Current Science
CS 5 (5), 4334-4349 (2025)
Received 12 June 2025 | **Revised 15 July** 2025 | **Accepted 22 August** 2025 | **Online Available 2 September** 2025
https://doi.org/10.52845/CS/2025-5-5-2

Current Science

Current Science

OPEN ACCESS

ISSN (O) 2795-8639

Original Article

Exploring Bidirectional Causal Relationships between Antibody Responses to Infectious Agents and Breast Cancer through Mendelian Randomization and Meta-Analyses

Yan Zhang[1], Changzhou Feng[2]*

[1]Department of Oncology, Shanghai Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China

[2]Department of Clinical Laboratory, The First People's Hospital of Lianyungang, The Affiliated Lianyungang Hospital of Xuzhou Medical University, The First Affiliated Hospital of Kangda College of Nanjing Medical University, Lianyungang Jiangsu, China

[*]Corresponding Author: Changzhou Feng

Abstract:
Utilizing a GWAS dataset for 46 antibody-mediated immune responses (AMIRs) defined by 13 infectious pathogens and four distinct breast cancer (BrCa) datasets, we applied bidirectional Mendelian randomization (MR) analysis to assess their causalities, supplemented by meta-analysis to resolve inconsistency. Forward MR demonstrated that Anti-human herpes virus (HHV) 6 IgG seropositivity (IVW: OR=1.01; 95%CI, 1.00–1.02), HHV-6 IE1A antibody levels (IVW: OR=1.04; 95%CI, 1.02–1.06) and Helicobacter pylori (Hp) VacA antibody levels (IVW: OR=1.02; 95%CI, 1.00–1.04) were positively associated with BrCa risk. In reverse MR, BrCa was positively associated with Polyomavirus 2 JC VP1 (JCPyV) antibody levels (IVW: OR=1.07; 95%CI, 1.04–1.11), and negatively with Anti-Hp IgG seropositivity (IVW: OR=0.93; 95%CI, 0.88–0.97) and Anti-varicella zoster virus (VZV) IgG seropositivity (IVW: OR=0.84; 95%CI, 0.79–0.90). Summary-data-based MR (SMR) analyses were performed to identify potential mediators. Functional analysis revealed the genes mapped to BrCa were enriched in microbial infection and immune-related pathways. These findings suggest HHV-6 and *Hp* are risk factors for BrCa and BrCa may increase susceptibility to JCPyV infection, which highlights the importance of preventing breast cancer by addressing pathogenic infections early and maintaining healthy microbiota.

Keywords: Mendelian randomization (MR); Infectious pathogens; Breast cancer (BrCa); Antibody-mediated immune responses (AMIRs)

Introduction

Breast cancer (BrCa), the most commonly diagnosed cancer worldwide in women, is a multifactorial disease with both genetic and modifiable risk factors[1]. Despite significant advances in diagnostics and therapeutics, BrCa-related mortality remains high globally. Thus, alongside the investigation of modifiable risk factors, the search for improved strategies for disease management continues[2]. In addition to well-established risk factors such as hormone-replacement therapy, alcohol intake, obesity, and physical inactivity[3], the potential causal relationships between infectious agents and BrCa have gained considerable attention in recent years[2, 4], particularly after studies a decade ago first revealed that the microbiome differs markedly between healthy and cancerous breast tissues[2, 5].

To date, previous studies have revealed few consistent patterns of perturbed microbial taxa. For example, Thompson *et al* found *E. coli* to be more abundant in normal breast adjacent tissue[6];

while Urbaniak *et al* reported higher levels of *E. coli* in breast cancerous tissue[7]. In addition, several infectious agents, including human mammary tumor virus[8], mouse mammary tumor virus (MMTV), *Chlamydia trachomatis* (*C. trachomatis*)[9], bovine leukemia virus[10], *Helicobacter pylori* (*Hp*)[11], Epstein-Barr virus (EBV)[8], human herpesvirus type-8[12], simian vacuolating virus 40[13], cytomegalovirus (CMV)[14], human papillomavirus (HPV) have been considered potential breast oncogenic factors[15]. However, other studies have demonstrated no etiological correlation between HPV and BrCa[16]. Similar inconsistence has been reported by Maryam Kadivar *et al* regarding the absence of EBV in BrCa[17]. Regarding human herpesvirus 6 (HHV-6), Annie Gravel *et al* found no evidence of an association with BrCa risk[18, 19]. In summary, the causal relationships between pathogenic infections and BrCa from observational studies remain currently divergent and inconclusive.

To some extent, this dispute is attributed to the extensive variation in study methods, sample sizes, and quality of evidence[20]. Additionally, the complexity of the causal association is further compounded by unmeasured confounding factors and potential reverse causation inherent in observational studies. The most notable confounder is that the presence of a perturbed microbiome in breast tumors is not proof of causation in itself [4, 21]. In contrast, the growth of particular microbial species could result from the unique nutrient supply or immunosuppression within the tumor environment[21]. Therefore, alternative methods are need to strengthen the causal inference between pathogenic infections and BrCa.

Mendelian randomization (MR) is a promising alternative method for inferring the causal relationship between an exposure and an outcome by using genetic variants as instrumental variables[22]. Since genetic variants are randomly allocated at birth, they are largely independent of self-selected behaviors and are established long before disease occurrence, thus minimizing confounders and reverse causality. Furthermore, most randomized controlled trials (RCTs) are time-consuming and expensive[23]. In contrast, a two-sample MR analysis allows genetic instrumental variables to be extracted from the summary statistics of large-scale, non-overlapping GWASs, effectively mimicking the design of RCTs[24], thereby reducing time and costs.

Using GWAS datasets for 46 antibody-mediated immune responses (AMIRs) involving 13 infectious pathogens[25] and four distinct BrCa datasets, we conducted two-sample bidirectional MR analysis to assess the causal associations between infectious agents and BrCa. Inverse variance weighted (IVW) and Bayesian Weighted MR (BWMR) methods were the primary approaches for determining the causality of the exposure on the outcome[26], supplemented by meta-analysis to resolve inconsistencies[27]. To further investigate the molecular mechanisms underlying the causal relationships, we mapped the instrumental variables (IVs) to genes for Functional Enrichment. Summary-data-based MR (SMR) analysis was then conducted to detect potential mediating genes[28]. In conclusion, MR harnessed the statistical power of large, pre-existing GWAS data to infer causal relationships between traits, enabling us to elucidate potential causal factors for the formulation of BrCa prevention strategies.

Material and Methods

Overview of the Study

This study intended to assess the causal associations between human pathogenic infections and BrCa. A two-sample bidirectional MR design was employed, utilizing genetic instrumental variable analysis based on summary-level data with SNPs as instruments for the risk factor. The methods of this study strictly adhered to the STROBE-MR guidelines. For causal estimates from MR studies to be valid, three assumptions must be met: (1) the genetic variants are strongly associated with the exposure, (2) the genetic variants are not associated with any potential confounder of the exposure–outcome association, and (3) the genetic variants do not affect the outcome through any pathway other than the exposure[29, 30]. The overall study workflow is summarized as a flowchart in **Figure 1**.

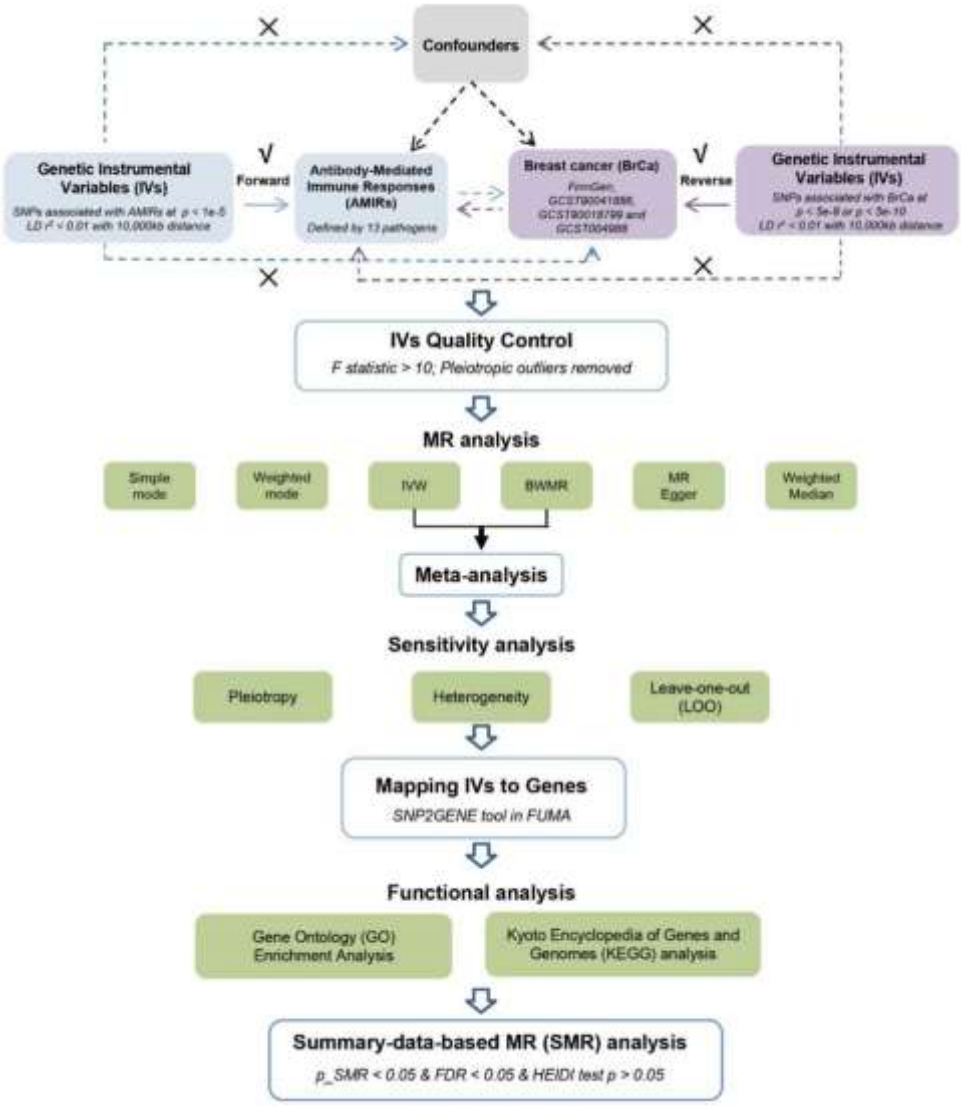

Figure 1: The overall study workflow of bidirectional MR analysis and SMR analysis.

Data Source (GWAS Data of the AMIRs and BrCa)

To explore the potential relationships between infectious agents and BrCa risk, we obtained the GWAS data of 46 AMIR phenotypes defined by 13 infectious pathogens from the UK Biobank cohort (UKB). In the GWAS datasets, 8735 individuals were incorporated for 15 seropositivity case–control phenotypes and a range of 276–8555 samples for 31 quantitative antibody measurement phenotypes[25]. To guarantee enough statistical power to identify associated loci, the GWAS data of pathogens with a seroprevalence of >15% of the 20 original pathogens were chosen.

To obtain reliable findings, BrCa-related single nucleotide polymorphisms (SNPs) were provided by four distinct GWAS datasets: (1) FinnGen Release 10 (C3_BREAST_EXALLC)[31]; (2) GCST004988[32]; (3) GCST90018799[33] and (4) GCST90041886[34]. An overview of the GWAS summary statistics is listed in **Table 1**.

Table 1: GWAS summary statistics utilized in this study

GWAS data	Phenotype	Samples (nCase/nControl)	SNPs	Population	Author
FinnGen		201,713 (18,786/182,927)	20,226,541	European	FinnGen

GCST004988	BrCa	139,274 (76,192/ 63,082)	11,080,744	European	Michailidou K
GCST90018799		337,280 (23,714/313,566)	24,103,144	European & East Asian	Sakaue S
GCST90041886		456,276 (10,152/ 446,124)	11,796,984	European	Jiang L
AMIRs	Defined by infectious agents	NA	NA	European	Butler-Laporte G

Genetic Instrumental Variables (IVs) selection

To recognize high-quality SNPs as IVs, a series of stringent selection criteria were implemented. Initially, SNPs robustly associated with exposures were screened at genome-wide significance ($p < 1 \times 10^{-5}$). Second, the recognized SNPs were clumped for linkage disequilibrium (LD) with a strict cutoff of $r^2 < 0.001$ within a window of 10,000 kb, and LD was estimated with the European samples from the 1000 Genome Project[35], using PLINK v1.9 in R software. Third, ambiguous or palindromic SNPs were removed. Furtherly, the F-statistic was calculated to mitigate the risk of weak instrument bias, applying a threshold of 10[36].

Given the higher susceptibility of cancer patients to certain infectious agents, we also enquired whether BrCa had a causal effect on infectious agents. SNPs ($p < 5 \times 10^{-8}$, $r^2 < 0.001$, kb = 10,000) from datasets of FinnGen, GCST90041886, and GCST90018799 were selected as IVs, while SNPs from GCST004988 were selected with a criterion of ($p < 5 \times 10^{-10}$, $r^2 < 0.001$, kb = 10,000). Subsequently, the same quality control procedures as those for the AMIRs were conducted.

Bidirectional two-sample MR Analysis

The approach of bidirectional two-sample MR analysis is presented in Figure 1. First, the GWAS data of exposure and outcome were harmonized using the selected IVs as matching indexes. Then, we removed SNPs associated with confounders that interfere with the associations between infectious agents and BrCa. We used the NHGRI-EBI GWAS Catalog (www.ebi.ac.uk/gwas), a Findable, Accessible, Interoperable and Reusable (FAIR) knowledgebase, to search if a list of SNPs (or SNPs in LD with those SNPs) have previously been significantly ($P < 1 \times 10^{-5}$) associated with a trait or disease[37]. According to the LDtrait Tool, SNPs were associated with alcohol intake, non-BrCa-related tumors, other pathogen infections, autoimmune diseases, and smoking behavior were removed (**Supplementary Table S1**).

We then performed the MR-PRESSO test to recognize pleiotropic outliers and remove candidate instruments. Cochran's Q test was employed to assess heterogeneity among the genetic variants (SNPs)[38]. Five complementary TSMR methodologies, including MR Egger, Weighted Median, IVW, Simple mode, and Weighted mode, were conducted in this MR to obtain valid causal inferences with some invalid IVs[39]. The IVW method, as the primary approach, showed the strongest ability to estimate causal relationships[40]. BWMR method, based on GWAS summary statistics, can efficiently infer the causality between a risk exposure factor and a trait or disease outcome[41]. MR-Egger method evaluates the causality through the slope coefficient of the Egger regression, which supplies a more robust evaluation free from ineffective IVs. Weighted median method can even protect against up to 50% of weak IVs[40]. Weighted mode method presents less bias and a lower type-I error rate to offer consistent estimates under the relaxed IV assumption[42]. To resolve inconsistent results from the four independent BrCa datasets, meta-analysis was subsequently performed to incorporate the results with a statistical significance via the IVW or BWMR method. We then conducted the leave-one-out (LOO) sensitivity to test whether the results were driven by a single SNP[43]. MR analyses were performed with "TwoSampleMR" packages in R version 4.3.1.

Mapping IVs to Genes

FUMA GWAS tool is a platform used to map, annotate and visualize GWAS results. To understand the mechanism of the bidirectional causal associations between infectious agents and BrCa, we entered the SNPs (IVs) that were significant in the meta-analysis results into the

FUMA GWAS tool. Hence, these lead SNPs of each phenotype were mapped to genes by adopting the SNP2GENE tool[44].

GO Function and KEGG Analysis of mapped gene

To categorize and evaluate mapped genes, Gene Ontology (GO) enrichment analysis was delivered based on cellular component (CC) and molecular function (MF) by "clusterProfiler" package in R studio. Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway involves the knowledge of the molecular interaction, reaction and relation networks for metabolism, genetic information processing, cellular processes, organismal systems, human diseases and drug development[45].

SMR Analysis of mapped gene

SMR analysis can be explained as an approach that combines GWAS and eQTL data to identify genes associated with complex traits by pleiotropy. The methodology can achieve a higher statistical power to test the expression level of a gene in relation to a phenotype than classical MR analyses. Cis-eQTLs (cis expression quantitative trait loci, cis-eQTLs) data was derived from the eQTLGen consortium (https://eqtlgen.org/). SMR test employs the default parameters recommended by the developers of the SMR software version 1.3.1. The HEIDI (heterogeneity-independent-instruments) test was applied to distinguish pleiotropy, where P-HEIDI > 0.05 indicated no pleiotropy. The false discovery rate (FDR) correction was used to resolve multiple testing concerns[46].

Results

The causal relationships of AMIRs on BrCa

Overview for Effect of AMIRs on BrCa

To explore the causal effect of infectious agents on BrCa, we extracted a total of 1147 SNPs (IVs) from 46 AMIRs phenotypes. After removing 15 SNPs related to confounders (**Supplementary Table S1**) and performing quality control, the numbers of harmonized SNPs in distinct GWAS data of BrCa, including the "FinnGen", "GCST004988", "GCST90018799" and "GCST90041886" studies, were 1019, 1097, 1091, and 1068, separately. The IVs mapping to the 46 AMIRs phenotypes ranged from 10 to 84 SNPs, with a minimum F statistic of 19.41, indicating the absence of weak instrument bias in the MR analysis (**Supplementary Table S2-5**).

Subsequently, the IVW method and BWMR analysis, as the primary methods in MR analysis, were conducted. As shown in **Figure 2**, the IVW method or BWMR results ($P < 0.05$) revealed 12 AMIRs with potential causal effects on BrCa, including *C. trachomatis*, *Hp*, EBV, HHV-6 and Human polyomavirus JCV (JCPyV). In contrast, antibody responses against Herpes simplex virus-1/2, Human CMV, Varicella zoster virus (VZV), *Toxoplasma gondii*, HHV-7, Human polyomavirus BKV, and Merkel cell polyomavirus showed no evidence of causality on BrCa.

Figure 2: The random-effect results of IVW and BWMR analysis for the causative effect of pathogenic AMIR phenotypes on the risk of BrCa.

Meta-analysis of the MR Results from four separate BrCa Datasets

Due to inconsistencies in the MR results across the four BrCa datasets, we applied meta-analysis tools to aggregate seven significant causative relationships identified by both IVW and BWMR methods ($P < 0.05$) (**Figure 3**). As illustrated in **Figure 4**, the results revealed three significant overlapping causal relationships: Anti-HHV-6 IgG seropositivity [IVW (OR 1.01; 95% CI, 1.00–1.02; $p < 0.01$); BWMR (OR 1.01; 95% CI, 1.00–1.02; $p < 0.01$)], HHV-6 IE1A antibody levels [IVW (OR 1.04; 95% CI, 1.02–1.06; $p < 0.001$); BWMR (OR 1.04; 95% CI, 1.02–1.07; $p < 0.001$)], and Hp VacA antibody levels [IVW (OR 1.02; 95% CI, 1.00–1.04; $p < 0.05$); BWMR (OR 1.02; 95% CI, 1.00–1.04; $p < 0.05$)], each showing a positive causal relationship on BrCa. As indicated in **Supplementary Tables S6–S9**, no evidence of pleiotropy or heterogeneity was observed among these three causal associations. The results of the LOO sensitivity analyses are presented in **Supplementary Figures S1–S12**.

Meta-analysis Results

			IVW				BWMR			
					Heterogeneity				Heterogeneity	
Group	Trait	Samples	OR (95% CI)	pval	I^2(%)	pval	OR (95% CI)	pval	I^2(%)	pval
Bacteria										
	Chlamydia trachomatis									
	Chlamydia trachomatis momp A antibody levels	964	1.00 (0.99 to 1.01)	0.91	62.3	0.05*	1.00 (0.99 to 1.01)	0.90	61	0.05
	Chlamydia trachomatis momp D antibody levels	1371	0.98 (0.97 to 1.00)	0.09	15.2	0.32	0.98 (0.96 to 1.00)	0.11	6	0.36
	Helicobacter pylori									
	Helicobacter pylori UREA antibody levels	2251	0.99 (0.97 to 1.00)	0.15	14.9	0.32	0.99 (0.97 to 1.00)	0.09	0	0.43
	Helicobacter pylori VacA antibody levels	1571	1.02 (1.00 to 1.04)	0.03*	40.3	0.17	1.02 (1.00 to 1.04)	0.01*	45.4	0.14
Herpesviridae										
	Human herpesvirus-6									
	Anti-human herpes virus 6 IgG seropositivity	8735	1.01 (1.00 to 1.02)	0.00**	17	0.31	1.01 (1.00 to 1.02)	0.01**	0	0.47
	Human herpes virus 6 IE1A antibody levels	6968	1.04 (1.02 to 1.06)	0.00***	0	0.66	1.04 (1.02 to 1.07)	0.00***	0	0.70
Polyomaviridae										
	Human polyomavirus JCV									
	Anti-polyomavirus 2 IgG seropositivity	8375	1.02 (1.00 to 1.03)	0.07	33.2	0.21	1.02 (1.00 to 1.04)	0.13	26.6	0.25

*p < 0.05, **p < 0.01, ***p < 0.001

Figure 3: Meta-analyses for the statistically significant causalities of AMIRs on BrCa. (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$)

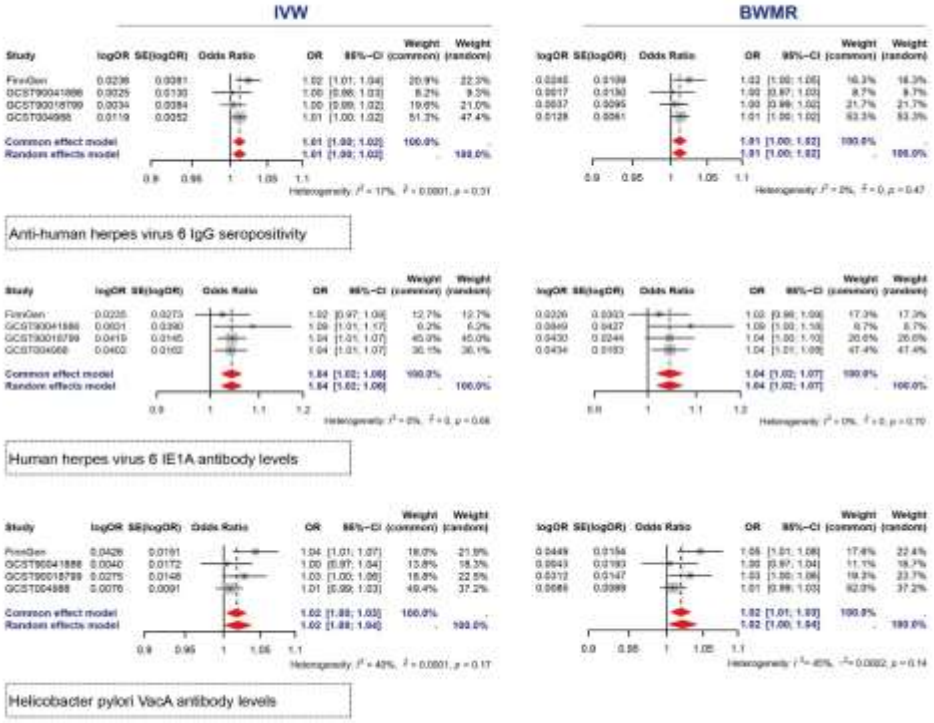

Figure 4: AMIRs of Significant causal effects on BrCa based on meta-analysis.

Mapping IVs of AMIRs to genes and SMR analysis

To explore the molecular mechanisms underlying the forward causal relationships, we evaluated the gene mappings of the IVs for each of the three significant exposures using the SNP2GENE tool. The 50 genes associated with IVs are detailed in **Supplementary Table S10**. After excluding genes with a HEIDI test ($p < 0.05$) and adjusting for FDR > 0.05, we identified ten genes with causal associations with BrCa, including TANGO2, RBP7, RNLS, DOCK3, LZIC, ITGA9, NMNAT1, ZBP1, BNC2 and ZWINT. Notably, BNC2, associated with IVs of HHV-6 IE1A antibody levels, was found to be associated with two distinct GWAS datasets of BrCa (**Figure 5**).

Figure 5: Forest plot of the mapped genes with a putative causal effect on BrCa via SMR analysis.

The causal relationships of BrCa on AMIRs

Overview of the causal effect of BrCa on AMIRs

To explore the causal relationships in the reverse direction, after removing 29 SNPs related to confounders (**Supplementary Table S11**) and performing quality control, the numbers of harmonized SNPs corresponding to the four independent GWAS datasets of BrCa ("FinnGen", "GCST004988", "GCST90018799" and "GCST90041886") were 1839, 3905, 1793 and 781, respectively. With a minimum F statistic of 29.80, the IVs for the 46 AMIRs ranged from 16 to 85 SNPs (**Supplementary Tables S12–S15**). Using these SNPs (IVs), we performed further MR to investigate the causal effect of BrCa on infectious agents.

Across IVW and BWMR approaches, we found no evidence for a causal effect of BrCa on antibodies against *C. trachomatis*, Herpes simplex virus-1/2, EBV, Human CMV, HHV-7, *Toxoplasma gondii*, Merkel cell polyomavirus, and Human polyomavirus BKV. However, the results demonstrated five suggestive and unique causal associations ($P < 0.05$) with infectious agents, including *Hp*, HHV-6, VZV, and Human polyomavirus BKV/JCV (**Figure 6**).

Figure 6: The random-effect of IVW and BWMR analyses of BrCa on the risk of the 46 AMIR phenotypes.

Meta-analysis of the MR Results for the Distinct Datasets of BrCa

To resolve inconsistencies of MR results among the four distinct BrCa datasets, meta-analyses of IVW and BWMR results were similarly conducted (**Figure 7**). The results demonstrated statistically significant overlapping causal relationships among the three associations presented in Figure 6. Specifically, BrCa was positively associated with Polyomavirus 2 JC VP1 antibody levels [IVW (OR 1.07; 95% CI, 1.04–1.11; $p < 0.001$); BWMR (OR 1.07; 95% CI, 1.03–1.11; $p < 0.001$)], and negatively associated with Anti-Hp IgG seropositivity [IVW (OR 0.93; 95% CI, 0.88–0.97; $p < 0.01$); BWMR (OR 0.93; 95% CI, 0.88–0.99; $p < 0.05$)] and Anti-VZV IgG seropositivity [IVW (OR 0.84; 95% CI, 0.79–0.90; $p < 0.001$); BWMR (OR 0.84; 95% CI, 0.78–0.90; $p < 0.001$)] (**Figure 8**). As indicated in Supplementary Tables S16-S19, no evidence of pleiotropy or heterogeneity was observed among these three causal associations. The results of the LOO sensitivity analyses are presented in **Supplementary Figures S13-S24**.

Meta-analysis Results

			IVW				BWMR			
					Heterogeneity				Heterogeneity	
Group	Trait	Samples	OR (95% CI)	pval	I^2(%)	pval	OR (95% CI)	pval	I^2(%)	pval
Bacteria										
	Helicobacter pylori									
	Anti-helicobacter pylori IgG seropositivity	8735	0.93 (0.88 to 0.97)	0.00**	0	0.51	0.93 (0.88 to 0.99)	0.02*	0	0.65
Herpesviridae										
	Varicella zoster virus									
	Anti-varicella zoster virus IgG seropositivity	8735	0.84 (0.79 to 0.90)	0.00***	0	0.53	0.84 (0.78 to 0.90)	0.00***	0	0.57
Polyomaviridae										
	Human polyomavirus JCV									
	Polyomavirus 2 JC VP1 antibody levels	5118	1.07 (1.04 to 1.11)	0.00***	0	0.56	1.07 (1.03 to 1.11)	0.00***	0	0.75

*p < 0.05, **p < 0.01, ***p < 0.001

Figure 7: Meta-analyses of the significant causal effects of BrCa on AMIRs. (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$)

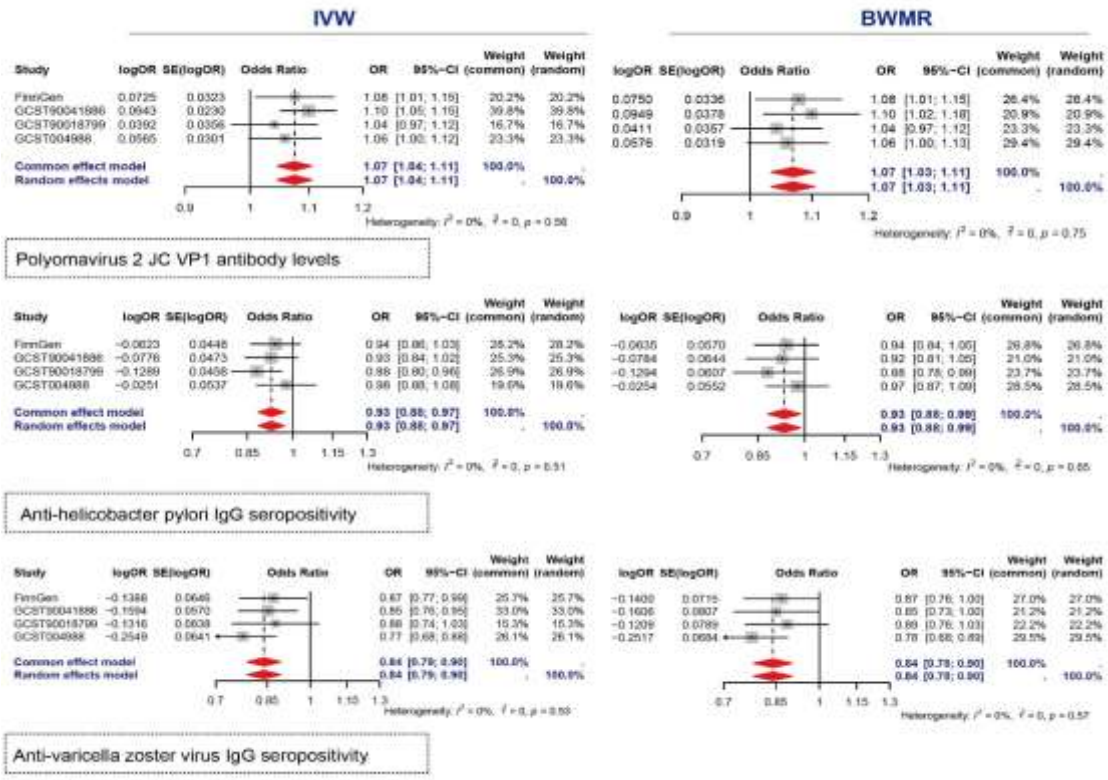

Figure 8: Significant causal effects of BrCa on AMIRs based on meta-analysis

Functional Enrichment and SMR Analyses

To obtain molecular insight into the reverse causality, the IVs (SNPs) of BrCa were mapped to genes using the SNP2GENE tool, with the 201 hit genes recorded in **Supplementary Table S20**. To further understand the biological mechanisms of these mapped genes, GO analyses based on cellular component (CC) and molecular function (MF), as well as KEGG analysis, were performed. As shown in Supplementary Table S21, the enriched functional pathways of these genes include nucleosome (GO:0000786), protein-DNA complex (GO:0032993), DNA packaging complex (GO:0044815), euchromatin (GO:0000791), structural constituent of chromatin (GO:0030527), sodium: phosphate symporter activity (GO:0005436), solute: monoatomic cation symporter activity (GO:0015294), solute: sodium symporter activity (GO:0015370), lipopeptide binding (GO:0071723), Toll-like receptor binding (GO:0035325), symporter activity (GO:0015293), NAD+ nucleosidase activity (GO:0003953), NAD(P)+ nucleosidase activity (GO:0050135), and NAD+ nucleotidase, cyclic ADP-ribose generating(GO:0061809) process. Additionally, KEGG analysis revealed that the genes mapped to BrCa were enriched in the Systemic lupus erythematosus pathway (p-adjust<0.05) (**Supplementary Table S21**). After excluding genes with a HEIDI test ($p < 0.05$) and adjusting for FDR > 0.05, SMR analysis identified 13 genes with putative causal effects on any of the four AMIRs, as listed in **Figure 9**.

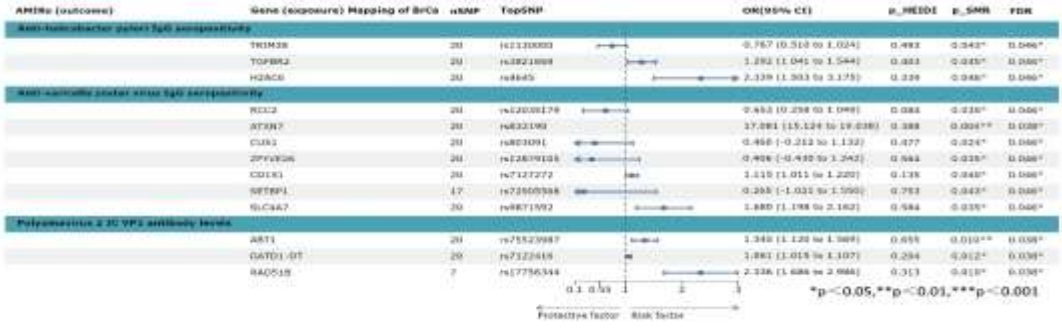

Figure 9: Forest plot of SMR analysis results of the causal effect of mapped genes on AMIR.

Discussion

BrCa is a heterogeneous disease with incidence and mortality rates differing across countries, attributed to a combination of genetic predisposition, lifestyle choices, and environmental factors. Recent observational studies have shown that pathogenic agents are related to breast oncogenesis; however, whether these associations are causal remains inconclusive. This study employed bidirectional two-sample MR analyses to systematically assess the causal associations between 13 infectious agents and BrCa.

In the forward MR analysis, our results showed three AMIRs with statistically significant evidence of potential causal effects on BrCa across meta-analysis. Specifically, Anti-human herpes virus 6 IgG seropositivity, Human herpes virus 6 IE1A antibody levels and Helicobacter pylori VacA antibody levels were positively correlated with BrCa risk, indicating that pathogenic agents such as HHV-6 and *Hp* are risk factors for BrCa. Meanwhile, although Anti-Chlamydia trachomatis tarp-D F2 antibody levels and Epstein-Barr virus EBNA-1 antibody levels are demonstrated to be positively related to BrCa risk in some GWAS datasets of BrCa, our meta-analysis revealed no statistical significance. Regarding previous studies that have reported *C. trachomatis*[9], EBV[8] and *Hp*[11] as being associated with BrCa, our results support *H. pylori* as a potential pathogenic agent. Additionally, our results provide evidence for HHV6 as a risk factor for BrCa, with no prior evidence of the causal association reported. In brief, these findings highlight the importance of preventing breast oncogenesis at the onset of these infections through the rational use of antibiotics, antiviral agents, and probiotics, as well as the need to maintain a healthy microbiota. However, due to the reliance on a single AMIR phenotype for each corresponding pathogen, these results should be interpreted with caution. More experimental and epidemiological evidence is needed to confirm these causal associations in the future.

Furthermore, SMR analysis of genes mapped by IVs of these AMIRs was conducted, suggesting the corresponding oncogenic pathogens may affect BrCa through these genes. Notably, among the genes related to AMIRs in our SMR analysis, BNC2 (rs12350739) was identified as a protective factor for BrCa; BNC2 was previously reported as a putative tumor suppressor gene in high-grade serous ovarian carcinoma[47] and a protective factor in cutaneous squamous cell carcinoma[48]. The protective role is the same as the genes of RBP7 and ZBP1, with RBP7 being reported as a tumor suppressor for breast cancer[49, 50] and ZBP1 essential for host defense against viruses[51]. Otherwise, DOCK3 and NMNAT1 were identified as risk factors for BrCa; NMNAT1 was demonstrated to prevent apoptosis of acute myeloid leukemia stem cells by governing Nuclear NAD^+ homeostasis[52]. Other genes in this study, having not been reported to be directly relevant to or contradictorily associated with BrCa[53, 54], may provide new clues for prospective explorations on the mechanism of action between infectious pathogens and BrCa.

Although much effort in this study was made to investigate the mechanism of action between infectious pathogens and BrCa, it remains uncertain. According to previous studies[8], multiple biological process are hypothesized for infection-mediated oncogenesis, includes affecting estrogen levels, lipid metabolism, genomic instability, proliferative signaling, immune regulation, modification of the tumor microenvironment and inflammatory response. Epstein–Barr nuclear antigen 1 (EBNA1), necessary to maintain stability of EBV particles and EBV replication, has been shown to enhance genomic instability, favoring mammary oncogenesis. *H. pylori* was reported to increase the risk of cancer not only for the prolonged immune activation and durative production of inflammatory factors[55], but also the expressed cagA and s1a vacA genes[56]. *C. trachomatis* infection was reported to be associated with the development of hormone-responsive breast cancer in females with high levels of IL-12[9]. Given the role of pathogenic infections in chronic inflammation, inflammation-mediated oncogenic processes and immune dysregulation, it is credible special microbes are relevant to the development of special cancers. More effort should focus on the underlying mechanisms linking infectious agents to breast cancer in the future for the development of precision medicine and personalized treatment strategies.

In reverse MR analysis, we demonstrated BrCa was positively correlated with Polyomavirus 2 JC VP1 antibody levels, while negatively with Anti-helicobacter pylori IgG seropositivity and Anti-varicella zoster virus IgG seropositivity, which suggests BrCa may increase the risk of JCPyV infection but a protective factor for infectious agents, like *Hp* and VZV. Hence, these results strengthen the credibility of the forward causal relationship.

To further explore the mechanism of the causal relationships between BrCa and infectious pathogens, we performed a GO, KEGG and SMR analysis based on BrCa's IV-mapped genes of significantly causal effect on AMIRs. The enriched functional pathways associated with these genes include five DNA packaging-associated cellular compartments (nucleosome (GO:0000786), protein-DNA complex (GO:0032993), DNA packaging complex (GO:0044815), euchromatin (GO:0000791), strctural constituent of chromatin (GO:0030527))[57], sodium: phosphate symporter activity (GO:0005436)[58], solute: monoatomic cation symporter activity (GO:0015294), lipopeptide binding (GO:0071723), solute:sodium symporter activity (GO:0015370), Toll-like receptor binding (GO:0035325)[59], NAD+ nucleosidase activity (GO:0003953)[60], NAD(P)+ nucleosidase activity (GO:0050135), cyclic ADP-ribose generating (GO:0061809), symporter activity (GO:0015293)[58]. All of them have been shown to play significant roles in microbial infection. Our KEGG analysis showed that the gene mapped of BrCa were enriched in Systemic lupus erythematosus process, which indicates autoimmune disorder and chronic inflammation. These findings suggest a causal relationship between BrCa and pathogen susceptibility through a common pathway. These genes provide important clues to design further functional studies to understand the mechanism whereby DNA variation leads to complex trait variation.

Undoubtedly, the results should be cautiously interpreted in combination with their own restrictions. First, only 13 infectious pathogens we used did not cover the diversity of infectious agents and the AMIRs data was derived from solo GWAS dataset, which may impair the study's accuracy. In addition, the BrCa patients from four distinct GWAS datasets, mainly confined to European populations, may not be a representative of the entire community. Third, with the minimal *F*-statistic values >10, suggesting the potential for weak instrumental bias to be low, there was likely some overlap between exposure and outcome participants. Fourth, assessments from MR studies of unrelated individuals can be biased for variable environmental and social factors, such as diverse gender structure and varied reproduction rate. Finally, we employed several complementary approaches to infer robust causal estimates, it still suffered pleiotropy and weak instrument bias, which could make the result inaccurate.

Conclusion

Collectively, this study employed MR, complemented with sensitivity analysis, to assess the bidirectional causal associations between infectious agents and BrCa. Through meta-analysis to resolve inconsistencies, our findings disclosed possible causal links with strong genetic evidence between infectious agents and BrCa, with HHV-6 and *H. pylori* identified as risk factors for BrCa. In the reverse direction, BrCa may increase susceptibility to JCPyV infection and confer relative immunity to infectious agents such as H. pylori and VZV. These insights will contribute to better prevention and intervention strategies at the pathogenic level for modifiable risks associated with BrCa.

Acknowledgments:

We expressed our sincere gratitude to all researchers and staff involved in the public GWAS data of this study. We also earnestly thank all the authors who contributed to this research.

Funding:

This work was supported by National Natural Science Foundation of China (grant number: 81972521)

Competing Interests

No potential conflicts of interest relevant to this article were declared.

Reference

1. Alibek, K., Kakpenova, A., Mussabekova, A., Sypabekova, M., & Karatayeva, N. (2013). Role of viruses in the development of breast cancer. *Infect Agent Cancer*, *8*, 32. https://doi.org/10.1186/1750-9378-8-32
2. Amusan, O. T., Wang, S., Yin, C., Koehler, H.

S., Li, Y., Tenev, T., Wilson, R., Bellenie, B., Zhang, T., Wang, J., Liu, C., Seong, K., Poorbaghi, S. L., Yates, J., Shen, Y., Upton, J. W., Meier, P., Balachandran, S., & Guo, H. (2024). RIPK1 is essential for Herpes Simplex Virus-triggered ZBP1-dependent necroptosis in human cells. *bioRxiv*. https://doi.org/10.1101/2024.09.17.613393

3. Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., Korbel, J. O., Marchini, J. L., McCarthy, S., McVean, G. A., & Abecasis, G. R. (2015). A global reference for human genetic variation. *Nature*, *526* (7571), 68-74. https://doi.org/10.1038/nature15393
4. Banerjee, S., Wei, Z., Tan, F., Peck, K. N., Shih, N., Feldman, M., Rebbeck, T. R., Alwine, J. C., & Robertson, E. S. (2015). Distinct microbiological signatures associated with triple negative breast cancer. *Sci Rep*, *5*, 15162. https://doi.org/10.1038/srep15162
5. Bønløkke, S., Blaakær, J., Steiniche, T., Høgdall, E., Jensen, S. G., Hammer, A., Balslev, E., Strube, M. L., Knakkergaard, H., & Lenz, S. (2018). Evidence of No Association Between Human Papillomavirus and Breast Cancer. *Front Oncol*, *8*, 209. https://doi.org/10.3389/fonc.2018.00209
6. Bowden, J., & Holmes, M. V. (2019). Meta-analysis and Mendelian randomization: A review. *Res Synth Methods*, *10*(4), 486-496. https://doi.org/10.1002/jrsm.1346
7. Burgess, S., Davey Smith, G., Davies, N. M., Dudbridge, F., Gill, D., Glymour, M. M., Hartwig, F. P., Kutalik, Z., Holmes, M. V., Minelli, C., Morrison, J. V., Pan, W., Relton, C. L., & Theodoratou, E. (2019). Guidelines for performing Mendelian randomization investigations: update for summer 2023. *Wellcome Open Res*, *4*, 186. https://doi.org/10.12688/wellcomeopenres. 15555.3
8. Burgess, S., Foley, C. N., Allara, E., Staley, J. R., & Howson, J. M. M. (2020). A robust and efficient method for Mendelian randomization with hundreds of genetic variants. *Nat Commun*, *11*(1), 376. https://doi.org/10. 1038/s41467-019-14156-4
9. Butler-Laporte, G., Kreuzer, D., Nakanishi, T., Harroud, A., Forgetta, V., & Richards, J. B. (2020). Genetic Determinants of Antibody-Mediated Immune Responses to Infectious Diseases Agents: A Genome-Wide and HLA Association Study. *Open Forum Infect Dis*, *7*(11), ofaa450. https://doi.org/10.1093/ofid/ofaa450
10. Canive, M., Badia-Bringué, G., Vázquez, P., Garrido, J. M., Juste, R. A., Fernandez, A., González-Recio, O., & Alonso-Hearn, M. (2022). A Genome-Wide Association Study for Tolerance to Paratuberculosis Identifies Candidate Genes Involved in DNA Packaging, DNA Damage Repair, Innate Immunity, and Pathogen Persistence. *Front Immunol*, *13*, 820965. https://doi.org/10.3389/fimmu.2022.820965
11. Cesaratto, L., Grisard, E., Coan, M., Zandonà, L., De Mattia, E., Poletto, E., Cecchin, E., Puglisi, F., Canzonieri, V., Mucignat, M. T., Zucchetto, A., Stocco, G., Colombatti, A., Nicoloso, M. S., & Spizzo, R. (2016). BNC2 is a putative tumor suppressor gene in high-grade serous ovarian carcinoma and impacts cell survival after oxidative stress. *Cell Death Dis*, *7*(12), e2526. https://doi.org/10.1038/cddis.2016.448
12. Chahal, H. S., Lin, Y., Ransohoff, K. J., Hinds, D. A., Wu, W., Dai, H. J., Qureshi, A. A., Li, W. Q., Kraft, P., Tang, J. Y., Han, J., & Sarin, K. Y. (2016). Genome-wide association study identifies novel susceptibility loci for cutaneous squamous cell carcinoma. *Nat Commun*, *7*, 12048. https://doi.org/10.1038/ncomms12048
13. Chitapanarux, T., Traisathit, P., Srikumm, P., Homkham, N., & Chitapanarux, I. (2024). Helicobacter pylori, Atrophic Gastritis, and Breast Cancer Risk: A Prospective Cohort Study with 8-year follow-up. *Research Square*. https://doi.org/10.21203/rs.3.rs-3657050/v1
14. Cohen, J. F., Chalumeau, M., Cohen, R., Korevaar, D. A., Khoshnood, B., & Bossuyt, P. M. (2015). Cochran's Q test was useful to assess heterogeneity in likelihood ratios in studies of diagnostic accuracy. *J Clin Epidemiol*, *68*(3), 299-306. https://doi.org/10.1016/j.jclinepi.2014. 09.005
15. Cummins, J., & Tangney, M. (2013). Bacteria and tumours: causative agents or opportunistic inhabitants? *Infect Agent Cancer*, *8*(1), 11. https://doi.org/10.1186/1750-9378-8-11
16. Deng, X., Luo, Y., Guan, T., & Guo, X. (2023). Identification of the Genetic Influence of SARS-CoV-2 Infections on IgA

Nephropathy Based on Bioinformatics Method. *Kidney Blood Press Res*, *48*(1), 367-384. https://doi.org/10.1159/000529687
17. Eliassen, E., Lum, E., Pritchett, J., Ongradi, J., Krueger, G., Crawford, J. R., Phan, T. L., Ablashi, D., & Hudnall, S. D. (2018). Human Herpesvirus 6 and Malignancy: A Review. *Front Oncol*, *8*, 512. https://doi.org/ 10.3389/fonc.2018.00512
18. Gao, A., Kouznetsova, V. L., & Tsigelny, I. F. (2020). Bovine leukemia virus relation to human breast cancer: Meta-analysis. *Microb Pathog*, *149*, 104417. https://doi.org/10. 1016/j. micpath.2020.104417
19. Geisler, J., Touma, J., Rahbar, A., Söderberg-Nauclér, C., & Vetvik, K. (2019). A Review of the Potential Role of Human Cytomegalovirus (HCMV) Infections in Breast Cancer Carcinogenesis and Abnormal Immunity. *Cancers (Basel)*, *11*(12). https:// doi.org/10.3390/cancers 11121842
20. Grant, A. J., & Burgess, S. (2024). A Bayesian approach to Mendelian randomization using summary statistics in the univariable and multivariable settings with correlated pleiotropy. *Am J Hum Genet*, *111*(1), 165-180. https://doi.org/10.1016/j.ajhg.2023.12.002
21. Gupta, I., Ulamec, M., Peric-Balja, M., Ramic, S., Al Moustafa, A. E., Vranic, S., & Al-Farsi, H. F. (2021). Presence of high-risk HPVs, EBV, and MMTV in human triple-negative breast cancer. *Hum Vaccin Immunother*, *17*(11), 4457-4466. https://doi.org/10.1080/ 21645515.2021. 1975452
22. Hachana, M., Trimeche, M., Ziadi, S., Amara, K., & Korbi, S. (2009). Evidence for a role of the Simian Virus 40 in human breast carcinomas. *Breast Cancer Res Treat*, *113*(1), 43-58. https://doi.org/10.1007/s10549-008-9901-z
23. Hariton, E., & Locascio, J. J. (2018). Randomised controlled trials - the gold standard for effectiveness research: Study design: randomised controlled trials. *Bjog*, *125*(13), 1716. https://doi.org/10.1111/1471-0528.15199
24. Hartwig, F. P., Davey Smith, G., & Bowden, J. (2017). Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol*, *46*(6), 1985-1998. https://doi.org/10.1093/ije/dyx102
25. Iob, E., Pingault, J. B., Munafò, M. R., Stubbs, B., Gilthorpe, M. S., Maihofer, A. X., & Danese, A. (2023). Testing the causal relationships of physical activity and sedentary behaviour with mental health and substance use disorders: a Mendelian randomisation study. *Mol Psychiatry*, *28*(8), 3429-3443. https://doi.org/10.1038/s41380-023-02133-9
26. Jiang, L., Zheng, Z., Fang, H., & Yang, J. (2021). A generalized linear mixed model association tool for biobank-scale data. *Nat Genet*, *53*(11), 1616-1621. https://doi.org/ 10.1038/s41588-021-00954-4
27. Kadivar, M., Monabati, A., Joulaee, A., & Hosseini, N. (2011). Epstein-Barr virus and breast cancer: lack of evidence for an association in Iranian women. *Pathol Oncol Res*, *17*(3), 489-492. https://doi.org/10. 1007 /s12253-010-9325-z
28. Karachalios, C., Petousis, S., Margioula-Siarkou, C., & Dinas, K. (2024). Human papillomaviruses and breast cancer: A systematic review and meta-analysis. *Oncol Lett*, *27*(2), 75. https://doi.org/10.3892/ol. 2023.14208
29. Kawamura, Y., Hashimoto, T., Miura, H., Kozawa, K., Yoshikawa, A., Ikeda, N., Yatsuya, H., Yasuoka, H., & Yoshikawa, T. (2020). Inherited chromosomally integrated human herpesvirus 6 and autoimmune connective tissue diseases. *J Clin Virol*, *132*, 104656. https://doi.org/10.1016/j.jcv.2020. 104656
30. Ki, M. R., Hwang, M., Kim, A. Y., Lee, E. M., Lee, E. J., Lee, M. M., Sung, S. E., Kim, S. H., Lee, H. S., & Jeong, K. S. (2014). Role of vacuolating cytotoxin VacA and cytotoxin-associated antigen CagA of Helicobacter pylori in the progression of gastric cancer. *Mol Cell Biochem*, *396*(1-2), 23-32. https://doi. org/10.1007/s11010-014-2138-8
31. Kurki, M. I., Karjalainen, J., Palta, P., Sipilä, T. P., Kristiansson, K., Donner, K. M., Reeve, M. P., Laivuori, H., Aavikko, M., Kaunisto, M. A., Loukola, A., Lahtela, E., Mattsson, H., Laiho, P., Della Briotta Parolo, P., Lehisto, A. A., Kanai, M., Mars, N., Rämö, J., . . . Palotie, A. (2023). FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature*, *613*(7944), 508-518. https://doi.org/ 10.1038/s41586-022-05473-8
32. Lawlor, D. A., Harbord, R. M., Sterne, J. A.,

Timpson, N., & Davey Smith, G. (2008). Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med*, *27*(8), 1133-1163. https://doi.org/10.1002/sim.3034

33. Li, H. N., Zheng, W. H., Du, Y. Y., Wang, G., Dong, M. L., Yang, Z. F., & Li, X. R. (2020). ZW10 interacting kinetochore protein may serve as a prognostic biomarker for human breast cancer: An integrated bioinformatics analysis. *Oncol Lett*, *19*(3), 2163-2174. https://doi.org/10.3892/ol.2020.11353
34. Li, N., Xiao, C., Li, Y., Zhang, Y., Lin, Y., Liu, Q., Tang, L., Xu, L., & Ren, Z. (2024). Association of Chlamydia trachomatis Infection With Breast Cancer Risk and the Modification Effect of IL-12. *Clin Breast Cancer*, *24*(7), e554-e559.e551. https://doi.org/10.1016/j.clbc.2024.05.003
35. Li, T., Shao, W., Wang, Y., Zhou, R., Yun, Z., He, Y., & Wu, Y. (2023). A two-sample mendelian randomization analysis investigates associations between gut microbiota and infertility. *Sci Rep*, *13*(1), 11426. https://doi.org/10.1038/s41598-023-38624-6
36. Lin, H., Han, Q., Wang, J., Zhong, Z., Luo, H., Hao, Y., & Jiang, Y. (2022). Methylation-Mediated Silencing of RBP7 Promotes Breast Cancer Progression through PPAR and PI3K/AKT Pathway. *J Oncol*, *2022*, 9039110. https://doi.org/10.1155/2022/9039110
37. Michailidou, K., Lindström, S., Dennis, J., Beesley, J., Hui, S., Kar, S., Lemaçon, A., Soucy, P., Glubb, D., Rostamianfar, A., Bolla, M. K., Wang, Q., Tyrer, J., Dicks, E., Lee, A., Wang, Z., Allen, J., Keeman, R., Eilber, U., . . . Easton, D. F. (2017). Association analysis identifies 65 new breast cancer risk loci. *Nature*, *551*(7678), 92-94. https://doi.org/ 10.1038/nature24284
38. Moss, S. F., & Blaser, M. J. (2005). Mechanisms of disease: Inflammation and the origins of cancer. *Nat Clin Pract Oncol*, *2*(2), 90-97; quiz 91 p following 113. https://doi.org/10.1038/ncponc0081
39. Mostovich, L. A., Prudnikova, T. Y., Kondratov, A. G., Loginova, D., Vavilov, P. V., Rykova, V. I., Sidorov, S. V., Pavlova, T. V., Kashuba, V. I., Zabarovsky, E. R., & Grigorieva, E. V. (2011). Integrin alpha9 (ITGA9) expression and epigenetic silencing in human breast tumors. *Cell Adh Migr*, *5*(5), 395-401. https://doi.org/10.4161/cam. 5.5.17949
40. Newton, R., Ziegler, J., Bourboulia, D., Casabonne, D., Beral, V., Mbidde, E., Carpenter, L., Reeves, G., Parkin, D. M., Wabinga, H., Mbulaiteye, S., Jaffe, H., Weiss, R., & Boshoff, C. (2003). The sero-epidemiology of Kaposi's sarcoma-associated herpesvirus (KSHV/HHV-8) in adults with cancer in Uganda. *Int J Cancer*, *103*(2), 226-232. https://doi.org/10.1002/ijc.10817
41. O'Connor, H., MacSharry, J., Bueso, Y. F., Lindsay, S., Kavanagh, E. L., Tangney, M., Clyne, M., Saldova, R., & McCann, A. (2018). Resident bacteria in breast cancer tissue: pathogenic agents or harmless commensals? *Discov Med*, *26*(142), 93-102.
42. Parida, S., & Sharma, D. (2020). Microbial Alterations and Risk Factors of Breast Cancer: Connections and Mechanistic Insights. *Cells*, *9*(5). https://doi.org/10.3390/cells9051091
43. Peerapen, P., & Thongboonkerd, V. (2023). Protein network analysis and functional enrichment via computational biotechnology unravel molecular and pathogenic mechanisms of kidney stone disease. *Biomed J*, *46*(2), 100577. https://doi.org/10.1016/j.bj.2023. 01. 001
44. Sakaue, S., Kanai, M., Tanigawa, Y., Karjalainen, J., Kurki, M., Koshiba, S., Narita, A., Konuma, T., Yamamoto, K., Akiyama, M., Ishigaki, K., Suzuki, A., Suzuki, K., Obara, W., Yamaji, K., Takahashi, K., Asai, S., Takahashi, Y., Suzuki, T., . . . Okada, Y. (2021). A cross-population atlas of genetic associations for 220 human phenotypes. *Nat Genet*, *53*(10), 1415-1424. https://doi.org/ 10.1038/s41588-021-00931-x
45. Schumacher, F. R., Al Olama, A. A., Berndt, S. I., Benlloch, S., Ahmed, M., Saunders, E. J., Dadaev, T., Leongamornlert, D., Anokian, E., Cieza-Borrella, C., Goh, C., Brook, M. N., Sheng, X., Fachal, L., Dennis, J., Tyrer, J., Muir, K., Lophatananon, A., Stevens, V. L., . . . Eeles, R. A. (2018). Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. *Nat Genet*, *50*(7), 928-936. https://doi.org/10.1038/s41588-018-0142-8
46. Shi, X., Jiang, Y., Kitano, A., Hu, T., Murdaugh, R. L., Li, Y., Hoegenauer, K. A., Chen, R., Takahashi, K., & Nakada, D.

(2021). Nuclear NAD(+) homeostasis governed by NMNAT1 prevents apoptosis of acute myeloid leukemia stem cells. *Sci Adv*, *7*(30). https://doi.org/10.1126/sciadv.abf3895
47. Singhal, A., & Cheng, C. Y. (2019). Host NAD+ metabolism and infections: therapeutic implications. *Int Immunol*, *31*(2), 59-67. https://doi.org/10.1093/intimm/dxy068
48. Skrivankova, V. W., Richmond, R. C., Woolf, B. A. R., Davies, N. M., Swanson, S. A., VanderWeele, T. J., Timpson, N. J., Higgins, J. P. T., Dimou, N., Langenberg, C., Loder, E. W., Golub, R. M., Egger, M., Davey Smith, G., & Richards, J. B. (2021). Strengthening the reporting of observational studies in epidemiology using mendelian randomisation (STROBE-MR): explanation and elaboration. *Bmj*, *375*, n2233. https://doi.org/10.1136/ bmj.n2233
49. Skrivankova, V. W., Richmond, R. C., Woolf, B. A. R., Yarmolinsky, J., Davies, N. M., Swanson, S. A., VanderWeele, T. J., Higgins, J. P. T., Timpson, N. J., Dimou, N., Langenberg, C., Golub, R. M., Loder, E. W., Gallo, V., Tybjaerg-Hansen, A., Davey Smith, G., Egger, M., & Richards, J. B. (2021). Strengthening the Reporting of Observational Studies in Epidemiology Using Mendelian Randomization: The STROBE-MR Statement. *Jama*, *326*(16), 1614-1621. https://doi.org/ 10.1001/jama.2021.18236
50. Sollis, E., Mosaku, A., Abid, A., Buniello, A., Cerezo, M., Gil, L., Groza, T., Güneş, O., Hall, P., Hayhurst, J., Ibrahim, A., Ji, Y., John, S., Lewis, E., MacArthur, J. A. L., McMahon, A., Osumi-Sutherland, D., Panoutsopoulou, K., Pendlington, Z., . . . Harris, L. W. (2023). The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. *Nucleic Acids Res*, *51*(D1), D977-d985. https://doi.org/10.1093/nar/gkac1010
51. Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*, *71*(3), 209-249. https://doi.org/10.3322/caac.21660
52. Takeda, K., & Akira, S. (2015). Toll-like receptors. *Curr Protoc Immunol*, *109*, 14.12. 11-14.12.10. https://doi.org/10.1002/0471142735.im1412s109
53. Thompson, K. J., Ingle, J. N., Tang, X., Chia, N., Jeraldo, P. R., Walther-Antonio, M. R., Kandimalla, K. K., Johnson, S., Yao, J. Z., Harrington, S. C., Suman, V. J., Wang, L., Weinshilboum, R. L., Boughey, J. C., Kocher, J. P., Nelson, H., Goetz, M. P., & Kalari, K. R. (2017). A comprehensive analysis of breast cancer microbiota and host gene expression. *PloS one*, *12*(11), e0188873. https://doi.org/10.1371/journal.pone.0188873
54. Toumazi, D., El Daccache, S., & Constantinou, C. (2021). An unexpected link: The role of mammary and gut microbiota on breast cancer development and management (Review). *Oncol Rep*, *45*(5). https://doi.org/10.3892/or.2021.8031
55. Urbaniak, C., Gloor, G. B., Brackstone, M., Scott, L., Tangney, M., & Reid, G. (2016). The Microbiota of Breast Tissue and Its Association with Breast Cancer. *Appl Environ Microbiol*, *82*(16), 5039-5048. https://doi.org/10.1128/aem.01235-16
56. Watanabe, K., Taskesen, E., van Bochoven, A., & Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. *Nat Commun*, *8*(1), 1826. https://doi.org/10.1038/s41467-017-01261-5
57. Wu, Y., Zeng, J., Zhang, F., Zhu, Z., Qi, T., Zheng, Z., Lloyd-Jones, L. R., Marioni, R. E., Martin, N. G., Montgomery, G. W., Deary, I. J., Wray, N. R., Visscher, P. M., McRae, A. F., & Yang, J. (2018). Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. *Nat Commun*, *9*(1), 918. https://doi.org/10.1038/s41467-018-03371-0
58. Yu, Y., Xu, Z., Zhou, H., Xu, R., Xu, J., Liu, W., Wu, Y., Qiu, Y., Zhang, G., Huang, X., & Chen, Y. (2024). RBP7 functions as a tumor suppressor in HR + breast cancer by inhibiting the AKT/SREBP1 pathway and reducing fatty acid. *Cancer Cell Int*, *24*(1), 118. https://doi.org/10.1186/s12935-024-03299-0
59. Zhao, J., Ming, J., Hu, X., Chen, G., Liu, J., & Yang, C. (2020). Bayesian weighted Mendelian randomization for causal inference based on summary statistics. *Bioinformatics*, *36*(5), 1501-1508. https://doi.org/10.1093/bioinformatics/btz749
60. Zhu, Z., Zhang, F., Hu, H., Bakshi, A.,

Robinson, M. R., Powell, J. E., Montgomery, G. W., Goddard, M. E., Wray, N. R., Visscher, P. M., & Yang, J. (2016). Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet*, *48*(5), 481-487. https://doi.org/10.1038/ng. 3538