Liu *et al. Breast Cancer Research* (2024) 26:172
https://doi.org/10.1186/s13058-024-01928-0

Breast Cancer Research

RESEARCH Open Access

Uncovering immune cell-associated genes in breast cancer: based on summary data-based Mendelian randomized analysis and colocalization study

Jingyang Liu[1†], Wen Sun[1†], Ning Li[1], Haibin Li[2,3], Lijuan Wu[1], Huan Yi[4*†], Jianguang Ji[5*†] and Deqiang Zheng[1*†]

Abstract

Background Breast cancer, which is the most prevalent form of cancer among women globally, encompasses various subtypes that demand distinct treatment approaches. The tumor microenvironment and immune response are of crucial significance in the development and progression of breast cancer. Nevertheless, there has been scant evidence concerning the genes within breast cancer - specific immune cells.

Methods We utilized summary data-based Mendelian randomization (SMR) to identify genes associated with breast cancer by utilizing expression quantitative trait loci (eQTL) datasets for 14 different immune cell types and genome-wide association studies (GWAS) for overall breast cancer and its subtypes. Furthermore, colocalization analysis was carried out to evaluate whether the observed association in SMR analyses is influenced by the same causal variant. Replication analysis and bulk RNA sequencing (bulkRNA-seq) analysis were employed to validate promising immune genes as potential drug targets.

Results After correcting for the rate of false discovery, we discovered a total of 17 genes in 9 immune cell types that were significantly associated with overall breast cancer and its subtypes. The genes *KCNN4*, *L3MBTL3*, *ZBTB38*, *MDM4*, and *TNFSF10* were identified in overall breast cancer and its subtypes. Colocalization analyses provided robust evidence in support of these associations. Notably, the *KCNN4* gene in non-classical MONOcytes (MONOnc) was further validated through replication analysis and bulkRNA-seq analysis.

†Jingyang Liu and Wen Sun contributed equally to this work.

†Deqiang Zheng, Jianguang Ji, and Huan Yi are equally contributing corresponding authors.

*Correspondence:
Huan Yi
yi.huan@fjmu.edu.cn
Jianguang Ji
Jianguangji@um.edu.mo
Deqiang Zheng
dqzheng@ccmu.edu.cn

Full list of author information is available at the end of the article

Conclusion In summary, our research has revealed a repertoire of genes within diverse immune cells associated with breast cancer. *KCNN4* gene in non-classical MONOcytes (MONOnc) exhibited a negative association with overall breast cancer and its subtypes, which was identified as a potential drug target for breast cancer, opening up new avenues for therapeutic interventions.

Keywords Breast cancer, Immune cell, Mendelian randomization, BulkRNA-seq, scRNA-seq, Drug target

Introduction

Breast cancer stands as the prevailing malignancy among women and the primary cause of cancer-related mortality, with a staggering 665,684 deaths in 2022 [1]. From a comprehensive clinical perspective [2], there are four subtypes of breast cancer requiring different treatment approaches: triple-negative, hormone receptor-negative and human epidermal growth factor receptor 2 (HER2)-positive, hormone receptor-positive and HER2-positive, hormone receptor-positive and HER2-negative breast cancer. Recent research has shown that breast cancer consists not only tumor cells, but also involves significant changes in the tumor microenvironment. These changes are now acknowledged as pivotal contributors to breast cancer initiation, advancement, and potential therapeutic targets [3]. The tumor microenvironment incorporates immune cells, soluble factors, and altered extracellular matrix. Harris et al. [4] found that women who took anti-inflammatory drugs twice a week or more for more than five years exhibited a 21% reduced risk of developing breast cancer, suggesting that it may be possible to mitigate breast cancer risk by regulating the function of immune cells. In addition, more evidence points out that strong infiltration of immune cells is often an indicator of chemotherapy response and positive prognosis of breast cancer [5, 6].

The immune system is generally categorized into two major components: innate immunity and acquired (or adaptive) immunity [7]. The innate immune system comprises various cell types, including phagocytes, neutrophils, macrophages, natural killer cells, basophils, eosinophils, and more. Adaptive immunity includes B and T cells [8]. Traditionally, breast cancer has been regarded as immunosilent, and characterized by lower tumor mutation burden (TMB) and immunogenicity compared to other tumor types. Studies indicate that 36% of basal-like tumors are classified with high TMB [9]. However, among these high-TMB tumors, only 24% exhibit significant immune invasion. These findings imply that only a minority of these tumors exhibit significant immunogenicity [9]. In addition, the immunogenicity of breast cancer exhibits significant heterogeneity, with distinct subtypes showing varying levels of immune infiltration [10]. However, precise gene expression in specific immune cells that influence the development of breast cancer and different subtypes has not been fully elucidated.

Mendelian randomization (MR) uses genetic variation as an instrumental variable to examine the causal influence of risk factors on health outcomes [11]. The instrumental variable (IV) method is put forth as an alternative statistical method to examine causality for expose-outcome associations while controlling for confounding factors [12]. Therefore, to better understand whether gene expression in different immune cell types influences breast cancer, we used pooled data to identify potential associations based on summary data-based Mendelian randomization (SMR).

Methods

This Mendelian randomization study utilized expressed quantitative trait loci (eQTL) data from Yazar et al. [13], and the genome-wide association study (GWAS) from the Breast Cancer Association Consortium (BCAC) [14] to explore the potential causal association between gene expression in specific immune cells and breast cancer. GWAS data from the FinnGen consortium was further employed to validate the observed findings. In addition, we performed differential gene expression analysis using the Bulk RNA sequencing (bulkRNA-seq) database to identify genes differentially expressed in breast cancer tissue. An overview of the study design is shown in Fig. 1.

Data sources for exposure

The eQTL data utilized in this study were obtained from the OneK1K cohort, as presented in the Yazar et al. article [13]. The cohort for this study consisted of 1.27 million peripheral blood mononuclear cells (PBMCs) obtained from 982 donors, with single-cell RNA sequencing (scRNA-seq) data utilized for analysis. The study identified 26,597 independent *cis*-eQTLs from 14 different immune cell types, including CD4+primary and central memory T cells (CD4NC), CD4+T cells with effector memory or central memory phenotype (CD4ET), SOX4-expressing CD4+T cells (CD4SOX4), and CD8+T cells with effector memory phenotype (CD8ET). CD8+infantilistic and central memory T cells (CD8NC), CD8+expressing S100B T cells (CD8S100B), natural killer cells (NK), NK recruitment cells (NKR), Plasma cells, memory B cells (Bmem), immature and naive B cells (Bin), classical MONOcytes (MONOc), non-classical MONOcytes (MONOnc), and dendritic cells (DC).

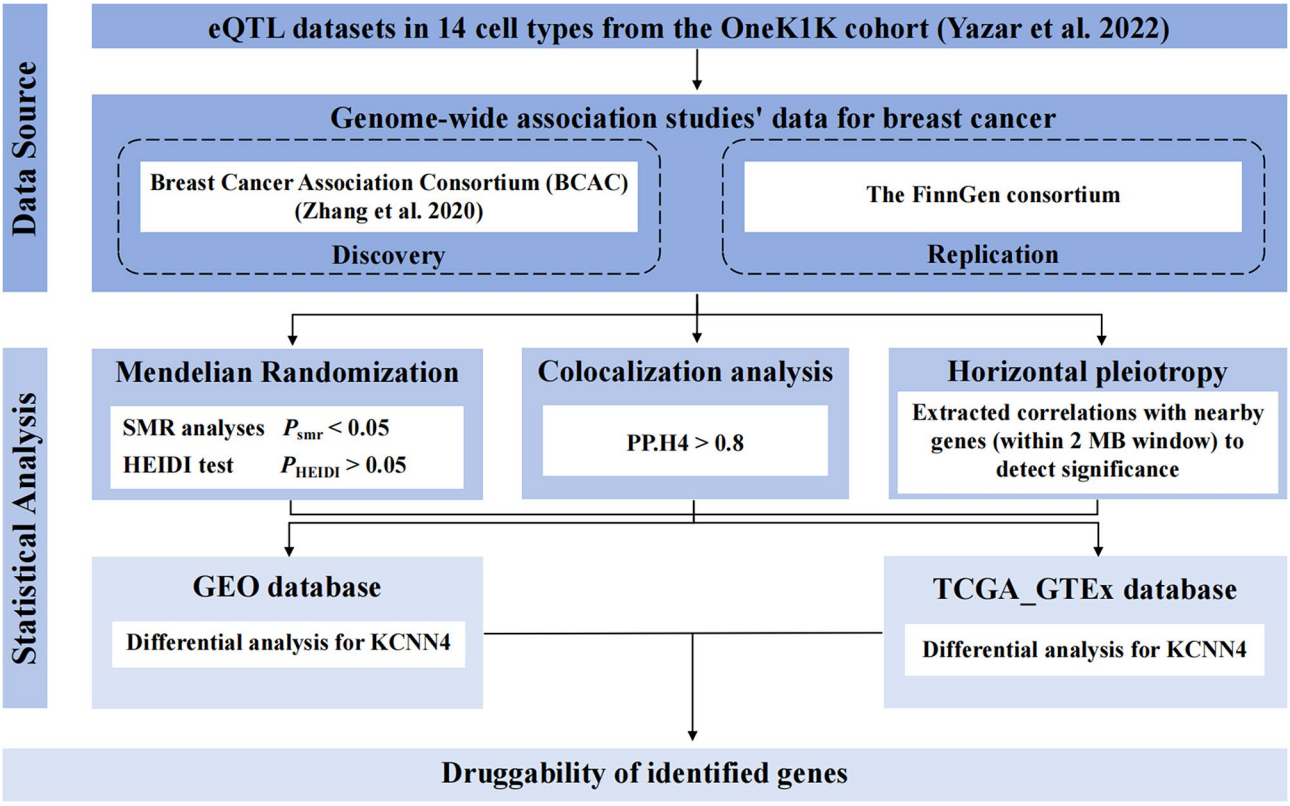

Fig. 1 Summary of study design, methods of analysis, and critical process. SMR, summary data-based Mendelian randomization; eQTL, expression quantitative trait loci; HEIDI test, heterogeneity in dependent instruments test; PP.H4, the posterior probability of hypothesis 4; GWAS, the genome-wide association study; GEO, the Gene Expression Omnibus; TCGA_GTEx, The Cancer Genome Atlas database combined with Genotype-Tissue Expression database

Data sources for the outcome

The Breast Cancer Association Consortium (BCAC) provided the largest available genome-wide association study (GWAS) summary data for breast cancer [14], This dataset includes a substantial number of cases (133,384) and controls (113,789), as well as 18,908 individuals of European ancestry with BRCA1 mutations from the Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA). Apart from studying overall breast cancer risk, BCAC also investigates subtype-specific characteristics. Breast cancer can be classified into different subtypes based on the presence or absence of hormone receptors (estrogen receptor (ER) and progesterone receptor (PR)), human epidermal growth factor receptor 2 (HER2) status, and tumor grade. The subtypes include luminal A-like (Luminal A), luminal B-like (Luminal B), luminal B/HER2-negative-like (Luminal B HER2Neg), HER2-enriched-like (HER2 Enriched), and triple-negative or basal-like (Triple Negative). In order to enhance the sensitivity for detecting associations with triple-negative subtypes, Zhang et al. [14] performed a meta-analysis of BCAC triple-negative cases and CIMBA BRCA1 mutation carriers, hereafter referred to as the BRCA-TN subtype.

In the replication analysis, we obtained published GWAS from the FinnGen Consortium for 18,786 breast cancer patients and 182,927 controls. The FinnGen study, a large-scale genomics initiative, has comprehensively examined more than 500,000 samples sourced from the Finnish biobank. This collaborative project encompasses research biobanks in Finland, various organizations, and international industry partners.

Data sources for bulk RNA sequencing data

We obtained microarray sequencing data GSE162228 from the Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo/). UCSC Xena is an integr ated website that aggregates bulk RNA-seq data from various cancer databases (http://xena.ucsc.edu/). We downloaded the breast cancer cohort from the GDC, the Cancer Genome Atlas (TCGA) Breast Cancer dataset, which includes 1,104 breast tumor samples and 113 normal breast tissue samples. Additionally, we used the Genotype-Tissue Expression (GTEx) database, which contains 179 normal breast tissue samples.

Selection of genetic instruments

In the SMR analysis, *cis*-eQTL genetic variants were utilized as instrumental variables (IVs) to assess gene expression. The cis region was defined as a 2 MB distance from the probe in both directions. Only genes with

at least one *cis*-eQTL having P_{eQTL} <5.0×10^{-8} were considered. Furthermore, single nucleotide polymorphisms (SNPs) with allele frequency differences greater than 0.2 between the eQTL data and the GWAS data were excluded. We permitted to exclude a maximum of 5% of SNPs based on allele frequency differences. The strength of the genetic instruments was calculated using the F-statistic (F-statistic=β^2/SE^2). If the F-statistic of the genetic variable is <10, it is considered as a weak genetic instrument and is thus excluded. Finally, the top-associated SNP with gene expression was selected as the genetic instrument. The screening of genetic instruments was performed by SMR software (version 1.3.1, https://yanglab.westlake.edu.cn/software/smr/#Download).

Statistical analyses

Summary data-based mendelian randomization and heterogeneity in dependent instrument test

SMR is a statistical approach that combines summary-level data from GWAS with data from eQTL studies to infer potential causal associations between genes and outcomes. Detailed methodologies have been previously described in the literature [15]. In brief, consider y as the outcome, x as the exposure, and z as the instrumental variable. By employing the two-step least-squares (2SLS) method, we estimate the size of the impact of exposure on the results with the formula:

$$b_{xy} = \frac{b_{zy}}{b_{zx}}$$

Where b_{zy} and b_{zx} represent the least-squares estimates of outcome and exposure with respect to the instrumental variable, b_{xy} is the effect size of exposure on outcome.

The heterogeneity in dependent instruments (HEIDI) test was introduced to examine whether there is a single causal variant underlying the association between gene expression and a trait. The HEIDI test was performed only when the number of SNPs reached or exceeded three. Additionally, SNPs that did not meet the criteria, such as those exhibiting high linkage disequilibrium (LD) (r^2>0.9) or weak LD (r^2<0.05), were excluded from the analysis. Associations with P_{HEIDI} < 0.05 indicated the presence of pleiotropy and were consequently excluded. The exclusion of SNPs and the HEIDI test were conducted using "smr" command in SMR software. False discovery rate (FDR) correction was applied across all immune cell types. The visualizations were generated using the online platform https://www.bioinformatics.com.cn.

Colocalization analysis

In order to disentangle the confounding effect of linkage disequilibrium [16]and examine the concordance between genes and breast cancer regarding a shared causal variant, we conducted colocalization analysis using the "coloc" R package. There are five hypotheses: Null hypothesis (H0): absence of a causal variant for either genes or breast cancer; Hypothesis 1 (H1): presence of single causal variant for genes only; Hypothesis 2 (H2): presence of single causal variant for breast cancer only; Hypothesis 3 (H3): presence of two independent causal variants for genes and breast cancer; Hypothesis 4 (H4): presence of a shared causal variant for genes and breast cancer. All significant genes of different cell types (P_{SMR} < 0.05, P_{HEIDI} > 0.05) were colocalized. The analysis employed default prior probabilities (PP.H1=1×10^{-4}, PP.H2=1×10^{-4}, PP.H3=1×10^{-5}). If the posterior probability for the colocalization hypothesis of H4 exceeds 80%, it suggests a high likelihood of colocalization between the gene and breast cancer.

Evaluate horizontal pleiotropy

To eliminate the effect of horizontal pleiotropy, we examined the correlations between each instrumental variable and all other genes in its vicinity (within a 2 MB window) to detect whether there was a significant correlation. If the genetic variation was associated with other genes and P<5×10^{-8}, SMR analysis was performed to assess the potential association between the expression of those genes and the outcome. If the data was insufficient to perform SMR analysis, the GWAS Catalog was utilized to explore the relationship between the adjacent genes and breast cancer. For those adjacent genes that do not meet either of these conditions, we believe that horizontal pleiotropy cannot be ruled out.

Identification of differential genes based on GEO and TCGA_GTEx databases

In the differential analysis, we employed the "limma" analysis method to identify genes that exhibited differential expression between 110 tumor samples and 23 normal samples obtained from GSE162228. A cutoff of P<0.05 and |log$_2$ fold change (FC)| > 0.5 was applied. To further explore the effect of differential gene expression on breast cancer, we combined the GTEx database with the TCGA database to compensate for the limited number of normal samples in the TCGA database, forming the TCGA_GTEx database. The batch effect was corrected using the "sva" package. TCGA_GTEx database was converted to log$_2$(counts+1). A Wilcoxon test [17] was used for the hypothesis tests, considering a significance level of P<0.05 as statistically significant. Boxplots were drawn with the R package "ggpubr".

Druggable genes identification

To assess the druggability of identified genes for breast cancer, we searched the identified gene in DrugBank,

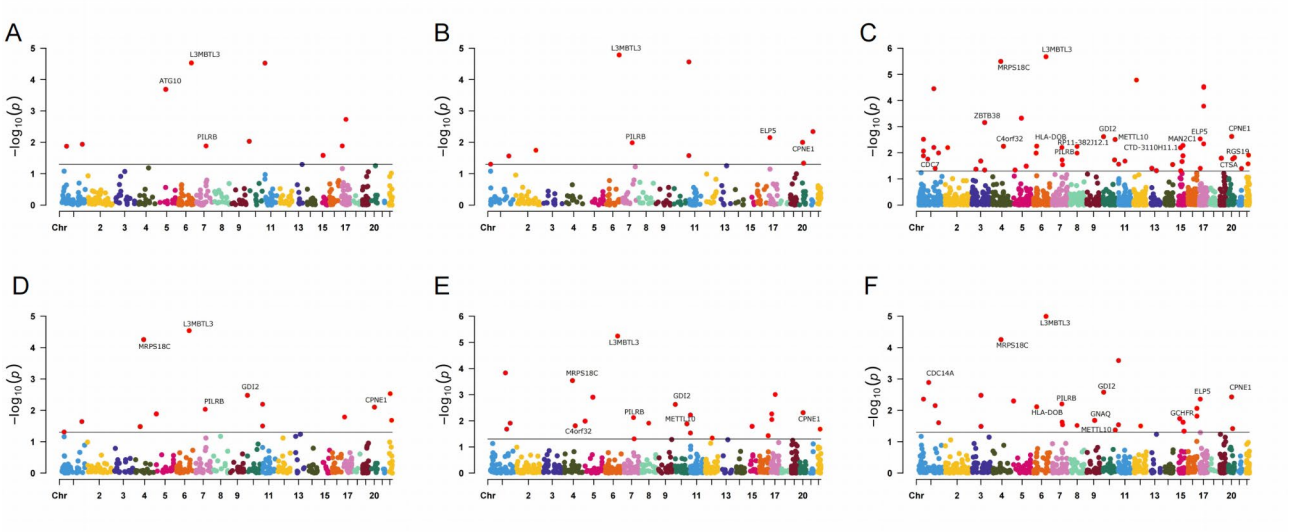

Fig. 2 Manhattan plots of significant results from summary data-based Mendelian randomization in overall breast cancer. **A** - **F** represents the associations between genes and overall breast cancer in Bin, Bmem, CD4nc, CD4et, CD8nc and CD8et cells, respectively. Black solid line indicates P-value threshold after false discovery rate is 0.05

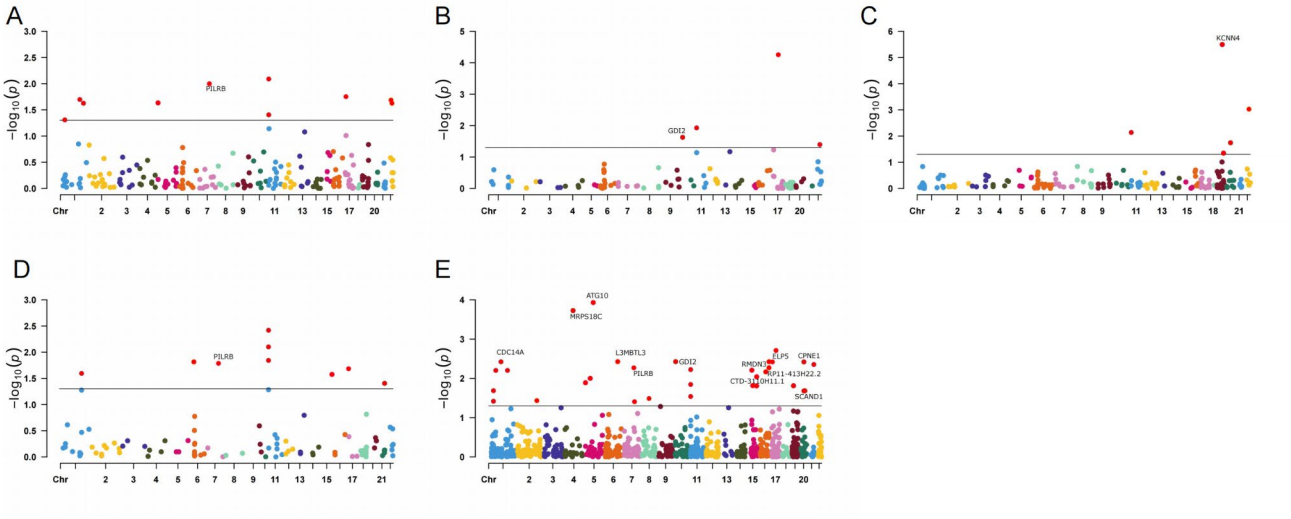

Fig. 3 Manhattan plots of significant results from summary data-based Mendelian randomization in overall breast cancer. **A** - **E** represents the associations between genes and overall breast cancer in CD8s100b, DC, MONOnc, NK and NKr cells, respectively. Black solid line indicates P-value threshold after false discovery rate is 0.05

Dependency Map, and the ChEMBL databases to identify the potential drugs that target this gene. Additionally, we conducted thorough searches of PubMed, Embase, Scopus, and Web of Science to review the research on these drugs associated with breast cancer.

Data were analyzed using R v4.2.3 and SMR v1.3.1 (https://yanglab.westlake.edu.cn/software/smr/#Download).

Results

SMR and HEIDI tests

Under a threshold of $FDR_{SMR} < 0.05$ and $P_{HEIDI} > 0.05$, we identified a total of 116 genes associated with overall breast cancer and its subtypes across 14 cell types (Figs. 2 and 3, Figures S1-S8). Although genes related to breast cancer were found in all cells, they were mainly concentrated in CD4nc, NKr, CD8et, CD8nc, and CD4et cells. The results from SMR and HEIDI are shown in Table S1-S14. The top 3 genes associated with overall breast cancer were *L3MBTL3* in CD4nc ($P_{FDR}=2.11\times10^{-6}$), *MRPS18C* in CD4nc ($P_{FDR}=3.19\times10^{-6}$) and *KCNN4* in MONOnc ($P_{FDR}=3.19\times10^{-6}$). For Luminal A breast cancer, the top 3 notable genes were *KCNN4* in MONOnc ($P_{FDR}=4.04\times10^{-5}$), *TNNT3* in CD4nc ($P_{FDR}=3.73\times10^{-4}$), *SCAMP3* in CD4nc ($P_{FDR}=4.48\times10^{-4}$). For Triple

Negative breast cancer, the top 3 significant genes were *MDM4* in CD4nc ($P_{FDR}=2.50\times10^{-7}$) and Bin ($P_{FDR}=5.78\times10^{-7}$), *RPS18* in Bmem ($P_{FDR}=6.07\times10^{-7}$). The top 3 genes associated with BRCA-TN breast cancer were *RPS18* in Bmem ($P_{FDR}=6.59\times10^{-9}$) and CD8nc ($P_{FDR}=6.59\times10^{-9}$), *MDM4* in CD4nc ($P_{FDR}=6.59\times10^{-9}$). No significant genes associated with breast cancer were found in the remaining subtypes. Table S15 presents significant genes related to breast cancer in each specific cell type.

Colocalization analysis

Based on the colocalization analysis in overall breast cancer and its subtypes (Table S1-S14), we identified a total of 39 genes in 14 different cell types that share causal variation with breast cancer and its subtypes, with a posterior probability more than 80% for the colocalization under hypothesis 4 (PP.H4>0.8). Specifically, we observed 24 genes associated with overall breast cancer across 11 immune cell types, 19 genes associated with Luminal A breast cancer across 12 immune cell types, 5 genes associated with Triple-Negative breast cancer across 14 immune cell types, and 5 genes associated with BRCA-TN breast cancer across 14 immune cell types.

Evaluation of horizontal pleiotropy

After the colocalization analysis, a total of 39 genes were included in the subsequent analysis. In order to eliminate the influence of horizontal pleiotropy, we extracted adjacent genes and tested their correlation with breast cancer. Detailed results are shown in Table S16. Finally, we found 11 genes related to overall breast cancers, namely *C4orf32* (in CD4nc, and CD8nc), *CDC7* (in CD4nc), *CTD-3110H11.1* (in CD4nc, and NKr), *CTSA* (in CD4nc), *GCHFR* (in CD8et), *GNAQ* (in CD8et), *HLA-DOB* (in CD4nc, and CD8et), *KCNN4* (in MONOnc), *L3MBTL3* (in Bin, Beme, CD4nc, CD4et, CD8nc, CD8et, and NKr), *RGS19* (in CD4nc), and *ZBTB38* (in CD4nc). The expression of genes *AKAP13* (in CD8et), *KCNN4* (in MONOnc), *L3MBTL3* (in Bin, Beme, CD4nc, CD4et, CD8nc, and CD8et), *MPRL42* (in CD4nc, and CD8et), *TNNT3* (in CD4nc), *YBEY* (in MONOc), and *ZBTB38* (in CD4nc, and CD8et) were associated with Luminal A breast cancer. Besides, *KCNN4* (in MONOnc), *MDM4* (in Beme, CD4et, and CD8et), *TNFSF10* (in CD4nc) were related to Triple-Negative and BRCA-TN breast cancer. The expression of gene *L3MBTL3* in Bin, Beme, CD4nc, CD4et, and CD8nc was also associated with BRCA-TN breast cancer.

Our conclusion, supported by the evidence from the comprehensive analysis mentioned above, is that 17 genes exhibited causal associations with breast cancer and its subtypes in 9 specific immune cell types.

Replication analysis

We conducted the same analysis using GWAS data provided by the FinnGen Consortium and eQTL data from immune cells as described above. After SMR analysis and HEIDI test, 73 genes were screened out from 14 immune cells. After colocalization analysis, 13 genes were shown to have a potential causal relationship with breast cancer. By extracting adjacent genes within 2 MB of these genes, we finally found causal relationships between the expression of genes *CNN2*, *FIBP*, and *KCNN4* with breast cancer. By comparing the BCAC Consortium and the FinnGen Consortium, we found that the *KCNN4* gene had a potential causal relationship with breast cancer in both GWAS databases, as shown in Table 1.

Identification of differential genes associated with breast cancer

We screened the gene sequencing data of GSE162228 and identified 2824 breast cancer-related differential expressed genes (DEGs) (Fig. 4A), including 1501 upregulated and 1323 down-regulated genes. We found the *KCNN4* gene exhibited different expression between tumor and normal tissues, consistent with the results of SMR analysis ($\log_2$FC = -0.735, $P=0.001$, Fig. 4B and C). As shown in Fig. 4D, the relative expression level of *KCNN4* in breast cancer samples was significantly lower than in normal samples ($P<0.05$). This result was consistent with the differential gene analysis results from the GEO database.

Druggability evaluation of the KCNN4

Among these 17 genes, we identified *KCNN4*, *L3MBTL3*, *MDM4*, *TNFSF10*, and *ZBTB38* were genes consistently associated with overall breast cancer and some subtypes.

Table 1 SMR analysis and colocalization analysis results of KCNN4 in discovery and replication databases

Consortium	Type	β	SE	*P*	FDR	P_{HEIDI}	PP.H4.abf
Breast Cancer Association Consortium	Overall	-2.42	0.41	3.59E-09	3.19E-06	4.41E-01	0.997
	LuminalA	-2.80	0.50	2.07E-08	4.04E-05	1.82E-01	0.992
	Triple-Negative	-3.18	0.76	2.81E-05	7.57E-03	9.54E-01	0.988
	BRCA1_Mutation	-2.47	0.59	2.62E-05	6.03E-03	5.47E-01	0.988
FinnGen Consortium	Overall	-1.79	0.53	6.38E-04	8.29E-02	7.84E-01	0.857

SMR, Summary data-based Mendelian randomization; β means the effect of genes on breast cancer using SMR method; SE means standard error; *P* means the P-value of the SMR analysis; FDR means P-value after false discovery rate; P_{HEIDI}, the P-value of the heterogeneity in dependent instruments; PP.H4.abf, the posterior probability of hypothesis 4

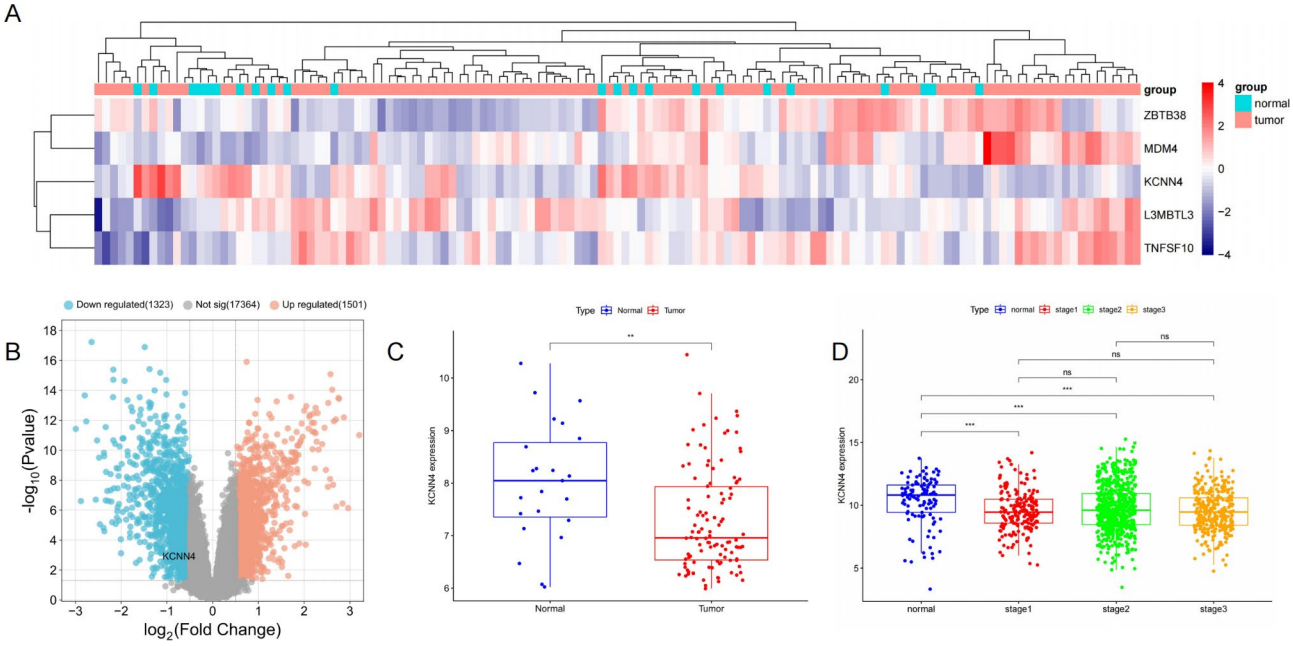

Fig. 4 Combination figure of BulkRNA-seq analysis. **A** Heat map of differential gene analysis in GEO database; **B** The volcano map of differential gene analysis in GEO database, blue represents down-regulated gene, orange represents up-regulated gene, and gray represents no significant gene; **C** Box plot of *KCNN4* gene expression(log_2counts) in normal tissues and breast cancer tissues in GEO database; **D** Box plot of *KCNN4* gene expression(log_2counts) in normal tissues and breast cancer tissues of different stages in TCGA_GTEx database

We conducted drug screening for five genes and reviewed the literature for relevant studies (Table S17). We found 21 drugs that target *KCNN4*, but none of them are specifically designed to treat breast cancer. However, it is worth noting that riluzole seems to have a preventive effect on breast cancer. We searched relevant literature and found that riluzole could induce apoptosis or inhibit the growth of breast cancer cells to a certain extent (Table S18) [18–22]. However, drugs targeting the other genes (*L3MBTL3*: Unc-1215, *MDM4*: Nutlin-3, *TNFSF10*: Onc201) have not yet been approved for clinical utilization.

Discussion

In this study, we employed MR analyses to investigate the genes within 14 immune cell types that are associated with breast cancer and its subtypes. Through SMR analysis (P_{FDR} <0.05 and P_{HEIDI} >0.05), colocalization with Hypothesis 4 posterior probability > 80% (PP.H4 > 0.8), and horizontal pleiotropy evaluation, we identified 17 genes across 9 immune cell types that showed significant associations with overall breast cancer and its subtypes. Among these 17 genes, we identified that *KCNN4*, *L3MBTL3*, *MDM4*, *TNFSF10*, and *ZBTB38* were genes consistently associated with overall breast cancer and some subtypes. In this study, we aimed to explore the causal associations of genes in various immune cell types with breast cancer and its subtypes through SMR and colocalization analyses, and indeed we identified a couple of strong reproducible associations among them. However, further replication analyses were needed to confirm these findings. *KCNN4* in MONOnc was also verified through the replication analysis and bulkRNA-seq analysis (GEO, TCGA_GTEx). As far as we know, our research is the first study to explore the genetic contributions of various genes across 14 immune cell types to breast cancer and its subtypes. The scRNA-seq has emerged as a valuable transcriptional profiling tool for defining cellular subpopulations and identifying cell-type-specific biomarkers and heterogeneity in various cancers, including breast cancer. We made use of extensive scRNA-seq data obtained from over a million peripheral blood mononuclear cells to examine the genetic effects on gene expression within each immune cell type.

KCNN4 (potassium calcium-activated channel subfamily N member 4), also known as SK4 and KCa3.1, is located in 19q13.31. Belonging to the family of K^+ channels activated by Ca^{2+} [23], $K_{Ca\ 3.1}$ channels are widely expressed in many organs and play a role in the migration, proliferation, and activation of blood cells and nervous system and vascular diseases [24–26]. In our SMR results, *KCNN4* exhibited a negative correlation with breast cancer. However, other studies indicate that high expression of *KCNN4* is related to tumor growth [27, 28] based on quantitative trait loci of gene expression, DNA methylation, and protein expression data [27, 28]. This may be because the K^+ channel is often dysregulated in cancer [29]. In particular, $K_{Ca\ 3.1}$ is upregulated during

tumor progression [30] and leads to malignancy. Previous studies have explored the association between KCNN4 and breast cancer by using mRNA expression data in whole blood, which could explain the potential inconsistency because our study examined the association by using single-cell transcriptomics within various immune cells. Recent evidence has suggested that ion channels expressed in tumor stromal cells, including immune cells, significantly contribute to the remodeling of the microenvironment and the progression of tumors. Studies have demonstrated that elevated levels of extracellular potassium (K^+) in the cancer microenvironment result in an increased intracellular K^+ concentration, which subsequently inhibits T cell function. However, the overexpression of $K_{V\ 1.3}$ and $K_{Ca\ 3.1}$ in T cells has been shown to enhance K^+ release, thereby restoring T cell function and reducing cancer growth [31, 32]. Therefore, the tumor lethality of T cells can be greatly amplified by using drugs that stimulate $K_{Ca\ 3.1}$ channels [32]. It has been reported that Riluzole can activate the SK4 channel [33], which provides an explanation for why Riluzole is capable of inhibiting the growth of breast cancer cells.

L3MBTL3, also referred to as *MBT1*, is a member of the MBT (malignant brain tumor) family [34]. Studies have demonstrated that the expression level of *L3MBTL3* varies between different cancers and normal tissues. However, the expression trends differ among various tumors, indicating that *L3MBTL3* may fulfill distinct functions in different tumor types [35]. Siddhartha et al. found that *L3MBTL3* was a common risk locus for breast and prostate cancers [36]. In addition, *L3MBTL3* variants were associated with colorectal, ER-negative breast, clear cell ovarian, and aggressive prostate cancers [37]. *ZBTB38*, a new member of the ZBTB family, is located on chromosome 3q23 and has eight exons. At present, studies specifically focusing on the relationship between *ZBTB38* and breast cancer are still lacking, but some studies have pointed out that the expression of *ZBTB38* is related to the development of prostate cancer. It was observed that elevated levels of *ZBTB38* could inhibit the proliferative and migratory capacities of prostate cancer cells [38], which was consistent with the results of another study [39].

In both triple-negative breast cancer and BRCA-TN breast cancer, we identified *TNFSF10* and *MDM4* as significant genes. *TNFSF10*, also known as *TRAIL*, is a member of the tumor necrosis factor (TNF) superfamily [40]. This ligand is classified as a homotrimeric type II transmembrane protein, which selectively triggers apoptosis in tumor cells or transformed cells, while demonstrating non-toxicity towards normal cells. Han et al. have used the mouse model to explore the impact of *TNFSF10*-deficient tumor cells on the modulation of immune cell infiltration and antitumor immune response [41]. Their findings revealed that decreased *TNFSF10* expression could result in an unfavorable tumor microenvironment (TME) and reduced responsiveness to chemotherapy and immunotherapy. The *MDM4* gene serves as an inhibitor of the p53 tumor suppressor protein [42]. Increased expression of the *MDM4* gene promotes breast cancer cells growth. Conversely, reducing *MDM4* expression can result in decreased cell numbers [43]. However, these findings contradict our research results, indicating a need for further investigation.

Our research presents several notable advantages. First, recognizing that gene function could vary by distinct cell type, we examined the causal relationship between gene expression in 14 different immune cell types and breast cancer, providing valuable insights into the cell-specific mechanisms underlying breast cancer. It should be noted that some genes were identified in different immune cells, which highlighted their potential combined action or synergy function on breast cancer. However, the effect sizes of the same genes on breast cancer and its subtypes showed a significant difference, underscoring the concept of tumor heterogeneity. Second, considering the various immunogenicity of different subtypes of breast cancer, we examined the overall breast cancer and its different subtypes separately. Third, we used replication analysis and bulkRNA-seq analysis for validation to enhance the robustness evidence of our studies. Finally, we employed various statistical analyses, including the HEIDI test and colocalization analysis, to ensure the validity and reliability of the results. A couple of limitations should be noted when we interpret our research findings. First, we only looked at European populations in both the discovery and validation databases, which means we cannot rule out the possibility that breast cancer may differ across ethnic groups. The incidence of breast cancer is higher in women of African descent than in women of European descent, which could contribute to growing health disparities in the field, especially as our goal is to identify druggable targets. Second, the small sample size of the eQTL data may have lower power to detect genes associated with breast cancer. Therefore, future studies including other more diverse populations and larger sample sizes were needed to make up these limitations.

Conclusion

Overall, we have successfully identified 17 genes that are associated with breast cancer and its subtypes across 9 different immune cell types. These genes exhibited varied effect sizes and distinct associations with breast cancer subtypes, offering insight into the potential role of immune cell-specific genes in cancer development and paving the way for targeted interventions or treatments in the future. Meanwhile, the identification of these genes opens up new avenues for the development of gene

- targeted drugs. By precisely targeting these genes, there is potential to design more effective and less toxic therapeutic agents, which could potentially enhance current breast cancer treatment approaches and offer additional options for patients to improve survival and quality of life.

Abbreviations

SMR	Summary data-based Mendelian randomization
eQTL	Expression quantitative trait loci
GWAS	Genome-wide association studies
bulkRNA-seq	Bulk RNA sequencing
HER2	Human epidermal growth factor receptor 2
TMB	Tumor mutation burden
MR	Mendelian randomization
IV	Instrumental variable
BCAC	The Breast Cancer Association Consortium
PBMCs	Peripheral blood mononuclear cells
scRNA-seq	Single-cell RNA sequencing
CD4NC	CD4 + primary and central memory T cells
CD4ET	CD4 + T cells with effector memory or central memory phenotype
CD4SOX4	SOX4-expressing CD4 + T cells
CD8ET	CD8 + T cells with effector memory phenotype
CD8NC	CD8 + infantilistic and central memory T cells
CD8S100B	CD8 + expressing S100B T cells
NK	Natural killer cells
NKR	NK recruitment cells
Bmem	Memory B cells
Bin	Immature and naive B cells
MONOc	Classical MONOcytes
MONOnc	Non-classical MONOcytes
DC	Dendritic cells
CIMBA	The Consortium of Investigators of Modifiers of BRCA1/2
ER	Estrogen receptor
PR	Progesterone receptor
GEO	The Gene Expression Omnibus
TCGA	The Cancer Genome Atlas
GTEx	The The Genotype-Tissue Expression
SNPs	Single nucleotide polymorphisms
2SLS	Two-step least-squares
HEIDI	The heterogeneity in dependent instruments
LD	Linkage disequilibrium
FDR	False discovery rate

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13058-024-01928-0.

Supplementary Material 1

Supplementary Material 2

Acknowledgements

We acknowledge and greatly appreciate the contributions of Seyhan Yazar et al. For providing the eQTL data used in this study, as well as the Breast Cancer Association Consortium and the FinnGen consortium for making their GWAS data available.

Author contributions

DZ, JJ and JL designed the research question; JL did the analyses; WS verified the analysis; JL drafted the manuscript; JL, WS, NL, HL, LW, HY, JJ, DZ revised the manuscript; DZ is responsible for the decision to submit the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by the Swedish Research Council (2021 – 01187) to Jianguang Ji, and the Beijing Municipal Health System Special Funds of High-Level Medical Personnel Construction (2022-3-042) to Deqiang Zheng.

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethical approval

All participating studies in the GWAS and eQTL data used in the article were approved by their appropriate ethics or institutional review board and all participants provided informed consent.

Competing interests

The authors declare no competing interests.

Author details

[1]Department of Epidemiology and Health Statistics, School of Public Health, Capital Medical University, Beijing, China
[2]Department of Cardiac Surgery, Beijing Chaoyang Hospital, Capital Medical University, Beijing, China
[3]Heart Center and Beijing Key Laboratory of Hypertension, Beijing Chaoyang Hospital, Capital Medical University, Beijing, China
[4]Department of Gynecologic Oncology, Fujian Maternity and Child Health Hospital College of Clinical Medicine for Obstetrics & Gynecology and Pediatrics, Fujian Medical University, Fuzhou, Fujian, China
[5]Faculty of Health Science, University of Macau, Taipa, Macao SAR, China

Received: 22 July 2024 / Accepted: 18 November 2024
Published online: 29 November 2024

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, Jemal A. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024;74(3):229–63.
2. Curigliano G, Burstein HJ, Winer EP, Gnant M, Dubsky P, Loibl S, et al. De-escalating and escalating treatments for early-stage breast cancer: the St. Gallen International Expert Consensus Conference on the primary therapy of early breast Cancer 2017. Ann Oncol. 2017;28(8):1700–12.
3. Soysal SD, Tzankov A, Muenst SE. Role of the Tumor Microenvironment in breast Cancer. Pathobiology. 2015;82(3–4):142–52.
4. Harris RE, Chlebowski RT, Jackson RD, Frid DJ, Ascenseo JL, Anderson G, et al. Breast cancer and nonsteroidal anti-inflammatory drugs: prospective results from the women's Health Initiative. Cancer Res. 2003;63(18):6096–101.
5. Kroemer G, Senovilla L, Galluzzi L, André F, Zitvogel L. Natural and therapy-induced immunosurveillance in breast cancer. Nat Med. 2015;21(10):1128–38.
6. Adams S, Gray RJ, Demaria S, Goldstein L, Perez EA, Shulman LN, et al. Prognostic value of tumor-infiltrating lymphocytes in triple-negative breast cancers from two phase III randomized adjuvant breast cancer trials: ECOG 2197 and ECOG 1199. J Clin Oncol. 2014;32(27):2959–66.
7. Tomar N, De RK. A brief outline of the immune system. Methods Mol Biol. 2014;1184:3–12.
8. Abbott M, Ustoyev Y. Cancer and the Immune System: the history and background of Immunotherapy. Semin Oncol Nurs. 2019;35(5):150923.
9. Thomas A, Routh ED, Pullikuth A, Jin G, Su J, Chou JW, et al. Tumor mutational burden is a determinant of immune-mediated survival in breast cancer. Oncoimmunology. 2018;7(10):e1490854.
10. Dieci MV, Miglietta F, Guarneri V. Immune infiltrates in breast Cancer: recent updates and clinical implications. Cells. 2021;10(2).
11. Bowden J, Holmes MV. Meta-analysis and mendelian randomization: a review. Res Synth Methods. 2019;10(4):486–96.
12. 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):3253–65.

13. Yazar S, Alquicira-Hernandez J, Wing K, Senabouth A, Gordon MG, Andersen S, et al. Single-cell eQTL mapping identifies cell type-specific genetic control of autoimmune disease. Science. 2022;376(6589):eabf3041.
14. Zhang H, Ahearn TU, Lecarpentier J, Barnes D, Beesley J, Qi G, et al. Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and subtype-specific analyses. Nat Genet. 2020;52(6):572–81.
15. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet. 2016;48(5):481–7.
16. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, Plagnol V. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet. 2014;10(5):e1004383.
17. Zhu X, You S, Du X, Song K, Lv T, Zhao H, Yao Q. LRRC superfamily expression in stromal cells predicts the clinical prognosis and platinum resistance of ovarian cancer. BMC Med Genomics. 2023;16(1):10.
18. Speyer CL, Bukhsh MA, Jafry WS, Sexton RE, Bandyopadhyay S, Gorski DH. Riluzole synergizes with paclitaxel to inhibit cell growth and induce apoptosis in triple-negative breast cancer. Breast Cancer Res Treat. 2017;166(2):407–19.
19. Speyer CL, Nassar MA, Hachem AH, Bukhsh MA, Jafry WS, Khansa RM, Gorski DH. Riluzole mediates anti-tumor properties in breast cancer cells independent of metabotropic glutamate receptor-1. Breast Cancer Res Treat. 2016;157(2):217–28.
20. Dolfi SC, Medina DJ, Kareddula A, Paratala B, Rose A, Dhami J, et al. Riluzole exerts distinct antitumor effects from a metabotropic glutamate receptor 1-specific inhibitor on breast cancer cells. Oncotarget. 2017;8(27):44639–53.
21. Speyer CL, Smith JS, Banda M, DeVries JA, Mekani T, Gorski DH. Metabotropic glutamate receptor-1: a potential therapeutic target for the treatment of breast cancer. Breast Cancer Res Treat. 2012;132(2):565–73.
22. Teh JL, Shah R, La Cava S, Dolfi SC, Mehta MS, Kongara S, et al. Metabotropic glutamate receptor 1 disrupts mammary acinar architecture and initiates malignant transformation of mammary epithelial cells. Breast Cancer Res Treat. 2015;151(1):57–73.
23. Ghanshani S, Wulff H, Miller MJ, Rohm H, Neben A, Gutman GA, et al. Up-regulation of the IKCa1 potassium channel during T-cell activation. Molecular mechanism and functional consequences. J Biol Chem. 2000;275(47):37137–49.
24. Chou CC, Lunn CA, Murgolo NJ. KCa3.1: target and marker for cancer, autoimmune disorder and vascular inflammation? Expert Rev Mol Diagn. 2008;8(2):179–87.
25. Sugunan S, Nampoothiri SS, Garg T, Krishnamurthy RG. Role of KCa3.1 channels in CNS diseases: a concise review. CNS Neurol Disord Drug Targets. 2016;15(10):1299–305.
26. Köhler R, Wulff H, Eichler I, Kneifel M, Neumann D, Knorr A, et al. Blockade of the intermediate-conductance calcium-activated potassium channel as a new therapeutic strategy for restenosis. Circulation. 2003;108(9):1119–25.
27. Zhang N, Li Y, Sundquist J, Sundquist K, Ji J. Identifying actionable druggable targets for breast cancer: mendelian randomization and population-based analyses. EBioMedicine. 2023;98:104859.
28. Ho PJ, Khng A, Tan BK, Khor CC, Tan EY, Lim GH et al. Characterizing the relationship between expression quantitative trait loci (eQTLs), DNA methylation quantitative trait loci (mQTLs), and breast Cancer risk variants. Cancers (Basel). 2024;16(11).
29. Pillozzi S, D'Amico M, Bartoli G, Gasparoli L, Petroni G, Crociani O, et al. The combined activation of K(ca)3.1 and inhibition of K(v)11.1/hERG1 currents contribute to overcome cisplatin resistance in colorectal cancer cells. Br J Cancer. 2018;118(2):200–12.
30. Muratori L, Petroni G, Antonuzzo L, Boni L, Iorio J, Lastraioli E, et al. hERG1 positivity and Glut-1 negativity identifies high-risk TNM stage I and II colorectal cancer patients, regardless of adjuvant chemotherapy. Onco Targets Ther. 2016;9:6325–32.
31. Eil R, Vodnala SK, Clever D, Klebanoff CA, Sukumar M, Pan JH, et al. Ionic immune suppression within the tumour microenvironment limits T cell effector function. Nature. 2016;537(7621):539–43.
32. Chandy KG, Norton RS, Immunology. Channelling potassium to fight cancer. Nature. 2016;537(7621):497–9.
33. Liu BS, Ferreira R, Lively S, Schlichter LC. Microglial SK3 and SK4 currents and activation state are modulated by the neuroprotective drug, riluzole. J Neuroimmune Pharmacol. 2013;8(1):227–37.
34. Xu T, Park SS, Giaimo BD, Hall D, Ferrante F, Ho DM, et al. RBPJ/CBF1 interacts with L3MBTL3/MBT1 to promote repression of notch signaling via histone demethylase KDM1A/LSD1. Embo j. 2017;36(21):3232–49.
35. Gan L, Yang C, Zhao L, Wang S, Ye Y, Gao Z. L3MBTL3 is a potential prognostic biomarker and correlates with Immune infiltrations in gastric Cancer. Cancers (Basel). 2023;16(1).
36. Kar SP, Beesley J, Amin Al Olama A, Michailidou K, Tyrer J, Kote-Jarai Z, et al. Genome-wide Meta-analyses of breast, ovarian, and prostate Cancer Association studies identify multiple new susceptibility loci Shared by at least two Cancer types. Cancer Discov. 2016;6(9):1052–67.
37. Toth R, Scherer D, Kelemen LE, Risch A, Hazra A, Balavarca Y, et al. Genetic variants in epigenetic pathways and risks of multiple cancers in the GAME-ON Consortium. Cancer Epidemiol Biomarkers Prev. 2017;26(6):816–25.
38. Ding G, Lu W, Zhang Q, Li K, Zhou H, Wang F, et al. ZBTB38 suppresses prostate cancer cell proliferation and migration via directly promoting DKK1 expression. Cell Death Dis. 2021;12(11):998.
39. de Dieuleveult M, Marchal C, Jouinot A, Letessier A, Miotto B. Molecular and clinical relevance of ZBTB38 expression levels in prostate Cancer. Cancers (Basel). 2020;12(5).
40. Wang S, El-Deiry WS. TRAIL and apoptosis induction by TNF-family death receptors. Oncogene. 2003;22(53):8628–33.
41. Han YJ, Zhang J, Hardeman A, Liu M, Karginova O, Romero R, et al. An enhancer variant associated with breast cancer susceptibility in black women regulates TNFSF10 expression and antitumor immunity in triple-negative breast cancer. Hum Mol Genet. 2023;32(1):139–50.
42. Wade M, Li YC, Wahl GM. MDM2, MDMX and p53 in oncogenesis and cancer therapy. Nat Rev Cancer. 2013;13(2):83–96.
43. Haupt S, Buckley D, Pang JM, Panimaya J, Paul PJ, Gamell C, et al. Targeting mdmx to treat breast cancers with wild-type p53. Cell Death Dis. 2015;6(7):e1821.

Publisher's note