Received: 19 March 2019 | Revised: 14 May 2019 | Accepted: 24 June 2019

DOI: 10.1002/mgg3.848

ORIGINAL ARTICLE

Molecular Genetics & Genomic Medicine Open Access WILEY

Evaluation of *GALNT16* polymorphisms to breast cancer risk in Chinese population

Huangfu Wu | **Guisheng He** | **Tao Song** | **Yazhen Zhang** | **Xiuxiu Chen** | **Huamin Chen** | **Wei Xiong** | **Chuanwei Sun** | **Chaoyang Zhao** | **Yunjing Chen**

Surgical Oncology, The Second Affiliated Hospital of Hainan Medical College, Haikou, Hainan Province, China

Correspondence
Huangfu Wu, #48 Baishuitang Road, Longhua District 570100, Haikou, Hainan Province, China.
Email: wuhuangfu12@163.com

Funding information
This study was supported by Research and Development Fund of the Second Affiliated Hospital of Hainan Medical College (2018-11).

Abstract

Background: Polypeptide N-acetylgalactosaminyltransferase 16 (*GALNT16*) is an N-acetylgalactosaminyltransferase gene that alters protein O-glycosylation, which plays a role in tumor development. This study aims to explore the association of eight *GALNT16* polymorphisms with susceptibility to breast cancer (BC).

Methods: This case–control study included 563 BC patients and 552 age-matched healthy controls from the Chinese Han population. The genotypes of *GALNT16* polymorphisms were detected using the Agena MassARRAY. The relationship between *GALNT16* polymorphisms and BC risk was evaluated using a chi-squared test with an odds ratio (OR) and 95% confidence intervals (CI) under five genetic models.

Results: We observed that rs2105269 and rs72625676 were associated with higher BC risk in younger patients with age ≤51 (rs2105269, codominant: $p = .006$; recessive: $p = .005$ additive: $p = .018$; and allele: $p = .012$; rs72625676, codominant: $p = .038$; recessive: $p = .037$). For rs1275678 polymorphism, there was a significantly decreased risk of BC among elder patients (codominant: $p = .017$; dominant: $p = .019$; additive: $p = .030$; and allele: $p = .029$). Further analysis by clinical characteristics showed rs2105269 was associated with tumor size and lymph node metastasis.

Conclusion: Our study suggests that *GALNT16* polymorphisms are associated with BC susceptibility in Chinese population.

KEYWORDS
breast cancer, *GALNT16*, polymorphism, susceptibility

1 | INTRODUCTION

Breast cancer (BC) is the most common cancer and the second leading cause of cancer-related death in women worldwide (Coleman et al., 2011; Jemal et al., 2011). According to data released by the World Health Organization (WHO), the newly diagnosed cancer cases showed in 2012 accounted for 25.2% of all female primary cancer. In China, the number of new cases of BC was recently reported as 187,000, ranking first in the incidence rate of female cancers and posing

Huangfu Wu and Guisheng He are joint first authors.

a serious threat to the health and quality of life of Chinese women (Fitzmaurice et al., 2017). It is widely known that multiple factors, especially genetic factors, contribute to BC susceptibility (Liu et al., 2017). Recent studies showed, critical genetic and epigenetic alterations in genes encoding glycosyltansferases can cause pathologic changes in several disease states, including cancer (Palamar, Shearston, Dawson, Mateu-Gelabert, & Ompad, 2016).

Polypeptide *N*-acetylgalactosaminyl transferases (*GALNTs*) are a large family of enzymes, which initiate the transfer of *N*-acetylgalactosamine (*GalNAc*) from UDP-GalNAc to the hydroxyl group of a serine or threonine residue have been associated with epithelial diseases (Chasman et al., 2008; Hussain, Nasir, & Al-Aama, 2013; Tian et al., 2015). The unusually large number of *GALNTs* is unique to O-glycosylation and the multiplicity of conserved isoforms in metazoan evolution suggests a need for cell- or tissue-specific isoforms (Bennett et al., 2012). To date, 20 *GALNT* family members have been identified in humans, and these isozymes have been shown to exhibit differential but overlapping substrate specificities and cell type-dependent expression patterns (Pratt et al., 2004; Wandall et al., 1997). Hence, the *GALNT* can be grouped into subfamilies. In subfamily Ib (*T2/T14/T16*), *GALNT2* is the only well-characterized isoform in the literature, but preliminary studies of these three enzymes show related functions (Bennett et al., 2012). And, it has been reported that downregulation of *GALNT2* promoted cell proliferation, migration, invasion, and tumor metastasis (Liu et al., 2016; Song et al., 2016; Yao-Ming et al., 2011). *GALNT14* promotes lung-specific BC metastasis by modulating self-renewal and interaction with the lung microenvironment (Song et al., 2016). Although, it has been identified that *GALNT16* significantly enriched for specific biological functions related to protein and lipid metabolism, insulin/IGF pathway-protein kinase B signaling cascade, prolactin signaling pathway, and AMPK signaling pathways, the functional roles of *GALNT16* on BC progression are poorly understood (Gao et al., 2017).

Thus, in order to assess the effect of single-nucleotide polymorphisms (SNPs) in *GALNT16* on BC risk, we conducted a case–control study to explore the association between eight SNPs of *GALNT16* and BC risk in Chinese Han population.

2 | METHODS

2.1 | Study subjects

A case–control study was performed with 563 BC patients randomly recruited from Shaanxi Provincial Cancer Hospital. All patients, who were Han Chinese, had confirmed by histopathological analysis. The exclusion criteria included

TABLE 1 Primers used for this study

SNP	First PCRP	Second PCRP	UEP_DIR	UEP_SEQ
rs2105269	ACGTTGGATGCGGGCTGCTTCGACATTTTG	ACGTTGGATGCCAAAGAACAGAAAGCAGCG	R	AGCGTCTTGCACAGA
rs61466740	ACGTTGGATGAGAAGTGGGCAGTCTCCACA	ACGTTGGATGTTCCAGGAAGACTCCTGGTG	R	ACGTTGGATGTTCCAGGAAGACTCCTGGTG
rs72722128	ACGTTGGATGGTGCTTAACTCTACAGAGCC	ACGTTGGATGGGAACTGCATTGGCTTCTTG	F	cctcGCTGCACCCCTTAAT
rs72625676	ACGTTGGATGTGTGGAGCGTAGGTCTCTG	ACGTTGGATGTGATCTGTCGCTTTGCTTGG	R	CGAGGTACGTTTGTCAC
rs745781	ACGTTGGATGTAGAAACAGGATCTTGCTA	ACGTTGGATGGGCTCACACCTATAATCCCG	F	ggccGGAATTTGAGACCCGCCTAG
rs1026385	ACGTTGGATGAACATAGAGGCCCTGAGCTG	ACGTTGGATGCTAGTCCTAGAGGGATGGTC	R	GGATCGCTGTGTGGAA
rs1275678	ACGTTGGATGCACTTATGTGTTGTGGGACG	ACGTTGGATGAAGAATCCCTGCTCCTAACG	R	ggggaCTAACGTCTCAGAGACATA
rs11623483	ACGTTGGATGTTCTTACAGTGGTCAGGCAG	ACGTTGGATGTTCTGCATCGCACTTTGCCC	R	CTTTGCCCCAGTCCC

Abbreviations: DIR, direction; PCRP, polymerase chain reaction primer; SEQ, sequence; SNP, single-nucleotide polymorphism; UEP, unextended mini sequencing primer.

patients who were diagnosed with other types of cancers and/or underwent radiotherapy or chemotherapy. And the control group was comprised of 552 unrelated and age-matched healthy individuals (without any underlying illnesses) from the same hospital. The methods were carried out in accordance with the World Medical Association Declaration of Helsinki, and the study was approved by the ethics committee of Shaanxi Provincial Cancer Hospital. After obtaining written informed consent, the data on clinical characteristics of patients, including tumor site, tumor size, lymph node metastasis, disease stage, Ki67 status, estrogen receptor (ER) status and progesterone receptor (PR) status, and human epidermal growth factor receptor (HER2) status, were collected from medical records.

2.2 | Genotyping assay

Peripheral blood of all subjects were collected in tubes containing EDTA and stored at −80°C. Then DNA was extracted using the GoldMag whole blood genomic DNA purification kit (GoldMag Co. Ltd., Xi′an, China) according to the manufacturer's protocol, and the quantity of DNA was measured by spectrometry (NanoDrop 2000 spectrophotometer, Thermo Scientific, Waltham, MA). Eight SNPs in *GALNT16* with a minor allele frequency (MAF)> 5% of the 1,000 Genomes Project data were selected in the present study. Agena MassARRAY Assay Design Software (version 3.0, Agena Bioscience, San Diego, CA) was used to design multiplexed SNP MassEXTEND assay. And Agena MassARRAY RS1000 was used to detect SNP genotyping (Fei et al., 2014; Xia et al., 2014, 2015). Primers of the eight SNPs are listed in Table 1. Data were analyzed with Agena Typer Software (version 4.0, Agena Bioscience, San Diego, CA).

2.3 | Statistical analysis

SPSS software (version 18.0, Chicago, IL) was used for statistical analyses of data. The Student *t* test or chi-squared test was used to examine the differences of basic parameters between two groups. Hardy–Weinberg equilibrium as well as the differences in allele frequencies between cases and controls were examined by chi-squared test for each SNP. The BC risk associated with genotypes was estimated by odds ratios (OR) with 95% confidence intervals (CI) for five different genetic models. We further performed haplotype analysis and linkage disequilibrium. For all test, a two-tailed *p*-value <.05 was considered statistically significant. HaploReg v4.1 (https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php) was conducted to predict the potential functions of the SNPs. And Gene Expression Profiling Interactive Analysis public database (GEPIA) (http://gepia.cancer-pku.cn/) was used to analyze *GALNT16* expression of BC and normal samples. Furthermore, we evaluated the relationship between prognostic significance of BC and the expression of *GALNT16* using Kaplan–Meier plotter (http://kmplot.com/analysis/).

TABLE 2 Characteristics of breast cancer cases and healthy controls

Characteristics	Cases (563)	Controls (552)	*p*-value
Age (years, mean ± *SD*)	52.05 ± 9.810	51.88 ± 9.849	.767
Age, years			
>51	297	295	
≤51	266	257	
Tumor site			
Left	226		
Right	219		
Absence	118		
Tumor size			
<2 cm	107		
≥2 cm	315		
Absence	141		
LN metastasis			
Negative	260		
Positive	275		
Absence	28		
Stage			
I-II	365		
III-IV	162		
Absence	36		
PR			
Negative	212		
Positive	341		
Absence	10		
ER			
Negative	161		
Positive	378		
Absence	24		
HER2 status			
Negative	91		
Positive	273		
Absence	199		
Ki67			
<10%	132		
≥10%	365		
Absence	66		

Abbreviations: CI, confidence interval; ER, estrogen receptor; HER2: human epidermal growth factor receptor 2; LN, lymph node; OR, odds ratio; PR, progesterone receptor.

TABLE 3 Basic information of candidate single-nucleotide polymorphism (SNPs) in the study

				Case (563)		Control (552)			
Gene	**SNP**	**Role**	**Position**	**MA**	**MAF**	**MA**	**MAF**	***p*-HWE**	**HaploReg**
GALNT16	rs2105269	intronic	69280517	A	0.3917	A	0.3702	.71	Selected eQTL hits
	rs61466740	intronic	69289592	C	0.2504	C	0.2291	.90	Enhancer histone marks, DNAse
	rs72722128	intronic	69293300	A	0.1421	A	0.1506	.87	Enhancer histone marks, DNAse, Motifs changed
	rs72625676	intronic	69307268	T	0.2709	T	0.2627	.91	Motifs changed
	rs745781	intronic	69312014	C	0.2096	C	0.2114	.52	Enhancer histone marks, Motifs changed, Selected eQTL hits
	rs1026385	intronic	69319346	G	0.0897	G	0.0833	.57	DNAse, Motifs changed
	rs1275678	intronic	69335147	A	0.1012	A	0.1205	.69	Enhancer histone marks, DNAse
	rs11623483	intronic	69345012	A	0.2651	A	0.2840	.92	Enhancer histone marks, DNAse, Motifs changed

Abbreviations: HWE, Hardy–Weinberg equilibrium; MA, minor allele; MAF, minor allele frequency; SNPs, single-nucleotide polymorphism.

3 | RESULTS

3.1 | Characteristics of the study subjects

Our study included 563 patients with BC and 552 healthy controls. The clinical and demographic characteristic of BC patients and controls are shown in Table 2. The cases and controls were matched by age (Student *t* test, $p = .767$). About 315 patients had a tumor size ≥2 cm and 275 patients had lymph node metastasis. Besides, patients with more than 10 percent of Ki67 are 365. The percentage of patients with PR-, ER-, and HER2-positive status was 60.57%, 67.14%, and 48.49%, respectively.

3.2 | Basic information and potential functions of the selected *GALNT16* SNPs

The detailed information of eight SNPs in *GALNT16* was listed in Table 3. For all SNPs, the MAFs were greater than 5% and the observed genotype frequencies complied with HWE in the control group. In addition, we annotate the functional elements of these selected SNPs using HaploRegv4.1. The results revealed that the SNPs in *GALNT16* were involved in the regulations related to selected eQTL hits, enhancer histones, DNAse, and motifs changed.

3.3 | Association of *GALNT16* SNPs with breast cancer susceptibility

The genotypic and allelic frequencies of *GALNT16* rs2105269, rs61466740, rs72722128, rs72625676, rs745781, rs1026385, rs1275678, and rs11623483 in cases and controls are shown in Table 4. We evaluated their associations with risk of BC by chi-squared test and OR and five genetic models (codominant, dominant, recessive, additive, and allele) were applied to assess the potential association by logistic regression adjusted for age (Zhou et al., 2014). However, we did not observe any significant association between the *GALNT16* polymorphisms and BC risk in all genetic models.

3.4 | Stratified analysis of *GALNT16* polymorphisms and BC risk

Then we performed a subgroup analysis regarding the effect of *GALNT16* polymorphisms, and rs2105269, rs72625676, and rs1275678 polymorphisms on BC according to age are displayed in Table 5. The results indicated that rs2105269 was associated with increased BC risk in the women with age ≤51 (codominant model: OR = 2.16, 95% CI = 1.25–3.17, $p = .006$; recessive model: OR = 2.08, 95% CI = 1.24–3.49, $p = .005$; additive model: OR = 1.35, 95% CI = 1.05–1.73, $p = .018$; and allele model: OR = 1.38, 95% CI = 1.07–1.79,

TABLE 4 Frequencies of *GALNT16* gene alleles and genotypes of BC patients and controls

Polymorphism	Genotype	Case	Control	OR (95% CI)	*p*-value
rs2105269					
Codominant	AA	89 (15.8%)	73 (13.25%)	1.25 (0.87–1.791)	.237
	AG	263 (46.71%)	262 (47.55%)	1.02 (0.79–1.32)	.875
	GG	211 (37.48%)	216 (39.20%)	1	
Dominant	AA-AG	352 (62.52%)	335 (60.80%)	1.07 (0.84–1.36)	.586
	GG	211 (37.48%)	216 (39.20%)	1	
Recessive	AA	89 (15.80%)	73 (13.25%)	1.23 (0.88–1.72)	.223
	AG-GG	474 (84.19%)	478 (86.75%)	1	
llele	A	441 (39.17%)	408 (37.02%)	1.10 (0.92–1.30)	.298
	G	685 (60.83%)	694 (62.98%)	1	
Additive				1.09 (0.92–1.30)	.312
rs61466740					
Codominant	CC	35 (6.22%)	28 (5.10%)	1.29 (0.77–2.17)	.337
	CT	212 (37.66%)	196 (35.64%)	1.12 (0.87–1.43)	.379
	TT	316 (56.13%)	326 (59.27%)	1	
Dominant	CC-CT	247 (43.87%)	224 (40.73%)	1.14 (0.90–1.45)	.284
	TT	316 (56.13%)	326 (59.27%)	1	
Recessive	CC	35 (6.21%)	28 (5.10%)	1.23 (0.74–2.06)	.419
	CT-TT	528 (93.78%)	522 (94.91%)	1	
Allele	C	282 (25.04%)	252 (22.91%)	1.12 (0.93–1.37)	.238
	T	844 (74.96%)	848 (77.09%)	1	
Additive				1.13 (0.93–1.37)	.234
rs72722128					
Codominant	AA	11 (1.95%)	13 (2.36%)	0.81 (0.36–1.83)	.612
	AG	138 (24.51%)	140 (25.40%)	0.94 (0.72–1.23)	.654
	GG	414 (73.53%)	398 (72.36%)	1	
Dominant	AA-AG	149 (26.47%)	153 (27.77%)	0.93 (0.71–1.21)	.583
	GG	414 (73.53%)	398 (72.23%)	1	
Recessive	AA	11 (1.95%)	13 (2.36%)	0.93 (0.73–1.17)	.533
	AG-GG	552 (98.05%)	538 (97.64%)	1	
Allele	A	160 (14.21%)	166 (15.06%)	0.93 (0.74–1.18)	.568
	G	966 (85.79%)	936 (84.94%)	1	
Additive				0.82 (0.37–1.85)	.639
rs72625676					
Codominant	TT	38 (6.75%)	37 (6.73%)	1.04 (0.64–1.67)	.887
	TC	229 (40.67%)	215 (39.09%)	1.07 (0.84–1.37)	.569
	CC	296 (52.58%)	298 (54.18%)	1	
Dominant	TT-TC	267 (47.42%)	252 (45.82%)	1.07 (0.84–1.35)	.581
	CC	296 (52.58%)	298 (54.18%)	1	
Recessive	TT	38 (6.75%)	37 (6.73%)	1.00 (0.63–1.61)	.986
	TC-CC	525 (93.25%)	513 (93.27%)	1	
Allele	T	305 (27.09%)	289 (26.27%)	1.04 (0.86–1.26)	.664
	C	821 (72.91%)	811 (73.73%)	1	
Additive				1.05 (0.86–1.26)	.652

(Continues)

TABLE 4 (Continued)

Polymorphism	Genotype	Case	Control	OR (95% CI)	*p*-value
rs745781					
Codominant	CC	25 (4.44%)	27 (4.90%)	0.91 (0.52–1.59)	.734
	CG	186 (33.04%)	179 (32.55%)	1.01 (0.78–1.30)	.930
	GG	352 (62.52%)	345 (62.73%)	1	
Dominant	CC-CG	211 (37.48%)	206 (37.39%)	1.00 (0.78–1.27)	.985
	GG	352 (62.52%)	345 (62.61%)	1	
Recessive	CC	25 (4.44%)	27 (4.90%)	0.90 (0.52–1.58)	.720
	CG-GG	538 (95.56%)	524 (95.10%)	1	
Allele	C	236 (20.96%)	233 (21.14%)	0.99 (0.81–1.21)	.915
	G	890 (79.04%)	869 (78.86%)	1	
Additive				0.99 (0.80–1.21)	.884
rs1026385					
Codominant	GG	4 (0.71%)	5 (0.91%)	0.80 (0.21–3.02)	.746
	GA	93 (16.52%)	82 (14.86%)	1.14 (0.82–1.57)	.442
	AA	466 (82.77%)	465 (84.24%)	1	
Dominant	GG-GA	97 (17.23%)	87 (15.76%)	1.12 (0.81–1.53)	.496
	AA	466 (82.77%)	465 (84.24%)	1	
Recessive	GG	4 (0.71%)	5 (0.91%)	0.79 (0.21–2.95)	.724
	GA-AA	559 (99.29%)	547 (99.09%)	1	
Allele	G	101 (8.97%)	92 (8.33%)	1.08 (0.81–1.46)	.593
	A	1,025 (91.03%)	1,012 (91.67%)	1	
Additive				1.01 (0.81–1.46)	.579
rs1275678					
Codominant	AA	5 (0.89%)	9 (1.63%)	0.53 (0.17–1.58)	.253
	AC	104 (18.47%)	115 (20.83%)	0.85 (0.64–1.15)	.296
	CC	454 (80.64%)	428 (77.54%)	1	
Dominant	AA-AC	109 (19.36%)	124 (22.46%)	0.54 (0.18–1.63)	.277
	CC	454 (80.64%)	428 (77.54%)	1	
Recessive	AA	5 (0.89%)	9 (1.63%)	0.82 (0.63–1.08)	.150
	AC-CC	558 (99.11%)	543 (98.37%)	1	
Allele	A	114 (10.12%)	133 (12.05%)	0.82 (0.63–1.07)	.148
	C	1,012 (89.88%)	971 (87.95%)	1	
Additive				0.82 (0.63–1.08)	.153
rs11623483					
Codominant	AA	41 (7.30%)	45 (8.17%)	0.85 (0.54–1.33)	.475
	AG	216 (38.43%)	223 (40.47%)	0.89 (0.70–1.15)	.384
	GG	305 (54.27%)	283 (51.36%)	1	
Dominant	AA-AG	257 (45.73%)	268 (48.64%)	0.89 (0.70–1.12)	.323
	GG	305 (54.27%)	283 (51.36%)	1	
Recessive	AA	41 (7.30%)	45 (8.17%)	0.89 (0.57–1.38)	.600
	AG-GG	521 (92.70%)	506 (91.83%)	1	
Allele	A	298 (26.51%)	313 (28.40%)	0.91 (0.75–1.10)	.317
	G	826 (73.49%)	789 (71.60%)	1	
Additive				0.91 (0.76–1.01)	.318

Note: All results are adjusted for age.

Abbreviations: BC, breast cancer; CI, confidence interval; OR, odds ratio.

TABLE 5 Stratified analysis of polymorphisms in *GALNT16* on BC risk by age

			Age >51 (297) versus controls (295)		Age ≤51 (266) versus controls (257)	
SNP	**Model**	**Allele/Genotype**	**OR (95% CI)**	***p*-value**	**OR (95% CI)**	***p*-value**
rs2105269	Codominant	AA	0.78 (0.47–1.29)	.326	2.16 (1.25–3.74)	**.006**
		AG	0.95 (0.66–1.36)	.783	1.08 (0.74–1.56)	.698
		GG	1		1	
	Dominant	AA-AG	0.91 (0.65–1.28)	.583	1.27 (0.90–1.80)	.172
		GG	1		1	
	Recessive	AA	0.80 (0.51–1.26)	.338	2.08 (1.24–3.49)	**.005**
		AG-GG	1		1	
	Allele	A	0.90 (0.72–1.14)	.393	1.38 (1.07–1.79)	**.012**
		G	1		1	
	Additive		0.90 (0.70–1.14)	.371	1.35 (1.05–1.73)	**.018**
rs72625676	Codominant	TT	0.59 (0.30–1.13)	.112	2.30 (1.05–5.03)	**.038**
		TC	1.10 (0.79–1.55)	.562	1.04 (0.72–1.48)	.853
		CC	1		1	
	Dominant	TT-TC	1.01 (0.73–1.39)	.967	1.15 (0.81–1.62)	.430
		CC	1		1	
	Recessive	TT	0.56 (0.29–1.06)	.076	2.27 (1.05–4.89)	**.037**
		TC-CC	1		1	
	Allele	T	0.91 (0.71–1.17)	.466	1.23 (0.93–1.63)	.145
		C	1		1	
	Additive		0.91 (0.70–1.18)	.477	1.24 (0.93–1.64)	.140
rs1275678	Codominant	AA	0.92 (0.18–4.61)	.917	0.32 (0.06–1.63)	.172
		AC	0.60 (0.40–0.91)	**.017**	1.26 (0.81–1.95)	.301
		CC	1		1	
	Dominant	AA-AC	0.62 (0.41–0.92)	**.019**	1.15 (0.75–1.75)	.521
		CC	1		1	
	Recessive	AA	1.01 (0.20–5.06)	.989	0.31 (0.06–1.55)	.154
		AC-CC	1		1	
	Allele	A	0.66 (0.46–0.96)	**.029**	1.04 (0.71–1.52)	.847
		C	1		1	
	Additive		0.66 (0.45–0.96)	**.030**	1.03 (0.71–1.51)	.865

Note: *p*-value < .05 was shown in bold.

Abbreviations: BC, breast cancer; CI, confidence interval; OR, odds ratio.

$p = .012$). Furthermore, compared with the wild genotype of rs72625676, we found a significantly increased risk of BC associated with the variant genotypes in two models (codominant model: OR = 2.30, 95% CI = 1.05–5.03, $p = .038$; recessive model: OR = 2.27, 95% CI = 1.05–4.89, $p = .037$) for the women whose ages are no more than 51. Nevertheless, rs1275678 had relationship with significantly decreasing the risk of BC in the subgroups of age >51 for genetic models (codominant model: OR = 0.60, 95% CI = 0.40–0.91, $p = .017$; dominant model: OR = 0.62, 95% CI = 0.41–0.92, $p = .019$; additive model: OR = 0.66, 95% CI = 0.46–0.96, $p = .030$; and allele model: OR = 0.66, 95% CI = 0.45–0.96, $p = .029$).

3.5 | Haplotype analysis of *GALNT16* polymorphisms and BC risk

As shown in Table S1, we did not find significant association between *GALNT16* haplotype and BC risk. We observed three blocks, they are block 1 (rs61466740 and rs72722128), block 2 (rs72625676, rs745781, and rs1026385), and block 3 (rs1275678 and rs11623483) (Figure 1). We further conducted haplotype analysis in age subgroups. In the subgroup of age >51, we found two blocks (block 1: rs61466740 and rs72722128, block 2: rs72625676 and rs745781) (Table S2 and Figure S1). For the individuals younger than 51 years

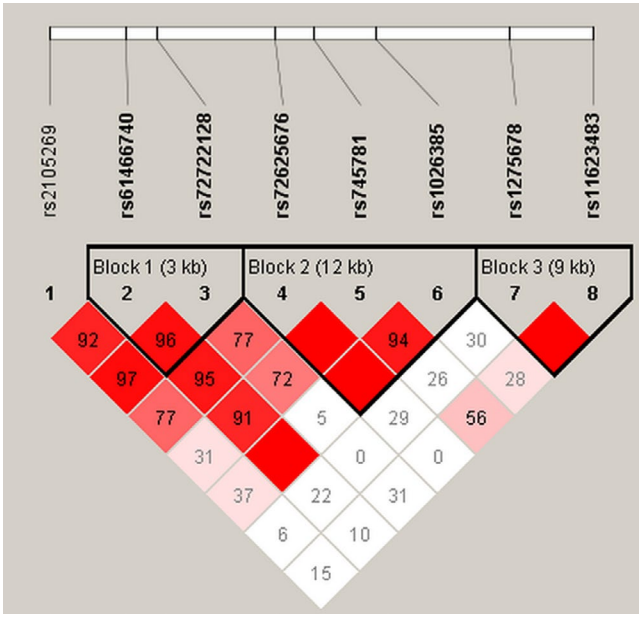

FIGURE 1 Haplotype block map for the *GALNT16* gene polymorphisms. Block 1 includes rs61466740 and rs72722128. Block 2 includes rs72625676, rs745781, and rs1026385. Block 3 includes rs1275678 and rs1162343. The numbers inside the diamonds indicate the D′ for pairwise analyses

old, $C_{rs72625676}G_{rs745781}A_{rs1026385}$ haplotype increased BC risk (p = .046) (Table S3). As shown in Figure S2, We observed two blocks (block 1: rs72625676, rs745781, and rs1026385; block 2: rs1275678 and rs11623483).

3.6 | Relationship between *GALNT16* SNPs and clinical features of BC patients

In order to identify the effect of *GALNT16* SNPs on different clinical characteristics of BC patients, we then analyzed the relationships between *GALNT16* polymorphisms and a series of clinicopatholoical parameters, such as tumor site/size, lymph node metastasis, and hormonal receptor status. As shown in Table 6, we found that the mutational genotype frequency of rs2105269 was significantly higher in patients with tumor size greater than 2cm (homozygote model: OR = 2.01, 95% CI = 1.00–4.03; heterozygote model: OR = 1.66, 95% CI = 1.03–2.69; dominant model: OR = 1.74, 95% CI = 1.11–2.73; additive model: OR = 1.48, 95% CI = 1.07–2.06; and allele model: OR = 1.47, 95% CI = 1.06–2.03) and lymph node metastasis (heterozygote model: OR = 1.59, 95% CI = 1.10–2.32). However, no significant association was detected in other clinical parameters of BC patients.

3.7 | Bioinformatics analysis of *GALNT16* expression and prognosis

Based on GEPIA dataset, *GALNT16* presented higher expression in BC tissues than in normal tissues (Figure S3). Then, the significantly association between *GALNT16* expression and BC prognosis was found according to Kaplan–Meier plotter (hazard ratio = 0.64; 95% CI = 0.55–0.75; p = 1.7e-08; Figure S4).

4 | DISCUSSION

Glycosylation is a posttranslational modification and is associated with various physiologic events. The aberrant expression of glycosyltransferase and the immature glycan structure of proteins and lipids are observed in the development and progression of cancers (Brockhausen, 1999; Fuster & Esko, 2005; Park, Katagiri, Chung, Kijima, & Nakamura, 2011; Park et al., 2010; Potapenko et al., 2010). Abnormalities of the glycan structure of glycoproteins are frequently observed in BC cells (Fuster & Esko, 2005; Park et al., 2011, 2010). In particular, the oncogenic roles of some cancer-specific glycosyltransferases had been identified and characterized previously. To further investigate the oncogenic role of aberrant glycosyltransferase expression, we attempted to identify the association of *GALNT16* polymorphisms and BC risk. In this case–control study, we successfully genotyped eight SNPs in the *GALNT16* and found that *GALNT16* polymorphisms are associated with BC susceptibility in the Chinese and may be involved in tumor progression.

GALNT16 (Polypeptide N-acetylgalactosaminyltransferase 16) is a protein coding gene, which catalyzes the initial reaction in O-linked oligosaccharide biosynthesis and transfers an N-acetyl-D-galactosamine residue to a serine or threonine residue on the protein receptor. An important paralog of this gene is *GALNT2*. Recent studies reported that *GALNT2* genetic polymorphisms were associated several cancers, including gastric adenocarcinoma, neuroblastoma, ovarian cancer, and BC (Gill et al., 2013; Liu et al., 2016; Terry et al., 2010; Wan-Ling et al., 2014). Although the overexpression of *GALNT2* involved in the cell growth of BC, only few researches are done on this field (Taisuke et al., 2014). Likewise, the overexpression of *GALNT14* plays a critical role in cell migration, invasion, and proliferation of BC by stimulating the epithelial mesenchymal transition of BC cells (Huanna et al., 2015). As Ib subfamily, *GALNT16* shares the same intron numbers (with minor variations in introns positions) with *GALNT2* and *GALNT14*. The currently available data on the enzymatic functions of GALNTs support the proposed subfamily classification. Moreover, due to the high sequence similarity of Ib subfamily, similar biological functions can be postulated.

Our study focused on the relationship of *GALNT16* and BC risk in Chinese Han populations and found that rs2105269 and rs72625676 polymorphisms were associated

TABLE 6 Associations between the *GALNT16* rs2105269 polymorphism and clinical characteristics of BC patients

	OR(95% CI)					
Variables	**Homozygote**	**Heterozygote**	**Dominant**	**Recessive**	**Additive**	**Allele**
Tumor site						
Left	1.53 (0.96–2.43)	1.07 (0.76–1.51)	1.17 (0.85–1.62)	1.47 (0.96–2.24)	1.20 (0.96–1.51)	1.19 (0.95–1.49)
Right	1.04 (0.64–1.71)	1.00 (0.71–1.4)	1.01 (0.73–1.39)	1.04 (0.66–1.65)	1.01 (0.80–1.28)	1.02 (0.81–1.28)
Tumor size						
<2 cm	1.00					
≥2 cm	**2.01 (1.00–4.03)**	**1.66 (1.03–2.69)**	**1.74 (1.11–2.73)**	1.53 (0.80–2.92)	**1.48 (1.07–2.06)**	**1.47 (1.06–2.03)**
LN metastasis						
Negative	1.00					
Positive	0.86 (0.51–1.45)	**1.59 (1.10–2.32)**	1.3730.96–1.95)	0.66 (0.41–1.08)	1.05 (0.82–1.34)	1.05 (0.82–1.34)
Stage						
I-II	1.00					
III-IV	1.03 (0.59–1.81)	1.06 (0.70–1.59)	1.05 (0.71–1.55)	1.00 (0.60–1.67)	1.02 (0.78–1.34)	1.03 (0.79–1.34)
PR						
Negative	1.00					
Positive	0.96 (0.57–1.62)	0.86 (0.59–1.25)	0.88 (0.62–1.26)	1.05 (0.65–1.69)	0.95 (0.74–1.22)	0.95 (0.74–1.22)
ER						
Negative	1.00					
Positive	1.14 (0.64–2.01)	1.04 (0.69–1.55)	1.06 (0.72–1.55)	1.12 (0.66–1.88)	1.06 (0.81–1.38)	1.06 (0.81–1.38)
HER2 status						
Negative	1.00					
Positive	0.94 (0.45–1.97)	1.01 (0.60–1.70)	0.99 (0.61–1.62)	0.93 (0.47–1.85)	0.98 (0.69–1.39)	0.98 (0.69–1.38)
Ki67						
<10%	1.00					
≥10%	0.93 (0.51–1.68)	1.05 (0.68–1.64)	1.02 (0.67–1.54)	0.90 (0.52–1.55)	0.98 (0.74–1.31)	0.99 (0.74–1.32)

Note: OR of significant association is presented in bold.

Abbreviations: BC, breast cancer; CI, confidence interval; ER, estrogen receptor; LN, lymph node; OR, odds ratio; PR, progesterone receptor.

with an increased BC risk in the women with age ≤51, and a relationship was found between the rs1275678 and BC subjects with age >51, which may predict rs1275678 is a protective factor. Furthermore, the A allele of rs2105269 was related with a larger size of tumor (≥2cm). It was also correlated with lymph node metastasis, indicating that patients with A allele of rs2105269 are more likely to have a worse prognosis. Our results update the previous studies, suggesting the critical to some SNPs could affect the susceptibility of *GALNT16*.

Some limitations could not be ignored in the study. First, choosing bias inevitably exists as this is a hospital-based, single-center study. Second, we did not analyze the impact of other risk factors such as lifestyle, family history, and menopausal status because of a lack of such data from both patients and controls. As our case–control study is the first research to elucidate on the association of *GALNT16* polymorphisms with BC risk, large sample size and further confirmation in other ethnic populations are needed.

5 | CONCLUSION

In summary, this case–control study indicates that the *GALNT16* polymorphisms are associated with BC susceptibility in the Chinese population and may be involved in tumor progression. Further functional studies and large population-based prospective studies are required to provide accurate evidence about the influence of *GALNT16* variants on BC.

ACKNOWLEDGMENTS

We thank the participants for taking part in this study. Specially, we sincerely thank all staff worked in Shaanxi Provincial Cancer Hospital for sample collection in our study.

CONFLICT OF INTEREST

All authors declare that they have no conflict of interests.

ORCID

Huangfu Wu https://orcid.org/0000-0001-9503-9773

REFERENCES

Bennett, E. P., Ulla, M., Henrik, C., Gerken, T. A., Fritz, T. A., & Tabak, L. A. (2012). Control of mucin-type O-glycosylation: A classification of the polypeptide GalNAc-transferase gene family. *Glycobiology*, *22*(6), 736–756. https://doi.org/10.1093/glycob/cwr182

Brockhausen, I. (1999). Pathways of O-glycan biosynthesis in cancer cells. *Biochimica et Biophysica Acta*, *1473*(1), 67–95. https://doi.org/10.1016/S0304-4165(99)00170-1

Chasman, D. I., Paré, G., Zee, R. Y. L., Parker, A. N., Cook, N. R., Buring, J. E., … Williams, P. T. (2008). Genetic loci associated with plasma concentration of LDL-C, HDL-C, triglycerides, ApoA1, and ApoB among 6382 Caucasian women in genome-wide analysis with replication. *Circulation Cardiovascular Genetics*, *1*(1), 21.

Coleman, M. P., Forman, D., Bryant, H., Butler, J., Rachet, B., Maringe, C., … Hatcher, J. (2011). Cancer survival in Australia, Canada, Denmark, Norway, Sweden, and the UK, 1995–2007 (the International Cancer Benchmarking Partnership): An analysis of population-based cancer registry data. *Lancet*, *377*(9772), 1149–1149.

Fitzmaurice, C., Allen, C., Barber, R. M., Barregard, L., Bhutta, Z. A., Brenner, H., … Naghavi, M. (2017). Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: A systematic analysis for the global burden of disease study. *JAMA Oncology*, *3*(4), 524. https://doi.org/10.1001/jamaoncol.2016.5688

Fuster, M. M., & Esko, J. D. (2005). The sweet and sour of cancer: Glycans as novel therapeutic targets. *Nature Reviews Cancer*, *5*(7), 526–542. https://doi.org/10.1038/nrc1649

Gao, Y., Jiang, J., Yang, S., Hou, Y., Liu, G. E., Zhang, S., … Sun, D. (2017). CNV discovery for milk composition traits in dairy cattle using whole genome resequencing. *BMC Genomics*, *18*(1), 265. https://doi.org/10.1186/s12864-017-3636-3

Gill, D. J., Tham, K. M., Chia, J., Wang, S. C., Steentoft, C., Clausen, H., … Bard, F. A. (2013). Initiation of GalNAc-type O-glycosylation in the endoplasmic reticulum promotes cancer cell invasiveness. *Proceedings of the National Academy of Sciences*, *110*(34), E3152–E3161. https://doi.org/10.1073/pnas.1305269110

Huanna, T., Tao, Z., Xiangfei, W., Longfei, A. N., Yuanyuan, X., Jianhua, W., … Chen, W. U. (2015). GALNT14 mediates tumor invasion and migration in breast cancer cell MCF-7. *Molecular Carcinogenesis*, *54*(10), 1159–1171. https://doi.org/10.1002/mc.22186

Hussain, M. R. M., Nasir, J., & Al-Aama, J. Y. (2013). Clinically significant missense variants in human GALNT3, GALNT8, GALNT12, and GALNT13 genes: intriguing in silico findings. *Journal of Cellular Biochemistry*, *115*(2), 313–327.

Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E., & Forman, D. (2011). Global cancer statistics. *CA: A Cancer Journal for Clinicians*, *61*(2), 69–90. https://doi.org/10.3322/caac.20107

Kang, F., Ma, W., Ma, X., Shao, Y., Yang, W., Chen, X., … Wang, J. (2014). Propranolol inhibits glucose metabolism and 18F-FDG uptake of breast cancer through posttranscriptional downregulation of hexokinase-2. *Journal of Nuclear Medicine*, *55*(3), 439–445. https://doi.org/10.2967/jnumed.113.121327

Liu, D. I., Wang, M., Tian, T., Wang, X.-J., Kang, H.-F., Jin, T.-B., … Dai, Z.-J. (2017). Genetic polymorphisms (rs10636 and rs28366003) in metallothionein 2A increase breast cancer risk in Chinese Han population. *Aging*, *9*(2), 547–555. https://doi.org/10.18632/aging.101177

Liu, S. Y., Shun, C. T., Hung, K. Y., Juan, H. F., Hsu, C. L., Huang, M. C., & Lai, I. R. (2016). Mucin glycosylating enzyme GALNT2 suppresses malignancy in gastric adenocarcinoma by reducing MET phosphorylation. *Oncotarget*, *7*(10), 11251–11262. https://doi.org/10.18632/oncotarget.7081

Palamar, J. J., Shearston, J. A., Dawson, E. W., Mateu-Gelabert, P., & Ompad, D. C. (2016). Nonmedical opioid use and heroin use in a nationally representative sample of us high school seniors. *Drug and Alcohol Dependence*, *158*, 132–138. https://doi.org/10.1016/j.drugalcdep.2015.11.005

Park, J. H., Katagiri, T., Chung, S., Kijima, K., & Nakamura, Y. (2011). Polypeptide N-acetylgalactosaminyltransferase 6 disrupts mammary acinar morphogenesis through O-glycosylation of fibronectin. *Neoplasia*, *13*(4), 320, IN329-326, IN310.

Park, J. H., Nishidate, T., Kijima, K., Ohashi, T., Takegawa, K., Fujikane, T., … Katagiri, T. (2010). Critical roles of mucin 1 glycosylation by transactivated polypeptide N-acetylgalactosaminyltransferase 6 in mammary carcinogenesis. *Cancer Research*, *70*(7), 2759–2769.

Potapenko, I. O., Haakensen, V. D., Lüders, T., Helland, Å., Bukholm, I., Sørlie, T., … Børresen-Dale, A.-L. (2010). Glycan gene expression signatures in normal and malignant breast tissue; possible role in diagnosis and progression. *Molecular Oncology*, *4*(2), 98–118. https://doi.org/10.1016/j.molonc.2009.12.001

Pratt, M. R., Hang, H. C., Hagen, K. G., Jason, R., Gerken, T. A., Tabak, L. A., & Bertozzi, C. R. (2004). Deconvoluting the functions of polypeptide N-alpha-acetylgalactosaminyltransferase family members by glycopeptide substrate profiling. *Chemistry & Biology*, *11*(7), 1009–1016. https://doi.org/10.1016/j.chembiol.2004.05.009

Song, K. H., Mi, S. P., Nandu, T. S., Gadad, S., Kim, S. C., & Kim, M. Y. (2016). GALNT14 promotes lung-specific breast cancer metastasis by modulating self-renewal and interaction with the lung microenvironment. *Nature Communications*, *7*, 13796. https://doi.org/10.1038/ncomms13796

Taisuke, M., Masato, K., Tetsuro, Y., Kazuma, K., Yasuo, M., Mitsunori, S., & Toyomasa, K. (2014). Involvement of B3GALNT2 overexpression in the cell growth of breast cancer. *International Journal of Oncology*, *44*(2), 427–434. https://doi.org/10.3892/ijo.2013.2187

Terry, K. L., Vitonis, A. F., Hernandez, D., Lurie, G., Song, H., Ramus, S. J., … Gentry-Maharaj, A. (2010). A polymorphism in the GALNT2 gene and ovarian cancer risk in four population based case-control studies. *International Journal of Molecular Epidemiology and Genetics*, *1*(4), 272–277.

Tian, E., Stevens, S. R., Guan, Y. U., Springer, D. A., Anderson, S. A., Starost, M. F., … Tabak, L. A. (2015). Galnt1 is required for normal heart valve development and cardiac function. *PLoS ONE*, *10*(1), e0115861. https://doi.org/10.1371/journal.pone.0115861

Wandall, H. H., Hassan, H., Mirgorodskaya, E., Kristensen, A. K., Roepstorff, P., Bennett, E. P., … Taylor-Papadimitriou, J. (1997). Substrate specificities of three members of the human UDP-N-acetyl-alpha-D-galactosamine: Polypeptide N-acetylgalactosaminyltransferase family, GalNAc-T1, -T2, and -T3. *Journal of Biological Chemistry*, *272*(38), 23503–23514.

Wan-Ling, H., Chih-Hsing, C., Yung-Ming, J., Meng-Yao, L., Yung-Li, Y., Shiann-Tarng, J., … Wen-Ming, H. (2014). GALNT2 suppresses malignant phenotypes through IGF-1 receptor and predicts favorable prognosis in neuroblastoma. *Oncotarget*, *5*(23), 12247–12259.

Wu, Y.-M., Liu, C.-H., Hu, R.-H., Huang, M.-J., Lee, J.-J., Chen, C.-H., … Huang, M.-C. (2011). Mucin glycosylating enzyme GALNT2 regulates the malignant character of hepatocellular carcinoma by modifying the EGF receptor. *Cancer Research*, *71*(23), 7270. https://doi.org/10.1158/0008-5472.CAN-11-1161

Xia, P., Jin, T., Geng, T., Sun, T., Li, X., Dang, C., … Sun, J. (2014). Polymorphisms in ESR1 and FLJ43663 are associated with breast cancer risk in the Han population. *Tumour Biology the Journal of the International Society for Oncodevelopmental Biology & Medicine*, *35*(3), 2187. https://doi.org/10.1007/s13277-013-1289-7

Xia, P., Li, B., Geng, T., Deng, Z., Dang, C., Chang, D., … Chen, C. (2015). FGFR2 gene polymorphisms are associated with breast cancer risk in the Han Chinese population. *American Journal of Cancer Research*, *5*(5), 1854–1861.

Zhou, L., He, N., Feng, T., Geng, T., Jin, T., & Chen, C. (2014). Association of five single nucleotide polymorphisms at 6q25.1 with breast cancer risk in northwestern China. *American Journal of Cancer Research*, *5*(8), 2467.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Wu H, He G, Song T, et al. Evaluation of *GALNT16* polymorphisms to breast cancer risk in Chinese population. *Mol Genet Genomic Med*. 2019;7:e848. https://doi.org/10.1002/mgg3.848