ARTICLE
Cancer Advances 2025;8:e25024. https://doi.org/10.53388/2025825024

Exploring and validating genetic associations between immune cells and breast cancer

Jian-Ying Pei[1#], Chun Zhang[2#], Jing-Ting Liu[1], Chong Zhang[1], Yi Wang[3], Yan Li[4*]

[1]Clinical Laboratory Center, Gansu Provincial Maternity and Child Care Hospital, Lanzhou 730050, China.
[2]Clinical Laboratory, Lanzhou University of Technology Hospital, Lanzhou 730050, China.
[3]The School of Public Health, Gansu University of Chinese Medicine, Lanzhou 730000, China.
[4]Department of biochemistry and molecular biology, Medical College of Northwest Minzu University, Lanzhou 730030, China.

[#]These authors contributed equally to this work and are co-first authors for this paper.
[*]**Correspondence to:** Yan Li, Department of biochemistry and molecular biology, Medical College of Northwest Minzu University, No. 1, Northwest Xincun, Lanzhou 730030, China. E-mail: 290192798@xbmu.edu.cn.

Abstract

Background: Observational epidemiology studies suggested the crucial role of immune cells in breast cancer (BC) development. The causalities of immunophenotypes with BC remain ambiguous. **Methods:** We performed a two-sample Mendelian randomization (MR) analysis to investigate the potential causalities between immunophenotype traits and BC, and validated the findings using the flow-cytometry data of lymphocytes from a case-control study. The genome-wide association studies (GWAS) data on immunological traits were taken from a public catalog for 731 immunophenotypes, and the GWAS data in 6 cohorts of BC, which included malignant neoplasm of breast and carcinoma in situ of breast, were retrieved from the FinnGen database. The case-control study was conducted at Gansu Provincial Maternity and Child Care Hospital from January 2024 to May 2025, and included 123 BC patients and 109 healthy controls. **Results:** After investigating genetically predicted immunophenotype biomarkers, we discovered 24 highly correlated immunophenotypes and 83 suggestive possible factors. The case-control study effectively confirmed the differences in absolute count, relative count, and MFI in some highly correlated lymphocytes between BC patients and healthy controls. The study underscored that the absolute counts of lymphocytes, T cells, and $CD45RA^{+}$ $CD8^{br}$ are related to a reduced risk of BC, but CCR7 on naive $CD8^{br}$ showed a promoting effect on the pathogenesis of BC. **Conclusion:** Our findings demonstrate a complex genetic predisposition linking immunophenotypes to BC and provide preliminary experimental support. This study highlights potential avenues for immunotherapy but warrants further validation in larger, multi-ethnic cohorts.

Keywords: breast cancer; immune cells; mendelian randomization; case-control study; causality

Highlights

1. This work bridges the gap by investigating the causative association between immune cells and breast cancer (BC).
2. The association between immunophenotype traits and BC is causative.
3. The highly correlative immunophenotypes of BC were verified in our cohort.

Author contributions

Zhang C and Liu JT conceived the idea for the paper and conducted the analysis. Liu JT and Zhang C wrote the manuscript. Pei JY and Wang Y obtained data and drafted the figure. Li Y and Pei JY supervised the study and reviewed the article. All authors read and approved the final manuscript.

Competing interests

The authors declare no conflicts of interest.

Acknowledgments

This research was funded by the Natural Science Foundation of Gansu Province (grant number: 24JRRA937 and 25JRRA1234), grants from the National Natural Science Foundation of China (grant number: 82360469), and Research Project for Introduced Talents of Northwest Minzu University (grant number: xbmuyjrc2023020). This study was conducted using the GWAS database and the FinnGen consortium. We would like to thank all participants, the GWAS and FinnGen consortium team, for their contribution.

Abbreviations

AC, absolute cell; BC, breast cancer; DCs, dendritic cells; DCIS, ductal carcinoma in situ; FDR, false discovery rate; GWAS, genome-wide association study; HLA, human leukocyte antigen; IVs, instrumental variables; IVW, inverse variance weighted; LD, linkage disequilibrium; MFI, median fluorescence intensity; MR, Mendelian randomization; NK, natural killer; OR, odds ratio; RC, relative cell; SNP, single nucleotide polymorphism; Tregs, regulatory T cells; ROC, receptor characteristics; AUC, area under the curve; TNBK, T cells, B cells and NK cells.

Citation

Pei JY, Zhang C, Liu JT, Zhang C, Wang Y, Li Y. Exploring and validating genetic associations between immune cells and breast cancer. *Cancer Adv*. 2025;8:e25024. doi: 10.53388/2025825024.

Peer review information

Cancer Advances thanks Hong-Yu Li and other anonymous reviewers for their contribution to the peer review of this paper.

Executive editor: Jian Jia.

Received: 29 August 2025; **Revised:** 28 September 2025; **Accepted:** 29 October 2025; **Available online:** 06 November 2025.

Introduction

Breast cancer (BC) was the most common cancer in women globally. According to the latest estimates of the global burden of cancer released by the World Health Organization (WHO)'s cancer agency, the International Agency for Research on Cancer, BC makes up nearly 2.3 million new cases, accounting for 11.6% of the total new cases, leading to 666,000 deaths (6.9%) [1]. Research indicates that BC risk was elevated by lifestyle factors such as being overweight, later age at menopause, advanced age at first birth, less breastfeeding, hormone-replacement therapy, oral contraceptive, and alcohol intake [2, 3]. In terms of primary prevention, losing weight, reducing in alcohol consumption, along with increasing physical activity and breastfeeding, may help lower the incidence of BC [4, 5]. Since there aren't many known modifiable risk factors for the condition, the goal of BC control has been to make early detection and timely, comprehensive cancer management. The fact has gradually come to be recognized that BC is a complicated disease that has various aetiologies, a unique development process with distinct molecular and phenotypic backgrounds, and clinical outcomes. Nevertheless, a thorough understanding of the BC risk factors remains insufficient.

The human immune system is one of the most intricate systems that performs a wide range of functions that help to regulate metabolic traits and maintain energy homeostasis. Recent literature has gradually suggested that the tumor microenvironment (TME) plays a critical role in BC [6–8]. The proportion of immune cells in breast tissue gradually rises from normal breast tissue to BC. Studies comparing the distribution of immune cells in BC and comparable normal breast tissue provide an illustration of this [7, 9]. Ruffell et al. evaluated the composition of tumor-infiltrating leukocytes in BC patients and identified that BC patients contained infiltrates dominated by T lymphocytes ($CD3\varepsilon^+$), with minor populations of natural killer (NK) cells ($CD3\varepsilon^-CD56^+NKG2D^+$) and B lymphocytes ($CD19/20^+HLA\text{-}DR^+CD3^-$). In contrast, the normal breast tissue showed more myeloid-lineage cells, such as macrophages ($CD14^{hi}CD11b^+HLA\text{-}DR^+$), mast cells ($Fc\varepsilon R1\alpha^+CD117^+CD11b^-CD49d^+$), and neutrophils ($CD15^+CD11b^+CD49d^-$) [10]. At the transcriptional level, Azizi et al. used single-cell RNA sequencing (scRNA-seq) to compare the distribution of immune cells in BC and paired normal tissue, and discovered an increase in the diversity of cell states, as seen by an increase in gene expression variation among T cells, monocytes, and NK cells [7]. The result of scRNA-seq analyses conducted by Ding et al. also showed that there were 4 major immune cell clusters (including T cells, B cells, NK cells, and myeloid cells) in BC, and T cells constituted the largest infiltrating immune cell cluster [11]. These subtypes can impact tumor growth directly by $CD4^+$ and $CD8^+$ cell-mediated cytotoxicity, or indirectly by stimulating or suppressing the immune system through secreted cytokines, growth factors, and other agents. Their distribution and features may also differ based on the BC subtype, their sensitivity to estrogen, the status of mutations, and the development of tertiary lymphoid structures [12]. Evidence showed that ductal carcinoma in situ (DCIS) had an increase in the density and extent of immune cell infiltration in comparison to normal breast tissue [13, 14]. The proportions of $FoxP3^+$ cells, $CD68^+$ and $CD68^+PCNA^+$ macrophages, $HLADR^+$ cells, $CD4^+$ T cells, $CD20^+$ B cells, and total tumor infiltrating lymphocytes (TILs) in high-grade DCIS are notably higher than in nonhigh-grade DCIS [15, 16]. A study also found that stromal FOXP3, PDL1, and dense TILs were poor prognostic factors for DCIS recurrence. Additionally, dense TILs were independently linked to poor outcomes for both invasive (HR = 2.1, P = 0.029) and all recurrences (HR = 7.0, P = 0.024) [17]. Notably, it was discovered that the TME was considerably enriched for not only promoting tumor growth but also antitumor activity, indicating the dynamic function that immunophenotypes play in both tumor promotion and suppression. The abnormal regulation in immune cells, therefore, is regarded as one of the triggers for BC. Typically, an antitumor immune response is often initiated by NK cells, $CD4^+$and $CD8^+$T cells, and dendritic cells (DCs), among other immune cells. But these cells can be recruited by various TME elements to enhance immune evasion and become pro-tumorigenic [18, 19]. For instance, cancer-associated fibroblasts (CAFs) secreted transforming growth factor-β (TGF-β), which can neutralize the antitumor response of NK cells, neutrophils, and macrophages [20]. However, the interpretability of our findings is limited by the inherent constraints of observational studies, including potential residual confounding and reverse causality, as well as a restricted sample size.

Leveraging genetic variants as instrumental variables (IVs), Mendelian randomization (MR) has become a widely adopted method for evaluating putative causal links between risk factors and diseases like cancer [21, 22]. This method provides a time- and cost-effective means to explore causation from observed associations, circumventing the need for randomized controlled trials or animal studies [23–27]. As genetic variants are randomly allocated during gametogenesis and genotypes are often unaffected by external settings, confounding factors, and reverse causality concerns can be substantially decreased in MR, which is often referred to as "nature's randomized trial" [28, 29]. Recent MR studies have shown the connection between exposure factors (such as obesity [30], tea consumption [25], periodontitis [31], blood metabolites [32], and depression [33]) and the risk of BC.

Considering the inadequate knowledge regarding the causal impact of peripheral blood immune cells on BC, this study intends to establish a causal association between immunophenotypes and BC. To further bolster our point of view, we verified the highly correlative immunophenotypes using lymphocyte flow-cytometry data from a case-control study. Our analysis offered an in-depth analysis of the relationships between circulating immunophenotypes and BC, which includes malignant neoplasm and carcinoma in situ of the breast. Our findings on the interactions between immune cells and BC may provide new insights into therapeutic targets and create a plethora of opportunities for the development of innovative treatment strategies.

Methods

Study design

Using a two-sample MR analysis, we investigated the causal relationship between the different subsets of BC and 731 immunophenotypes. Firstly, we obtained the public genome-wide association studies (GWAS) summary data, which included BC and 731 immunophenotypes. Three core assumptions were made to reduce the impact of bias on the results when choosing IVs (Figure 1a): (1) The IVs must have a strong correlation with the exposure; (2) The IVs cannot be linked to any other confounders; and (3) The IVs selected should act on the outcome only through exposure, without going through any other pathways [34, 35]. Second, the causative link between peripheral immune cells and BC was assessed using a two-sample MR analysis. In the end, peripheral blood was obtained from patients with BC as well as healthy volunteers. Using flow cytometry, the quantification and phenotypic identification of highly correlative immunophenotypes in malignant neoplasms of the breast were subsequently validated (Figure 1b). The outline of the analytical framework is shown in Figure 1.

Exposure and outcome data sources

The single nucleotide polymorphisms (SNPs) data of 731 peripheral blood immunophenotypes were retrieved from the GWAS Catalog (GCST90001391-GCST90002121) conducted by Valeria Orrù et al from 3,757 European individuals [36]. We characterized a total of 731 immunophenotypic traits, which were profiled across seven panels and four measurement types. Supplementary Table 1 contains the full details. The R10 version of the FinnGen database provided the GWAS data for the BC. In order to acquire more extensive and detailed causal relationships, we acquired GWAS data on BC as results from 6 cohorts, which included malignant neoplasm of breast and carcinoma in situ of breast. Table 1 offers the comprehensive details of the datasets used in this study. All GWAS utilized in this analysis were publicly available and had obtained written informed consent from participants with

approval from their respective institutional review boards. Importantly, there was no sample overlap between the exposure and outcome datasets.

Instrumental variable selection

To guarantee the precision and efficacy of the link between immune cells and the risk of BC, we employed a set of criteria to identify IVs that met the three MR analysis assumptions. The following are the procedures that we employed to select IVs:

Firstly, we selected genetic instruments for the immunophenotypes and HDP based on a genome-wide significance threshold of $P < 5 \times 10^{-8}$. We set the linkage disequilibrium (LD) of SNPs that satisfied "$r^2 < 0.001$ and kb > 10,000" to ensure the independence of the aggregated SNPs, as the presence of LD could lead to bias. Additionally, we calculated the F-statistic to assess the IV correlations' strength and chose SNPs with an F-statistic > 10 for inclusion in this MR analysis in order to prevent weak instrumental bias (MR Assumptions 1) [37]. Subsequently, SNPs exhibiting a noteworthy correlation with BC ($P < 1 \times 10^{-5}$) were excluded (MR Assumptions 2). Thirdly, we used the PhenoScanner website to search for and remove SNPs that corresponded to confounders (MR Assumptions 3) [38]. A total of 2,622 independent SNPs remained as the final genetic instruments for subsequent MR analyses of immunophenotypes.

MR analysis

Numerous statistical methods, such as inverse variance weighted (IVW) [39], MR Egger [40], Weighted median [41], Weighted mode [42], were used in MR analysis. To ensure result correctness, the IVW method was adopted as the principal analysis in this study, which is frequently utilized in MR research owing to its exceptional accuracy in effect estimation. Heterogeneity was evaluated using Cochran's Q test [43]. The potential bias resulting from horizontal pleiotropy was evaluated and corrected for using MR Egger intercept analysis and leave-one-out analysis [44]. The Benjamini-Hochberg correction, which controls the false discovery rate (FDR), was used to adjust multiple testing in order to account for the potential for increasing the overall type I error during multiple comparisons. Immunophenotypes with an adj.*P* value < 0.05 and have causal links with 6 cohorts of BC were defined as "highly correlative" factors, whereas those showing a *P* value < 0.05 but 0.05 < adj.*P* value < 0.2 were defined as "suggestive possible" factors. We used the MR Steiger directionality test to examine if exposure was directionally causative for the result [45]. The "Two Sample MR" in R software 4.3.1 was used to conduct all of the analysis [46].

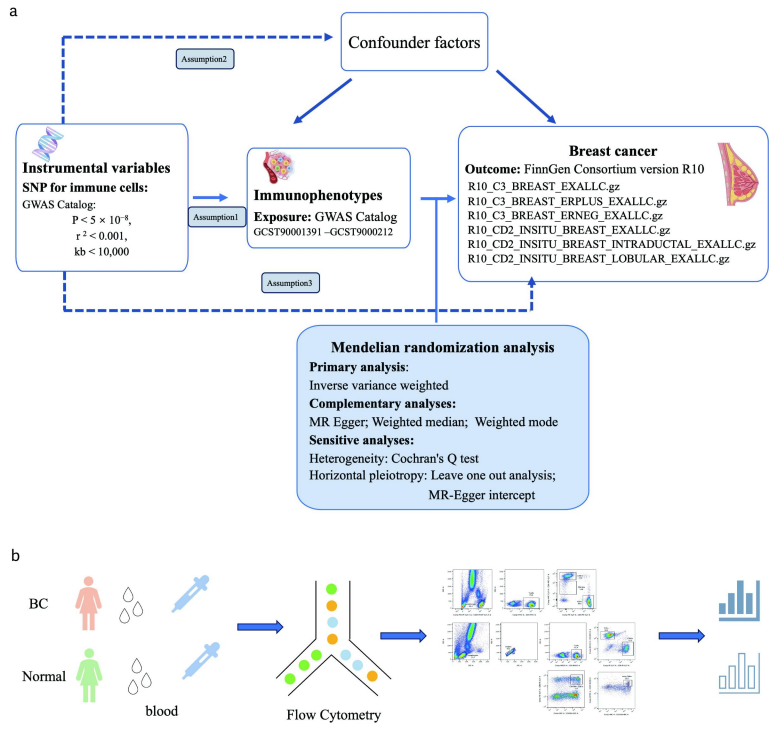

Figure 1 The workflow of this study design.
(a) Flow chart of Mendelian randomization analysis; (b) Flow chart of the validation analytic workflow. GWAS, genome-wide association study; SNPs, single nucleotide polymorphisms; BC, breast cancer.

Table 1 The GWAS datasets used for analyses

Trait/Disease	Data type	Consortium	Sample size	Case	Control
Immunophenotypes	Exposure	GWAS Catalog	3,757	–	–
Malignant neoplasm of breast	Outcome	Finngen	201,713	18786	182,927
Malignant neoplasm of breast, HER-positive	Outcome	Finngen	193,082	10404	182,678
Malignant neoplasm of breast, HER-negative	Outcome	Finngen	188,866	6188	182,678
Carcinoma in situ of breast	Outcome	Finngen	183,004	241	182,763
Carcinoma in situ of breast, intraductal	Outcome	Finngen	185,172	2390	182,782
Carcinoma in situ of breast, lobular	Outcome	Finngen	183,098	184	182,914

GWAS, genome-wide association study; HER, human epidermal growth factor receptor.

Clinical validation of Mendelian randomization

Study population. The study was conducted at the Gansu Provincial Maternity and Child Care Hospital from January 2024 to May 2025. We enrolled 123 patients with BC who displayed histologically confirmed BC but no history of malignant disease and had not yet undergone any anticancer therapy. The control group consisted of 109 matched healthy female donors with no pathology declared, no current medication except contraceptives or menopausal hormone therapy. Peripheral blood samples (2 mL) were collected both from healthy donors and from patients on the morning of their surgery using Vacutainer EDTA tube and prepared for determining the immune profile within 24 h. This study involves human participants and complies with the Declaration of Helsinki's ethical criteria and authorized by the Ethics Committees of Gansu Provincial Maternity and Child Care Hospital [(2023) GSFY-No.58]. Informed consent was signed by each patient before they took part in the research.

Flow cytometry data acquisition and analysis. The immunophenotypes selected for flow cytometry validation were prioritized based on the strength of their MR association (*P*-value), the magnitude of the effect size, their established biological relevance in cancer immunology, and the practical availability of specific antibodies. We selected immunophenotypes for FCM validation based on criteria: (1) strength of MR association (using the "highly correlative" adj.*P* value < 0.05) ; (2) biological plausibility and relevance to BC based on literature; (3) practical availability of specific antibodies for our flow cytometry panel.

Peripheral blood immunophenotypes were stained with EDTA-whole blood in this study. In brief, 100 µl of peripheral blood was mixed with monoclonal antibodies and incubated in the dark at room temperature for 15 min. Staining protocols are presented in detail in Supplementary Table 2. Following that, samples were then lysed with 1×FACS lysing solution (BD Biosciences) at room temperature for 15 min. Before acquisition, the lysed samples were washed and centrifuged twice at 300 g for 5 min to get rid of extra antibody, lysed RBCs, and platelets. Flow cytometry was performed on the leftover cell pellet by resuspending it in 0.5 mL of PBS. FACSLyricTM flow cytometers (BD Biosciences, USA) were used to acquire the data, and BD FASCSuitTM software (BD Biosciences) was used for analysis. The panel of immunophenotypic characterization was assessed according to the research conducted by Orrù et al [36]. To improve consistency in data processing, the same researcher manually gated the cell populations.

Statistical analysis

In the case-control study, we used an independent samples t-test and a Mann-Whitney U test to assess differences in the proportions of immune cells in healthy controls versus BC patients. If the data passed the normality test, an independent samples t-test was performed; if the data failed to pass the normality test, a Mann-Whitney U test was used. $P < 0.05$ was considered statistically significant. The statistical significance of differences between patient groups was performed using R software (v.4.2.2) package "ggpubr" and "ggplot2". $P < 0.05$ was denoted by *, $P < 0.01$ by **, $P < 0.001$ by ***, $P < 0.0001$ by ****. Additionally, the ROC curve model was performed to explore the performance of absolute count, relative count, and MFI to distinguish BC patients from normal females. The ROC curves were generated using R software (v.4.2.2) package "pROC" [47].

Results

Causal association between immunophenotypes and BC risk

In our exploration of immunophenotypes' causal effects on BC, 2622 SNPs associated with immunophenotypes were found in this study, after the selection of IVs. We utilized the IVW method as the principal analysis in a two-sample MR analysis. In total, 83 suggestive possible immunophenotype traits (124 pairs, *P* value <0.05, $0.05<$ adj.*P* value <0.2, Figure 2, Supplementary Table 3) were associated with BC in 6 cohorts. After conducting multiple test adjustments via the FDR method, 24 highly correlative immunophenotypes (26 pairs, adj.*P* value < 0.05, Figure 3) were detected. When the highly correlative immunophenotypes were classified in 7 panels, 8 pairs belonged to B cells, 6 from T cells, B cells and NK cells (TBNK), 5 from Treg, 3 from the maturation stages of T cells, and 4 from monocytes. For illustrating the trends in immunophenotypes under the four MR methods, scatter plots were created (Supplementary Figure 1).

The IVW method indicated a causal association of 35 immunophenotypes with malignant neoplasm of breast, among which 16 traits were highly correlative factors (Figure 3). Compared to other panels, the B cell panel demonstrated the greatest number of significant correlations for BC. CD20 was the one whose expression was most frequently seen in the immune cells of our study. The MFI of CD20 on B cell (OR = 0.878, 95% CI: 0.823–0.937, P = 7.89E-05), IgD^+ $CD24^-$ (OR = 0.904, 95% CI: 0.848–0.964, P = 0.002), IgD^+ $CD38^{br}$ (OR = 0.902, 95% CI: 0.852–0.954, P = 3.64E-04), IgD^+ $CD38^{dim}$ (OR = 0.926, 95% CI: 0.882–0.972, P = 0.002), naive-mature B cell (OR = 0.92, 95% CI: 0.869–0.974, P = 0.004), IgD^+ (OR = 0.912, 95% CI: 0.86–0.967, P = 0.002), and transitional B cell (OR = 0.891, 95% CI: 0.825–0.963, P = 0.004) were negative association with BC. Similarly, the increase of $CD8^{dim}$ %T cell (OR = 0.868, 95% CI: 0.777–0.97, P = 0.013), lymphocyte AC (OR = 0.822, 95% CI: 0.745–0.908, P = 1.05E-04), T cell AC (OR = 0.827, 95% CI: 0.752–0.91, P = 9.49E-05), $CD8^{dim}$ %leukocyte (OR = 0.861, 95% CI: 0.765–0.968, P = 0.013), CD28 on $CD45RA^+$ $CD4^+$ (OR = 0.899, 95% CI: 0.827–0.978, P = 0.013), and CD28 on $CD28^+$ $CD45RA^+$ $CD8^{br}$ (OR = 0.921, 95% CI: 0.865–0.982, P = 0.011) were associated with a reduced risk of BC in TBNK and Treg panel. Conversely, CCR7 on naive $CD8^{br}$ (OR = 1.128, 95% CI: 1.027–1.239, P = 0.012) and CD40 on monocytes (OR = 1.038, 95% CI: 1.01–1.066, P = 0.007) had an association with an elevated risk of BC. In HER+BC, we found clear associations between T cells, Tregs, B cells, and the occurrence of BC. Three highly correlative traits, including Lymphocyte AC (OR = 0.816, 95% CI: 0.718–0.928, P = 0.002), T cell AC (OR = 0.813, 95% CI: 0.719–0.921, P = 0.001), and CD20 on IgD^+ $CD38^{br}$ (OR = 0.891, 95% CI: 0.827–0.959, P = 0.002), were identified. In addition to these, 24 suggestive possible immunophenotypes were found. $CD8^{br}$ and $CD8^{dim}$ %leukocyte (OR = 1.267, 95% CI: 1.065–1.509, P = 0.008), CD4 on EM $CD4^+$ (OR = 1.091, 95% CI: 1.009–1.18, P = 0.029), and CD45RA on naive $CD8^{br}$ (OR = 1.128, 95% CI: 1.021–1.247, P = 0.018) may increase the risk of HER+BC. When MR analysis was performed with 731 immunophenotypes as exposure and HER–BC as outcome, we did not find any immunological characteristics at a significance level of 0.05, even after multiple test adjustments via the FDR method. However, at a significance threshold of 0.20, we found 31 suggestive possible factors, such as B cells, Tregs, monocytes, and dendritic cells.

To obtain a more thorough and coherent conclusion, we involved 3 cohorts with carcinoma in situ of the breast, including intraductal and lobular carcinoma in situ of the breast. By FDR correction (adj.P < 0.05), we found 7 highly correlative immunophenotypes linked to lobular carcinoma in situ of the breast (Figure 3). Four traits were confirmed to be positively correlated with BC risk: CD40 on $CD14^+$ $CD16^-$ monocyte (OR = 1.414, 95% CI: 1.075–1.86, P = 0.013), CD40 on $CD14^+$ $CD16^+$ monocyte (OR = 1.431, 95% CI: 1.117–1.833, P = 0.005), CD40 on monocytes (OR=1.472, 95%CI:1.151–1.883, P = 0.002), CD8 on $CD28^+$ $CD45RA^+$ $CD8^{br}$ (OR = 1.452, 95% CI: 1.092–1.931, P = 0.01). On the contrary, $CD25^{hi}$ %$CD4^+$ (OR = 0.379, 95% CI: 0.174–0.823, P = 0.014), CD25 on $CD45RA^-$ CD4 not Treg (OR = 0.36, 95% CI: 0.166–0.781, P = 0.01), and CD45RA on naive $CD8^{br}$ (OR = 0.395, 95% CI: 0.192–0.815, P = 0.012) were confirmed to be protective factors against lobular carcinoma in situ of breast. However, there were no significant highly correlative factors (adj.P < 0.05) found in intraductal carcinoma in situ of the breast. Nonetheless, we discovered 9 suggestive possible variables, including B cells, Tregs, and NKT cell, at a significance level of 0.20.

Sensitivity analysis

Sensitivity analysis was done to evaluate the robustness of the results

between the identified immune cells and BC because the IVW approach is prone to weak instrumental bias. The robustness of our results was demonstrated by the lack of heterogeneity detected by MR-Egger (Supplementary Figure 2) and Cochran's Q test (Supplementary Table 4). The stability of the data was further demonstrated using scatterplots (Supplementary Figure 1) and funnel plots (Supplementary Figure 3). Furthermore, Supplementary Figure 4 presents the result of the leave-one-out analysis. All *P*-values were greater than 0.05 and the intercepts of the MR-Egger regression did not substantially differ from 0, suggesting that there was no horizontal pleiotropy. Supplementary Table 5 also revealed no indication of reverse causality in the MR Steiger directionality test.

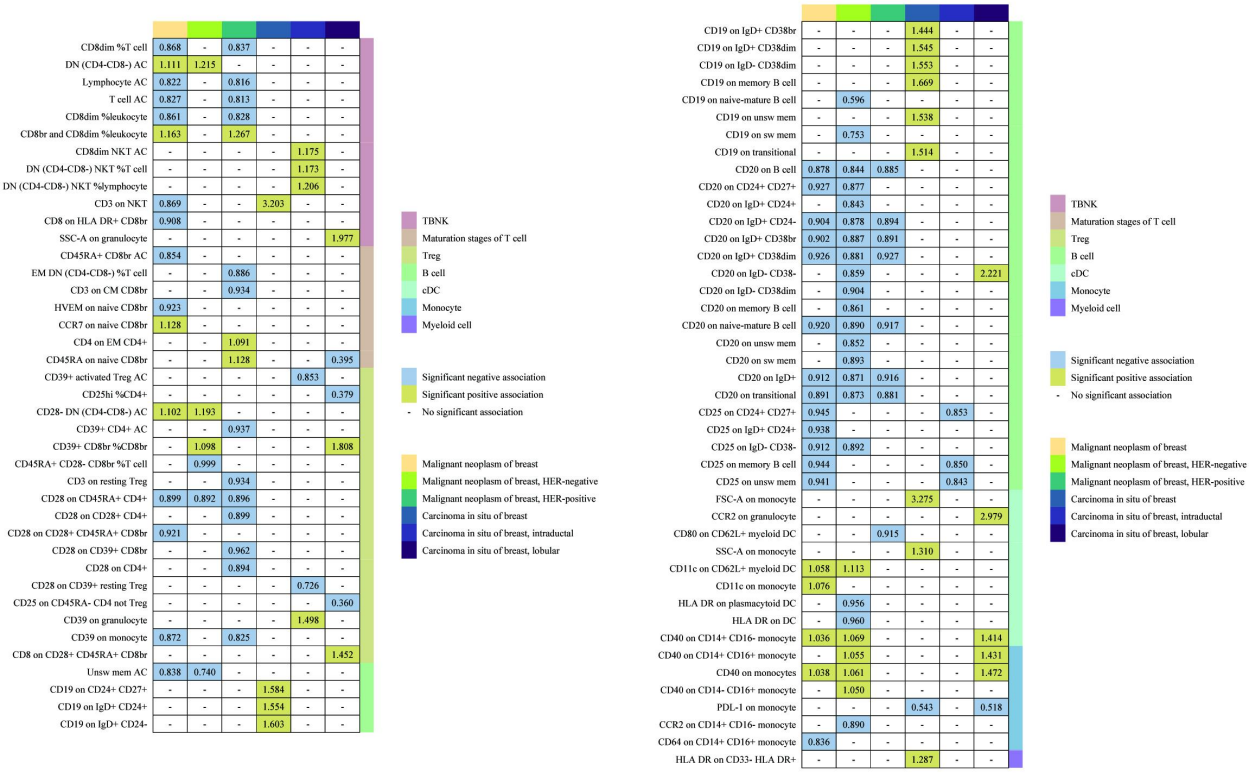

Figure 2 Summary of associations of genetically predicted immunophenotype traits with BC.
The numbers in the box are the odds ratios for associations of exposure for breast cancer. BC, breast cancer.

Outcome	Immunophenotypes Trait	SNP(n)	adj.P	OR(95% CI)
Malignant neoplasm of breast	CD45RA+ CD8br AC	3	0.025	0.854 (0.769 to 0.949)
	CD8dim %T cell	2	0.046	0.868 (0.777 to 0.970)
	Lymphocyte AC	2	**0.003**	0.822 (0.745 to 0.908)
	T cell AC	2	**0.003**	0.827 (0.752 to 0.910)
	CD8dim %leukocyte	2	0.046	0.861 (0.765 to 0.968)
	CD20 on B cell	4	**0.003**	0.878 (0.823 to 0.937)
	CD20 on IgD+ CD24−	4	0.020	0.904 (0.848 to 0.964)
	CD20 on IgD+ CD38br	5	**0.007**	0.902 (0.852 to 0.954)
	CD20 on IgD+ CD38dim	7	0.020	0.926 (0.882 to 0.972)
	CD20 on naive−mature B cell	6	0.025	0.920 (0.869 to 0.974)
	CD20 on IgD+	5	0.020	0.912 (0.860 to 0.967)
	CD20 on transitional	2	0.025	0.891 (0.825 to 0.963)
	CD28 on CD45RA+ CD4+	2	0.046	0.899 (0.827 to 0.978)
	CD28 on CD28+ CD45RA+ CD8br	3	0.046	0.921 (0.865 to 0.982)
	CCR7 on naive CD8br	2	0.046	1.128 (1.027 to 1.239)
	CD40 on monocytes	5	0.039	1.038 (1.010 to 1.066)
Malignant neoplasm of breast, HER-positive	Lymphocyte AC	2	0.044	0.816 (0.718 to 0.928)
	T cell AC	2	0.044	0.813 (0.719 to 0.921)
	CD20 on IgD+ CD38br	5	0.044	0.891 (0.827 to 0.959)
Carcinoma in situ of breast, lobular	CD25hi %CD4+	2	0.042	0.379 (0.174 to 0.823)
	CD25 on CD45RA− CD4 not Treg	2	0.042	0.360 (0.166 to 0.781)
	CD40 on CD14+ CD16− monocyte	4	0.042	1.414 (1.075 to 1.860)
	CD40 on CD14+ CD16+ monocyte	5	0.042	1.431 (1.117 to 1.833)
	CD40 on monocytes	5	0.042	1.472 (1.151 to 1.883)
	CD45RA on naive CD8br	2	0.042	0.395 (0.192 to 0.815)
	CD8 on CD28+ CD45RA+ CD8br	8	0.042	1.452 (1.092 to 1.931)

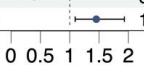

Figure 3 The significant causal effect of immunophenotypes on BC.
OR, odds ratio; Cl, confidence interval; SNP, single nucleotide polymorphism; AC, absolute cell; HER, human epidermal growth factor receptor; CD, cluster of differentiation; BC, breast cancer.

Validation of mendelian randomization

Baseline characteristic

The study involved 232 participants, of whom 123 were BC patients and 109 were healthy females (Table 2). The mean age of the BC group was 53.02 ± 8.81 years, and the mean age of the control group was 51.72 ± 8.03 years. There were no significant differences between the two groups in terms of age, body mass index, age at menarche, breastfeeding, number of births, and menopausal status ($P > 0.05$).

Immunophenotype traits are associated with the risk of BC. Mendelian randomization was used to confirm the direction of the causal links, and then we could use the case-control study to validate our findings. Using multiparametric flow cytometry, we verified the highly correlative immunophenotype traits. We investigated the TBNK, maturation stages of T cell, Treg, and B cell panels in patients with BC and healthy controls. Figure 4 displays the gating strategies of flow cytometry plots for highly correlative immunophenotypes.

Based on flow-cytometry data of lymphocytes from peripheral blood, we observed that BC patients had lower absolute counts of lymphocytes (P = 1.5E-08, Figure 5A) and T cells (P = 9.4E-11, Figure 5B). However, there was no difference in the levels of $CD8^{dim}$ T cells between BC patients and healthy individuals (Supplementary Figure 5A, 5B). Next, we conducted an investigation of maturation stages of T cell and Treg panel. The results revealed that the CCR7 level on naive $CD8^{br}$ T cell ($CD45RA^+CCR7^+$) of BC patients was significantly higher than that of healthy females (P = 4.5E-09, Figure 5D), and $CD45RA^+CD8^{br}$ T cell was lower (P = 1.2E-07, Figure 5C). In Treg panel, the percentage of $CD45RA^+CD4^+$ and $CD28^+CD45RA^+CD8^{br}$ was lower in BC patients compared with healthy controls, but the difference was not statistically significant (Supplementary Figure 5C). Meanwhile, the MFI of CD28 on $CD45RA^+$ $CD4^+$ (P = 0.021) and $CD28^+CD45RA^+$ $CD8^{br}$ (P = 4.7E-05) T cells were significantly higher than that of healthy females (Figure 5E). At last, we checked the proportion of B cell subsets and CD20 expression on B cells. The BC patients had a lower amount of IgD^+ $CD38^{br}$ (P = 6.7E-15, Figure 5F) and transitional B cell (P = 0.008, Figure 5F). Additionally, there was a significant increase in the CD20 level on IgD^+ (P = 6.7E-04), IgD^+ $CD24^-$ (P = 0.0014), and IgD^+CD38^{dim} (P = 0.0015) B cells in BC patients when compared to the control group (Figure 5G).

The study employed the analysis of receptor characteristics (ROC) to facilitate discrimination between the BC patients and the healthy control group. The absolute count, relative count, and MFI can be relevant parameters to distinguish BC patients from normal females, according to the area under the ROC curve (AUC). As Figure 5H illustrates, the AUC demonstrated acceptable accuracy for the absolute count of lymphocytes (AUC = 71.548%), T cells (AUC = 74.648%), and $CD45RA^+CD8^+$T cells (AUC = 70.146%). 1,545 was found to be the best cutoff value for T cell absolute count with a sensitivity of 61.468% and a specificity of 71.544% to distinguish the BC patients and the healthy control group. The best cutoff values of lymphocyte absolute count and $CD45RA^+CD8^+$T cell absolute count were 943.47 and 132.79 separately (Supplementary Table 6). Analyzing the relative count of IgD^+CD38^{br} B cell produced analogous results; at the cutoff value of 2.63, the AUC reached with 94.495% sensitivity and 60.163% specificity. Additionally, at the cutoff value of 35018.5, the AUC for CCR7 on naive $CD8^+$ T cells ($CD45RA^+CCR7^+$) was 72.320%, indicating 90.826% sensitivity and 57.724% specificity (Figure 5H, Supplementary Table 6).

Table 2 Characteristics of BC cases and controls

Characteristics	BC (n = 123)	Controls (n = 109)	*P*
Age (years), mean ± SD	53.02 ± 8.81	51.72 ± 8.03	0.244
Body mass index (kg/m2), mean ± SD	23.96 ± 2.60	23.88 ± 2.42	0.810
Age at menarche (years), mean ± SD	13.97 ± 0.85	13.90 ± 0.89	0.550
Breastfeeding, n (%)			
No	7 (5.69)	7 (6.42)	> 0.999
Yes	116 (94.31)	102 (93.58)	
Number of births, n (%)			
0	7 (5.69)	7 (6.42)	0.982
1	37 (30.08)	34 (31.19)	
2	39 (31.71)	35 (32.11)	
≥ 3	40 (32.52)	33 (30.28)	
Menopausal status, n (%)			
Premenopausal	66 (53.66)	65 (59.63)	0.433
Postmenopausal	57 (46.34)	44 (40.37)	
Hypertension diagnosis, n (%)	20 (16.26)	8 (7.34)	0.06
Type II diabetes diagnosis, n (%)	7 (5.69)	5 (4.59)	0.935
History of cancer,n (%)	0 (0)	0 (0)	-
Ever smokers, n (%)	4 (3.25)	1 (0.92)	0.442
Second-hand smoke exposure, *n* (%)	69 (56.10)	61 (55.69)	> 0.999
Alcohol consumption, n (%)	3 (2.44)	1 (0.92)	0.701

BC, breast cancer.

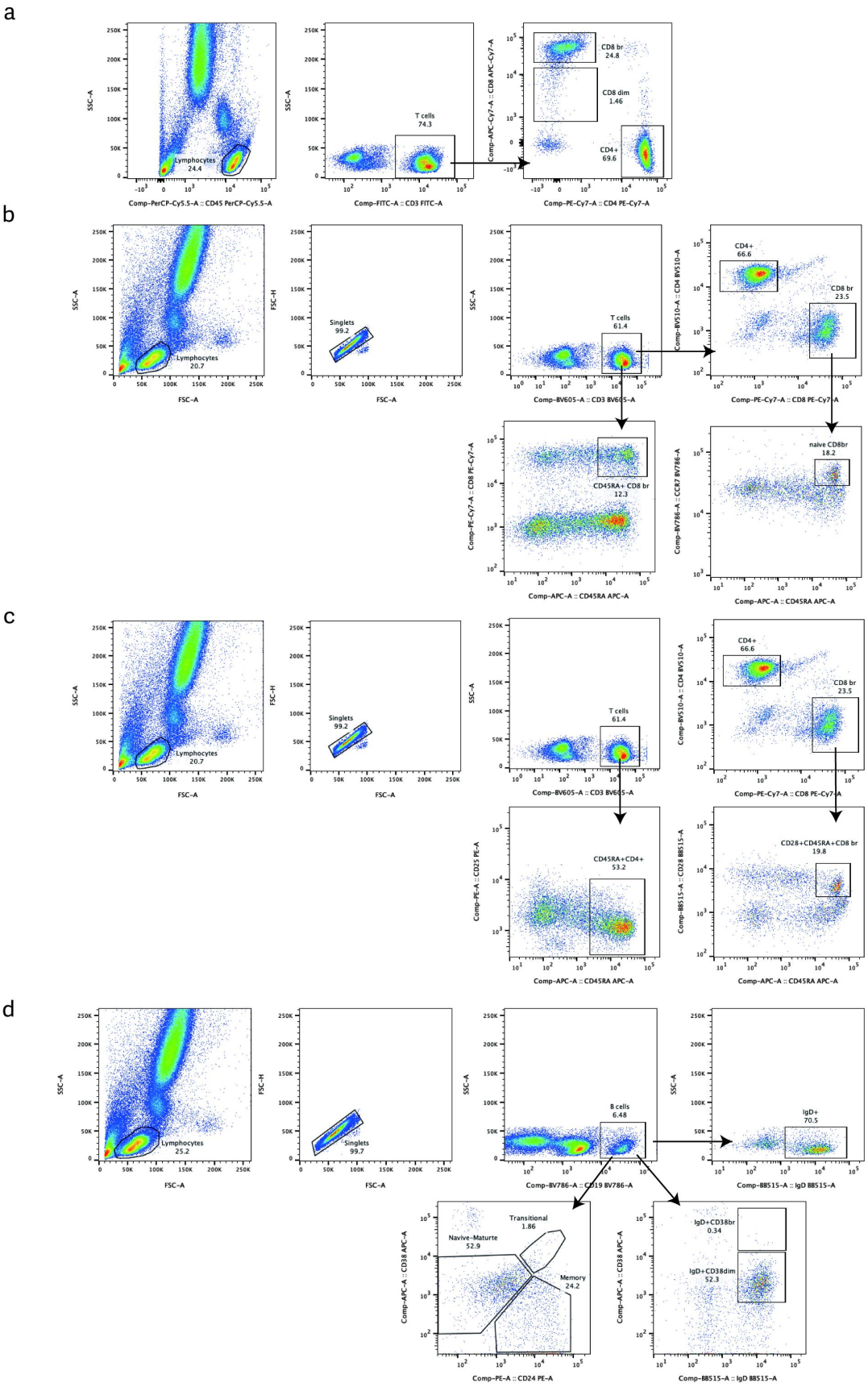

Figure 4 Flow cytometry gating strategy of this study.
(a) TBNK panel; (b) maturation stages of T cell panel; (c) Treg panel; (d) B cell panel. Treg, regulatory T cells; TBNK, T cells, B cells and NK cells.

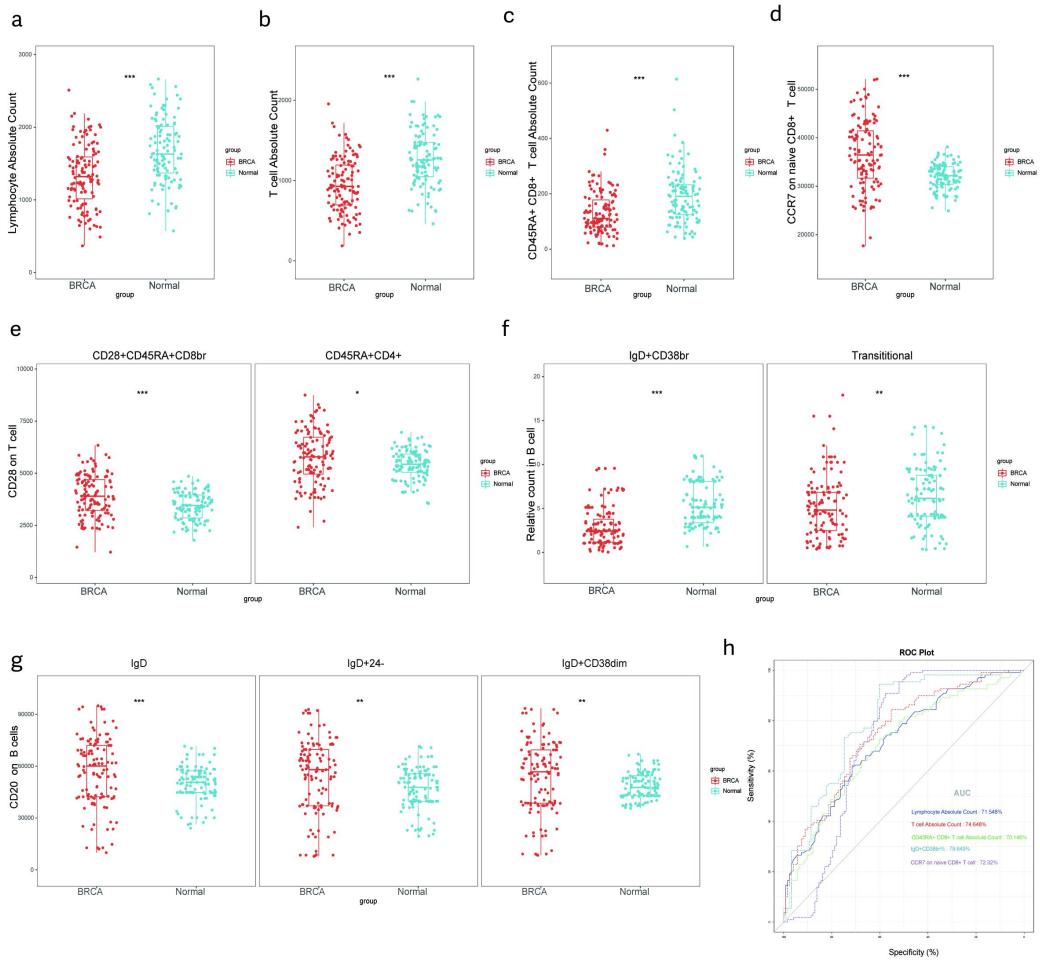

Figure 5 Differences in immunophenotypes distribution in healthy controls and BRCA patients.
The horizontal line in the boxes denotes the median value (50th percentile), while the box contains the 25th to 75th percentiles of the dataset. The whiskers mark the 5th and 95th percentiles, and the values of individuals were marked with dots. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$. (a) lymphocyte absolute count; (b) T cell absolute count; (c) CD45RA$^+$ CD8^{br} absolute count; (d) CCR7 on naive CD8^{br}; (e) CD28 on T cells; (f) B cell relative count; (g) CD20 on naive-mature B cell; (h) receiver operating characteristic curves for prediction capacity of absolute count, relative count, and MFI. BRCA, breast cancer; AUC, area under the curve; MFI, median fluorescence intensity.

Discussion

BC is a heterogeneous disease in which heterogeneous tumor cells, infiltrating immune cells, and cancer-associated fibroblasts shape the microenvironment of BC [48]. The immunological systems mediating dual functions in tumors have been the subject of extensive study for decades. It is commonly known that the immune system not only protects against tumor development by eliminating immunogenic tumor variations. It also promotes tumor growth by modifying tumor immunogenicity through a process called “immunoediting” [49], which involves both genetic and environmental factors. In order to improve the completeness and precision of the study, we used MR analysis in conjunction with a case-control study to evaluate the causal links between immunological imbalance and BC. The MR analysis covered six BC cohorts (including malignant neoplasm of breast and carcinoma in situ of breast). Through taking advantage of the most extensive GWAS datasets currently accessible for peripheral blood immunophenotyping, our work substantially expands the breadth of previous research and improves understanding of the relationships between immune cells and BC [36]. Drawing from an enormous variety of publicly available genetic data, we evaluated genetically predicted biomarkers of immunophenotypes and discovered 83 (124 pairs) suggestive possible ($0.05 <$ adj.P value < 0.2) immunophenotype traits and 24 (26 pairs) highly correlative (adj.P value < 0.05) ones linked to BC. We used flow cytometry to corroborate the outcome in a case-control study after MR analysis had established the causal relationship between the highly correlative factors and BC. The differences in absolute count, relative count, and MFI in some highly correlative lymphocytes between BC patients and healthy controls were effectively confirmed, and the performance of these parameters mentioned above was evaluated by drawing the ROC curve and calculating AUC. To facilitate public comprehension of the outstanding results, we have additionally plotted a schematic summary graph in Figure 6. This study provides interesting conjectures about the molecular functions of these immunological markers in BC. Overall, the investigation yielded more evidence than we had anticipated for the comprehensive and consistent etiology of immunophenotype in BC.

Tumor infiltrates contain a heterogeneous population of immune cells frequently dominated by T cells [8, 10]. As demonstrated by a recent study with the single-cell RNA sequencing, cells are the maximal immune cell cluster infiltrated in BC. According to Ding et al., there was an enrichment of exhausted CD8$^+$ T cells in triple negative breast cancer (TNBC), and a portion of the cytotoxic CD8$^+$ T cells seemed to be in a transition state towards exhaustion, suggesting limited tumor-cytotoxic activity [11].

In our study, patients with BC had a significantly lower amount of T cells and CD45RA$^+$CD8^{br} than the normal control group, which was in line with a previous study [50]. The study demonstrated that BC patients had a tendency for having more CD45RO$^+$T cells and less

CD45RA$^+$T cells. Additionally, tumor burden may be related to this tendency [50]. As we know, naive CD45RA$^+$T cells are mature T cells that have not been activated by antigens, while memory CD45RO$^+$T cells are a type of T cell that produce a faster and stronger immune response after antigen stimulation. The presence of tumor cells continuously may cause CD45RA$^+$T cells to become CD45RO$^+$T cells, which would change the trend between these two T-cell subgroups [50]. C-C chemokine receptor type 7 (CCR7) is a member of the C-C chemokine receptor family, expressed on the membrane of endothelial and immune cells [51]. By MR analysis, we found that CCR7 on naive CD8^{br} (OR = 1.128, 95% CI: 1.027–1.239, P = 0.012) had an association with an elevated risk of BC the result of the case-control study also revealed that CCR7 level on naive CD8^{br} T cell (CD45RA$^+$CCR7$^+$) of BC patients was significantly higher than that of healthy females. The result suggested a causal promoting effect of CCR7 expression on naive CD8$^+$ T cells. CCR7 is a central regulator of T-cell migration to lymphoid organs [52, 53]. One plausible mechanism is that higher CCR7 expression might sequester naive T cells within lymph nodes, impairing their recruitment to the tumor site and thus facilitating immune evasion. This hypothesis is supported by studies showing that altered T-cell trafficking is a feature of the immunosuppressive microenvironment [54, 55]. Research has revealed a favorable correlation between elevated levels of CCR7 expression and larger tumor size, deeper lymphatic invasion, and lower survival rate, which indicates that CCR7 is a potential biomarker to predict lymph node metastasis in BC [56, 57]. Another investigation discovered that cytoplasmic CCR7 expression is linked to a more aggressive tumor and a higher risk of tumor recurrence in patients with TNBC [58].

Cancer was inhibited by the tumor-specific T-cell response, which was impacted by both costimulatory and coinhibitory signals [59]. Cluster of differentiation 28 (CD28), as a crucial T-cell costimulatory molecule, CD28 enhanced the activation and proliferation of T cells when combined with its ligand B7-1 (CD80) or B7-2 (CD86) [60]. By the IVW method in MR analysis, we found that the increase of CD28 on CD45RA$^+$ CD4$^+$ (OR = 0.899, 95% CI: 0.827–0.978, P = 0.013), and CD28$^+$ CD45RA$^+$ CD8^{br} (OR = 0.921, 95% CI: 0.865–0.982, P = 0.011) led to decreased BC incidence rates. However, in our case-control study, the MFI of CD28 on CD45RA$^+$ CD4$^+$ and CD28$^+$ CD45RA$^+$ CD8^{br} T cells were significantly higher than that of healthy females. Clearly, the findings of the MR analysis and the case-control study were not consistent. We tried to explain away our conflicting findings. This may reflect the critical distinction between a lifelong genetic predisposition and the acquired state of T cells in established cancer. MR infers lifelong, developmental effects of genetic predisposition. A genetically determined higher baseline of CD28 expression might confer a long-term protective effect against BC initiation by maintaining a robust naive T cell pool. In contrast, FCM measures the current state in established patients. The elevated CD28 in patients could be a reactive, compensatory mechanism of the tumor immune microenvironment or a marker of specific T cell subsets. Chronic antigen exposure in tumors drives T cell exhaustion, a dysfunctional state characterized by the altered expression of costimulatory molecules, including CD28 [61, 62]. Therefore, the elevated CD28 we measured in patients could be a marker of a specific exhausted or activated subset, a consequence of the disease rather than its cause. The distinct populations in the two studies could be a possible explanation. In this study, the MR analysis explored the genetic association in the European group (Sardinians/Finnish), whereas the case-control study investigated the differences in immunophenotype in the Asian population (Chinese). There could be genetic heterogeneity among various groups in BC. Additionally, in comparison to European populations, the sample size of the Chinese population (123 cases and 109 controls) may not be large enough. The study conducted by Yan et al. indicated a positive association between the rs3116496 polymorphism of the CD28 (genotype CC vs. genotype TT) and the elevated risk of BC (OR = 2.15, P = 0.044) in Chinese women [60]. This result somewhat validates the reliability of our research findings, which showed that CD28 on T cells played an important role in BC.

Tumor-infiltrating lymphocytes have a significant impact on tumor growth and the immunotherapy response. Whereas B cells' function is less understood in relation to T cells, when it comes to solid tumors. Pursuant to recent views, B cells may have protumor or antitumor effects, depending on their phenotype, the TME in which they are found, and the antibodies they release [63]. In our study, the B cell panel had 7 highly correlative immunophenotypes (8 pairs), which is the largest number among other panels. CD20 was the one whose expression was most frequently seen in the immune cells of our study. The result of MR analysis revealed that CD20 on B cell, IgD$^+$ CD24$^-$, IgD$^+$ CD38^{br}, IgD$^+$ CD38^{dim}, naive-mature B cell, IgD$^+$, and transitional B cell were associated with the risk of BC. Additionally, there was a significant difference in the CD20 level on IgD$^+$, IgD$^+$CD24, and IgD$^+$CD38^{dim} B cells between BC patients and the control group. There is disagreement about the prognostic value of plasma cells and B cells in BC: some research indicated a negative correlation with the BC prognosis [64], while other studies indicated a positive correlation [65]. According to a study, invasive BC have higher levels of tumor-infiltrating B cells (TIL-B) than normal breast tissues, and this difference is especially obvious in extensively infiltrated tumors. Furthermore, there is a significant correlation between TIL-B and the histological grade, proliferating cell nuclear antigen (Ki67), estrogen receptor (ER), and progesterone receptor (PR) status [66]. Schmidt et al. found that the B-cell metagene was correlated with survival and showed an independent prognostic information in carcinomas with high proliferative activity (HR, 0.66; 95% CI, 0.46–0.97) [67]. The results of Schnellhardt et al.'s research demonstrated that CD20$^+$ cells mainly infiltrated in tumor epithelium and stroma. High densities of CD20$^+$ B cells in both the invading front (IF) and core tumor (CT) regions were correlated with lower disease-free survival (DFS) [68].

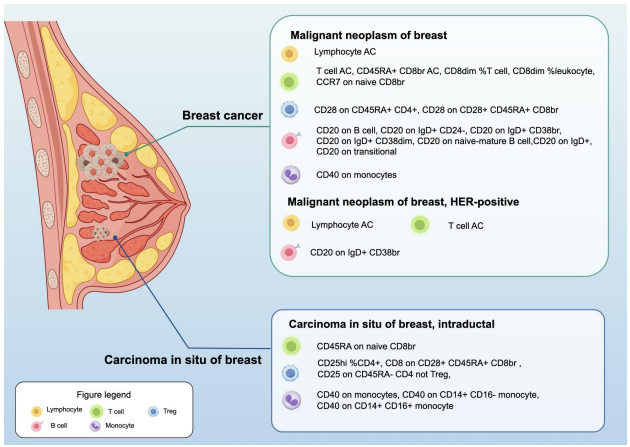

Figure 6 A schematic summary graph for the highly correlative immunophenotypes (created with figdraw)

In this study, we genetically predicted the causality of immune cell phenotypes on BC and validated part of the highly correlative immunophenotypes with flow cytometry data in a case-control study. This study presents several notable advantages. The primary advantages of this study were its extensive and representative set of immune cell phenotypes, its rigorous MR study design, which included strict criteria for screening IVs and strong causal association inferences from the main analysis that were backed by multiple sensitivity analyses. Secondly, by using GWAS summary data from six BC cohorts with a sizable sample size, we were able to obtain adequate power to clarify causal relationships. Besides, compared to comparable observational studies, our methodology worked more statistically efficiently owing to the high sample size and thorough coverage of immune cells. Thirdly, we focused on the heterogeneity of highly correlative immunological traits between BC patients and healthy controls, and the validation of the causal relationships was performed by multiparametric flow cytometry in a case-control study to ensure the validity of our findings. According to our results, these immunologic factors offer potential indicators to serve as useful biomarkers for the early diagnosis and treatment of BC in clinics and provide tentative understanding into the potential relationship between immune cell signatures and BC risk. The findings pave the way for additional research aimed at deciphering the processes by which these immune cells impact BC and creating targeted treatments.

However, our study does have several drawbacks. First of all, our MR findings are based on European GWAS data, while the FCM validation was conducted in a Chinese-based validation cohort. Genetic backgrounds and allele frequencies of IVs can differ across ethnicities, which might lead to discrepancies in causal estimates and validation outcomes, necessitating future multi-ethnic studies [69]. Additionally, we were only able to validate a subset of the significant MR hits due to practical constraints. The reasons for non-validation of other signals could include false positives in the MR analysis, insufficient power of our validation study to detect small effects, or the fact that some genetically influenced immunophenotypes may not be readily measurable in peripheral blood. The small sample size in our case-control study, which should be increased in future studies, in order to validate absolute count, relative count, and MFI of immunophenotype as biomarkers for early cancer detection by multiparametric flow cytometry. Future studies with larger, independent cohorts are warranted to evaluate the diagnostic or prognostic potential of these immunophenotypic signatures. Finally, because our validation samples were collected prospectively and there was insufficient prognostic data available for BC, our study was unable to be further examined in the context of clinical prognosis or treatment response. Given this, we anticipate that other research will examine and confirm our findings in the future.

While our ROC analysis showed promising AUC values, its clinical utility remains exploratory due to the small sample size. The diagnostic performance of these immune markers should be benchmarked against established serum biomarkers like CA15-3 and CEA in future larger studies. Beyond single markers, the future clinical translation may lie in developing a multi-parameter model that integrates the most informative immunophenotypes identified here. The logical next steps would be: (1) technical validation of these markers in a large, independent cohort with pre-diagnostic samples to assess their predictive potential; (2) integration with existing clinical and molecular data to evaluate incremental value; and (3) if validated, exploration as a liquid biopsy tool for monitoring disease progression or treatment response.

Conclusion

In conclusion, our investigation revealed intriguing details regarding the potential causality between immunophenotypes and BC, and the causal links were partially validated with peripheral blood lymphocyte subpopulations as categorized using flow cytometry. Utilizing the MR analysis and multiparametric flow cytometry, we identified highly correlative associations between immune signatures and the risk of BC, such as lymphocyte AC, T cell AC, $CD45RA^{+}$ $CD8^{br}$ AC, and CCR7 on naive $CD8^{br}$. These insights provide fresh perspectives for further studies aimed at understanding the mechanisms through which these immune cells influence BC and developing targeted therapies.

Declaration

Availability of data and materials

Summary data used for this study can be accessed through the following links. Any additional information required to reanalyze the data reported in this paper is available from the lead contact (Table 3).

Table 3 Details of core tools and resources in this study

Resource	Source	Identifier
Deposited data		
GWAS summary statistics for 731 immunophenotypes	Orrù et al. [36]	https://gwas.mrcieu.ac.uk
GWAS summary statistics for BC	The FinnGen study	https://www.finngen.fi/en
Software and algorithms		
R	R project	https://www.r-project.org
Two sample MR	Hemani et al. [46]	https://mrcieu.github.io/TwoSampleMR/
FlowJo™ v10.8 software	BD Life Sciences	N/A
BD FASCSuite™ software	BD Life Sciences	N/A

GWAS, genome-wide association study.

References

1. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2024;74(3):229–263. Available at: http://doi.org/10.3322/caac.21834
2. Armenta-Guirado BI, González-Rocha A, Mérida-Ortega Á, López-Carrillo L, Denova-Gutiérrez E. Lifestyle Quality Indices and Female Breast Cancer Risk: A Systematic Review and Meta-Analysis. *Adv Nutr.* 2023;14(4):685–709. Available at: http://doi.org/10.1016/j.advnut.2023.04.007
3. Fakhri N, Chad MA, Lahkim M, et al. Risk factors for breast cancer in women: an update review. *Med Oncol.* 2022;39(12):197. Available at: http://doi.org/10.1007/s12032-022-01804-x
4. Duggan C, Dvaladze A, Rositch AF, et al. The Breast Health Global Initiative 2018 Global Summit on Improving Breast Healthcare Through Resource-Stratified Phased Implementation: Methods and overview. *Cancer.* 2020;126(S10):2339–2352. Available at: http://doi.org/10.1002/cncr.32891

5. Lauby-Secretan B, Scoccianti C, Loomis D, et al. Breast-Cancer Screening — Viewpoint of the IARC Working Group. *N Engl J Med.* 2015;372(24):2353–2358. Available at: http://doi.org/10.1056/NEJMsr1504363
6. Pal B, Chen Y, Vaillant F, et al. A single-cell RNA expression atlas of normal, preneoplastic and tumorigenic states in the human breast. *EMBO J.* 2021;40(11):e107333. Available at: http://doi.org/10.15252/embj.2020107333
7. Azizi E, Carr AJ, Plitas G, et al. Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. *Cell.* 2018;174(5):1293–1308.e36. Available at: http://doi.org/10.1016/j.cell.2018.05.060
8. Wagner J, Rapsomaniki MA, Chevrier S, et al. A Single-Cell Atlas of the Tumor and Immune Ecosystem of Human Breast Cancer. *Cell.* 2019;177(5):1330–1345.e18. Available at: http://doi.org/10.1016/j.cell.2019.03.005
9. Hussein MR, Hassan HI. Analysis of the mononuclear inflammatory cell infiltrate in the normal breast, benign proliferative breast disease, in situ and infiltrating ductal breast carcinomas: preliminary observations. *J Clin Pathol.* 2006;59(9):972–977. Available at: http://doi.org/10.1136/jcp.2005.031252
10. Ruffell B, Au A, Rugo HS, Esserman LJ, Hwang ES, Coussens LM. Leukocyte composition of human breast cancer. *Proc Natl Acad Sci U S A.* 2011;109(8):2796–2801. Available at: http://doi.org/10.1073/pnas.1104303108
11. Ding S, Qiao N, Zhu Q, et al. Single-cell atlas reveals a distinct immune profile fostered by T cell-B cell crosstalk in triple negative breast cancer. *Cancer Commun.* 2023;43(6):661–684. Available at: http://doi.org/10.1002/cac2.12429
12. Goff SL, Danforth DN. The Role of Immune Cells in Breast Tissue and Immunotherapy for the Treatment of Breast Cancer. *Clin Breast Cancer.* 2021;21(1):e63–e73. Available at: http://doi.org/10.1016/j.clbc.2020.06.011
13. Tower H, Ruppert M, Britt K. The Immune Microenvironment of Breast Cancer Progression. *Cancers.* 2019;11(9):1375. Available at: http://doi.org/10.3390/cancers11091375
14. Thompson E, Taube JM, Elwood H, et al. The immune microenvironment of breast ductal carcinoma in situ. *Mod Pathol.* 2016;29(3):249–258. Available at: http://doi.org/10.1038/modpathol.2015.158
15. Kim M, Chung YR, Kim HJ, Woo JW, Ahn S, Park SY. Immune microenvironment in ductal carcinoma in situ: a comparison with invasive carcinoma of the breast. *Breast Cancer Res.* 2020;22(1):32. Available at: http://doi.org/10.1186/s13058-020-01267-w
16. Campbell MJ, Baehner F, O'Meara T, et al. Characterizing the immune microenvironment in high-risk ductal carcinoma in situ of the breast. *Breast Cancer Res Treat.* 2016;161(1):17–28. Available at: http://doi.org/10.1007/s10549-016-4036-0
17. Toss MS, Abidi A, Lesche D, et al. The prognostic significance of immune microenvironment in breast ductal carcinoma in situ. *Br J Cancer.* 2020;122(10):1496–1506. Available at: http://doi.org/10.1038/s41416-020-0797-7
18. Labani-Motlagh A, Ashja-Mahdavi M, Loskog A. The Tumor Microenvironment: A Milieu Hindering and Obstructing Antitumor Immune Responses. *Front Immunol.* 2020;11:940. Available at: http://doi.org/10.3389/fimmu.2020.00940
19. Zhao H, Wu L, Yan G, et al. Inflammation and tumor progression: signaling pathways and targeted intervention. *Signal Transduct Target Ther.* 2021;6(1):263. Available at: http://doi.org/10.1038/s41392-021-00658-5
20. Bierie B, Moses HL. Transforming growth factor beta (TGF-β) and inflammation in cancer. *Cytokine Growth Factor Rev.* 2010;21(1):49–59. Available at: http://doi.org/10.1016/j.cytogfr.2009.11.008
21. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet.* 2014;23(R1):R89–R98. Available at: http://doi.org/10.1093/hmg/ddu328
22. Evans DM, Davey Smith G. Mendelian Randomization: New Applications in the Coming Age of Hypothesis-Free Causality. *Annu Rev Genomics Hum Genet.* 2015;16(1):327–350. Available at: http://doi.org/10.1146/annurev-genom-090314-050016
23. Sekula P, Del Greco MF, Pattaro C, Köttgen A. Mendelian Randomization as an Approach to Assess Causality Using Observational Data. *J Am Soc Nephrol.* 2016;27(11):3253–3265. Available at: http://doi.org/10.1681/ASN.2016010098
24. Li J, Bai H, Qiao H, et al. Causal effects of COVID-19 on cancer risk: A Mendelian randomization study. *J Med Virol.* 2023;95(4):e28722. Available at: http://doi.org/10.1002/jmv.28722
25. Deng Y, Ge W, Xu H, Zhang J. A Mendelian randomization study of the effect of tea intake on breast cancer. *Front Nutr.* 2022;9:956969. Available at: http://doi.org/10.3389/fnut.2022.956969
26. Zhu M, Ma Z, Zhang X, et al. C-reactive protein and cancer risk: a pan-cancer study of prospective cohort and Mendelian randomization analysis. *BMC Med.* 2022;20(1):301. Available at: http://doi.org/10.1186/s12916-022-02506-x
27. Mingdong W, Xiang G, Yongjun Q, Mingshuai W, Hao P. Causal associations between gut microbiota and urological tumors: a two-sample mendelian randomization study. *BMC Cancer.* 2023;23(1):854. Available at: http://doi.org/10.1186/s12885-023-11383-3
28. Grover S, Del Greco MF, Stein CM, et al. Mendelian Randomization. *Methods Mol Biol.* 2017;1666:581–628. Available at: http://doi.org/10.1007/978-1-4939-7274-6_29
29. Ference BA, Holmes MV, Smith GD. Using Mendelian Randomization to Improve the Design of Randomized Trials. *Cold Spring Harb Perspect Med.* 2021:a040980. Available at: http://doi.org/10.1101/cshperspect.a040980
30. Hao Y, Xiao J, Liang Y, et al. Reassessing the causal role of obesity in breast cancer susceptibility: a comprehensive multivariable Mendelian randomization investigating the distribution and timing of exposure. *Int J Epidemiol.* 2022;52(1):58–70. Available at: http://doi.org/10.1093/ije/dyac143
31. Ding M, Zhang Z, Chen Z, Song J, Wang B, Jin F. Association between periodontitis and breast cancer: two-sample Mendelian randomization study. *Clin Oral Investig.* 2023;27(6):2843–2849. Available at: http://doi.org/10.1007/s00784-023-04874-x
32. Wang Y, Liu F, Sun L, et al. Association between human blood metabolome and the risk of breast cancer. *Breast Cancer Res.* 2023;25(1):9. Available at: http://doi.org/10.1186/s13058-023-01609-4
33. Wu X, Zhang W, Zhao X, et al. Investigating the relationship between depression and breast cancer: observational and genetic analyses. *BMC Med.* 2023;21(1):170. Available at: http://doi.org/10.1186/s12916-023-02876-w
34. Emdin CA, Khera AV, Kathiresan S. Mendelian Randomization. *JAMA.* 2017;318(19):1925. Available at: http://doi.org/10.1001/jama.2017.17219
35. Boef AGC, Dekkers OM, le Cessie S. Mendelian randomization studies: a review of the approaches used and the quality of reporting. *Int J Epidemiol.* 2015;44(2):496–511. Available at: http://doi.org/10.1093/ije/dyv071
36. Orrù V, Steri M, Sidore C, et al. Complex genetic signatures in immune cells underlie autoimmunity and inform therapy. *Nat*

Genet. 2020;52(10):1036–1045. Available at: http://doi.org/10.1038/s41588-020-0684-4
37. Burgess S, Thompson SG. Avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol.* 2011;40(3):755–764. Available at: http://doi.org/10.1093/ije/dyr036
38. Staley JR, Blackshaw J, Kamat MA, et al. PhenoScanner: a database of human genotype–phenotype associations. *Bioinformatics.* 2016;32(20):3207–3209. Available at: http://doi.org/10.1093/bioinformatics/btw373
39. Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. *Stat Methods Med Res.* 2015;26(5):2333–2355. Available at: http://doi.org/10.1177/0962280215597579
40. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol.* 2017;32(5):377–389. Available at: http://doi.org/10.1007/s10654-017-0255-x
41. 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(4):304–314. Available at: http://doi.org/10.1002/gepi.21965
42. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol.* 2017;46(6):1985–1998. Available at: http://doi.org/10.1093/ije/dyx102
43. Bowden J, Del Greco M F, Minelli C, et al. Improving the accuracy of two-sample summary-data Mendelian randomization: moving beyond the NOME assumption. *Int J Epidemiol.* 2018;48(3):728–742. Available at: http://doi.org/10.1093/ije/dyy258
44. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol.* 2015;44(2):512–525. Available at: http://doi.org/10.1093/ije/dyv080
45. Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet.* 2017;13(11):e1007081. Available at: http://doi.org/10.1371/journal.pgen.1007081
46. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife.* 2018;7:e34408. Available at: http://doi.org/10.7554/eLife.34408
47. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics.* 2011;12(1):77. Available at: http://doi.org/10.1186/1471-2105-12-77
48. Baxevanis CN, Fortis SP, Perez SA. The balance between breast cancer and the immune system: Challenges for prognosis and clinical benefit from immunotherapies. *Semin Cancer Biol.* 2021;72:76–89. Available at: http://doi.org/10.1016/j.semcancer.2019.12.018
49. Schreiber RD, Old LJ, Smyth MJ. Cancer Immunoediting: Integrating Immunity's Roles in Cancer Suppression and Promotion. *Science.* 2011;331(6024):1565–1570. Available at: http://doi.org/10.1126/science.1203486
50. Wang WH, Xu HY, Zhao ZM, Zhang GM, Lin FW. Dynamic and significant changes of T-cell subgroups in breast cancer patients during surgery and chemotherapy. *Int Immunopharmacol.* 2018;65:279–283. Available at: http://doi.org/10.1016/j.intimp.2018.09.039
51. Legler DF, Uetz-von Allmen E, Hauser MA. CCR7: Roles in cancer cell dissemination, migration and metastasis formation. *Int J Biochem Cell Biol.* 2014;54:78–82. Available at: http://doi.org/10.1016/j.biocel.2014.07.002
52. Jaeger K, Bruenle S, Weinert T, et al. Structural Basis for Allosteric Ligand Recognition in the Human CC Chemokine Receptor 7. *Cell.* 2019;178(5):1222–1230.e10. Available at: http://doi.org/10.1016/j.cell.2019.07.028
53. Förster R, Davalos-Misslitz AC, Rot A. CCR7 and its ligands: balancing immunity and tolerance. *Nat Rev Immunol.* 2008;8(5):362–371. Available at: http://doi.org/10.1038/nri2297
54. Deng Y, Tan C, Huang S, et al. Site-Specific Polyplex on CCR7 Down-Regulation and T Cell Elevation for Lymphatic Metastasis Blocking on Breast Cancer. *Adv Healthc Mater.* 2022;11(22):e2201166. Available at: http://doi.org/10.1002/adhm.202201166
55. Dolina JS, Van Braeckel-Budimir N, Thomas GD, Salek-Ardakani S. $CD8^+$ T Cell Exhaustion in Cancer. *Front Immunol.* 2021;12:715234. Available at: http://doi.org/10.3389/fimmu.2021.715234
56. Liu Y, Ji R, Li J, et al. Correlation effect of EGFR and CXCR4 and CCR7 chemokine receptors in predicting breast cancer metastasis and prognosis. *J Exp Clin Cancer Res.* 2010;29(1):16. Available at: http://doi.org/10.1186/1756-9966-29-16
57. Cabioglu N, Yazici MS, Arun B, et al. CCR7 and CXCR4 as Novel Biomarkers Predicting Axillary Lymph Node Metastasis in T1 Breast Cancer. *Clin Cancer Res.* 2005;11(16):5686–5693. Available at: http://doi.org/10.1158/1078-0432.CCR-05-0014
58. Gurgel DC, Wong DVT, Bandeira AM, et al. Cytoplasmic CCR7 (CCR7c) immunoexpression is associated with local tumor recurrence in triple-negative breast cancer. *Pathol Res Pract.* 2020;216(12):153265. Available at: http://doi.org/10.1016/j.prp.2020.153265
59. Allison JP, Hurwitz AA, Leach DR. Manipulation of costimulatory signals to enhance antitumor T-cell responses. *Curr Opin Immunol.* 1995;7(5):682–686. Available at: http://doi.org/10.1016/0952-7915(95)80077-8
60. Yan Y, Zhang X. The association between CD28 gene rs3116496 polymorphism and breast cancer risk in Chinese women. *Biosci Rep.*2017;37(6):BSR20170884. Available at: http://doi.org/10.1042/BSR20170884
61. Esensten JH, Helou YA, Chopra G, Weiss A, Bluestone JA. CD28 Costimulation: From Mechanism to Therapy. *Immunity.* 2016;44(5):973–988. Available at: http://doi.org/10.1016/j.immuni.2016.04.020
62. Palermo B, Franzese O, Frisullo G, et al. CD28/PD1 co-expression: dual impact on $CD8^+$ T cells in peripheral blood and tumor tissue, and its significance in NSCLC patients' survival and ICB response. *J Exp Clin Cancer Res.* 2023;42(1):287. Available at: http://doi.org/10.1186/s13046-023-02846-3
63. Sharonov GV, Serebrovskaya EO, Yuzhakova DV, Britanova OV, Chudakov DM. B cells, plasma cells and antibody repertoires in the tumour microenvironment. *Nat Rev Immunol.* 2020;20(5):294–307. Available at: http://doi.org/10.1038/s41577-019-0257-x
64. Mohammed ZMA, Going JJ, Edwards J, Elsberger B, McMillan DC. The relationship between lymphocyte subsets and clinico-pathological determinants of survival in patients with primary operable invasive ductal breast cancer. *Br J Cancer.* 2013;109(6):1676–1684. Available at: http://doi.org/10.1038/bjc.2013.493
65. Iglesia MD, Vincent BG, Parker JS, et al. Prognostic B-cell Signatures Using mRNA-Seq in Patients with Subtype-Specific Breast and Ovarian Cancer. *Clin Cancer Res.* 2014;20(14):3818–3829. Available at: http://doi.org/10.1158/1078-0432.CCR-13-3368
66. Garaud S, Buisseret L, Solinas C, et al. Tumor-infiltrating B cells signal functional humoral immune responses in breast cancer. *JCI Insight.* 2019;4(18):e129641. Available at: http://doi.org/10.1172/jci.insight.129641

67. Schmidt M, Böhm D, von Törne C, et al. The Humoral Immune System Has a Key Prognostic Impact in Node-Negative Breast Cancer. *Cancer Res.* 2008;68(13):5405–5413. Available at: http://doi.org/10.1158/0008-5472.CAN-07-5206
68. Schnellhardt S, Erber R, Büttner-Herold M, et al. Tumour-Infiltrating Inflammatory Cells in Early Breast Cancer: An Underrated Prognostic and Predictive Factor. *Int J Mol Sci.* 2020;21(21):8238. Available at: http://doi.org/10.3390/ijms21218238
69. Locke AE, Steinberg KM, Chiang CWK, et al. Exome sequencing of Finnish isolates enhances rare-variant association power. *Nature.* 2019;572(7769):323–328. Available at: http://doi.org/10.1038/s41586-019-1457-z