

RESEARCH ARTICLE

Analysis of the autoimmune regulator (*AIRE*) gene variant rs2075876 (G/A) association with breast cancer susceptibility

Manal S. Fawzy^{1,2}  | Eman A. Toraih^{3,4} 

¹Department of Biochemistry, Faculty of Medicine, Northern Border University, Arar, Saudi Arabia

²Department of Medical Biochemistry and Molecular Biology, Faculty of Medicine, Suez Canal University, Ismailia, Egypt

³Department of Surgery, School of Medicine, Tulane University, New Orleans, Louisiana, USA

⁴Genetics Unit, Department of Histology and Cell Biology, Faculty of Medicine, Suez Canal University, Ismailia, Egypt

Correspondence

Manal S. Fawzy, Department of Medical Biochemistry, Faculty of Medicine, Northern Border University, P.O.Box: 1321, Arar 91431, Saudi Arabia.
Email: manal2_khashana@ymail.com

Funding information

The Deanship of Scientific Research, Northern Border University, Grant/Award Number: MED-2018-3-9-F-7821

Abstract

Background: Recently, unexpected autoimmune regulator (*AIRE*) implication in the scenario of several cancers, including breast cancer (BC), has emerged. This study aims to explore for the first time the possible association between *AIRE* gene rs2075876 G>A variant and BC risk in a sample of the Middle East population.

Method: In this case-control study, we genotyped *AIRE* rs2075876 G>A variant in 200 unrelated patients with BC and 340 cancer-free controls using a real-time allelic discrimination polymerase chain reaction. Odds ratios (ORs) and 95% confidence intervals (CIs) were applied to estimate the strength of association under several genetic models. In silico analysis of *AIRE* was also executed.

Results: The minor allele (A) frequency of the specified variant accounted for 0.28 in the controls. G/G homozygote was significantly more frequent among patients (94%) compared to controls (66%) ($P < .001$). After adjusting confounding variables, individuals with A allele conferred protection against developing BC under allelic model (OR = 0.33, 95% CI = 0.20-0.55), recessive model (OR = 0.25, 95% CI = 0.10-0.60), dominant model (OR = 0.12, 95% CI = 0.05-0.29), and homozygote comparison (OR = 0.20, 95% CI = 0.08-0.50). In silico analysis revealed *AIRE* enrichment in several cancer-related pathways. Kaplan-Meier plotter for the cancer databases showed association of *AIRE* expression with prognosis in triple-negative BC (HR = 2.44, 95% CI = 1.44-4.15, log-rank P -value < .001).

Conclusion: The *AIRE* rs2075876 G>A variant showed association with BC risk in the study population. Further large-scale replication studies in different ethnicity are warranted to confirm the findings.

KEYWORDS

AIRE, Breast cancer, real-time genotyping, single nucleotide polymorphism

1 | INTRODUCTION

According to the global cancer statistics, breast cancer (BC) is the most commonly diagnosed female cancer worldwide.^{1,2} It is

a heterogeneous disease that can be classified into distinct subgroups: luminal A [estrogen receptor (ER)⁺ and/or progesterone (PR)⁺/human epidermal receptor 2 (HER2)⁻], luminal B (ER⁺ and/or PR⁺/HER2⁺), (ER⁻ and PR⁻/HER2⁺), or triple-negative (ER⁻, PR⁻, and

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2020 The Authors. *Journal of Clinical Laboratory Analysis* Published by Wiley Periodicals LLC.

HER2⁺) subtypes according to the expression of particular clusters of co-expressed genes.³ Despite the breast cancer etiology has not been completely elucidated, it is thought to be multifactorial, with implication of both environmental and genetic factors.⁴

The immune system serves as an important natural barrier to cancer development. Innate and adaptive responses are carefully orchestrated through soluble and membrane-bound receptors to eliminate precancerous cells and control neoplastic progression.⁵ Avoiding immune destruction has been considered a hallmark of cancer.^{6,7} It is therefore of great interest to identify immune genes that influence susceptibility to breast cancer.⁸

One of the best characterized genes that play a crucial role in autoimmunity is the protein-coding autoimmune regulator (*AIRE*) gene (Gene ID: 326). This gene encodes a transcriptional regulator that plays an essential role to promote the negative selection of autoreactive T cells in the thymus by regulating the expression of a wide array of self-antigens that have the commonality of being tissue-restricted in their expression pattern in the periphery, called tissue-restricted antigens (<https://www.genecards.org/cgi-bin/carddisp.pl?gene=AIRE>). *AIRE* protein, in addition, was shown to have E3 ubiquitin ligase activity, which mediates proteasome pathway crucial for developing self-tolerance, cell cycle, DNA repair, and cell trafficking.⁹ Interestingly, not only is *AIRE* highly expressed in medullary thymic epithelial cells (mTECs), but there is an experimental evidence of its expression in peripheral lymphoid organs.¹⁰⁻¹² Gardner et al, have reported the extra-thymic *AIRE*-expressing cells (eTACs) to be a distinct “bone marrow-derived antigen-presenting CD45^{low}” population that induce functional inactivation of effectors cells (CD4⁺) T cells mediating tolerance through a mechanism involving deficient co-stimulation.¹³ Very recently, *AIRE* expression was reported in some solid cancers, including BC cells.¹³⁻¹⁵ Bianchi et al, for the first time, have detected an unexpected involvement of *AIRE* in the BC scenario.¹⁵ They reported *AIRE* expression in BC and an association with prognosis, which provides compelling evidence for the relevance of this unusual transcription factor as a new marker for BC.

Single nucleotide polymorphisms (SNPs) are the most common genetic variations, which influence inter-individual predisposition to breast carcinogenesis and prognosis.¹⁶ SNP association analysis has provided valuable information about the genetic susceptibility of BC.¹⁷ The human *AIRE* gene is located in the chromosome 21q22.3 region. Polymorphisms in this gene were previously shown to be genetically associated with autoimmune diseases, such as rheumatoid arthritis and autoimmune polyendocrine syndrome type 1 among others.^{18,19} The pathogenic variant rs2075876, caused by a point mutation (c.653-387G>A) in intron 5 (chr21:44289270), is located within exons forming the domain in the middle of protein chain, which functions in chromatin-dependent transcriptional regulation, nuclear localization, and oligomer formation.²⁰

As no case-control study of *AIRE* gene variant rs2075876 in sporadic BC has been reported and given the association of *AIRE* expression with breast cancer prognosis reported by Bianchi et al¹⁵ and confirmed by our in silico analysis (Figure 1), the authors were

inspired to examine for the first time the possible association between rs2075876 variant and BC risk in a sample of Middle East population.

2 | MATERIALS AND METHODS

2.1 | Study participants

In this case-control study, a total of 540 unrelated Egyptian women were enrolled (340 controls and 200 cases with primary BC). The cases were consecutively recruited from the Oncology Diagnostic Unit of Suez Canal University, Ismailia, Egypt. Patients had a histologically confirmed primary BC with relevant stage assigned according to the TNM classification of the Sixth Edition of the American Joint Committee on Cancer (AJCC) Manual for Staging of Cancer²¹ and graded by the Nottingham grading system.²² The participants who have pre-surgical treatment of chemotherapy, radiotherapy, endocrine therapy, or biotherapy were excluded. The unrelated healthy control group was randomly selected from the blood bank donors based on a negative history of any disease interfering with the study and has no family history of breast cancer or any other neoplastic diseases. All participants also had no history of any chronic diseases (eg, heart disease, renal problems, and diabetes mellitus). All relevant clinical and pathological information was revised and completed from the medical files including cancer type, grade and stage at the time of diagnosis, presence of lymph node metastases, and ER, PR, and HER2 expression in breast cancer tissue. The study was conducted in accordance with the guidelines in the Declaration of Helsinki and was approved by the ethics committee of the researchers' university. The informed consent was obtained from all participants before taking part.

2.2 | Blood sampling

Venous blood samples were collected from all participants under a complete aseptic condition on plain and trisodium ethylene-diamine-tetra-acetic acid (EDTA) (1 mg/mL) tubes. The former tubes were centrifuged at 2500 g for 30 minutes at 4°C, and the serum was divided into aliquots and kept frozen at -80°C until analysis of any required laboratory test that is not available in the patients' medical records. The EDTA tubes were used for genotyping studies.

2.3 | *AIRE* mutation analysis in cancer databases

The mutation spectrum of the *AIRE* gene in cancer was identified from The Cancer Genome Atlas (TCGA, <https://cancergenome.nih.gov/>), which contains genome maps from >30 cancer types. The role of the *AIRE* gene in cancer hallmarks was determined using Cancer Hallmarks Analytics Tool (CHAT) (<http://chat.lionproject.net/>), a web browser for classifying cancer-related text and text mining-associated biological processes from PubMed articles.²³ The prognostic value of the *AIRE* gene in BC patients was explored in Kaplan-Meier

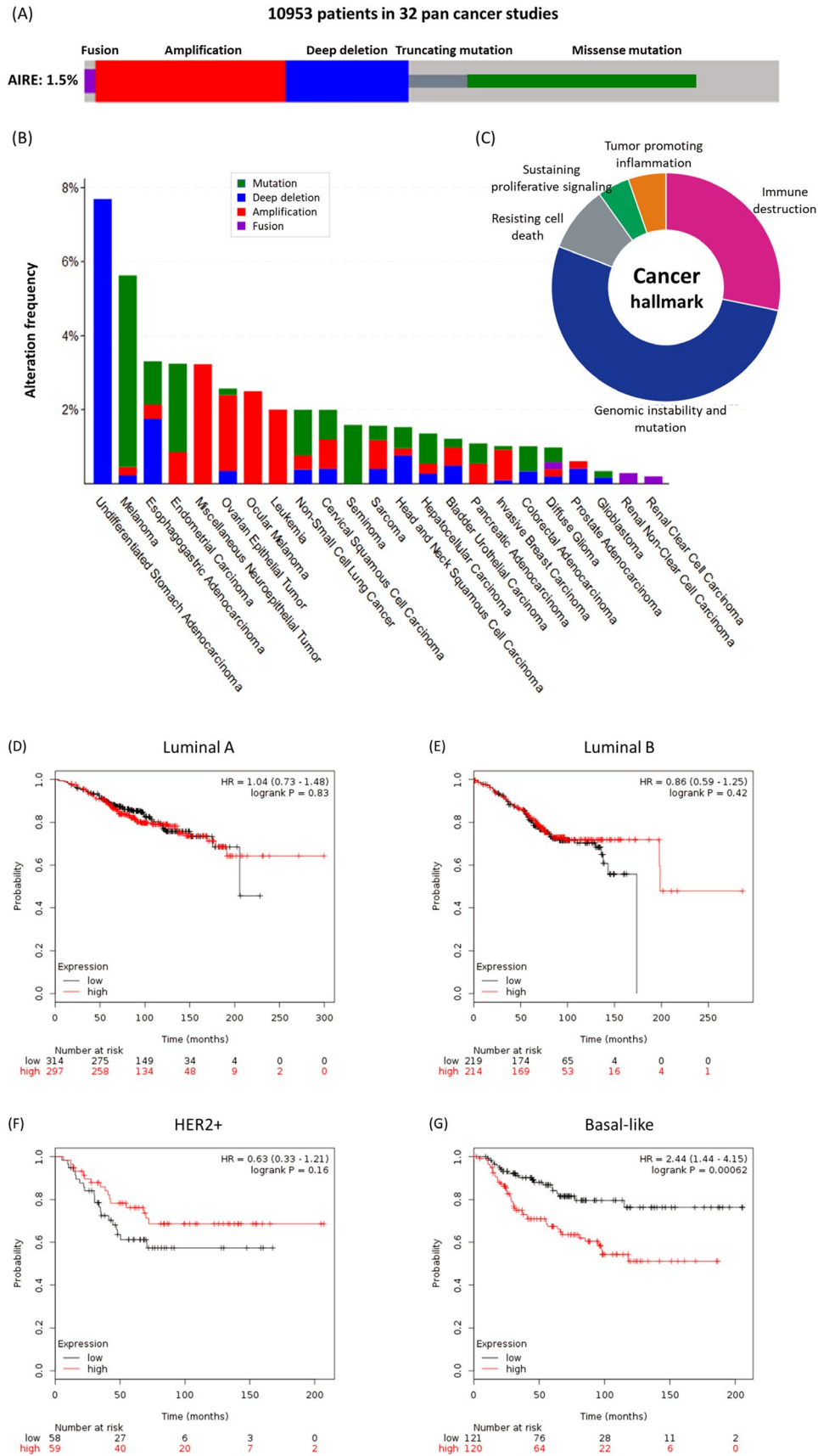


FIGURE 1 Prognostic role of *AIRE* gene in cancer. A, Alteration frequency of the *AIRE* gene in 32 pan-cancer studies in The Cancer Genome Atlas (TCGA). B, Type of somatic mutations and copy number variations. C, Enrichment of the *AIRE* gene in cancer hallmarks. D-G, Kaplan-Meier estimates for overall survival analysis in different molecular intrinsic subtypes

Plotter (<https://kmpplot.com/analysis/>), which includes gene chip and RNA-seq data from microarray and TCGA data.²⁴ The following thirty-three datasets were included: GSE11121 (N = 200), GSE12093 (N = 136), GSE12276 (N = 204), GSE1456 (N = 159), GSE16391 (N = 55), GSE6446 (N = 120), GSE16716 (N = 47), GSE17705 (N = 196), GSE17907 (N = 54), GSE9195 (N = 77), GSE18728 (N = 61), GSE19615 (N = 115), GSE20194 (N = 45), GSE20271 (N = 96), GSE2034 (N = 286), GSE20685 (N = 327), GSE20711 (N = 90), GSE21653 (N = 240), GSE2063 (N = 99), GSE26971 (N = 276), GSE2990 (N = 102), GSE31448 (N = 71), GSE31519 (N = 67), GSE32646 (N = 115), GSE3494 (N = 251), GSE37946 (N = 41), GSE41998 (N = 279), GSE42568 (N = 121), GSE45255 (N = 139), GSE4611 (N = 153), GSE5327 (N = 58), GSE6532 (N = 82), and GSE7390 (N = 198). Multivariate analysis was performed with proportionate hazards assumption. Cutoff value for gene expression was set at the median, and the following variables were included in the model MK167, ESR1, and HER2 status. Next, stratified analysis by molecular subtyping was performed.

2.4 | Allelic discrimination of rs2075876 G/A variant

Genomic DNA was extracted from the buffy coat using QIAamp DNA Blood Mini kit (Catalog No. 51 104; Qiagen) in accordance with the manufacturer's instructions. NanoDrop ND-1000 (NanoDrop Technologies, Inc) was used for assessment of DNA concentration and purity. Genotyping of *AIRE* gene rs2075876 (G/A) variant was identified using StepOne™ Real-Time Polymerase Chain Reaction (Applied Biosystems, USA). PCRs were run in a total volume of 20 μ L containing genomic DNA (20 ng) with nuclease-free water up to 9 μ L, TaqMan SNP Genotyping Assay Mix (1 μ L) (Applied Biosystems, assay ID C_15863141_20), TaqMan Genotyping PCR Master Mix (10- μ L).

The thermocycler conditions were followed: denaturation at 95°C for 10 minutes, followed by 40 cycles of 90°C for 15 seconds and 60°C for 1 minutes.²⁵ Genotyping was conducted blind to the case/control status. Ten percent of the samples were randomly selected and re-genotyped with 100% reproducibility in genotyping recall (Figure 2).

2.5 | Statistical analysis

Data were analyzed using SPSS (IBM SPSS Statistics for Windows, version 26.0., IBM Corp.). Chi-square and Fisher's exact tests were used for qualitative variables. Allele and genotype frequencies were calculated as previously described.²⁶ The Hardy-Weinberg equilibrium (HWE) was estimated using the Online Encyclopedia for Genetic Epidemiology (OEGE) software (<http://www.oege.org/software/hwe-mr-calc.shtml>). Adjusted odds ratios (ORs) with a 95% confidence interval (CI) using logistic regression models were calculated for multiple genetic association models.²⁷ The study power estimation using G power-3 software (<http://www.gpower.hhu.de/>) showed that with the specified case-control study design and allowable error rate (alpha error = 0.05), the specified sample size for each group can give 99% study power. A two-tailed *P*-value of .05 was considered statistically significant.

3 | RESULTS

3.1 | Baseline characteristics of the study population

The mean age of controls was 46.1 ± 10.8 and that of BC patients was 45.3 ± 12.3 ($P = .557$). Demographic, clinical, and pathological characteristics are represented in Table 1.

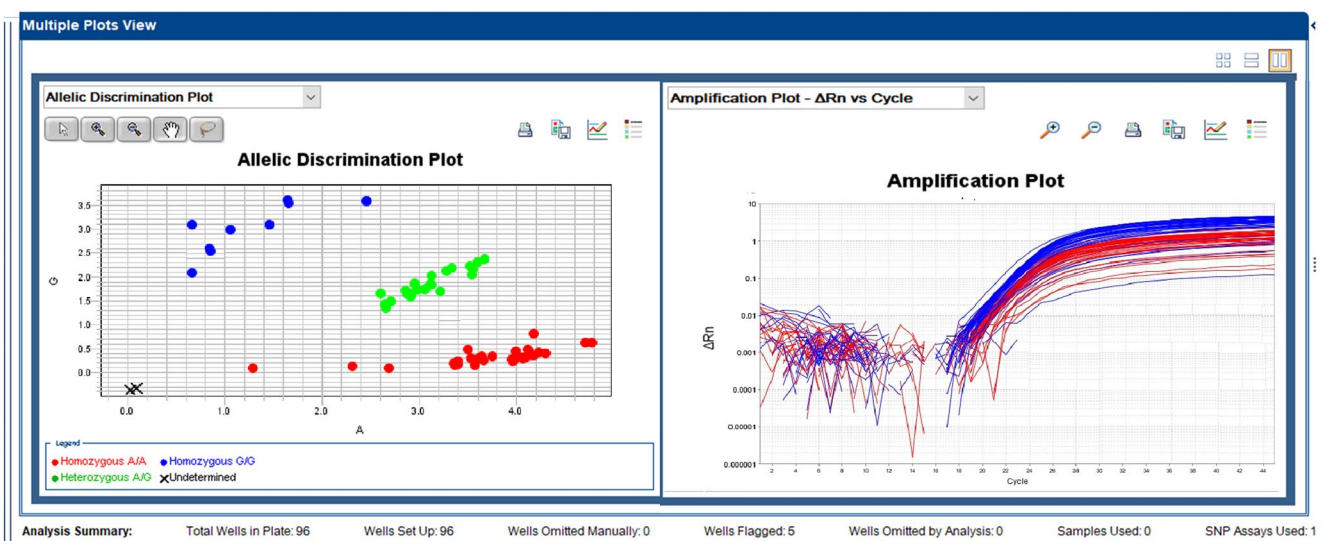


FIGURE 2 Allele discrimination plot of one allele (FAM probe) versus the other allele (VIC probe) of the real-time PCR run (SNP Assay: rs2075876 G/A) (left panel) and the amplification plot for each sample, which represents fluorescence from both probes FAM/VIC (right panel). Alleles as wild-type homozygous G/G (upper left blue cluster), heterozygous allele G/A (middle green cluster), mutant homozygous A/A (lower right red cluster), and the negative controls (x) that represent the background fluorescence from samples with no template DNA were retrieved by allele calling software (SDS version 1.3.1; Applied Biosystems)

TABLE 1 Demographic and clinical characteristics of breast cancer patients

Characteristics	n	%
Risk factors		
FH of cancer		
Negative	140	70.0
Positive	60	30.0
Degree of relation		
First degree	28	14.0
Second degree	24	12.0
Third degree	8	4.0
Menarche		
Late	72	36.0
Early	128	64.0
Menopause		
Early	90	45.0
Late	10	5.0
OCP usage		
Negative	160	80.0
Positive	40	20.0
Parity		
Multipara	172	86.0
Nullipara	28	14.0
Late first gravida		
Negative	192	96.0
Positive	8	4.0
Clinical presentation		
Symptoms		
Mastalgia	62	31.0
Breast mass	164	82.0
Skin lesions	32	16.0
Nipple change	36	18.0
Axillary pain	14	7.0
Axillary mass	36	18.0
Side		
Left	74	37.0
Right	383	191.5
Site		
LOQ	18	9.0
UIQ	38	19.0
UOQ	96	48.0
Retro	48	24.0
No of masses		
Single	152	76.0
Multiple	48	24.0

(Continues)

TABLE 1 (Continued)

Characteristics	n	%
Pathological analysis		
HPD		
Ductal	76	38.0
Lobular	52	26.0
Medullary	28	14.0
Mucinous	20	10.0
Tubular	12	6.0
Metaplastic	12	6.0
Grade		
G2	128	64.0
G3	72	36.0
LVI		
Negative	104	52.0
Positive	96	48.0
T stage		
T2	100	50.0
T3	44	22.0
T4B	40	20.0
T4D	16	8.0
N stage		
N0	56	28.0
N1	68	34.0
N2	64	32.0
N3	12	6.0
M stage		
M0	88	44.0
M1	84	42.0
Mx	28	14.0
Site of metastasis		
Bone	36	18.0
Liver	12	6.0
Lung	30	15.0
Skin infiltration		
Negative	160	80.0
Positive	40	20.0
Receptor status		
ER/PR		
Negative	80	40.0
Positive	120	60.0
HER2+		
Negative	164	82.0
Positive	36	18.0

(Continues)

TABLE 1 (Continued)

Characteristics	n	%
Molecular subtype		
Luminal A	96	48.0
Luminal B	24	12.0
HER2+	12	6.0
TNBC	68	34.0
IHPI score		
Good	120	60.0
Moderate	68	34.0
Poor	12	6.0
Prognostic factors		
Clinical stage		
IIA	36	18.0
IIB	28	14.0
IIIA	24	12.0
IIIB	28	14.0
IV	84	42.0
NPI score		
Good	100	50.0
Poor	100	50.0
ESMO score		
Low risk	76	38.0
High risk	124	62.0
Recurrence		
Negative	96	48.0
Positive	104	52.0
DFS		
Prolonged	96	48.0
Short	104	52.0

Note: Data are presented as number and percentage. Chi-square test was applied.

Abbreviations: FH, family history; OCP, oral contraceptive pills; HRT, hormonal replacement therapy; HPD, histopathological diagnosis; LVI, lymphovascular infiltration; ER/PR, estrogen and progesterone receptor; HER2⁺, positive for human epidermal growth factor receptor 2/neu receptor; IHPI, Immunohistochemical Prognostic Index estimated based on the three receptor status (HER2, ER, and PR); NPI, Nottingham Prognostic Index, calculated as $[0.2 \times \text{tumor size in cm}] + \text{tumor grade [1-3]} + \text{lymph node stage [1-3, according to stages A-C]}$; ESMO, European Society of Medical Oncology; DFS, disease-free survival.

3.2 | Prognostic value of *AIRE* from the cancer databases

Screening 10 953 cancer patients in thirty-two TCGA pan-cancer studies revealed an alteration of the *AIRE* gene in 1.5% of cases (Figure 1A). Somatic mutations included amplification, deep deletions, fusion events, and mutations in various types of cancer (Figure 1B). According to CHAT database, *AIRE* is enriched in immune destruction, genome instability and mutation, resisting cell

death, sustaining proliferative signaling, and tumor-promoting inflammation (Figure 1C). KM survival curves demonstrated an association of *AIRE* expression level with prognosis in triple-negative basal-like molecular subtype of BC (HR = 2.44, 95% CI = 1.44-4.15, log-rank *P*-value < .001) (Figure 1G).

3.3 | Genotyping of *AIRE* polymorphism

AIRE gene variant rs2075876 is located within intron 5 and caused by substitution of G to A (c.653-387G>A). Genotype frequencies of rs2075876 were deviated from HWE in the study population (ie, a marked decrease in the heterozygous frequency of the overall study population was observed; hence, the observed genotype frequency was not in line with the expected one) (Supplementary Table S1). MAF (A allele) accounted for 0.28 in controls. According to 1000 Genome Project (<https://www.ncbi.nlm.nih.gov/variation/tools/1000genomes/>), this figure was significantly different from that of Africans (MAF = 0.35) and East Asians (MAF = 0.37). On comparing between patients and controls, G/G homozygote was significantly more frequent among patients (94%) compared to controls (66%) (*P* < .001). In contrast, the other homozygote A/A was less represented among cases, accounting for 6% of cases, while heterozygote genotype G/A was absent in patients (Figure 3).

3.4 | Genetic association model for disease risk assessment

After adjusting confounding variables, individuals with A allele conferred protection against developing BC under allelic model (OR = 0.33, 95% CI = 0.20-0.55), recessive model (OR = 0.25, 95% CI = 0.10-0.60), dominant model (OR = 0.12, 95% CI = 0.05-0.29), and homozygote comparison (OR = 0.20, 95% CI = 0.08-0.50) (Figure 3D). Given the low frequency of A/A in patients (6%) and absence of heterozygote genotype, multiple regression analysis was not feasible. Table 2 shows no significant associations between the homozygous minor allele *AIRE* variant (A/A) with the clinicopathological features of BC patients.

4 | DISCUSSION

Given the newly emerged implication of several *AIRE* gene variations in multiple cancers (Figure 1B), and as there is no association analysis of the study *AIRE* variant with BC risk has been documented, this study aimed to investigate whether *AIRE* rs2075876 G>A polymorphism is associated with the development of BC in our population. Here, we identified that the study variant is associated with BC susceptibility under several genetic association models.

As a transcriptional and post-transcriptional regulator in medullary thymic epithelial cells and peripheral tissues, *AIRE* plays a substantial role in "thymocyte education" and "negative selection" by controlling

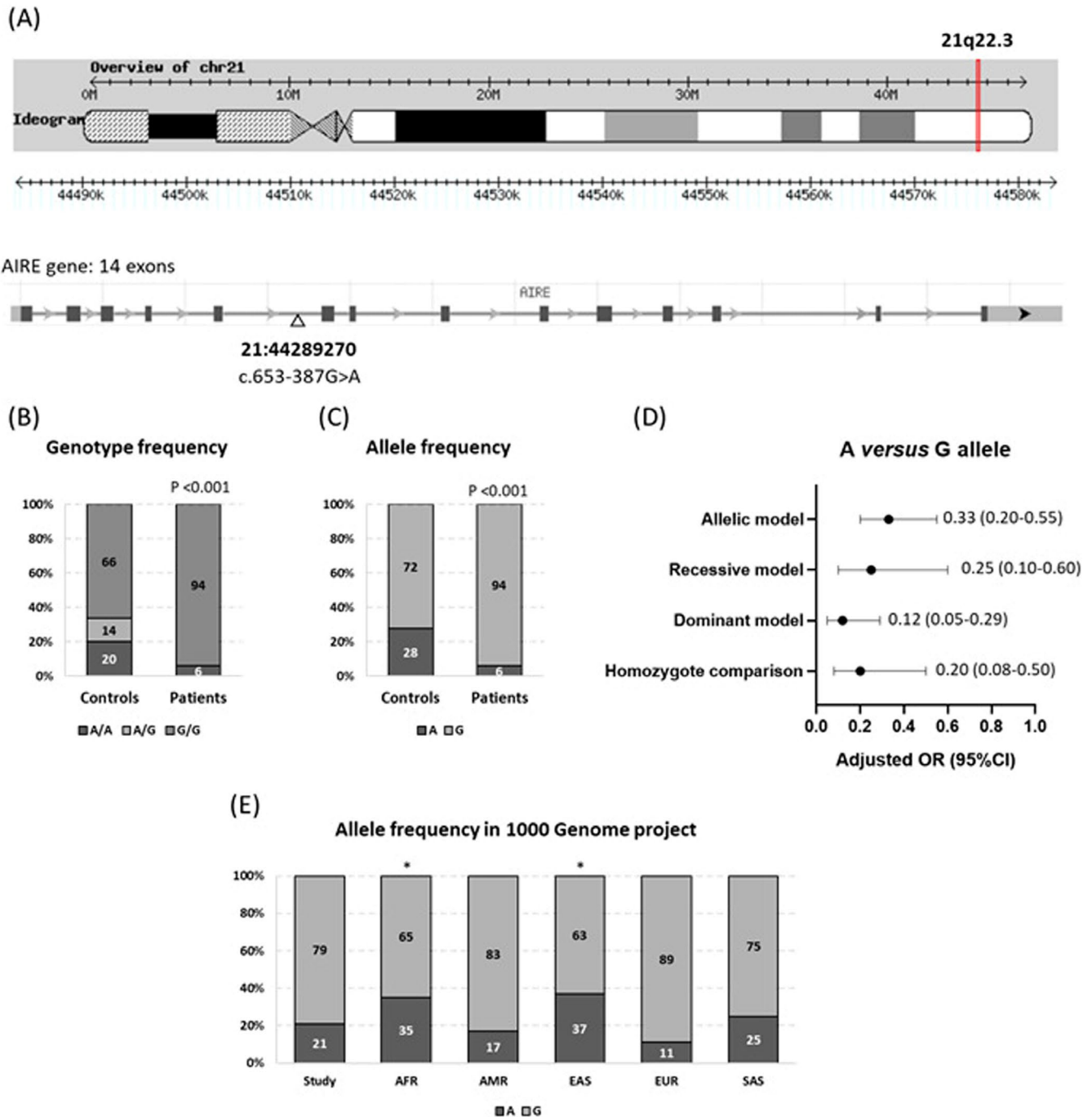


FIGURE 3 Genotype and allele frequencies of the AIRE variant in the study population. A, Chromosomal and genomic location of the AIRE rs2075876 (G/A) variant within intron 5. B, Genotype frequency of rs2075876 variant. Chi-square test was applied. C, Allele frequency of rs2075876 variant. Chi-square test was applied. D, Genetic association models and breast cancer risk. OR, odds ratio; 95% CI, 95% confidence interval. Adjusted by age, family history, early menarche, late menopause, nullipara, late first pregnancy, absent breastfeeding, and hormonal treatment. E, Comparison between the study allele frequency and that in other populations. AMR, American; AFR, African; EAS, East Asian; EUR, European; SAS, South Asians. (*) Significant from allele frequency identified in the study

the peripheral antigen expression in thymus.²⁸ Accumulating evidence indicated that AIRE protein is a non-conventional transcriptional factor as it does not bind directly to any DNA motifs; however, it is implied in releasing the stalled RNA polymerase II at the promoter regions of the target genes to induce gene expression.²⁹ Our in silico analysis revealed that this transcriptional regulator

can significantly interact with several protein partners implied in many physiological/pathological cellular processes with high confidence level (Figure 4).³⁰ These proteins, including the AIRE, have been enriched in nucleosome assembly, chromatin silencing, interleukin-7-mediated signaling pathway, telomere organization, regulation of immune system process, response to stress and stimulus,

Characteristics		A/A (n = 12)	G/G (n = 188)	P value
Age	Mean ± SD	48.0 ± 15.2	45.1 ± 12.3	.27
FH		8 (66.7)	172 (70.2)	.85
		4 (33.3)	56 (29.8)	
Skin infiltration	No	10 (83.3)	150 (79.8)	.83
	Yes	2 (16.7)	38 (20.2)	
Metastasis	M0	4 (33.3)	84 (44.7)	.86
	M1	6 (50)	78 (41.5)	
	Mx	2 (16.7)	26 (13.8)	
Clinical stage	IIA	2 (16.7)	34 (18.1)	.88
	IIB	0 (0)	28 (14.9)	
	IIIA	2 (16.7)	22 (11.7)	
	IIIB	2 (16.7)	26 (13.8)	
	IV	6 (50)	78 (41.5)	
NPI	Good	4 (33.3)	96 (51.1)	.67
	Poor	8 (66.7)	92 (48.9)	
ESMO	Low risk	4 (33.3)	72 (38.3)	.80
	High risk	8 (66.7)	116 (61.7)	
ER/PR	Negative	2 (16.7)	78 (41.5)	.39
	Positive	10 (83.3)	110 (58.5)	
HER2 ⁺	Negative	10 (83.3)	154 (81.9)	.93
	Positive	2 (16.7)	34 (18.1)	
Molecular subtype	Luminal A	8 (66.7)	88 (46.8)	.67
	Luminal B	2 (16.7)	22 (11.7)	
	HER2 ⁺	0 (0)	12 (6.4)	
	TNBC	2 (16.7)	66 (35.1)	
IHP	Good	10 (83.3)	110 (58.5)	.37
	Moderate	2 (16.7)	66 (35.1)	
	Poor	0 (0)	12 (6.4)	
Recurrence	Negative	4 (33.3)	92 (48.9)	.45
	Positive	8 (66.7)	96 (51.1)	
Survival	Alive	8 (66.7)	156 (83)	.29
	Dead	4 (33.3)	32 (17)	
DFS	Mean ± SD	12.5 ± 7.8	14.3 ± 5.8	.45
OS	Mean ± SD	16.0 ± 5.4	16.5 ± 4.2	.75

Note: Data are presented as mean ± SD or number (percentage). The t test and two-sided chi-square tests were used.

Abbreviations: FH, family history; NPI, Nottingham Prognostic Index; calculated as $[0.2 \times \text{tumor size in cm}] + \text{tumor grade [1-3]} + \text{lymph node stage [1-3, according to stages A-C]}$; ESMO, European Society of Medical Oncology; ER/PR, estrogen and progesterone receptors; HER2⁺, positive for human epidermal growth factor receptor 2/neu receptor; IHP, Immunohistochemical Prognostic Index estimated based on the three receptor status (HER2, ER, and PR); DFS, disease-free survival; OS, overall survival.

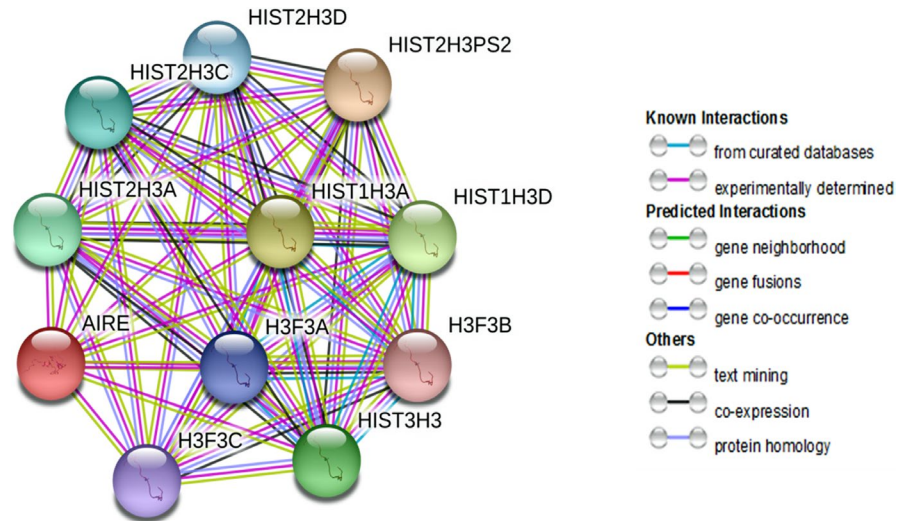
regulation of transcription and RNA metabolic process, and positive regulation of cell growth among other identified processes (Table S2). KEGG (Kyoto Encyclopedia of Genes and Genomes) enrichment analysis also revealed that AIRE is involved in “transcriptional misregulation in cancer” pathway (hsa05202) (Table S2). Interestingly,

all the above-mentioned pathways and processes were identified to play essential roles in BC etiology, progression, and/or outcome,³¹⁻³⁹ which add evidence to AIRE implication in the BC.

Several AIRE variants were previously shown to be associated with many autoimmune diseases,^{13,13,13,13} including rheumatoid

TABLE 2 Association between clinicopathological characteristics of BC patients and AIRE rs2075876 G>A variant homozygous genotypes

FIGURE 4 STRING analysis of the AIRE protein (accession#: CAA08759) with its experimentally validated functional partners. All the proteins encoded by a single, protein-coding gene locus are represented as network nodes. Protein-protein associations (ie, contribute to a shared function) are represented by colored edges, which indicate the type of interaction evidence that is explained in the figure key (right side). HIST2H3PS2, histone cluster 2 H3 pseudogene; HIST1H3A, histone cluster 1 H3 family member a; H3F3A, H3 histone, family 3A. Source, STRING version 11 (<https://string-db.org/>)³⁰



arthritis,^{18,40} vitiligo,⁴¹ systemic sclerosis,⁴² Addison disease,⁴³ systemic lupus erythematosus,⁴⁴ and alopecia areata.²⁵ However, a limited number of association studies for rs2075876 G>A polymorphism were identified.^{25,40,45-47} Furthermore, no previous study explored the relationship of the specified study variant with the BC.

As rs2075876 G>A is an intronic variant with low pathogenicity prediction score (6.657) based on the online CADD PHRED score prediction tool for single or combined variants <https://cadd.gs.washington.edu/snv>,⁴⁸ it is difficult to explain the exact biological function of this polymorphism. However, the non-coding variants at the transcription level may disrupt the transcription factor binding sites, the splice sites, and other functional sites as proposed by Chen et al⁵ and supported recently by Conteduca et al.⁴⁹ According to the GEO (Gene Expression Omnibus) database mining, Terao et al revealed a statistically significant correlation between the risk allele (A) of rs2075876 and AIRE down-regulation.⁴⁵ Furthermore, the specified variant may be in linkage disequilibrium with distinct genetic markers, which synergistically influence disease/cancer risk that deserve future functional studies to evaluate the role of this putative variant in disease risk.

It is worth noting that despite the genotype frequencies of the study variant deviated from the HWE, no standard guidelines for rejecting SNPs that depart from HWE have been developed.⁵⁰ In practice, all variants for which HWE *P*-values decrease below a pre-determined threshold should be checked manually for genotyping quality.⁵¹ As the current genotyping of this study was based on a high-quality real-time TaqMan allelic discrimination PCR assay with genotype recall rate (100%), and the specified quality control measurements were applied throughout the methodology, the current variant departure from HWE suggested a violation of HWE assumptions in the study groups rather than technical genotyping errors.^{50,51}

Although this study has enough power to detect the association of the specified polymorphism with BC risk, some potential limitations merit careful consideration: first, as the patients and controls recruited from one center may not be representative of the general population. Second, with a deficit in the number of observed

heterozygotes as well as the presence of few cases carrying A/A genotype, this could potentially affect HWE congruence. Thus, our findings should be confirmed by large-scale multicenter studies in different populations. Third, the authors have to confirm that despite there was a significant association with BC risk conferred by an individual locus, this might have a small impact on the disease risk as genes act multiplicatively. However, this small impact could be included in more advanced multiplicative polygenic risk scores.⁵² Individually, these risk alleles confer a very small increase in BC risk but their joint effect with other lifestyle risk factors could “improve risk-reduction and screening strategies by targeting those most likely to benefit” as concluded by Lakeman et al.⁵³ Fourth, the underlying mechanisms of the implication of this SNP in BC were not investigated. Further functional analyses are required to explore the molecular mechanisms by which this variant could participate in BC susceptibility.

In conclusion, our study provided evidence that rs2075876 G>A variant of the AIRE gene may contribute to BC susceptibility. It is worth noting that in combination with other genetic and non-genetic data, this variant might be helpful in refining the BC patients profile. The efficiency of population-based preventive programs such as screening mammography could be improved in a near future by targeting women who are at the greatest risk for BC according to their genotyping profile and other environmental risk factors.

ACKNOWLEDGMENTS

This study was supported by research funding from the Deanship of Scientific Research (grant no. MED-2018-3-9-F-7821), Northern Border University (NBU), Arar, Saudi Arabia.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

ORCID

Manal S. Fawzy  <https://orcid.org/0000-0003-1252-8403>

Eman A. Toraih  <https://orcid.org/0000-0001-9267-3787>

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394-424.
2. Li NA, Deng Y, Zhou L, et al. Global burden of breast cancer and attributable risk factors in 195 countries and territories, from 1990 to 2017: results from the Global Burden of Disease Study 2017. *J Hematol Oncol*. 2019;12(1):140.
3. Kennecke H, Yerushalmi R, Woods R, et al. Metastatic behavior of breast cancer subtypes. *J Clin Oncol*. 2010;28(20):3271-3277.
4. Song M, Lee KM, Kang D. Breast cancer prevention based on gene-environment interaction. *Mol Carcinog*. 2011;50(4):280-290.
5. Chen S, Zhang Q, Shen L, et al. Investigation of CD28 gene polymorphisms in patients with sporadic breast cancer in a Chinese Han population in Northeast China. *PLoS ONE*. 2012;7(10):e48031.
6. Grivennikov SI, Greten FR, Karin M. Immunity, inflammation, and cancer. *Cell*. 2010;140(6):883-899.
7. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646-674.
8. Nagarajan D, McArdle SEB. *Immune landscape of breast cancers*. *Biomedicines*. 2018;6:1.
9. Abramson J, Giraud M, Benoist C, Mathis D. Aire's partners in the molecular control of immunological tolerance. *Cell*. 2010;140(1):123-135.
10. Gardner JM, DeVoss JJ, Friedman RS, et al. Deletional tolerance mediated by extrathymic Aire-expressing cells. *Science*. 2008;321(5890):843-847.
11. Poliani PL, Kisand K, Marrella V, et al. Human peripheral lymphoid tissues contain autoimmune regulator-expressing dendritic cells. *Am J Pathol*. 2010;176(3):1104-1112.
12. Gardner J, Metzger T, McMahon E, et al. Extrathymic Aire-expressing cells are a distinct bone marrow-derived population that induce functional inactivation of CD4(+) T cells. *Immunity*. 2013;39(3):560-572.
13. Bruserud Ø, Costea D-E, Laakso S, et al. Oral tongue malignancies in autoimmune polyendocrine syndrome type 1. *Front Endocrinol (Lausanne)*. 2018;9:463.
14. Kalra R, Bhagyaraj E, Tiwari D, et al. AIRE promotes androgen-independent prostate cancer by directly regulating IL-6 and modulating tumor microenvironment. *Oncogenesis*. 2018;7(5):43.
15. Bianchi F, Sommariva M, De Cecco L, et al. Expression and prognostic significance of the autoimmune regulator gene in breast cancer cells. *Cell Cycle*. 2016;15(23):3220-3229.
16. Fanale D, Amodeo V, Corsini LR, Rizzo S, Bazan V, Russo A. Breast cancer genome-wide association studies: there is strength in numbers. *Oncogene*. 2012;31(17):2121-2128.
17. Zhang B, Beeghly-Fadiel A, Long J, Zheng W. Genetic variants associated with breast-cancer risk: comprehensive research synopsis, meta-analysis, and epidemiological evidence. *Lancet Oncol*. 2011;12(5):477-488.
18. Bérczi B, Gerencsér G, Farkas N, et al. Association between AIRE gene polymorphism and rheumatoid arthritis: a systematic review and meta-analysis of case-control studies. *Sci Rep*. 2017;7(1):14096.
19. Faiyaz-Ul-Haque M, Bin-Abbas B, Al-Abdullatif A, et al. Novel and recurrent mutations in the AIRE gene of autoimmune polyendocrinopathy syndrome type 1 (APS1) patients. *Clin Genet*. 2009;76(5):431-440.
20. Husebye ES, Anderson MS, Kampe O. Autoimmune polyendocrine syndromes. *N Engl J Med*. 2018;378(12):1132-1141.
21. Singletary SE, Connolly JL. Breast cancer staging: working with the sixth edition of the AJCC Cancer Staging Manual. *CA Cancer J Clin*. 2006;56(1):37-47; quiz 50-1.
22. Rakha EA, El-Sayed ME, Lee AHS, et al. Prognostic significance of Nottingham histologic grade in invasive breast carcinoma. *J Clin Oncol*. 2008;26(19):3153-3158.
23. Baker S, Ali I, Silins I, et al. Cancer Hallmarks Analytics Tool (CHAT): a text mining approach to organize and evaluate scientific literature on cancer. *Bioinformatics*. 2017;33(24):3973-3981.
24. Nagy A, Lanczky A, Menyhart O, Gyorffy B. Validation of miRNA prognostic power in hepatocellular carcinoma using expression data of independent datasets. *Sci Rep*. 2018;8(1):9227.
25. Toraih EA, Ameen HM, Hussein MH, et al. Association of autoimmune regulator gene Rs2075876 variant, but not gene expression with alopecia areata in males: a case-control study. *Immunol Invest*. 2020;49(1-2):146-165.
26. Fawzy MS, Hussein MH, Abdelaziz EZ, Yamany HA, Ismail HM, Toraih EA. Association of microRNA-196a2 variant with response to short-acting beta2-agonist in COPD: an Egyptian pilot study. *PLoS ONE*. 2016;11(4):e0152834.
27. Toraih EA, Fawzy MS, Mohammed EA, Hussein MH, El-Labban MM. MicroRNA-196a2 biomarker and targetome network analysis in solid tumors. *Mol Diagn Ther*. 2016;20(6):559-577.
28. Bruserud O, Oftedal BE, Wolff AB, Husebye ES. AIRE-mutations and autoimmune disease. *Curr Opin Immunol*. 2016;43:8-15.
29. Giraud M, Yoshida H, Abramson J, et al. Aire unleashes stalled RNA polymerase to induce ectopic gene expression in thymic epithelial cells. *Proc Natl Acad Sci USA*. 2012;109(2):535-540.
30. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. 2019;47(D1):D607-D613.
31. Ye X, Franco AA, Santos H, Nelson DM, Kaufman PD, Adams PD. Defective S phase chromatin assembly causes DNA damage, activation of the S phase checkpoint, and S phase arrest. *Mol Cell*. 2003;11(2):341-351.
32. Corpet A, De Koning L, Toedling J, et al. Asf1b, the necessary Asf1 isoform for proliferation, is predictive of outcome in breast cancer. *EMBO J*. 2011;30(3):480-493.
33. deAlmeida BP, Apolonio JD, Binnie A, Castelo-Branco P. Roadmap of DNA methylation in breast cancer identifies novel prognostic biomarkers. *BMC Cancer*. 2019;19(1):219.
34. Al-Rawi MA, Rmali K, Watkins G, Mansel RE, Jiang WG. Aberrant expression of interleukin-7 (IL-7) and its signalling complex in human breast cancer. *Eur J Cancer*. 2004;40(4):494-502.
35. Samavat H, Xun X, Jin A, Wang R, Koh WP, Yuan JM. Association between prediagnostic leukocyte telomere length and breast cancer risk: the Singapore Chinese Health Study. *Breast Cancer Res*. 2019;21(1):50.
36. Edechi CA, Ikeogu N, Uzonna JE, Myal Y. Regulation of immunity in breast cancer. *Cancers (Basel)*. 2019;11(8):1080.
37. Antonova L, Aronson K, Mueller CR. Stress and breast cancer: from epidemiology to molecular biology. *Breast Cancer Res*. 2011;13(2):208.
38. Jha S, Rajagopalan D, Hora S, Jadhav SP. Breast Cancer: From transcriptional control to clinical outcome. *Breast Cancer*. 2017;453. <https://doi.org/10.5772/66888>
39. Li L, Fu L, Wang K. Integrated bioinformatics analysis the function of RNA binding proteins (RBPs) and their prognostic value in breast cancer. *Front Pharmacol*. 2019;10:140. <https://doi.org/10.3389/fphar.2019.00140>
40. Xu YS, Jiang XJ, Chen JM. A single nucleotide polymorphism of AIRE gene located in the 21q22.3 increases the risk of rheumatoid arthritis. *Oncotarget*. 2017;8(42):71556-71562.
41. Tazi-Ahni R, McDonagh AJ, Wengraf DA, et al. The autoimmune regulator gene (AIRE) is strongly associated with vitiligo. *Br J Dermatol*. 2008;159(3):591-596.

42. Ferrera F, Rizzi M, Spreccacenero B, et al. AIRE gene polymorphisms in systemic sclerosis associated with autoimmune thyroiditis. *Clin Immunol*. 2007;122(1):13-17.
43. Eriksson D, Bianchi M, Landegren N, et al. Common genetic variation in the autoimmune regulator (AIRE) locus is associated with autoimmune Addison's disease in Sweden. *Sci Rep*. 2018;8(1):8395.
44. Montufar-Robles I, Robles-Garnica JC, Cadena-Sandoval D, et al. The AIRE Ser196Ser synonymous variant is a risk factor for systemic lupus erythematosus. *Cell Immunol*. 2019;346:103986.
45. Terao C, Yamada R, Ohmura K, et al. The human AIRE gene at chromosome 21q22 is a genetic determinant for the predisposition to rheumatoid arthritis in Japanese population. *Hum Mol Genet*. 2011;20(13):2680-2685.
46. Shao S, Li XR, Cen H, Yin ZS. Association of AIRE polymorphisms with genetic susceptibility to rheumatoid arthritis in a Chinese population. *Inflammation*. 2014;37(2):495-499.
47. Feng ZJ, Zhang SL, Wen HF, Liang Y. Association of rs2075876 polymorphism of AIRE gene with rheumatoid arthritis risk. *Hum Immunol*. 2015;76(4):281-285.
48. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res*. 2019;47(D1):D886-D894.
49. Conteduca G, Fenoglio D, Parodi A, et al. AIRE polymorphism, melanoma antigen-specific T cell immunity, and susceptibility to melanoma. *Oncotarget*. 2016;7(38):60872-60884.
50. Lewis CM, Knight J. Introduction to genetic association studies. *Cold Spring Harb Protoc*. 2012;2012(3):297-306.
51. Esser C, Tomluk J. Reporting Hardy-Weinberg tests in case-control studies: reasons for caution but not for panic reactions. *J Invest Dermatol*. 2005;124(5):1082-1083.
52. Mavaddat N, Michailidou K, Dennis J, et al. Polygenic risk scores for prediction of breast cancer and breast cancer subtypes. *Am J Hum Genet*. 2019;104(1):21-34.
53. Lakeman IMM, Schmidt MK, vanAsperen CJ, Devilee P. Breast cancer susceptibility—towards individualised risk prediction. *Curr Genet Med Rep*. 2019;7(2):124-135.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Fawzy MS, Toraih EA. Analysis of the autoimmune regulator (AIRE) gene variant rs2075876 (G/A) association with breast cancer susceptibility. *J Clin Lab Anal*. 2020;34:e23365. <https://doi.org/10.1002/jcla.23365>