Identifying new biomarkers and potential therapeutic targets for breast cancer through the integration of human plasma proteomics: a Mendelian randomization study and colocalization analysis

Jingshuang Song[1,2] and Huawei Yang[1,3]*

[1]Department of Breast Surgery, Guangxi Medical University Cancer Hospital, Nanning, China,
[2]Department of Breast and Thyroid Surgery, Affiliated Hospital of Guilin Medical University,
Guilin, China, [3]Laboratory of Breast Cancer Diagnosis and Treatment Research of Guangxi
Department of Education, Guangxi Medical University, Nanning, China

Background: The proteome is a crucial reservoir of targets for cancer treatment. While some targeted therapies have been developed, there are still significant challenges in early diagnosis and treatment, highlighting the need to identify new biomarkers and therapeutic targets for breast cancer. Therefore, we conducted a comprehensive proteome-wide Mendelian randomization (MR) study to identify novel biomarkers and potential therapeutic targets for breast cancer.

Methods: Protein quantitative trait locus (pQTL) data were extracted from two published plasma proteome-wide association studies. Genetic variants associated with breast cancer were obtained from the Breast Cancer Association Consortium, which included 133,384 cases and 113,789 controls, and the Finnish cohort study, comprising 18,786 cases and 182,927 controls. We employed summary-based MR and colocalization methods to identify potential drug targets for breast cancer, which were subsequently validated using a two-sample MR approach. Finally, a protein-protein interaction (PPI) network was constructed to detect interactions between the identified proteins and existing cancer drug targets.

Results: Gene-predicted levels of ten proteins were associated with breast cancer risk. Decreased levels of CASP8, DDX58, CPNE1, ULK3, PARK7, and BTN2A1, as well as increased levels of TNFRSF9, TNXB, DNPH1, and TLR1, were linked to an elevated risk of breast cancer. Among these, CASP8 and DDX58 were supported by tier-one evidence, while CPNE1, ULK3, PARK7, and TNFRSF9 received tier-two evidence support. The remaining proteins, TNXB, BTN2A1, DNPH1, and TLR1, were supported by tier-three evidence. CASP8, DDX58, CPNE1, ULK3, PARK7, and TNFRSF9 have already been identified as targets in drug development and potential therapeutic targets for breast cancer treatment. Additionally, ULK3 showed promise as a prognostic biomarker for breast cancer.

Conclusions: The present study identified several novel potential drug targets and biomarkers for breast cancer, providing new insights into its diagnosis and treatment. The integration of PPI and druggability evaluations enhances the prioritization of these therapeutic targets, paving the way for future drug development efforts.

1 Introduction

Breast cancer is one of the most prevalent malignancies among women worldwide, with an incidence that continues to rise. In 2023, the United States alone had some 300,000 new cases, accounting for approximately 15.32% of all newly diagnosed cancers. Simultaneously, around 43,000 deaths due to breast cancer were recorded, constituting 7.2% of all cancer-related mortalities. Breast cancer profoundly affects patients' quality of life and overall health (1, 2), and despite the advancements in treatment modalities, significant challenges, including the inadequacy of early diagnosis, the unpredictability of treatment outcomes, and the development of drug resistance, remain (3). As a result, identifying novel biomarkers and therapeutic targets has become a critical focus of contemporary breast cancer research.

Proteomics is a high-throughput technology that can reflect normal physiological processes and cancer pathobiology. Researchers can discover novel cancer-associated biomarkers by analyzing protein expression profiles in tumor tissues or body fluids, offering theoretical support for personalized patient treatment (4). Previous observational studies have identified specific circulating proteins associated with breast cancer risk (5–8); however, reverse causality or confounding factors may obscure the conclusions drawn from traditional observational research.

Mendelian randomization (MR) is a method employed to estimate causal effects within specific hypothetical contexts based on the principle that genes are randomly assigned from parents to offspring during gametogenesis and conception. Unlike traditional observational studies, MR is not susceptible to the biases of reverse causality or confounding (9). Consequently, several studies using the MR approach have uncovered various circulating proteins linked to breast cancer risk. For example, Jia and colleagues employed a two-sample MR approach to assess the association between 1,142 proteins and breast cancer risk, identifying 22 proteins linked to this risk (10). Similarly, Shu et al. utilized the same method to examine 2,994 proteins, uncovering 56 associated with breast cancer risk (11). They further explored the relationship between 1,890 circulating proteins and various breast cancer subtypes, identifying 98 proteins significantly associated with one or more subtypes (12). Additionally, Mälarstig and colleagues used a two-sample MR approach to identify five proteins potentially causally linked to breast cancer (13). However, these studies often relied on a single analytical method or faced limitations in protein coverage and sample size, generating inconsistent findings that have hindered a comprehensive understanding of the relationship between protein expression and breast cancer risk. Two recent studies have further advanced this field by employing bidirectional MR and colocalization analyses to systematically explore potential drug targets between plasma proteins and breast cancer. Colocalization analysis effectively distinguishes causal relationships from linkage disequilibrium (LD) within the genome, thereby enhancing the reliability of the results. This multifaceted approach offers crucial insights into identifying potential drug targets for breast cancer and lays a solid foundation for future drug development (14, 15).

The present study utilized Summary-based Mendelian Randomization (SMR), an advanced extension of traditional MR methods. The SMR approach integrates independent genome-wide association study (GWAS) summary data with quantitative trait locus (QTL) data, thereby prioritizing potential causal genes identified in GWAS. Unlike conventional MR methods, SMR can more accurately distinguish potential causal associations from LD within the genome, yielding more reliable causal inferences (16). By combining SMR with colocalization analysis, we systematically investigated the relationship between the human plasma proteome and breast cancer risk. This innovative approach enabled us to address some of the limitations of previous studies, providing more robust supporting evidence (17). Given the limitations of evidence from a single methodological approach, we further employed a two-sample MR approach for validation,

Abbreviations: MR, Mendelian randomization; SMR, Summarized data-based Mendelian Randomization; GWAS, genome-wide association studies; QTL, Quantitative Trait Locus; pQTL, protein quantitative trait locus; LD, linkage disequilibrium; IVW, inverse variance weighting; BCAC, Breast Cancer Association Consortium; SNPs, single nucleotide polymorphisms; FDR, false discovery rate; OR, odds ratio; CI, confidence interval; PPI, protein-protein interaction.

systematically assessing the potential of proteins as novel biomarkers and therapeutic targets for breast cancer. Future research should integrate multi-omics data, including expression quantitative trait locus (eQTL) and methylation quantitative trait locus (mQTL), by combining SMR and two-sample MR methods, as such an approach could offer new perspectives on the molecular mechanisms of breast cancer and provide critical insights for identifying targets in personalized therapy.

2 Materials and methods

The present research adhered to the guidelines outlined in the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) (Supplementary STROBE-MR checklist table) (18).

2.1 Study design

Both preliminary and validation analyses were conducted in the present study. Protein quantitative trait locus (pQTL) data from two large-scale proteomic studies were used, and the SMR method was employed to evaluate the association between proteome and breast cancer. Positive findings from this initial assessment were subjected to Bayesian colocalization analysis. For proteins that met the criteria after colocalization analysis, the causal relationship with breast cancer was further validated using a two-stage MR framework, which included discovery and replication phases and was supplemented by sensitivity analyses. The MR analysis adhered rigorously to the following fundamental assumptions: (i) relevance assumption: single nucleotide polymorphisms (SNPs) are significantly associated with the exposure (protein expression levels); (ii) independence assumption: SNPs are independent of confounding factors, meaning they are not associated with variables that influence both the exposure and the disease outcome; (iii) exclusion restriction assumption: SNPs affect breast cancer risk exclusively through protein expression levels and not through other pathways (9) (Figure 1).

2.2 Proteomic data source

The pQTL data for the proteomic studies were obtained from research conducted by Ferkingstad et al. and Sun et al. The former study assessed the plasma protein levels in 35,559 Icelandic individuals using the SomaScan multiplex aptamer assay. They profiled 4,719 proteins and identified pQTL data for 18,084 protein quantitative trait loci (19). Sun et al. performed proteomic profiling of plasma samples from 54,219 participants in the UK Biobank using the antibody-based Olink Explore 3072 PEA technology. They profiled 2,923 distinct proteins and yielded pQTL data for 14,287 protein quantitative trait loci (20) (Supplementary Table S1).

2.3 Study population

The genetic data relevant to breast cancer were obtained from the Breast Cancer Association Consortium (BCAC) and the FinnGen Biobank. The BCAC consortium combined three datasets: iCOGS (38,349 cases and 37,818 controls), OncoArray (80,125 cases and 58,383 controls), and additional GWAS studies (14,910 cases and 17,588 controls), resulting in a total of 133,384 breast cancer cases and 113,789 controls (21). The FinnGen Biobank dataset includes 18,786 cases and 182,927 controls (22). In order to ensure the robustness of the study and minimize bias, the datasets used were exclusively derived from populations of European ancestry.

This study utilized publicly available databases, with all participant involvement ethically approved by their respective review boards and informed consent obtained from all subjects in the original studies (Supplementary Table S2).

2.4 Instrumental variable selection

The selection of genetic instruments for pQTL analysis adhered to the following criteria: first, genome-wide significant associations with a p-value of $<5\times10^{-8}$ were established, and SNPs significantly associated with any protein were extracted; an r² value of 0.001 and a distance of 10,000 kb were used to exclude SNPs in linkage disequilibrium (LD) (23, 24). Second, the F statistic ($F = (\frac{R^2}{1-R^2})(\frac{n-k-1}{k}), R^2 = 2 \times (1 - MAF) \times MAF \times (\frac{\beta}{sd})^2$) was calculated to evaluate the strength of the association between SNPs and IVs, with an F > 10 indicating a sufficiently robust association to effectively mitigate bias from weak IVs (25). Third, SNPs within 1 Mb of the transcription start site of genes encoding proteins were classified as cis pQTL, whereas those outside this region were categorized as trans pQTL. Due to the considerable pleiotropy associated with trans pQTL, only cis pQTL were selected as IVs for this study (26).

2.5 Statistical analysis

The preliminary analysis utilized SMR, an extension of the MR concept, to investigate whether the effect sizes of SNPs on phenotypes are mediated by gene expression. This approach prioritizes GWAS hits for genes, facilitating subsequent functional investigations. These methodologies are applicable to various molecular quantitative trait loci data, including DNA methylation quantitative trait loci and pQTL. The present study employed SMR software with default settings via the command line for the analysis. The effect size (β) of variants indicated the direction of protein expression changes (16). The p-values from the results were adjusted using false discovery rate (FDR) correction, with associations possessing a PFDR < 0.05 considered statistically significant (27).
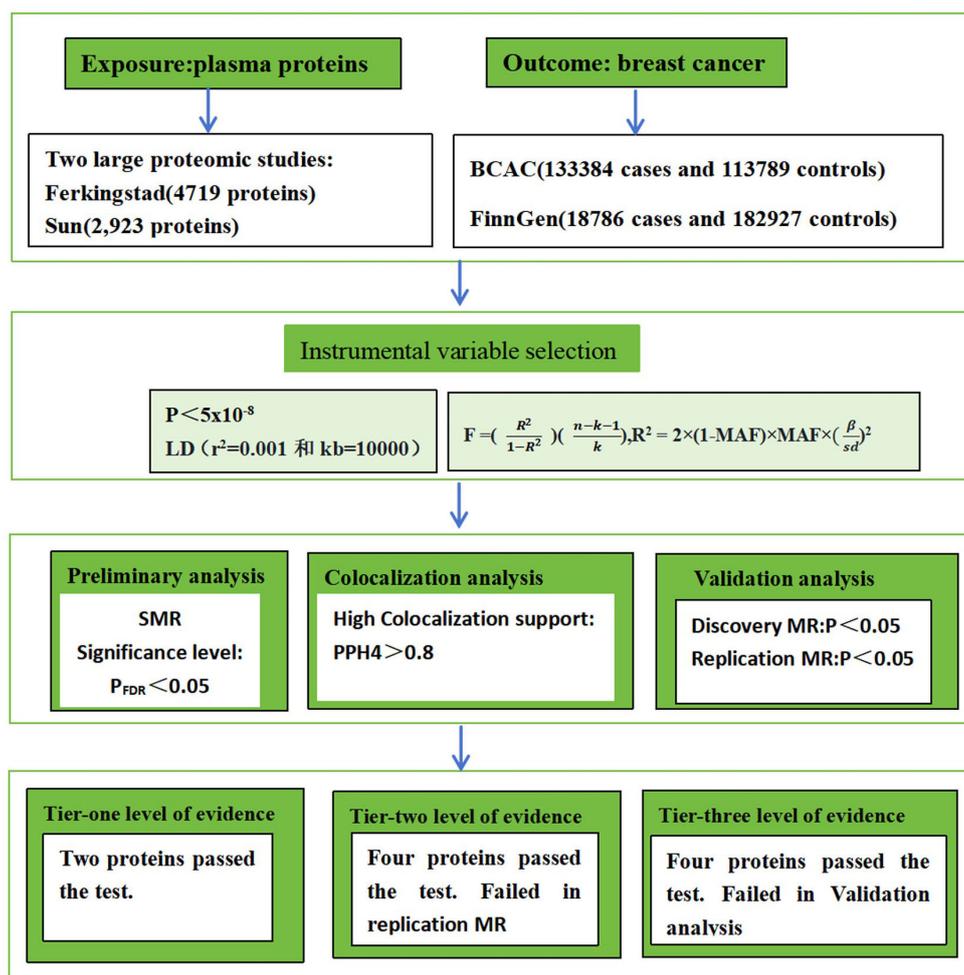
FIGURE 1
The workflow of study design. BCAC stands for Breast Cancer Association Consortium. LD stands for Linkage Disequilibrium. MR stands for Mendelian randomization. SMR stands for Summary-data-based MR.

2.6 Colocalization analysis

In order to optimize outcomes, Bayesian colocalization analysis was used to evaluate whether protein expression and breast cancer are influenced by the same causal variant, thereby discerning confounding effects due to linkage LD. Bayesian colocalization analysis involved five hypotheses: H0, no significant association exists between protein expression and breast cancer with any SNP locus within a genomic region; H1, protein expression is significantly associated with SNP loci within a genomic region; H2, breast cancer is significantly associated with SNP loci within a genomic region; H3, protein expression, and breast cancer are significantly associated with SNP loci within a genomic region, but driven by different causal variants; and H4, protein expression and breast cancer are significantly associated with SNP loci within a genomic region, driven by the same causal variant (28). Colocalization analysis was conducted on all SNPs within ±500 kb of gene start sites, utilizing default parameters: P1 = 1×10^{-4} (prior probability of SNP association with protein), P2 = 1×10^{-4} (prior probability of SNP association with breast cancer), and P12 =

1×10^{-5} (prior probability of SNP association with both protein expression and breast cancer). Posterior probabilities were used to assess the support for each hypothesis, with a posterior probability of PP.H4 > 80% considered compelling evidence of colocalization (17).

2.7 Validation analysis

Validation analysis was conducted using a two-stage (discovery and replication) MR approach. In the Two-Sample MR analysis, the inverse variance weighted (IVW) method (29), the weighted median method (30), and the MR-Egger method (31) were employed as the primary analytical techniques. The IVW method, which has the highest statistical power, assumes the absence of an intercept term and that all genetic variants are valid IVs. On the other hand, the MR-Egger method accounts for the presence of an intercept term, though its testing efficiency may be less precise compared to the IVW method. As a complement to MR-Egger, the weighted median method allows for including some invalid

variants, provided that at least half of them are valid IVs. In both the discovery and replication stages of MR, individual protein-level data (https://www.decode.com) were acquired, and separate two-sample MR analyses for breast cancer were conducted. A P value < 0.05 indicated a statistically significant association.

2.8 Sensitivity analysis

Cochran's Q test was employed to evaluate heterogeneity among genetic variants. If the P-value of Cochran's Q test was < 0.05, a random effects model was used for MR analysis; otherwise, a fixed effects model was applied (32). Additionally, the MR-Egger and MR-PRESSO methods were used to detect the presence of horizontal pleiotropy. By identifying and correcting for pleiotropy, MR-PRESSO can reduce bias caused by pleiotropy and provide more reliable causal effect estimates. Moreover, MR-PRESSO offers correction methods and evaluates the robustness of causal estimates through sensitivity analysis (33). We also used forest plots to assess the causal effect of each SNP and compared these with the causal estimates from the IVW and MR-Egger methods. A leave-one-out analysis was conducted by removing each SNP individually to evaluate whether a single variant drives the association between the exposure and outcome variables (34).

Ultimately, the selected proteins that met the criteria were categorized into three tiers based on the strength of evidence. Tier one included proteins meeting all standards in SMR, discovery MR, replication MR, and colocalization. Tier two included proteins meeting standards in SMR, discovery MR, and colocalization. Tier three included proteins meeting standards in SMR and colocalization.

The SMR tool version 0.1.3 (https://yanglab.westlake.edu.cn/software/SMR/#Overview) was utilized. All MR analyses were performed in R software version 4.3.1, using the Two Sample MR (version 0.5.9), MR-PRESSO (version 1.0), dplyr (version 1.1.3),

circlize (version 0.4.15), stringr (version 1.5.0), ComplexHeatmap (version 2.15.4), and coloc (version 5.2.3) packages.

3 Results

3.1 Preliminary SMR and colocalization results

Following thorough IV processing, 7,981 cis pQTLs were utilized for SMR analysis, detecting significant associations with breast cancer susceptibility across 64 proteins (PFDR < 0.05) (Supplementary Table S3).

Next, colocalization analysis was conducted separately for these proteins in relation to breast cancer. The results indicated that 10 proteins, CPNE1, TNXB, ULK3, CASP8, BTN2A1, PARK7, DNPH1, DDX58, TNFRSF9, and TLR1, had PP.H4 results of > 80% (Table 1; Figure 2).

3.2 Validation analysis

In the discovery phase of MR, CPNE1 exhibited an inverse association with breast cancer risk (IVW: odds ratio (OR) = 0.94, 95% confidence interval (CI): 0.90-0.98, P = 0.005); ULK3 demonstrated a negative correlation with breast cancer risk (IVW: OR = 0.78, 95% CI: 0.65-0.93, P = 0.005); CASP8 revealed a negative correlation with breast cancer risk (IVW: OR = 0.83, 95% CI: 0.77-0.91, P = 2.74E-05); PARK7 manifested a negative correlation with breast cancer risk (IVW: OR = 0.84, 95% CI: 0.74-0.95, P = 0.005); DDX58 indicated an inverse association with breast cancer risk (IVW: OR = 0.84, 95% CI: 0.72-0.98, P = 0.023); and TNFRSF9 exhibited a positive association with breast cancer risk (IVW: OR = 1.18, 95% CI: 1.07-1.30, P =7.10E-04). However, TNXB, BTN2A1,

TABLE 1 The summary of SMR results for the ten proteins that meet colocalization criteria with breast cancer.

Protein	Protein full name	SMR			Colocalization
		Beta	P_{FDR}	OR(95%CI)	PP.H4
CASP8	caspase 8	-0.15	8.34E-05	0.86(0.81-0.91)	0.97
DDX58	DExD/H-box helicase 58 (also known as RIG-I, Retinoic acid-Inducible Gene I)	-0.12	5.06E-04	0.89(0.85-0.93)	0.99
CPNE1	Copine 1	-0.04	5.34E-03	0.96(0.94-0.98)	0.90
ULK3	unc-51 like kinase 3	-0.37	0.01	0.69(0.59-0.81)	0.97
PARK7	Parkinsonism associated deglycase	-0.19	5.06E-04	0.83(0.77-0.90)	0.93
TNFRSF9	TNF receptor superfamily member 9	0.19	0.00	1.21(1.11-1.32)	0.82
TNXB	tenascin XB	0.04	0.04	1.04(1.02-1.07)	0.81
BTN2A1	butyrophilin subfamily 2 member A1	-0.14	3.06E-05	0.87(0.83-0.92)	0.97
DNPH1	2'-deoxynucleoside 5'-phosphate N-hydrolase 1	0.09	0.00	1.10(1.05-1.15)	0.94
TLR1	toll like receptor 1	0.10	2.56E-09	1.10(1.07-1.13)	0.85

SMR, Summary-data-based Mendelian randomization, PP.H4 values were all higher than 0.80 under priors (p12 = 1e−5) and windows (± 500kb).
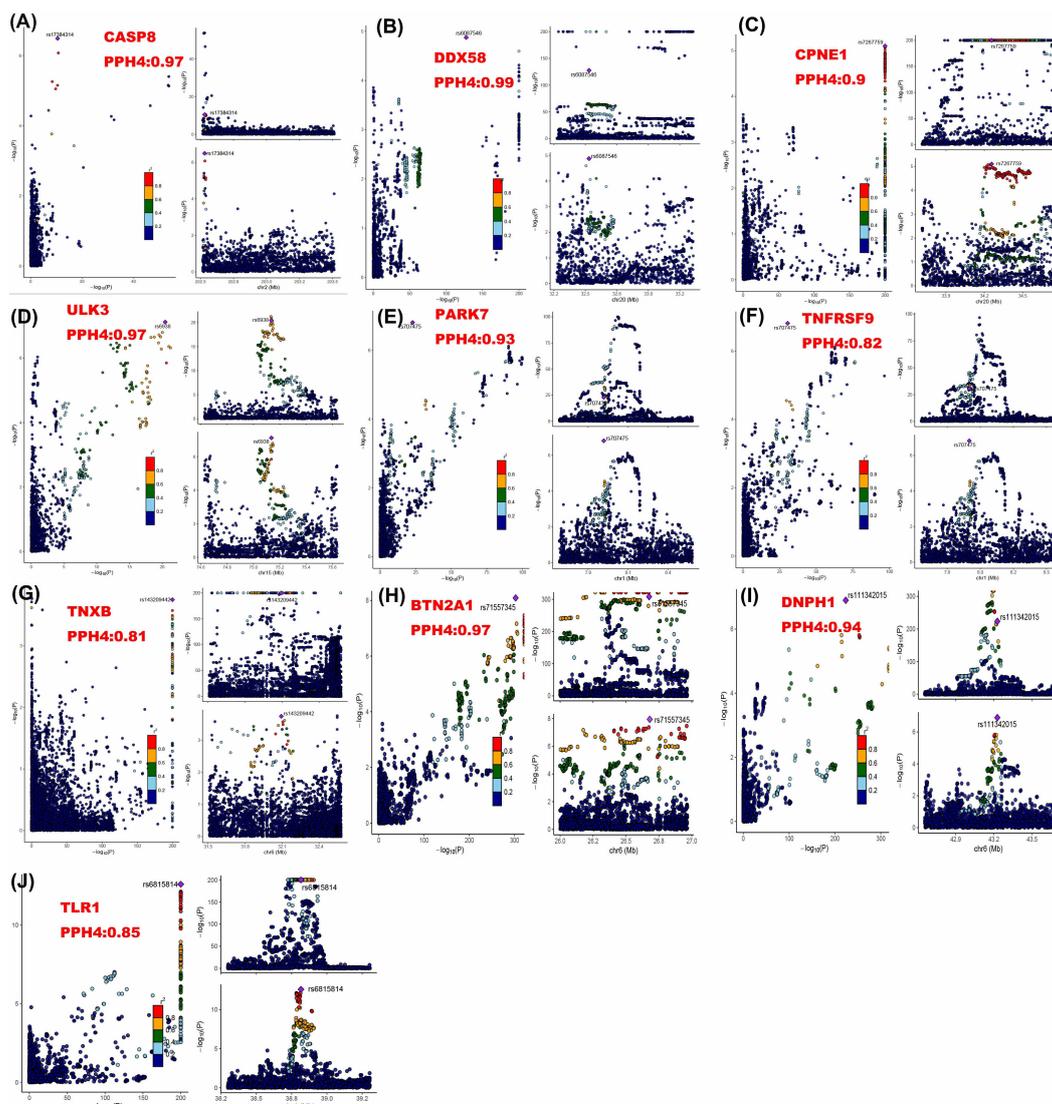
FIGURE 2
Regional association plots for colocalization analysis of ten proteins with breast cancer risk. The lead SNP is shown as a purple diamond. SNPs within ±500 kb of the protein quantitative trait locus were included; p12 = 1e-5, prior probability a SNP is associated with both protein and breast cancer.

DNPH1, and TLR1 had no significant correlation with breast cancer (P > 0.05 for all) (Figure 3; Supplementary Table S4).

In the replication stage of MR, CASP8 continued to show a negative correlation with breast cancer risk (IVW: OR = 0.75, 95% CI: 0.62-0.89, P = 0.001), and DDX58 maintained an inverse association with breast cancer risk (IVW: OR = 0.80, 95% CI: 0.67-0.94, P = 0.007). However, CPNE1, ULK3, PARK7, and TNFRSF9 did not replicate (P > 0.05 for all) (Figure 3; Supplementary Table S4).

3.3 Sensitivity analysis

Both MR-Egger regression and IVW methods detected heterogeneity in CASP8, DNPH1, and ULK3, prompting the use of a random effects model for MR analysis. Meanwhile, MR-Egger

and MR-PRESSO methods detected no horizontal pleiotropy (Figure 3). To ensure the stability of the study results, the symmetrical distribution of SNPs in the funnel plot was confirmed. A leave-one-out analysis was performed to assess the influence of individual SNPs on the results, revealing no significant impact from any single SNP. Additionally, to gain a more comprehensive understanding of the data, scatter plots were generated to illustrate the causal relationships between proteins and breast cancer (Supplementary Figure S1-S6).

3.4 PPI and drug evaluation

The STRING database was used to construct a protein-protein interaction (PPI) network to elucidate the connections among TLR1, CASP8, and DDX58 proteins, as well as the interactions

Exposure	Main.outcome	Method	Pvalue	OR(95%CI)	Heterogeneity.Test.P	Pleio.P	MR.PRESSO..Global.Test.P
CASP8	Breast cancer(BCAC)	IVW(random effect model)	0.00	0.83(0.77 to 0.91)	0.03	0.12	0.4
		MR Egger	0.34	0.93(0.83 to 1.05)	0.24		
		Weighted median	0.00	0.85(0.80 to 0.90)			
CASP8	Breast cancer(FinnGen)	IVW	0.00	0.75(0.62 to 0.89)	0.36	0.63	0.43
		MR Egger	0.20	0.64(0.35 to 1.19)	0.29		
		Weighted median	0.03	0.77(0.61 to 0.98)			
DDX58	Breast cancer(BCAC)	IVW	0.02	0.84(0.72 to 0.98)	0.33	0.95	0.37
		MR Egger	0.56	0.82(0.45 to 1.49)	0.21		
		Weighted median	0.03	0.81(0.66 to 0.98)			
DDX58	Breast cancer(FinnGen)	IVW	0.01	0.80(0.67 to 0.94)	0.75	0.55	0.72
		MR Egger	0.58	0.89(0.61 to 1.31)	0.7		
		Weighted median	0.02	0.78(0.64 to 0.96)			
CPNE1	Breast cancer(BCAC)	IVW	0.01	0.94(0.90 to 0.98)	0.47	0.4	0.51
		MR Egger	0.17	0.96(0.90 to 1.01)	0.45		
		Weighted median	0.13	0.96(0.90 to 1.01)			
CPNE1	Breast cancer(FinnGen)	IVW	0.83	0.99(0.88 to 1.10)	0.22	0.86	0.24
		MR Egger	0.99	1.00(0.84 to 1.20)	0.17		
		Weighted median	0.53	0.96(0.84 to 1.09)			
ULK3	Breast cancer(BCAC)	IVW(random effect model)	0.01	0.78(0.65 to 0.93)	0.05	0.51	0.14
		MR Egger	0.38	0.44(0.11 to 1.83)	0.05		
		Weighted median	0.02	0.85(0.73 to 0.98)			
ULK3	Breast cancer(FinnGen)	IVW	0.08	0.82(0.65 to 1.03)	0.26	0.44	0.33
		MR Egger	0.54	1.94(0.33 to 11.40)	0.26		
		Weighted median	0.08	0.79(0.61 to 1.02)			
PARK7	Breast cancer(BCAC)	IVW	0.01	0.84(0.74 to 0.95)	0.16	0.43	0.22
		MR Egger	0.95	1.02(0.64 to 1.62)	0.15		
		Weighted median	0.03	0.85(0.74 to 0.98)			
PARK7	Breast cancer(FinnGen)	IVW	0.50	1.07(0.89 to 1.22)	0.92	0.89	0.9
		MR Egger	0.74	1.11(0.62 to 2.00)	0.84		
		Weighted median	0.71	1.04(0.83 to 1.31)			
TNFRSF9	Breast cancer(BCAC)	IVW	0.00	1.18(1.07 to 1.30)	0.35	0.44	0.38
		MR Egger	0.52	1.08(0.86 to 1.37)	0.32		
		Weighted median	0.02	1.17(1.03 to 1.33)			
TNFRSF9	Breast cancer(FinnGen)	IVW	0.84	0.98(0.84 to 1.15)	0.87	0.77	0.88
		MR Egger	0.87	1.03(0.73 to 1.45)	0.79		
		Weighted median	0.68	0.96(0.79 to 1.17)			
TNXB	Breast cancer(BCAC)	IVW	0.65	0.99(0.96 to 1.03)	0.18	0.52	0.25
		MR Egger	0.46	0.98(0.94 to 1.03)	0.16		
		Weighted median	0.48	0.99(0.95 to 1.03)			
BTN2A1	Breast cancer(BCAC)	IVW	0.49	0.97(0.90 to 1.05)	0.75	0.26	0.78
		MR Egger	0.20	0.88(0.73 to 1.06)	0.82		
		Weighted median	0.49	0.97(0.88 to 1.06)			
DNPH1	Breast cancer(BCAC)	IVW(random effect model)	0.57	0.95(0.79 to 1.14)	0.02	0.09	0.09
		MR Egger	0.16	1.28(1.03 to 1.60)	0.87		
		Weighted median	0.77	1.02(0.90 to 1.15)			
TLR1	Breast cancer(BCAC)	IVW	0.09	1.11(0.99 to 1.24)	0.72	0.73	0.81
		MR Egger	0.30	1.14(0.95 to 1.37)	0.56		
		Weighted median	0.14	1.10(0.97 to 1.26)			

P<0.05 was considered statistically significant

0   0.5   1   1.5
protective factor   risk factor

FIGURE 3
Validation analysis. The discovery MR is derived from the BCAC database, while the replication MR is derived from the FinnGen database.

involving CASP8 and PARK7 (Figure 4). Drug evaluation, conducted through platforms such as DGIdb4.0 (35) and DrugBank5.0 (36), identified CASP8, DDX58, CPNE1, ULK3, PARK7, and TNFRSF9 as notable targets for pharmacological exploration. Currently, therapeutic agents targeting CASP8 include bardoxolone, under investigation for lymphoma and solid tumor management; bryostatin 1, explored for HIV infection and Alzheimer's disease intervention; AN-9, scrutinized for liver cancer, lung cancer, melanoma, and leukemia; trichostatin A; oleandrin, assessed for lung cancer therapy and chemotherapy-induced adverse effects. Pharmaceutical candidates targeting DDX58 include INARIGIVIR SOPROXIL, which is used as an immunomodulator and antiviral agent. Theophylline is a medication tailored for CPNE1 and used to mitigate symptoms associated with reversible airflow obstruction in conditions such as asthma, chronic obstructive pulmonary disease, and other pulmonary ailments. Fostamatinib, a spleen tyrosine kinase inhibitor, targets ULK3, providing therapeutic relief for chronic immune thrombocytopenia following alternative interventions. Therapeutic strategies aimed at PARK7 include copper, a transition metal present in various supplements and vitamins, and intravenous infusion solutions used in total parenteral nutrition. TNFRSF9 is targeted by Urelumab, currently under investigation for its efficacy against leukemia, multiple myeloma, malignant tumors, solid tumors, and B-cell non-Hodgkin's lymphoma (Supplementary Table S5).

4 Discussion

The present study used publicly available large-sample pQTL and GWAS databases to analyze the causal relationship between 2,385 proteins and breast cancer, identifying 10 proteins associated with breast cancer risk. Among these, decreased levels of CASP8, DDX58, CPNE1, ULK3, PARK7, and BTN2A1 were found, alongside increased levels of TNFRSF9, TNXB, DNPH1, and TLR1. CASP8 and DDX58 exhibit the highest evidence strength, while CPNE1, ULK3, PARK7, and TNFRSF9 show secondary evidence strength. TNXB, BTN2A1, DNPH1, and TLR1 demonstrated the strength of the tertiary evidence. Druggability assessments prioritized six protein biomarkers already developed as drug targets for various chronic diseases or cancers, suggesting their potential repurposing as therapeutic targets for breast cancer.

FIGURE 4
The Protein-protein interaction (PPI) network of proteins identified by proteome-wide Mendelian randomization. Lines represent interactions between proteins. Green line indicates gene neighborhood and predicted interaction; Blue line indicates known interaction from curated databases; Fuchsia line indicates known interaction that is experimentally determined. Black line indicates co-expression. Data information was from STRING database.

Our study identified several candidate proteins previously linked to breast cancer, including CASP8, DDX58, CPNE1, PARK7, BTN2A1, TNFRSF9, TNXB, DNPH1, and TLR1, with CASP8 and DDX58 supported by the most robust evidence. CASP8, also known as caspase 8, is a critical initiator enzyme in the apoptosis pathway, which is vital in regulating programmed cell death (37, 38). Beyond its apoptotic function, CASP8 influences various cellular signaling pathways involved in inflammatory and immune responses (39). In oncology, CASP8 is recognized as a significant tumor suppressor gene, with aberrant expression or dysfunction linked to the onset, invasion, and metastasis. Experimental evidence indicates that CASP8 induces PD-L1 degradation by upregulating TNFAIP3 (A20) expression, and reduced CASP8 expression may predict sensitivity to anti-PD-L1/PD-1 immunotherapy (40). Previous studies have also suggested that the CASP8 D302H polymorphism decreases breast cancer risk associated with *BRCA1* and *BRCA2* mutations, delaying cancer onset (41). A preliminary study on Iranian breast cancer patients further reported significantly decreased CASP8 expression (42), which is consistent with our findings. Our preliminary SMR analysis and validation MR analysis both supported the protective role of elevated CASP8 levels against breast cancer risk (Table 1; Figure 3). Notably, a meta-analysis further substantiated the role of CASP8 in cancer susceptibility. This study assessed the association between CASP8 rs3834129 and rs1045485 polymorphisms with the risk of breast cancer and other malignancies, revealing that these polymorphisms significantly reduced the risk of breast cancer and several other cancers, particularly in Asian and Caucasian populations (43). These findings provide compelling evidence

supporting the protective role of elevated CASP8 protein levels against breast cancer risk, underscoring the importance of CASP8 as a tumor suppressor gene in breast cancer. Additionally, the potential of CASP8 as a therapeutic target in other cancers, such as liver cancer, lung cancer, melanoma, and leukemia, has been explored, highlighting its broader applicability in cancer treatment (Supplementary Table S5). These studies emphasize the relevance of CASP8 in drug development, bolstering the case for considering CASP8 as a potential target in breast cancer therapy. While our results strongly support the protective effect of elevated CASP8 protein levels against breast cancer risk, the literature presents contrasting findings. For instance, a prospective observational study indicated that increased CASP8 levels might be associated with poorer prognosis in patients with metastatic breast cancer (44). This discrepancy could be due to differences in study populations, such as the distinction between metastatic patients and those with early-stage breast cancer, which may involve significant variations in disease progression and immune response. Alternatively, it may stem from differing research methodologies or analytical strategies. These divergences further underscore the importance of exploring the role of CASP8 across different cancer stages and subtypes.

DDX58, also known as RIG-I, is a critical intracellular pattern recognition receptor pivotal in immune responses. Cao et al. have shown that deficiencies in RIG-I contribute to chemotherapy resistance in triple-negative breast cancer by impeding apoptosis mediated through type I IFN signaling. They also found that patients with diminished DDX58 expression have lower rates of achieving pathological complete response and exhibit poorer prognosis (45). Additionally, studies focusing on innate immune strategies for activating breast cancer cells and the tumor microenvironment have shown that RIG-I activation within breast tumors enhances tumor-infiltrating lymphocytes while diminishing tumor growth and metastasis (46). These findings underscore the robust immunogenicity and therapeutic potential of RIG-I agonists when delivered to tumors, particularly in the context of less immunogenic breast cancers (46). Consistent with these observations, a previous study demonstrated that the active metabolite of tamoxifen (TAM), 4-hydroxytamoxifen (4-OH-TAM), regulates the expression of multiple genes, including the upregulation of DDX58 in estrogen receptor-positive (ER+) breast cancer MCF-7 cells. This research revealed that DDX58 and other genes were upregulated following 4-OH-TAM treatment, underscoring its role in both estrogen receptor-dependent and independent pathways (47). Our study indicates that lower levels of DDX58 protein were associated with an increased risk of breast cancer, which is consistent with previous foundational research (Table 1; Figure 3). Herein, we provided robust genetic evidence supporting the protective role of DDX58 against breast cancer risk. Currently, DDX58 is under investigation for its potential use as an immune modulator and antiviral agent, indicating its promise as a novel therapeutic target for breast cancer (Supplementary Table S5).

In the present study, CPNE1, PARK7, and TNFRSF9 were supported by secondary evidence strength. CPNE1 (Copine-1) is a calcium-binding protein with crucial roles in cellular signal transduction, adhesion, and apoptosis (48). Our research indicated

that decreased circulating levels of CPNE1 are associated with an increased risk of breast cancer, which is consistent with findings by Ren et al. (14). However, multiple studies have also shown that CPNE1 promotes aerobic glycolysis and metastasis in triple-negative breast cancer (TNBC) through the PI3K/AKT/HIF-1α signaling pathway, thereby accelerating tumor progression (49). Additionally, other research has demonstrated that CPNE1 is overexpressed in TNBC tissues and cell lines, closely associated with tumor size, distant metastasis, and the survival rates of TNBC patients. CPNE1 also promotes tumorigenesis and radioresistance in TNBC cells by activating the AKT signaling pathway (50). Although most foundational studies suggest that CPNE1 has a pro-tumor role in cancer progression, our study and Ren's research, employing MR, provide robust evidence from a causal perspective that CPNE1 may have a protective role in breast cancer. Our findings indicated a negative association between CPNE1 and breast cancer risk (OR: 0.94, 95% CI: 0.90-0.98), and Ren's study yielded similar results (OR: 0.96, 95% CI: 0.94-0.98). This discovery suggests that a reduction in CPNE1 levels may increase the risk of breast cancer, which contradicts the tumor-promoting role of CPNE1 supported by conventional basic research. Therefore, while existing research predominantly focuses on the oncogenic role of CPNE1 in cancer, our and Ren's study provide causal evidence through MR analysis, revealing the potential protective function of CPNE1. This causal insight offers a new perspective on CPNE1 as a potential therapeutic target in breast cancer and suggests that future research should further explore the dual mechanisms of CPNE1 to gain a more comprehensive understanding of its role in breast cancer progression.

PARK7, also known as DJ-1 protein, exhibits findings similar to those of CPNE1. PARK7 is widely expressed intracellularly and is involved in regulating cellular responses to oxidative stress, protecting mitochondrial function, maintaining cellular redox balance, and inhibiting apoptosis (51). The present study suggests that decreased circulating levels of PARK7 are associated with an increased risk of breast cancer (Figure 3), which is in line with findings by Wang et al. (52). In their retrospective study, Tsuchiya and colleagues demonstrated that DJ-1 protein expression in invasive ductal carcinoma (IDC) tissues was lower than in adjacent non-cancerous epithelial tissues despite higher mRNA levels. Among IDC patients, lower DJ-1 protein expression was significantly associated with shorter disease-free survival (P = 0.015) and overall survival (P = 0.020) (53). However, an observational study indicated that DJ-1 is upregulated in HR+ breast cancer and significantly correlates with poor prognosis (54). These discrepancies may stem from differences in the breast cancer molecular subtypes used in our analysis compared to traditional epidemiological studies, or they may underscore limitations in adjusting for confounding factors and reverse causation in traditional epidemiological research. In summary, the relationship between PARK7 and breast cancer risk remains inconclusive. PARK7 is supported by secondary evidence strength in our study, suggesting its potential as a therapeutic target for breast cancer. However, further experimental studies are needed to clarify the directionality of the associations between PARK7 and breast cancer.

TNFRSF9, also known as 4-1BB, is a protein that belongs to the tumor necrosis factor receptor superfamily and has a critical role in immune regulation, particularly in the activation and proliferation of T cells. When the 4-1BB receptor binds with its ligand, 4-1BBL, it triggers various signaling pathways, including AKT, NF-kB, and MAPK, promoting T cell proliferation, survival, and function (55). Currently, monoclonal antibodies targeting 4-1BB, such as urelumab and utomilumab, have been used in the treatment of B-cell non-Hodgkin lymphoma, lung cancer, breast cancer, soft tissue sarcoma, and other solid tumors (56). Our findings suggest that elevated levels of TNFRSF9 protein increase the risk of breast cancer (Figure 3), supported by secondary evidence strength. Our research enhances the genetic evidence linking TNFRSF9 elevation to an increased risk of breast cancer. Moreover, in their study, Harao et al. demonstrated that 4-1BB-enhanced expansion of CD8+ TILs can significantly promote the growth of these T cells within TNBC tumors. This approach can be used to identify immunogenic mutations within autologous TNBC tumor tissues. These findings underscore the potential application of 4-1BB in immunotherapy and offer a novel perspective on adoptive immunotherapy for TNBC (57). Our research further corroborates the role of TNFRSF9 in breast cancer and provides new genetic evidence supporting its potential as a therapeutic target.

ULK3 has emerged from our study as a novel prognostic biomarker for breast cancer. ULK3, or Unc-51 like autophagy activating kinase 3, primarily regulates the autophagy pathway within cells (58). Previous studies have identified the upregulation of ULK3 in squamous cell carcinoma of the skin and head and neck (59), and its potential as a prognostic biomarker in colon cancer (60, 61). However, there are currently no basic experimental research reports on the association between ULK3 and breast cancer. A preliminary SMR analysis conducted in the present study indicated a negative correlation between ULK3 and the risk of breast cancer (Table 1). This finding was confirmed by validation MR analyses (Figure 3). Notably, fostamatinib, a spleen tyrosine kinase inhibitor used for chronic immune thrombocytopenia after other treatments, targets ULK3 (Supplementary Table S5), which shows promise as a new therapeutic target for breast cancer treatment and as a prognostic biomarker.

In the present study, preliminary SMR analysis suggested potential causal relationships between TNXB, BTN2A1, DNPH1, CGA, TLR1, and breast cancer; however, these associations were not validated in the MR analysis, resulting in only tertiary evidence support, which indicates insufficient strength of evidence. Further research is necessary to confirm their associations with breast cancer.

This study has several strengths. Firstly, we systematically explored the relationship between plasma protein levels and breast cancer risk using a two-stage proteome-wide MR design. This approach benefits from a large sample size and comprehensive coverage. It also mitigates the risks of reverse causation and confounding biases. Secondly, our study included preliminary and validation studies, encompassing discovery and replication MR analyses, thereby providing robust evidence for our findings. Thirdly, we employed colocalization methods to minimize false positives arising from LD and horizontal pleiotropy. Additionally, our study samples were drawn from European populations, reducing potential biases related to racial differences in research outcomes. Fourthly, PPI and druggability assessments offer insights into the potential pathogenic roles of candidate proteins in breast cancer, aiding in the prioritization of druggable targets. Notably,

proteins such as CASP8 and DDX58, which are already targeted for other diseases, exhibit promising potential as novel therapeutic targets for breast cancer.

However, the present study also has several limitations. Firstly, we did not investigate the relationship between circulating proteins and specific breast cancer subtypes due to the absence of cross-validated databases. This limitation underscores the necessity for future research to thoroughly explore the roles of these proteins across different breast cancer subtypes. Secondly, our study samples were exclusively from European populations, potentially limiting the generalizability of our findings to other ethnic groups. Further studies are needed to determine the applicability of these results across diverse racial populations. Thirdly, while we identified causal associations between relevant proteins and breast cancer, we could not conduct comprehensive biological experiments due to financial constraints. Future research could incorporate animal models and cell line experiments by addressing this gap to provide more robust evidence supporting our findings.

5 Conclusion

Using the MR combined colocalization method, we identified several plasma proteins associated with breast cancer risk, notably CASP8 and DDX58, which are promising targets for developing screening biomarkers and therapeutic drugs for breast cancer. Based on our findings, future experimental and clinical studies are essential to assess the efficacy and validate the potential of these candidate drugs.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

JS: Data curation, Formal analysis, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. HY: Conceptualization, Funding acquisition, Project administration, Supervision, Validation, Writing – original draft, Writing – review & editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fendo.2024.1449668/full#supplementary-material

References

1. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. *CA Cancer J Clin*. (2023) 73:17–48. doi: 10.3322/caac.21763

2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. (2021) 71:209–49. doi: 10.3322/caac.21660

3. McDonald ES, Clark AS, Tchou J, Zhang P, Freedman GM. Clinical diagnosis and management of breast cancer. *J Nucl Med*. (2016) 57 Suppl 1:9S–16S. doi: 10.2967/jnumed.115.157834

4. Hanash SM, Pitteri SJ, Faca VM. Mining the plasma proteome for cancer biomarkers. *Nature*. (2008) 452:571–9. doi: 10.1038/nature06916

5. Peila R, Rohan TE. Circulating levels of biomarkers and risk of ductal carcinoma in *situ* of the breast in the UK Biobank study. *Int J Cancer*. (2024) 154:1191–203. doi: 10.1002/ijc.34795

6. Thomas CE, Dahl L, Byström S, Chen Y, Uhlén M, Mälarstig A, et al. Circulating proteins reveal prior use of menopausal hormonal therapy and increased risk of breast cancer. *Transl Oncol*. (2022) 17:101339. doi: 10.1016/j.tranon.2022.101339

7. Byström S, Eklund M, Hong MG, Fredolini C, Eriksson M, Czene K, et al. Affinity proteomic profiling of plasma for proteins associated to area-based mammographic breast density. *Breast Cancer Res*. (2018) 20:14. doi: 10.1186/s13058-018-0940-z

8. Monson KR, Goldberg M, Wu HC, Santella RM, Chung WK, Terry MB. Circulating growth factor concentrations and breast cancer risk: a nested case-control study of IGF-1, IGFBP-3, and breast cancer in a family-based cohort. *Breast Cancer Res*. (2020) 22:109. doi: 10.1186/s13058-020-01352-0

9. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet*. (2014) 23:R89–98. doi: 10.1093/hmg/ddu328

10. Jia G, Yang Y, Ping J, Xu S, Liu L, Guo X, et al. Identification of target proteins for breast cancer genetic risk loci and blood risk biomarkers in a large study by integrating genomic and proteomic data. *Int J Cancer*. (2023) 152:2314–20. doi: 10.1002/ijc.34472

11. Shu X, Bao J, Wu L, Long J, Shu XO, Guo X, et al. Evaluation of associations between genetically predicted circulating protein biomarkers and breast cancer risk. *Int J Cancer*. (2020) 146:2130–8. doi: 10.1002/ijc.32542

12. Shu X, Zhou Q, Sun X, Flesaker M, Guo X, Long J, et al. Associations between circulating proteins and risk of breast cancer by intrinsic subtypes: a Mendelian randomisation analysis. *Br J Cancer*. (2022) 127:1507–14. doi: 10.1038/s41416-022-01923-2

13. Mälarstig A, Grassmann F, Dahl L, Dimitriou M, McLeod D, Gabrielson M, et al. Evaluation of circulating plasma proteins in breast cancer using Mendelian randomisation. *Nat Commun*. (2023) 14:7680. doi: 10.1038/s41467-023-43485-8

14. Ren F, Jin Q, Liu T, Ren X, Zhan Y. Proteome-wide mendelian randomization study implicates therapeutic targets in common cancers. *J Trans Med*. (2023) 21:646. doi: 10.1186/s12967-023-04525-5

15. Sun J, Luo J, Jiang F, Zhao J, Zhou S, Wang L, et al. Exploring the cross-cancer effect of circulating proteins and discovering potential intervention targets for 13 site-specific cancers. *J Natl Cancer Inst*. (2024) 116:565–73. doi: 10.1093/jnci/djad247

16. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet*. (2016) 48:481–7. doi: 10.1038/ng.3538

17. Giambartolomei C, Vukcevic D, SChadt EE, Franke L, Hingorani AD, Wallace C, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet*. (2014) 10:e1004383. doi: 10.1371/journal.pgen.1004383

18. Cuschieri S. The STROBE guidelines. *Saudi J Anaesth*. (2019) 13:S31–4. doi: 10.4103/sja.SJA_543_18

19. Ferkingstad E, Sulem P, Atlason BA, Sveinbjornsson G, Magnusson MI, Styrmisdottir EL, et al. Large-scale integration of the plasma proteome with genetics and disease. *Nat Genet*. (2021) 53:1712–21. doi: 10.1038/s41588-021-00978-w

20. Sun BB, Chiou J, Traylor M, Benner C, Hsu YH, Richardson TG, et al. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature*. (2023) 622:329–38. doi: 10.1038/s41586-023-06592-6

21. Zhang H, Ahearn TU, Lecarpentier J, Barnes D, Beesley J, Qi G, et al. Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and subtype-specific analyses. *Nat Genet*. (2020) 52:572–81. doi: 10.1038/s41588-020-0609-2

22. Kurki MI, Karjalainen J, Palta P, Sipilä TP, Kristiansson K, Donner K, et al. FinnGen: Unique genetic insights from combining isolated population and national health register data. *medRxiv*. (2022) 03:1–56. doi: 10.1101/2022.03.03.22271360

23. Luo J, Le Cessie S, Blauw GJ, Franceschi C, Noordam R, van Heemst D. Systemic inflammatory markers in relation to cognitive function and measures of brain atrophy: a Mendelian randomization study. *GeroScience*. (2022) 44:2259–70. doi: 10.1007/s11357-022-00602-7

24. Burgess S, Davey Smith G, Davies NM, Dudbridge F, Gill D, Glymour MM, et al. Guidelines for performing Mendelian randomization investigations: update for summer 2023. *Wellcome Open Res*. (2023) 4:186. doi: 10.12688/wellcomeopenres.15555.3

25. Burgess S, Thompson SG; CRP CHD Genetics Collaboration. Avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol*. (2011) 40:755–64. doi: 10.1093/ije/dyr036

26. Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, et al. Genomic atlas of the human plasma proteome. *Nature*. (2018) 558:73–9. doi: 10.1038/s41586-018-0175-2

27. Shi X, Wei T, Hu Y, Wang M, Tang Y. The associations between plasma soluble Trem1 and neurological diseases: a Mendelian randomization study. *J Neuroinflamm*. (2022) 19:218. doi: 10.1186/s12974-022-02582-z

28. Burgess S, Bowden J, Fall T, Ingelsson E, Thompson SG. Sensitivity analyses for robust causal inference from Mendelian randomization analyses with multiple genetic variants. *Epidemiology*. (2017) 28:30–42. doi: 10.1097/EDE.0000000000000559

29. Sun J, Zhao J, Jiang F, Wang L, Xiao Q, Han F, et al. Identification of novel protein biomarkers and drug targets for colorectal cancer by integrating human plasma proteome with genome. *Genome Med*. (2023) 15:75. doi: 10.1186/s13073-023-01229-9

30. Slob EAW, Burgess S. A comparison of robust Mendelian randomization methods using summary data. *Genet Epidemiol*. (2020) 44:313–29. doi: 10.1002/gepi.22295

31. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol*. (2016) 40:304–14. doi: 10.1002/gepi.21965

32. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol*. (2017) 32:377–89. doi: 10.1007/s10654-017-0255-x

33. Greco M FD, Minelli C, Sheehan NA, Thompson JR. Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. *Stat Med*. (2015) 34:2926–40. doi: 10.1002/sim.6522

34. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet*. (2018) 50:693–8. doi: 10.1038/s41588-018-0099-7

35. Freshour SL, Kiwala S, Cotto KC, Coffman AC, McMichael JF, Song JJ, et al. Integration of the Drug-Gene Interaction Database (DGIdb 4.0) with open crowdsource efforts. *Nucleic Acids Res*. (2021) 49:D1144–51. doi: 10.1093/nar/gkaa1084

36. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res*. (2018) 46: D1074–82. doi: 10.1093/nar/gkx1037

37. Fritsch M, Günther SD, Schwarzer R, Albert MC, Schorn F, Werthenbach JP, et al. Caspase-8 is the molecular switch for apoptosis, necroptosis and pyroptosis. *Nature*. (2019) 575:683–7. doi: 10.1038/s41586-019-1770-6

38. Newton K, Wickliffe KE, Maltzman A, Dugger DL, Reja R, Zhang Y, et al. Activity of caspase-8 determines plasticity between cell death pathways. *Nature*. (2019) 575:679–82. doi: 10.1038/s41586-019-1752-8

39. Lehle AS, Farin HF, Marquardt B, Michels BE, Magg T, Li Y, et al. Intestinal inflammation and dysregulated immunity in patients with inherited caspase-8 deficiency. *Gastroenterology*. (2019) 156:275–8. doi: 10.1053/j.gastro.2018.09.041

40. Zou J, Xia H, Zhang C, Xu H, Tang Q, Zhu G, et al. Casp8 acts through A20 to inhibit PD-L1 expression: The mechanism and its implication in immunotherapy. *Cancer Sci*. (2021) 112:2664–78. doi: 10.1111/cas.14932

41. Palanca Suela S, Esteban Cardeñosa E, Barragán González E, de Juan Jiménez I, Chirivella González I, Segura Huerta A, et al. CASP8 D302H polymorphism delays the age of onset of breast cancer in BRCA1 and BRCA2 carriers. *Breast Cancer Res Treat*. (2010) 119:87–93. doi: 10.1007/s10549-009-0316-2

42. Aghababazadeh M, Dorraki N, Javan FA, Fattahi AS, Gharib M, Pasdar A. Downregulation of Caspase 8 in a group of Iranian breast cancer patients - A pilot study. *J Egypt Natl Canc Inst*. (2017) 29:191–5. doi: 10.1016/j.jnci.2017.10.001

43. Hashemi M, Aftabi S, Moazeni-Roodi A, Sarani H, Wiechec E, Ghavami S. Association of CASP8 polymorphisms and cancer susceptibility: A meta-analysis. *Eur J Pharmacol*. (2020) 881:173201. doi: 10.1016/j.ejphar.2020.173201

44. Gunnarsdottir FB, Bendahl PO, Johansson A, Benfeitas R, Rydén L, Bergenfelz C, et al. Serum immuno-oncology markers carry independent prognostic information in patients with newly diagnosed metastatic breast cancer, from a prospective observational study. *Breast Cancer Res*. (2023) 25:29. doi: 10.1186/s13058-023-01631-6

45. Cao S, Long X, Xiao L, Zhang P, Shen M, Chen F, et al. DDX58 deficiency leads to triple negative breast cancer chemotherapy resistance by inhibiting Type I IFN-mediated signalling apoptosis. *Front Oncol*. (2024) 14:1356778. doi: 10.3389/fonc.2024.1356778

46. Elion DL, Jacobson ME, Hicks DJ, Rahman B, Sanchez V, Gonzales-Ericsson PI, et al. Therapeutically active RIG-I agonist induces immunogenic tumor cell killing in breast cancers. *Cancer Res*. (2018) 78:6183–95. doi: 10.1158/0008-5472.CAN-18-0730

47. Fang Q, Yao S, Luo G, Zhang X. Identification of differentially expressed genes in human breast cancer cells induced by 4-hydroxyltamoxifen and elucidation of their pathophysiological relevance and mechanisms. *Oncotarget*. (2017) 9:2475–501. doi: 10.18632/oncotarget.23504

48. Tang H, Zhu J, Du W, Liu S, Zeng Y, Ding Z, et al. CPNE1 is a target of miR-335-5p and plays an important role in the pathogenesis of non-small cell lung cancer. *J Exp Clin Cancer Res*. (2018) 37:131. doi: 10.1186/s13046-018-0811-6

49. Cao J, Cao R, Liu Y, Dai T. CPNE1 mediates glycolysis and metastasis of breast cancer through activation of PI3K/AKT/HIF-1α signaling. *Pathology Res practice*. (2023) 248:154634–4. doi: 10.1016/j.prp.2023.154634

50. Shao Z, Ma X, Zhang Y, Sun Y, Lv W, He K, et al. CPNE1 predicts poor prognosis and promotes tumorigenesis and radioresistance via the AKT singling pathway in triple-negative breast cancer. *Mol Carcinog*. (2020) 59:533–44. doi: 10.1002/mc.23177

51. Kim RH, Peters M, Jang Y, Shi W, Pintilie M, Fletcher GC, et al. DJ-1, a novel regulator of the tumor suppressor PTEN. *Cancer Cell*. (2005) 7:263–73. doi: 10.1016/j.ccr.2005.02.010

52. Wang Y, Yi K, Chen B, Zhang B, Gao J. Elucidating the susceptibility to breast cancer: an in-depth proteomic and transcriptomic investigation into novel potential plasma protein biomarkers. *Front Mol Biosci*. (2024) 10. doi: 10.3389/fmolb.2023.1340917

53. Tsuchiya B, Iwaya K, Kohno N, Kawate T, Akahoshi T, Matsubara O, et al. Clinical significance of DJ-1 as a secretory molecule: retrospective study of DJ-1 expression at mRNA and protein levels in ductal carcinoma of the breast. *Histopathology*. (2012) 61:69–77. doi: 10.1111/j.1365-2559.2012.04202.x

54. Xie Y, Li Y, Yang M. DJ-1: A potential biomarker related to prognosis, chemoresistance, and expression of microenvironmental chemokine in HR-positive breast cancer. *J Immunol Res*. (2023) 2023:1–15. doi: 10.1155/2023/5041223

55. Sanmamed MF, Etxeberría I, Otano I, Melero I. Twists and turns to translating 4-1BB cancer immunotherapy. *Sci Trans Med*. (2019) 11:eaax4738. doi: 10.1126/scitranslmed.aax4738

56. Shen X, Zhang R, Nie X, Yang Y, Hua Y, Peng Lü. 4-1BB targeting immunotherapy: mechanism, antibodies, and chimeric antigen receptor T. *Cancer biotherapy radiopharmaceuticals*. (2023) 38:431–44. doi: 10.1089/cbr.2023.0022

57. Harao M, Forget MA, Roszik J, Gao H, Babiera GV, Krishnamurthy S, et al. 4-1BB-enhanced expansion of CD8+ TIL from triple-negative breast cancer unveils mutation-specific CD8+ T cells. *Cancer Immunol Res*. (2017) 5:439–45. doi: 10.1158/2326-6066.CIR-16-0364

58. Young ARJ, Narita M, Ferreira M, Kirschner K, Sadaie M, Darot JF, et al. Autophagy mediates the mitotic senescence transition. *Genes Dev*. (2009) 23:798–803. doi: 10.1101/gad.519709

59. Goruppi S, Clocchiatti A, Bottoni G, Di Cicco E, Ma M, Tassone B, et al. he ULK3 kinase is a determinant of keratinocyte self-renewal and tumorigenesis targeting the arginine methylome. *Nat Commun*. (2023) 14:887. doi: 10.1038/s41467-023-36410-6

60. Xu J, Dai S, Yuan Y, Xiao Q, Ding K. A prognostic model for colon cancer patients based on eight signature autophagy genes. *Front Cell Dev Biol*. (2020) 8. doi: 10.3389/fcell.2020.602174

61. Wang Y, Lin K, Xu T, Wang L, Fu L, Zhang G, et al. Development and validation of prognostic model based on the analysis of autophagy-related genes in colon cancer. *Aging*. (2021) 13:19028–47. doi: 10.18632/aging.v13i14