang YJ, Chung CM, Chen JW, Pan WH. Genome-wide gene-based association study. BMC Proc. 2009. https://doi.org/10.1186/1753-6561-3-S7-S135.
7. Yazdani A, Yazdani A, Mendez-Giraldez R, Samiei A, Kosorok MR, Schaid DJ. From classical mendelian randomization to causal networks for systematic integration of multi-omics. Front Genet. 2022;13: 990486.
8. Akiyama M. Multi-omics study for interpretation of genome-wide association study. J Hum Genet. 2021;66(1):3–10.
9. Kim MS, Song M, Kim B, et al. Prioritization of therapeutic targets for dyslipidemia using integrative multi-omics and multi-trait analysis. Cell Rep Med. 2023;4(9): 101112.

10. Zhu Z, Zhang F, Hu H, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet. 2016;48(5):481–7.
11. McRae AF, Marioni RE, Shah S, et al. Identification of 55,000 replicated DNA methylation QTL. Sci Rep. 2018;8(1):17605.
12. Ferkingstad E, Sulem P, Atlason BA, et al. Large-scale integration of the plasma proteome with genetics and disease. Nat Genet. 2021;53(12):1712–21.
13. Auton A, Brooks LD, Durbin RM, et al. A global reference for human genetic variation. Nature. 2015;526(7571):68–74.
14. Chen J, Ruan X, Sun Y, et al. Multi-omic insight into the molecular networks of mitochondrial dysfunction in the pathogenesis of inflammatory bowel disease. EBioMedicine. 2024;99: 104934.
15. Su WM, Gu XJ, Dou M, et al. Systematic druggable genome-wide Mendelian randomisation identifies therapeutic targets for Alzheimer's disease. J Neurol Neurosurg Psychiat. 2023;94(11):954–61.
16. Storm CS, Kia DA, Almramhi MM, et al. Finding genetically-supported drug targets for Parkinson's disease using Mendelian randomization of the druggable genome. Nat Commun. 2021;12(1):7342.
17. Yoo M, Shin J, Kim J, et al. DSigDB: drug signatures database for gene set analysis. Bioinformatics. 2015;31(18):3069–71.
18. Medunjanin S, Putzier M, Nöthen T, et al. DNA-PK: gatekeeper for IKKγ/NEMO nucleocytoplasmic shuttling in genotoxic stress-induced NF-kappaB activation. Cell Mol Life Sci. 2020;77(20):4133–42.
19. Davidson SI, Liu Y, Danoy PA, et al. Association of STAT3 and TNFRSF1A with ankylosing spondylitis in Han Chinese. Ann Rheum Dis. 2011;70(2):289–92.
20. Cortes A, Hadler J, Pointon JP, et al. Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. Nat Genet. 2013;45(7):730–8.
21. Hao Y, Créquer-Grandhomme A, Javier N, et al. Structures and mechanism of human glycosyltransferase β1,3-N-acetylglucosaminyltransferase 2 (B3GNT2), an important player in immune homeostasis. J Biol Chem. 2021;296: 100042.
22. Wang CM, Jan Wu YJ, Lin JC, Huang LY, Wu J, Chen JY. Genetic effects of B3GNT2 on ankylosing spondylitis susceptibility and clinical manifestations in Taiwanese. J Formos Med Assoc. 2022;121(7):1283–94.
23. Keidel S, Chen L, Pointon J, Wordsworth P. ERAP1 and ankylosing spondylitis. Curr Opin Immunol. 2013;25(1):97–102.
24. Castro-Dopico T, Dennison TW, Ferdinand JR, et al. Anti-commensal IgG drives intestinal inflammation and type 17 immunity in ulcerative colitis. Immunity. 2019;50(4):1099-114.e10.
25. Lee YH, Bae SC, Song GG. FCGR2A, FCGR3A, FCGR3B polymorphisms and susceptibility to rheumatoid arthritis: a meta-analysis. Clin Exp Rheumatol. 2015;33(5):647–54.
26. Harley JB, Alarcón-Riquelme ME, Criswell LA, et al. Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. Nat Genet. 2008;40(2):204–10.
27. Khor CC, Davila S, Breunis WB, et al. Genome-wide association study identifies FCGR2A as a susceptibility locus for Kawasaki disease. Nat Genet. 2011;43(12):1241–6.
28. Liu JZ, van Sommeren S, Huang H, et al. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. Nat Genet. 2015;47(9):979–86.
29. Ivanova M, Manolova I, Miteva L, Gancheva R, Stoilov R, Stanilova S. Genetic variations in the IL-12B gene in association with IL-23 and IL-12p40 serum levels in ankylosing spondylitis. Rheumatol Int. 2019;39(1):111–9.
30. Wong RH, Wei JC, Huang CH, et al. Association of IL-12B genetic polymorphism with the susceptibility and disease severity of ankylosing spondylitis. J Rheumatol. 2012;39(1):135–40.
31. Zhu KJ, Zhu CY, Shi G, Fan YM. Meta-analysis of IL12B polymorphisms (rs3212227, rs6887695) with psoriasis and psoriatic arthritis. Rheumatol Int. 2013;33(7):1785–90.
32. Zhong Z, Su G, Kijlstra A, Yang P. Activation of the interleukin-23/interleukin-17 signalling pathway in autoinflammatory and autoimmune uveitis. Prog Retin Eye Res. 2021;80: 100866.
33. Elizondo DM, Brandy NZD, da Silva RLL, et al. Allograft inflammatory factor-1 governs hematopoietic stem cell differentiation into cDC1 and monocyte-derived dendritic cells through IRF8 and RelB in vitro. Front Immunol. 2019;10:173.
34. Elizondo DM, Andargie TE, Yang D, Kacsinta AD, Lipscomb MW. Inhibition of allograft inflammatory factor-1 in dendritic cells restrains CD4(+) T Cell effector responses and induces CD25(+)Foxp3(+) T regulatory subsets. Front Immunol. 2017;8:1502.
35. Piotrowska K, Słuczanowska-Głabowska S, Kurzawski M, et al. Over-expression of allograft inflammatory factor-1 (AIF-1) in patients with rheumatoid arthritis. Biomolecules. 2020. https://doi.org/10.3390/biom10071064.
36. Kimura M, Kawahito Y, Obayashi H, et al. A critical role for allograft inflammatory factor-1 in the pathogenesis of rheumatoid arthritis. J Immunol. 2007;178(5):3316–22.
37. Alkassab F, Gourh P, Tan FK, et al. An allograft inflammatory factor 1 (AIF1) single nucleotide polymorphism (SNP) is associated with anticentromere antibody positive systemic sclerosis. Rheumatology (Oxford). 2007;46(8):1248–51.
38. Zheng W, Rao S. Knowledge-based analysis of genetic associations of rheumatoid arthritis to inform studies searching for pleiotropic genes: a literature review and network analysis. Arthritis Res Ther. 2015;17(1):202.
39. Kamatani Y, Matsuda K, Ohishi T, et al. Identification of a significant association of a single nucleotide polymorphism in TNXB with systemic lupus erythematosus in a Japanese population. J Hum Genet. 2008;53(1):64–73.
40. Blaho VA, Galvani S, Engelbrecht E, et al. HDL-bound sphingosine-1-phosphate restrains lymphopoiesis and neuroinflammation. Nature. 2015;523(7560):342–6.
41. Zhu W, He X, Cheng K, et al. Ankylosing spondylitis: etiology, pathogenesis, and treatments. Bone Res. 2019;7:22.
42. Braun J, Kiltz U, Heldmann F, Baraliakos X. Emerging drugs for the treatment of axial and peripheral spondyloarthritis. Expert Opin Emerg Drugs. 2015;20(1):1–14.
43. Tian C, Shu J, Shao W, Zhou Z, Guo H, Wang J. Efficacy and safety of IL inhibitors, TNF-α inhibitors, and JAK inhibitors in patients with ankylosing spondylitis: a systematic review and Bayesian network meta-analysis. Ann Transl Med. 2023;11(4):178.
44. Reveille JD, Hirsch R, Dillon CF, Carroll MD, Weisman MH. The prevalence of HLA-B27 in the US: data from the US national health and nutrition examination survey, 2009. Arthritis Rheum. 2012;64(5):1407–11.
45. Jiang L, Wang M, Lin S, et al. A quantitative proteome map of the human body. Cell. 2020;183(1):269-83.e19.

Publisher's Note

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