



## OPEN ACCESS

EDITED BY  
Zhangsheng Yu,  
Shanghai Jiao Tong University, China

REVIEWED BY  
Sriraman Devarajan,  
Dasman Diabetes Institute, Kuwait  
Xiaorong Yang,  
Qilu Hospital, Shandong University,  
China  
Jingru Yu,  
Karolinska Institutet (KI), Sweden

\*CORRESPONDENCE  
Fangmeng Fu,  
✉ [ffm@fjmu.edu.cn](mailto:ffm@fjmu.edu.cn)  
Haomin Yang,  
✉ [haomin.yang@ki.se](mailto:haomin.yang@ki.se)

<sup>†</sup>These authors have contributed equally  
to this work

SPECIALTY SECTION  
This article was submitted to Statistical  
Genetics and Methodology,  
a section of the journal  
Frontiers in Genetics

RECEIVED 27 July 2022  
ACCEPTED 09 December 2022  
PUBLISHED 04 January 2023

CITATION  
Yu X, Zhang Y, Lin Y, Zou S, Zhu P,  
Song M, Fu F and Yang H (2023), The  
association between plasma  
chemokines and breast cancer risk and  
prognosis: A mendelian  
randomization study.  
*Front. Genet.* 13:1004931.  
doi: 10.3389/fgene.2022.1004931

COPYRIGHT  
© 2023 Yu, Zhang, Lin, Zou, Zhu, Song,  
Fu and Yang. This is an open-access  
article distributed under the terms of the  
[Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is  
permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original  
publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or  
reproduction is permitted which does  
not comply with these terms.

# The association between plasma chemokines and breast cancer risk and prognosis: A mendelian randomization study

Xingxing Yu<sup>1†</sup>, Yanyu Zhang<sup>1†</sup>, Yuxiang Lin<sup>2,3,4</sup>, Shuqing Zou<sup>1</sup>,  
Pingxiu Zhu<sup>1</sup>, Mengjie Song<sup>1</sup>, Fangmeng Fu<sup>2,3,4\*</sup> and  
Haomin Yang<sup>1\*</sup>

<sup>1</sup>Department of Epidemiology and Health Statistics, School of Public Health, Fujian Medical University, Fuzhou, China, <sup>2</sup>Department of Breast Surgery, Fujian Medical University Union Hospital, Fuzhou, China, <sup>3</sup>Department of General Surgery, Fujian Medical University Union Hospital, Fuzhou, China, <sup>4</sup>Breast Cancer Institute, Fujian Medical University, Fuzhou, China

**Background:** Despite the potential role of several chemokines in the migration of cytotoxic immune cells to prohibit breast cancer cell proliferation, a comprehensive view of chemokines and the risk and prognosis of breast cancer is scarce, and little is known about their causal associations.

**Methods:** With a two-sample Mendelian randomization (MR) approach, genetic instruments associated with 30 plasma chemokines were created. Their genetic associations with breast cancer and its survival by molecular subtypes were extracted from the recent genome-wide association study of 133,384 breast cancer cases and 113,789 controls, with available survival information for 96,661 patients. We further tested the associations between the polygenic risk score (PRS) for chemokines and breast cancer in the UK Biobank cohort using logistic regression models, while the association with breast cancer survival was tested using Cox regression models. In addition, the association between chemokine expression in tumors and breast cancer survival was also analyzed in the TCGA cohort using Cox regression models.

**Results:** Plasma CCL5 was causally associated with breast cancer in the MR analysis, which was significant in the luminal and HER-2 enriched subtypes and further confirmed using PRS analysis (OR = 0.94, 95% CI = 0.89–1.00). A potential causal association with breast cancer survival was only found for plasma CCL19, especially for ER-positive patients. Although not replicated in the UK Biobank, we still found an inverse association between CCL19 expression in tumors and breast cancer overall and relapse-free survival in the TCGA cohort (HR = 0.58, 95% CI = 0.35–0.95).

**Conclusion:** We observed an inverse association between genetic predisposition to CCL5 and breast cancer, while CCL19 was associated with breast cancer survival. These associations suggested the potential of these chemokines as tools for breast cancer prevention and treatment.

## KEYWORDS

breast cancer, chemokine, mendelian randomization, molecular subtypes of breast cancer, survival

## Introduction

Breast cancer is one of the most common malignant tumors, which seriously affects the physical and mental health of women worldwide and accounts for 24.2% and 15.0% of the total number of new cancers and deaths in women, respectively (Bray et al., 2018). Although breast cancer has caused a considerable global burden, its etiology is not yet completely understood. Currently, the known breast cancer risk factors include age, sex, race, family history, genetic mutations, early menarche/late menopause and hormone replacement therapy, while evidence for the causal effect of inflammation in breast cancer carcinogenesis is still limited (Winters et al., 2017; Deshmukh et al., 2019). To prevent the incidence of breast cancer and find potential therapeutic targets, we need to better understand the role of inflammatory biomarkers in breast cancer.

At the first stage of malignant evolution, chronic inflammation may play an important role during the initiation, development and prognosis of breast cancer (Danforth, 2021), in which chemokines might regulate the migration of immune cells toward the tumor and act as biomarkers for breast cancer. Studies have shown that CCL-4, -7, -8, -11, -15, -16, -19, -22, -23, -24, -25 and CXCL-5, -8, -9, and -16 were increased in breast cancer compared with normal adjacent breast tissue (Abrahamsson et al., 2017), and genetic polymorphisms near genes encoding CCL-2, -4, -5, -20, -22, CXCL-10, -12 and XCL-1 were found to be associated with the risk and prognosis of breast cancer (Eskandari-Nasab et al., 2014; Li et al., 2014; Jafarzadeh et al., 2015; Jafarzadeh et al., 2016; Stephens et al., 2017; He and Zhang, 2018; Hu et al., 2018; Chou et al., 2020). However, traditional observational studies are susceptible to reverse causation and residual confounding (Smith and Ebrahim, 2003). Contrary to previous studies, Bowen Chen et al. (2020) used several bioinformatics analysis tools to analyze the expression data of CC chemokines in patients with breast cancer and yielded inconsistent results. Therefore, it is still a challenge to fully understand the causal association between chemokines and breast cancer.

Mendelian randomization (MR) uses genetic variation as an instrumental variable (IV) to infer the association between exposure factors and disease outcomes. Because the genotypes are presumed to be randomly allocated in the process of gamete formation, MR analyses are largely free of reverse causality and less sensitive to confounding factors, which overcomes the limitations of conventional observational studies (Smith and Ebrahim, 2003; Davies et al., 2018).

In the present study, we implemented a two-sample MR design to investigate the potential causal effect of plasma chemokines on the risk and prognosis of breast cancer and to analyze their molecular subtypes to provide more credible evidence

on the potential role of chemokines in the development of breast cancer. We further used the UK Biobank and TCGA cohorts to reaffirm our findings.

## Methods

### Study population

Mendelian randomization analysis was used to assess the causal association between plasma chemokines and breast cancer. In the first step, the associations were tested using a two-sample approach with summary statistics from genome-wide association studies (GWAS), which does not need individual-level data (Pierce and Burgess, 2013). Genetic instrumental variables for plasma chemokines were created by identifying single nucleotide polymorphisms (SNPs) associated with 41 plasma chemokines in a recent GWAS on the human plasma proteome at the GWAS significant level ( $p < 5 \times 10^{-8}$ ) (Sun et al., 2018). From the 5,965 SNPs reported in the original study, 91 SNPs for 30 plasma chemokines were selected after excluding SNPs with linkage disequilibrium (LD threshold:  $r^2 < 0.01$ ). The SNPs used in this study and their corresponding plasma chemokines and proportion of variance explained are listed in [Supplementary Table S1](#). We used  $R^2$  and F statistics to estimate the proportion of variance explained by SNPs, with the following formulas

$$R^2 = \frac{2\beta^2 * (MAF) * (1-MAF)}{2\beta^2 * (MAF) * (1-MAF) + (se(\beta))^2 * 2N * MAF * (1-MAF)}$$

And  $F = \frac{R^2 * (N-1-k)}{(1-R^2) * k}$ , in which  $\beta$  is the coefficient for the SNP, MAF is the minor allele frequency,  $se(\beta)$  is the standard error of the coefficient, N is the sample size and k is the number of SNPs used.

GWAS summary statistics for breast cancer overall and by molecular subtypes (luminal A, luminal B, Her-2 enriched and triple negative) were downloaded from the Breast Cancer Association Consortium (Zhang et al., 2020), with 133,384 breast cancer cases and 113,789 controls. The GWAS summary statistics for breast cancer survival were also downloaded from BCAC, including 96,661 patients with 7,697 breast cancer mortality cases; among them, 64,171 were ER-positive patients, and 16,172 were ER-negative (Escala-Garcia et al., 2019). Among the 91 selected SNP, 7 SNPs were removed during the harmonization process of the dataset, because they were palindromic with intermediate allele frequencies ([Supplementary Table S1](#)), leaving 84 SNPs in the final analysis. Heterogeneity among genetic variants was quantified by  $I^2$  statistic ([Supplementary Table S2](#)). All studies involved in the GWAS on the human plasma proteome and the Breast Cancer Association Consortium were approved by the

relevant research ethics committee, and all participants provided informed consent.

For those identified plasma chemokines in the two sample Mendelian randomization studies for breast cancer, Mendelian randomization was further applied to validate the findings in UK Biobank. Briefly, the UK Biobank is a prospective cohort comprising approximately 502,505 participants (55% women) of various ethnicities, aged 40–69 at recruitment between 2006 and 2010. At enrolment, information on the participants' health-related factors and socio-demographic characteristics and biological samples were collected in 22 assessment centers around United Kingdom.

Blood samples from women in the UK Biobank were genotyped using a custom UK Biobank Axiom Array, comprising 825,927 SNPs. Details of the array design, sample handling, quality control processes and imputation of non-genotyped variants are described elsewhere (Bycroft et al., 2018). Breast cancer cases were obtained through linkage with the national cancer registries in the United Kingdom and updated until 2015. The ICD-10 code C50 and ICD-9 code 174 were used to identify breast cancer cases, while the controls were women without a diagnosis of breast cancer. To assess whether breast cancer is associated with a genetic predisposition to plasma chemokines, we selected the same genome-wide significant SNPs used in the two-sample MR. For all individuals, a weighted polygenic risk score (PRS) was calculated using the following formula:

$$\text{PRS} = \beta_1 x_1 + \beta_2 x_2 + \beta_k x_k + \beta_n x_n$$

where  $\beta$  is the per-allele log odds ratio (OR) of the breast cancer-associated risk allele for SNP  $k$ ,  $x_k$  is the number of alleles for the same SNP (0, 1, 2), and  $n$  is the total number of breast cancer SNPs included in the profile. For the analyses, PRS was categorized into tertiles.

To further explore the effect of the identified chemokines associated with breast cancer survival, we explored the association between PRS of the chemokines and breast cancer mortality in the UK Biobank. After that, 1,077 women in the TCGA breast cancer cohort were followed for their overall and breast cancer survival. Gene expression data of the chemokines in tumor tissues were also downloaded and categorized into tertiles.

## Statistical analysis

Several methods were used to estimate the potential causal effect of plasma chemokines on breast cancer and its prognosis. The main analysis was based on a fixed-effects meta-analysis using an inverse-variance weighting model (IVW), in which  $\beta$  and  $se$  were calculated using the following formulas  $\beta_{IVW} = \frac{\sum_{i=1}^N \beta_{exp,i} \beta_{out,i} \sigma_{out,i}^2}{\sum_{i=1}^N \beta_{exp,i} \sigma_{out,i}^2}$  and  $se(\beta_{IVW}) = \frac{1}{\sum_{i=1}^N \beta_{exp,i} \sigma_{out,i}^2}$ :  $N$  is the total number of SNPs,  $\beta_{exp,i}$  is the coefficient of the  $i$ -th SNP on the exposure,  $\beta_{out,i}$  is the coefficient of the  $i$ -th SNP on the outcome, and  $\sigma_{out,i}$  is the standard error of  $\beta_{out,i}$ . If only one SNP was identified for the chemokine, the Wald ratio method was used by

calculating the ratio of  $\beta_{exp}$  and  $\beta_{out}$ , while  $se$  could be approximately estimated using the Delta method (Pagoni et al., 2019). We further performed sensitivity analysis to test horizontal pleiotropy using weighted median and Egger regression methods. The weighted median estimator is consistent with the IVW estimation even when up to 50% of the SNPs are invalid instrumental variables, while MR-Egger regression simply adds an intercept term of pleiotropy into the IVW regression model (Bowden et al., 2016).

To test the key assumption of no horizontal pleiotropy in MR, we further analyzed the effect of plasma chemokines on potential risk factors for breast cancer, e.g., age at menarche, age at menopause, age at smoking initiation, and BMI, using their corresponding GWAS summary statistics (Wood et al., 2016; Liu et al., 2019; Elsworth et al., 2020).

To validate the discovered potential causal association between plasma chemokines and breast cancer, we used logistic regression models to calculate the ORs of breast cancer among UK biobank women by tertiles of PRS for plasma chemokines. These associations were adjusted for age, the first ten principal components, and the assessment center. We additionally adjusted age at diagnosis, year of breast cancer diagnosis in the Cox regression model to validate the association between PRS for chemokines and breast cancer survival in the UK biobank. We also used the Cox model to estimate the hazard ratios of overall survival and relapse-free survival in the TCGA cohort by tertiles of gene expression of the chemokines in tumor tissues and adjusted for age at diagnosis, estrogen, progesterone, and tumor stage. Missingness in the tumor characteristics were categorized with missing indicator.

All statistical tests were two-sided and performed using R 3.6.3 and Stata 15.1.

## Results

In the main analysis, the selected SNPs explained 0.9%–36.3% of the variance in plasma chemokines, and the corresponding F-statistic ranged from 29.7 to 1,879.3 (Supplementary Table S1), suggesting no evidence of weak instrument bias. A higher plasma CCL5 was inversely associated with breast cancer (OR = 0.89, 95% CI: 0.85 to 0.94,  $p < 0.001$ ), which was primarily attributed to the luminal A and luminal B subtypes (OR for luminal A = 0.88, 95% CI: 0.82 to 0.94; OR for luminal B = 0.80, 95% CI: 0.69–0.94). Plasma growth-regulated alpha protein and Ck-beta-8 were also associated with breast cancer, although the  $p$ -value did not survive the Bonferroni correction for multiple testing (Figure 1; Supplementary Table S3).

In the subtype-specific analysis, plasma CCL19 was associated with luminal A breast cancer, while CXCL16 was associated with HER-2-enriched breast cancer. In addition, plasma CCL16 was positively associated with luminal A breast cancer but inversely associated with triple-negative breast cancer. Sensitivity analysis using MR Egger and weighted median methods yielded similar results (Supplementary Table S3).

**TABLE 1** The association between polygenic risk scores for chemokines and breast cancer in the UK Biobank.

Chemokine	No.	No. of cases	Or (95% CI)	p-Value
<b>Growth-regulated alpha protein</b>				
Tertile 1	131,908	7,461	1.00 (REF)	
Tertile 2	91,150	5,184	1.02 (0.98–1.06)	0.251
Tertile 3	41,129	2,115	0.97 (0.92–1.02)	0.194
<b>C-C motif chemokine 5</b>				
Tertile 1	136,038	7,636	1.00 (REF)	
Tertile 2	98,239	5,503	0.98 (0.95–1.02)	0.273
Tertile 3	29,910	1,621	<b>0.94 (0.89–1.00)</b>	0.045
<b>Ck-beta-8-1</b>				
Tertile 1	91,937	5,073	1.00 (REF)	
Tertile 2	121,826	6,799	0.97 (0.93–1.00)	0.073
Tertile 3	50,424	2,888	0.99 (0.94–1.04)	0.664

The model was adjusted for age at recruitment, assessment center and top 10 principal genetic components. Statistically significant results with  $p$ -value < 0.05 are bolded.

To validate the findings in the two-sample MR analysis, 264,187 women who had available genotyped data from the UK Biobank were involved in the analysis, with 14,760 breast cancer patients. In multivariable logistic regression analysis, a higher polygenic risk scores for plasma CCL5 was inversely associated with breast cancer (OR for the highest tertile = 0.94, 95% CI = 0.89–1.00,  $p = 0.045$ ). However, no statistically significant associations for plasma growth-regulated alpha protein, Ck-beta-8 and breast cancer were found. (Table 1).

In the analysis of breast cancer survival, we found a potential causal effect of plasma CCL19 on breast cancer survival (HR = 0.85, 95% CI = 0.75–0.96,  $p < 0.01$ ), especially for ER-positive breast cancer patients (HR = 0.82, 95% CI = 0.70–0.97) (Figure 2; Supplementary Table S4). However, when we further investigated the association between PRS of CCL19 and breast cancer mortality in the UK Biobank cohort, no statistically significant association was found between them (Table 2). Despite this, in the TCGA breast cancer patient cohort, CCL19 expression in tumor tissues was associated with increased overall and relapse-free survival (HR for the last tertile vs. first tertile = 0.58, 95% CI = 0.35–0.95). This association was also observed for ER-positive patients (Table 3).

MR estimates for the effects of plasma chemokines on breast cancer risk factors indicated little evidence of horizontal pleiotropy (Supplementary Table S5).

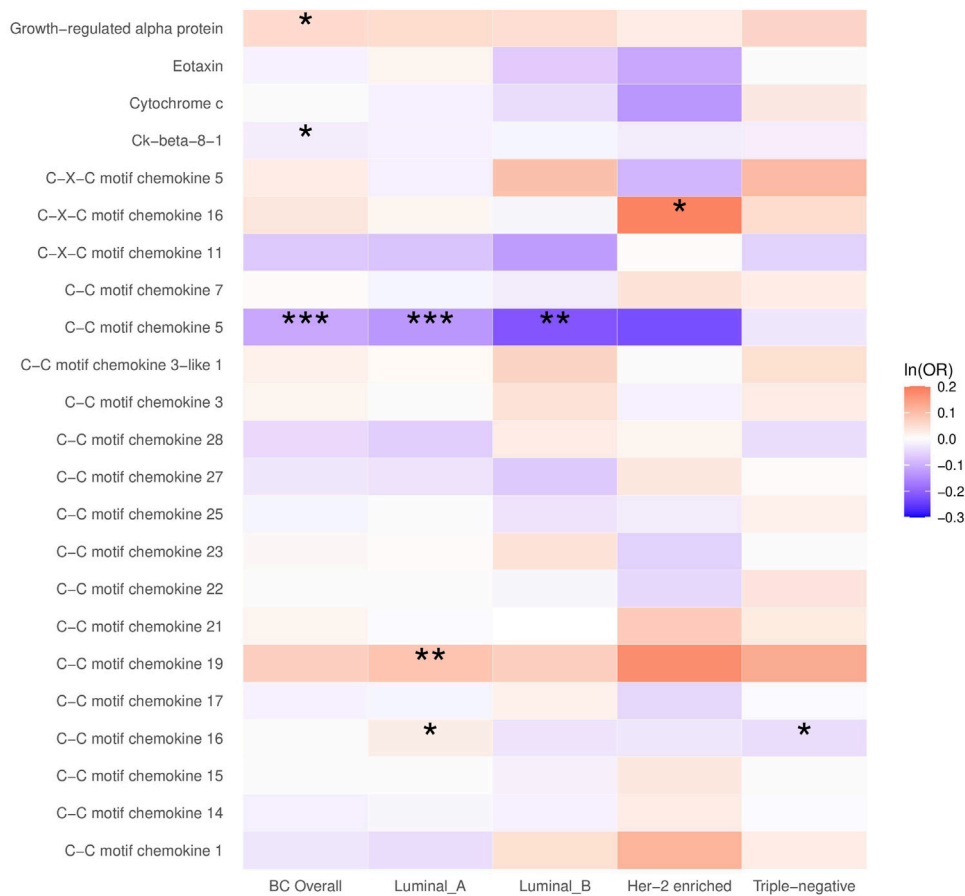
## Discussion

In this Mendelian randomization study of 30 plasma chemokines, we found an inverse association between genetically predicted plasma CCL5 and breast cancer in both the two-sample and one-sample analyses. This association was

more pronounced in luminal and HER-2-enriched breast cancer patients. An effect on cancer survival was only found for plasma CCL19, especially for ER-positive patients. In addition, we also found an inverse association between CCL19 expression and breast cancer mortality using tumor tissue samples.

As key proteins in immune cell migration, chemokines have been studied for their effect on inflammation-associated cancers. One previous Mendelian randomization analysis on cytokines found an association between genetically predicted CCL2, CCL4, and GRO $\alpha$  and breast cancer. As the SNPs associated with plasma CCL2 and CCL4 did not exceed the GWAS significance threshold, we were not able to replicate these findings in our study. However, both our study and a previous study confirmed the association between GRO $\alpha$  and breast cancer, reaffirming the role of GRO $\alpha$  in breast cancer carcinogenesis and probably through the effect of Th17 cells (Ma et al., 2018).

Our analysis found an inverse association between CCL5 and breast cancer, with a similar effect size in a recent phenome-wide Mendelian randomization (Zheng et al., 2020). We further suggested that the association was driven by luminal and HER-2-enriched subtypes of breast cancer. CCL5 might play an antitumor role as it is negatively related to the number of tumor associated macrophages (Fujimoto et al., 2009), which are the major components of the breast tumor stroma and contribute to tumor growth and development. In addition, many autoimmune diseases are associated with increased levels of CCL5 and inversely associated with breast cancer risk (Wadström et al., 2020; Zeng et al., 2022). All the above evidence supported the role of plasma CCL5 in breast tumor carcinogenesis. Considering its strong correlation with the pathogenesis of luminal and HER-2 enriched breast cancer, CCL5 might be used as a



**FIGURE 1**

The causal association between plasma chemokines and breast cancer risk overall and by molecular subtype. OR: Odds ratio. Statistical significance is indicated by \*\*\* ( $p < 0.001$ ), \*\* ( $p < 0.01$ ) and \* ( $p < 0.05$ ). The estimates are based on the fixed effect of the inverse variance weighted method or Wald ratio method for meta-analyzing individual SNP results. ORs show the risk of breast cancer overall and by subtypes per SD increase of ranked-inverse normalized concentration in exposure to plasma chemokines.

potential biomarker for breast cancer screening and risk stratification.

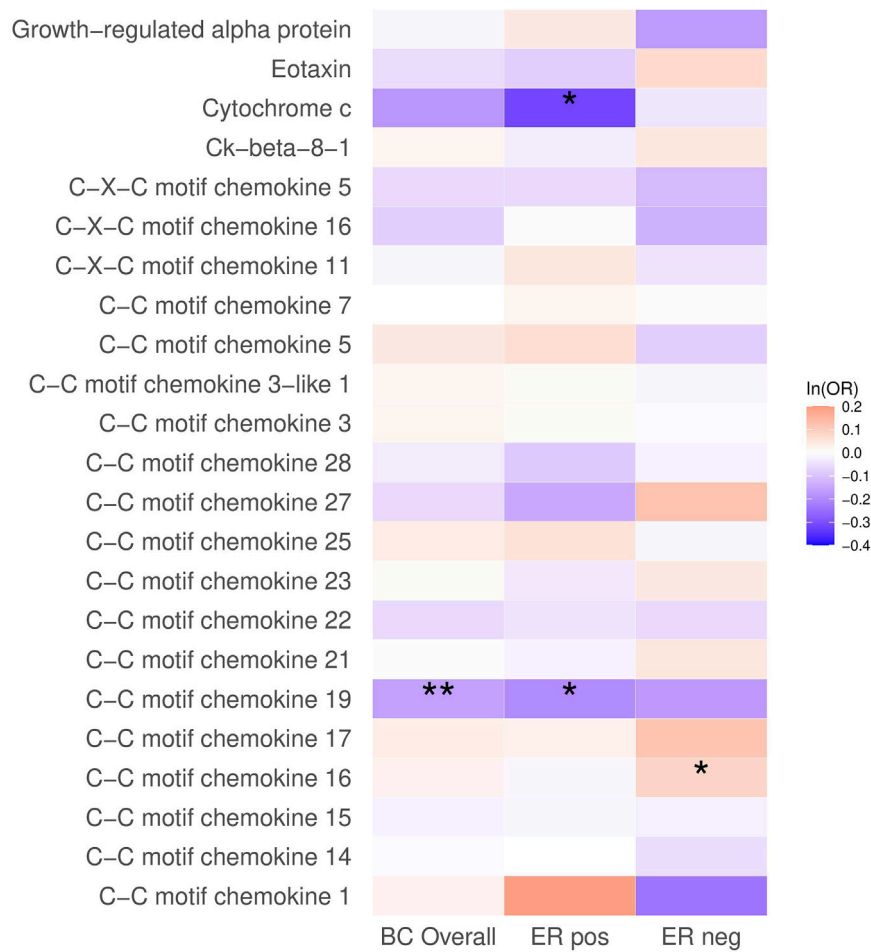
In addition, both genetically predicted plasma CCL19 and CCL19 gene expression in tumors was associated with increased survival of breast cancer in our study. As downregulation of the CCL19/CCR7 axis has been observed in metastatic breast cancer compared to the primary tumor (Szekely et al., 2018) and CCL19 suppresses angiogenesis and proliferation of several other types of cancers (Xu et al., 2018; Zhou et al., 2020), it is biologically plausible that CCL19 inhibits breast cancer metastasis through similar AIM2 and Met/ERK/Elk-1/HIF-1 $\alpha$ /VEGF-A pathways (Chen et al., 2006) and therefore increases patient survival. Since we did not replicate the association between CCL19 and breast cancer in the UK biobank, this finding should be interpreted with caution.

Interestingly, we also found that many of the associations between chemokines and breast cancer or its survival were more pronounced in luminal A- or ER-positive cancers, suggesting that

the effect of chemokines might be related to estrogen-related carcinogenesis (Lim and Lin, 2021). However, we still found a potential causal association between CCL16 and triple-negative breast cancer. Further study is therefore needed to confirm this association and test its potential role as a biomarker or treatment target for triple-negative breast cancer.

The strength of our study included the use of the most recent and largest summary statistics for breast cancer with subtype-specific information and a comprehensive dataset for plasma chemokines. Our study had several limitations. As no sex-specific GWAS for chemokines has been reported, we assumed that the sex difference in the chemokine GWAS was small and that the original GWAS had already adjusted for age, sex and principal components. In addition, as both the BCAC and the UK biobank dataset included prevalent breast cancer patients, it is possible that the association between chemokines and breast cancer may suffer from survivorship bias, and may not represent patients with more aggressive





**FIGURE 2**

The causal association between plasma chemokines and breast cancer survival. HR: hazard ratio. Statistical significance is indicated by \*\* ( $p < 0.01$ ) and \* ( $p < 0.05$ ). The estimates are based on the fixed effect of the inverse variance weighted method or Wald ratio method for meta-analyzing individual SNP results. HRs show the risk of breast cancer overall and by subtypes per SD increase of ranked-inverse normalized concentration in exposure to plasma chemokines.

**TABLE 2 The association between polygenic risk scores for CCL19 and breast cancer mortality in the UK Biobank.**

Chemokine	No. of cases	No. of mortality	HR (95% CI)	p-value
C-C motif chemokine 19				
Tertile 1	6,014	342	1.00 (REF)	
Tertile 2	3,980	228	1.01 (0.86–1.20)	0.868
Tertile 3	4,766	283	1.05 (0.90–1.23)	0.541
Standardized continuous	14,760	853	1.01 (0.94–1.08)	0.871

The model was adjusted for age at diagnosis, year of breast cancer diagnosis, assessment centers and top 10 principal genetic components.

tumors. However, results by different molecular subtypes still indicated some associations with HER-2 enriched and triple negative cancer. Finally, CCL19 can be secreted by cells other than cancer cells, while in our study only CCL19 in tumor

tissues were used for the validation of its association with breast cancer survival. The generalizability of the findings is also affected by the use of European ancestry, and the results should be replicated in other populations.

**TABLE 3** The association between CCL19 and survival of breast cancer patients in TCGA cohort.

	Total no.	Overall survival			Relapse free survival		
		No. of cases	HR (95% CI)	<i>p</i> -value	No. of cases	HR (95% CI)	<i>p</i> -value
<b>Breast cancer overall</b>							
Tertile 1	358	64	1.00 (REF)		39	1.00 (REF)	
Tertile 2	358	46	<b>0.64 (0.43–0.96)</b>	0.029	23	<b>0.49 (0.29–0.85)</b>	0.011
Tertile 3	357	43	<b>0.66 (0.43–1.00)</b>	0.051	34	<b>0.58 (0.35–0.95)</b>	0.032
<b>ER + breast cancer</b>							
Tertile 1	264	46	1.00 (REF)		26	1.00 (REF)	
Tertile 2	275	33	<b>0.52 (0.32–0.84)</b>	0.008	19	<b>0.51 (0.27–0.97)</b>	0.039
Tertile 3	249	22	<b>0.44 (0.26–0.77)</b>	0.003	17	<b>0.43 (0.22–0.82)</b>	0.010
<b>ER- breast cancer</b>							
Tertile 1	72	12	1.00 (REF)		10	1.00 (REF)	
Tertile 2	72	10	0.92 (0.38–2.24)	0.854	4	0.40 (0.12–1.38)	0.149
Tertile 3	94	20	1.09 (0.49–2.44)	0.832	16	0.81 (0.33–1.98)	0.648

The model was adjusted for age at diagnosis, ER, PR, and stage. Statistically significant results with *p*-value < 0.05 are bolded.

## Conclusion

In conclusion, we found that lower plasma CCL5 was causally associated with breast cancer, while plasma and tumor-derived CCL19 might be associated with increased survival of breast cancer patients. Both findings were more pronounced in ER-positive patients, indicating the potential use of CCL5 and CCL19 as biomarkers or drug targets in future studies for breast cancer prevention and treatment.

## Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: Summary statistics for pQTLs of plasma chemokines can be downloaded from the IEU open GWAS project (<https://gwas.mrcieu.ac.uk/>, with the GWAS ID from prot-a-387 to prot-a-411 and prot-a-738 to prot-a-755). Summary statistics for breast cancer and survival by subtypes are available on the BCAC website. (<https://bcac.ccge.medschl.cam.ac.uk/bcacdata/oncoarray/oncoarray-and-combined-summary-result/gwas-summary-associations-breast-cancer-risk-2020/> and <https://bcac.ccge.medschl.cam.ac.uk/bcacdata/oncoarray/oncoarray-and-combined-summary-result/gwas-summary-results-survival-2019/>). Data from the UK Biobank (<http://www.ukbiobank.ac.uk/>) are available to all researchers upon making an application. Part of this research was conducted using the UK Biobank Resource under Application 61083.

## Ethics statement

The studies involving human participants were reviewed and approved by the Institutional Review Board of Fujian Medical University (No. 2020KY055). The patients/participants provided their written informed consent to participate in this study.

## Author contributions

XY and HY had full access to all data and take responsibility for the integrity of the data and the accuracy of the analysis. Study concept and design: HY acquisition, analysis, and interpretation of data: XY, SZ, and PZ Drafting of the manuscript: XY, MS, and YZ Critical revision of the manuscript for important intellectual content: HY, XY, YL, and FF Statistical analysis: XY and YZ Funding: HY.

## Funding

HY is supported by the Natural Science Foundation of Fujian Province (Grant no: 2021J01721), the Startup Fund for High-level Talents of Fujian Medical University (Grant no: XRCZX2020007), Startup Fund for Scientific Research, Fujian Medical University (Grant no: 2019QH1002) and Laboratory Construction Program of Fujian Medical University (Grant no: 1100160208). The sponsor of the study had no role in the study

design, data collection, data analysis, data interpretation, or writing of the report.

## Acknowledgments

We want to acknowledge the participants and investigators of the BCAC consortium.

## 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.

## References

- Abrahamsson, A., Rzepecka, A., and Dabrosin, C. (2017). Equal pro-inflammatory profiles of CCLs, CXCLs, and matrix metalloproteinases in the extracellular microenvironment *in vivo* in human dense breast tissue and breast cancer. *Front. Immunol.* 8, 1994. doi:10.3389/fimmu.2017.01994
- Bowden, J., Davey Smith, G., Haycock, P. C., and Burgess, S. (2016). Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. *Genet. Epidemiol.* 40 (4), 304–314. doi:10.1002/gepi.121965
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68 (6), 394–424.
- Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., et al. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature* 562 (7726), 203–209. doi:10.1038/s41586-018-0579-z
- Chen, B., Zhang, S., Li, Q., Wu, S., He, H., and Huang, J. (2020). Bioinformatics identification of CCL8/21 as potential prognostic biomarkers in breast cancer microenvironment. *Biosci. Rep.* 40 (11), BSR20202042. doi:10.1042/BSR20202042
- Chen, I-F., Ou-Yang, F., Hung, J-Y., Liu, J-C., Wang, H., Wang, S-C., et al. (2006). AIM2 suppresses human breast cancer cell proliferation *in vitro* and mammary tumor growth in a mouse model. *Mol. Cancer Ther.* 5 (1), 1–7. doi:10.1158/1535-7163.MCT-05-0310
- Chou, W. C., Hsiung, C. N., Chen, W. T., Tseng, L. M., Wang, H. C., Chu, H. W., et al. (2020). A functional variant near XCL1 gene improves breast cancer survival via promoting cancer immunity. *Int. J. Cancer* 146 (8), 2182–2193. doi:10.1002/ijc.32855
- Danforth, D. N. (2021). The role of chronic inflammation in the development of breast cancer. *Cancers* 13 (15), 3918. doi:10.3390/cancers13153918
- Davies, N. M., Holmes, M. V., and Davey Smith, G. (2018). Reading mendelian randomisation studies: A guide, glossary, and checklist for clinicians. *BMJ* 362, k601. doi:10.1136/bmj.k601
- Deshmukh, S. K., Srivastava, S. K., Poojarla, T., Dyess, D. L., Holliday, N. P., Singh, A. P., et al. (2019). Inflammation, immunosuppressive microenvironment and breast cancer: Opportunities for cancer prevention and therapy. *Ann. Transl. Med.* 7 (20), 593. doi:10.21037/atm.2019.09.68
- Elsworth, B., Lyon, M., Alexander, T., Liu, Y., Matthews, P., Hallett, J., et al. (2020). The MRC IEU OpenGWAS data infrastructure. *bioRxiv* 2020, 244293. doi:10.1101/2020.08.10.244293
- Escala-Garcia, M., Guo, Q., Dörk, T., Canisius, S., Keeman, R., Dennis, J., et al. (2019). Genome-wide association study of germline variants and breast cancer-specific mortality. *Br. J. Cancer* 120 (6), 647–657. doi:10.1038/s41416-019-0393-x
- Eskandari-Nasab, E., Hashemi, M., Ebrahimi, M., Amininia, S., Bahari, G., Mashhadi, M. A., et al. (2014). Evaluation of CCL5 -403 G>A and CCR5 Δ32 gene polymorphisms in patients with breast cancer. *Cancer Biomark.* 14 (5), 343–351. doi:10.3233/CBM-140411
- Fujimoto, H., Sangai, T., Ishii, G., Ikehara, A., Nagashima, T., Miyazaki, M., et al. (2009). Stromal MCP-1 in mammary tumors induces tumor-associated

## 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/fgene.2022.1004931/full#supplementary-material>

macrophage infiltration and contributes to tumor progression. *Int. J. Cancer* 125 (6), 1276–1284. doi:10.1002/ijc.24378

He, S., and Zhang, X. (2018). The rs1024611 in the CCL2 gene and risk of gynecological cancer in asians: A meta-analysis. *World J. Surg. Oncol.* 16 (1), 34. doi:10.1186/s12957-018-1335-4

Hu, G. N., Tzeng, H. E., Chen, P. C., Wang, C. Q., Zhao, Y. M., Wang, Y., et al. (2018). Correlation between CCL4 gene polymorphisms and clinical aspects of breast cancer. *Int. J. Med. Sci.* 15 (11), 1179–1186. doi:10.7150/ijms.26771

Jafarzadeh, A., Fooladseresht, H., Minaee, K., Bazrafshani, M. R., Khosravimashizi, A., Nemati, M., et al. (2015). Higher circulating levels of chemokine CCL22 in patients with breast cancer: Evaluation of the influences of tumor stage and chemokine gene polymorphism. *Tumour Biol. J. Int. Soc. Oncodevelopmental Biol. Med.* 36 (2), 1163–1171. doi:10.1007/s13277-014-2739-6

Jafarzadeh, A., Fooladseresht, H., Nemati, M., Assadollahi, Z., Sheikhi, A., Ghaderi, A., et al. (2016). Higher circulating levels of chemokine CXCL10 in patients with breast cancer: Evaluation of the influences of tumor stage and chemokine gene polymorphism. *Cancer Biomark.* 16 (4), 545–554. doi:10.3233/CBM-160596

Li, J., Lindstrom, L. S., Foo, J. N., Rafiq, S., Schmidt, M. K., Pharoah, P. D., et al. (2020). 2q36.3 is associated with prognosis for oestrogen receptor-negative breast cancer patients treated with chemotherapy. *Nat. Commun.* 5, 4051. doi:10.1038/ncomms5051

Lim, C. L., and Lin, V. C. L. (2021). Estrogen markedly reduces circulating low-density neutrophils and enhances pro-tumoral gene expression in neutrophil of tumour-bearing mice. *BMC cancer* 21 (1), 1017. doi:10.1186/s12885-021-08751-2

Liu, M., Jiang, Y., Wedow, R., Li, Y., Brazel, D. M., Chen, F., et al. (2019). Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat. Genet.* 51 (2), 237–244. doi:10.1038/s41588-018-0307-5

Ma, K., Yang, L., Shen, R., Kong, B., Chen, W., Liang, J., et al. (2018). Th17 cells regulate the production of CXCL1 in breast cancer. *Int. Immunopharmacol.* 56, 320–329. doi:10.1016/j.intimp.2018.01.026

Pagoni, P., Dimou, N. L., Murphy, N., and Stergiakouli, E. (2019). Using Mendelian randomisation to assess causality in observational studies. *Evid. Based Ment. Health* 22 (2), 67–71. doi:10.1136/ebmental-2019-300085

Pierce, B. L., and Burgess, S. (2013). Efficient design for mendelian randomization studies: Subsample and 2-sample instrumental variable estimators. *Am. J. Epidemiol.* 178 (7), 1177–1184. doi:10.1093/aje/kwt084

Smith, G. D., and Ebrahim, S. (2003). Mendelian randomization: Can genetic epidemiology contribute to understanding environmental determinants of disease? *Int. J. Epidemiol.* 32 (1), 1–22. doi:10.1093/ije/dyg070

Stephens, K. E., Levine, J. D., Aouizerat, B. E., Paul, S. M., Abrams, G., Conley, Y. P., et al. (2017). Associations between genetic and epigenetic variations in cytokine genes and mild persistent breast pain in women



following breast cancer surgery. *Cytokine* 99, 203–213. doi:10.1016/j.cyto.2017.07.006

Sun, B. B., Maranville, J. C., Peters, J. E., Stacey, D., Staley, J. R., Blackshaw, J., et al. (2018). Genomic atlas of the human plasma proteome. *Nature* 558 (7708), 73–79. doi:10.1038/s41586-018-0175-2

Szekely, B., Bossuyt, V., Li, X., Wali, V. B., Patwardhan, G. A., Frederick, C., et al. (2018). Immunological differences between primary and metastatic breast cancer. *Ann. Oncol.* 29 (11), 2232–2239. doi:10.1093/annonc/mdy399

Wadström, H., Pettersson, A., Smedby, K. E., and Askling, J. (2020). Risk of breast cancer before and after rheumatoid arthritis, and the impact of hormonal factors. *Ann. Rheumatic Dis.* 79 (5), 581–586. doi:10.1136/annrheumdis-2019-216756

Winters, S., Martin, C., Murphy, D., and Shokar, N. K. (2017). Breast cancer epidemiology, prevention, and screening. *Prog. Mol. Biol. Transl. Sci.* 151, 1–32. doi:10.1016/bs.pmbts.2017.07.002

Wood, A. R., Tyrrell, J., Beaumont, R., Jones, S. E., Tuke, M. A., Ruth, K. S., et al. (2016). Variants in the FTO and CDKAL1 loci have recessive effects on risk of obesity and type 2 diabetes, respectively. *Diabetologia* 59 (6), 1214–1221. doi:10.1007/s00125-016-3908-5

Xu, Z., Zhu, C., Chen, C., Zong, Y., Feng, H., Liu, D., et al. (2018). CCL19 suppresses angiogenesis through promoting miR-206 and inhibiting Met/ERK/Elk-1/HIF-1 $\alpha$ /VEGF-A pathway in colorectal cancer. *Cell. Death Dis.* 9 (10), 974. doi:10.1038/s41419-018-1010-2

Zeng, Z., Lan, T., Wei, Y., and Wei, X. (2022). CCL5/CCR5 axis in human diseases and related treatments. *Genes. & Dis.* 9 (1), 12–27. doi:10.1016/j.gendis.2021.08.004

Zhang, H., Ahearn, T. U., Lecarpentier, J., Barnes, D., Beesley, J., Qi, G., et al. (2020). Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and subtype-specific analyses. *Nat. Genet.* 52 (6), 572–581. doi:10.1038/s41588-020-0609-2

Zheng, J., Haberland, V., Baird, D., Walker, V., Haycock, P. C., Hurle, M. R., et al. (2020). Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. *Nat. Genet.* 52 (10), 1122–1131. doi:10.1038/s41588-020-0682-6

Zhou, R., Sun, J., He, C., Huang, C., and Yu, H. (2020). CCL19 suppresses gastric cancer cell proliferation, migration, and invasion through the CCL19/CCR7/AIM2 pathway. *Hum. Cell.* 33 (4), 1120–1132. doi:10.1007/s13577-020-00375-1