Breast Cancer (2019) 26:562–572
https://doi.org/10.1007/s12282-019-00956-5

ORIGINAL ARTICLE

Impact of germinal center-associated nuclear protein polymorphisms on breast cancer risk and prognosis in a Japanese population

Haruru Kotani[1,2,3] · Hidemi Ito[4,5] · Kazuhiko Kuwahara[6,7] · Kiyotaka Kuzushima[7] · Hiroji Iwata[1] · Nobuyuki Tsunoda[2] · Masato Nagino[2] · Keitaro Matsuo[3,8]

Received: 29 November 2018 / Accepted: 14 February 2019 / Published online: 27 February 2019

Abstract

Background Germinal center-associated nuclear protein (GANP) is a phosphoprotein involved in mRNA export and regulation of DNA recombination. Although GANP expression in human breast cancer tissue is associated with breast cancer prognosis, the association between the genetic background of GANP and susceptibility and prognosis of breast cancer is unclear.

Methods We selected 694 breast cancer cases and 1376 age- and menopausal status-matched non-cancer controls from the Hospitable-based Epidemiologic Research Program, conducted at Aichi Cancer Center between 2001 and 2005. We evaluated the impact of two polymorphisms at the *GANP* locus (rs2839178 and rs11702450) on the susceptibility and prognosis of breast cancer. Reference alleles were defined as the A allele for rs2839178 and G allele for rs11702450.

Results The GG genotype of rs2839178 was statistically significantly associated with breast cancer risk (odds ratio [OR] 0.48, 95% confidence interval [CI] 0.30–0.76, $P=0.002$). In prognostic analysis, compared to those with AA genotype of rs2839178, patients with AG or GG genotypes had longer disease-free survival (DFS) (hazard ratio [HR] 0.71, 95% CI 0.49–1.04 and HR 0.42, 95% CI 013–1.42, respectively, P for trend = 0.04). eQTL analysis indicated that association with rs2839178 can be explained by the effect of rs2839173 on expression of GANP/MCM3AP.

Conclusions The G allele of rs2839178 at the *GANP* locus was significantly associated with reduced breast cancer risk and longer DFS in breast cancer patients, showing a consistent direction in the association between susceptibility and clinical outcome. GANP is, therefore, important for the occurrence and progression of sporadic breast cancer.

Keywords Breast cancer risk · Breast cancer prognosis · *GANP*/*MCM3AP* · Polymorphism

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12282-019-00956-5) contains supplementary material, which is available to authorized users.

✉ Hidemi Ito
hidemi@aichi-cc.jp

1 Department of Breast Oncology, Aichi Cancer Center Hospital, 1-1 Kanokoden, Chikusa-ku, Nagoya 464-8681, Japan

2 Division of Surgical Oncology, Department of Surgery, Nagoya University Graduate School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya 466-8550, Japan

3 Division of Cancer Epidemiology and Prevention, Aichi Cancer Center Research Institute, 1-1 Kanokoden, Chikusa-ku, Nagoya 464-8681, Japan

4 Division of Cancer Information and Control, Aichi Cancer Center Research Institute, 1-1 Kanokoden, Chikusa-ku, Nagoya 464-8681, Japan

5 Division of Cancer Descriptive Epidemiology, Nagoya University Graduate School of Medicine, 5 Tsurumai-cho, Showa-ku, Nagoya 466-8550, Japan

6 Department of Diagnostic Pathology, Fujita Health University Graduate School of Medicine, 1-98 Dengakugakubo, Kutsukake-cho, Toyoake 470-1192, Japan

7 Division of Immunology, Aichi Cancer Center Research Institute, 1-1 Kanokoden, Chikusa-ku, Nagoya 464-8681, Japan

8 Department of Epidemiology, Nagoya University Graduate School of Medicine, 5 Tsurumai-cho, Showa-ku, Nagoya 466-8550, Japan

Introduction

Breast cancer is the most common cancer among women worldwide, with about 1.67 million new cases diagnosed in 2012 [1]. Despite the generally good prognosis of breast cancer patients, more than 520,000 patients died from the disease in 2012, making it the fifth-leading cause of cancer death that year [1]. Although several risk factors such as no parity, later live birth, and obesity [2, 3], and genetic factors like *BRCA1* and *BRCA2* have been identified [4, 5], the cause of breast cancer has not been fully elucidated.

Germinal center-associated nuclear protein (GANP), also known as minichromosome maintenance complex component 3-associated protein (MCM3AP), is a nuclear phosphoprotein involved in mRNA export and the regulation of DNA recombination [6, 7]. The coding region of GANP is located on human chromosome 21q22.3. *GANP* reportedly has tumor-suppressive effects in glioblastomas [8]; however, it is controversial, because reverse effect is also observed in the other tumors [9–11]. GANP expression was recently shown to be significantly decreased in human breast cancer tissue compared to normal breast tissue. Furthermore, lower GANP expression is an independent risk factor for poor prognosis of the disease [12]. Homozygous *Ganp*-deficient mice are embryonic lethal [13], and aged female *Ganp*-heterozygous mice develop mammary gland tumors that express similar profiles for estrogen receptor (ER), progesterone receptor (PgR), human epidermal growth factor receptor type 2 (HER2), and Ki-67 to those observed in humans [12]. Moreover, tissue-specific targeting of the *Ganp* gene causes the generation of mammary gland tumors in both *Ganp*-homodeficient and *Ganp*-heterodeficient mice [12]. These results suggest that *GANP* plays an important role in the genesis and progression of breast cancer. We, therefore, hypothesized that genetic polymorphisms in *GANP* may be associated with the risk of breast cancer.

Here, we conducted a case–control study to investigate whether single-nucleotide polymorphisms (SNPs) at the *GANP* locus are associated with breast cancer risk in a Japanese population. Furthermore, we evaluated the prognostic impact of *GANP* polymorphisms among breast cancer patients.

Materials and methods

Subjects

The details of this case–control study were described in our previous study [14]. Subjects were selected from the database of the Hospital-based Epidemiologic Research Program at Aichi Cancer Center (HERPACC), conducted at Aichi Cancer Center Hospital (ACCH) in Nagoya, Japan. Case subjects were 697 female breast cancer patients (354 premenopausal, 343 postmenopausal) without a history of breast cancer and who initially visited ACCH between January 2001 and November 2005. Control subjects were 1394 age- and menopause status-matched females (708 premenopausal and 686 postmenopausal) who visited the hospital without a history of cancer during the same period. The case-to-control subject ratio was 1:2. The framework of HERPACC has been described in detail elsewhere [15, 16]. Briefly, 23,408 HERPACC-enrolled first-visit outpatients were asked to provide blood samples, as well as information on lifestyle factors using a self-administered questionnaire, which was checked by a trained interviewer. Approximately 95% of eligible subjects completed the questionnaire and 55% provide blood samples [14]. The general lifestyle of non-cancer outpatients at our hospital was previously shown to be in accordance with that of a general randomly selected population from the electoral roll in Nagoya City, Aichi Prefecture, confirming the feasibility of their use as controls in epidemiological studies [17]. This study was approved by the Institutional Ethics Review Board of ACCH, and all participants provided written informed consent.

Evaluation of environmental factors

The questionnaire included items on height and weight, menopausal status, parity and lactation, drinking and smoking habits, exercise, past medical history, family history of cancer, and referral pattern to our hospital. This information was obtained at interview about 1 year before the onset of symptoms for symptomatic patients or at interview for asymptomatic patients.

Body mass index (BMI) was calculated as the weight divided by the height squared (kg/m^2). Drinking habit was categorized into never, former, and current light, moderate, and heavy drinking. Light drinking was defined as the consumption of less than 5 g of ethanol per day; moderate drinking as between 5 and less than 23 g per day; and heavy drinking as 23 g or more per day. Smoking habit was categorized into never, former, and current smoking of <20 and ≥ 20 pack-years. Former drinkers and smokers were defined as those who had quit drinking or smoking at least 1 year previously. Subjects were considered to have experienced lactation if they had had a lactation period of more than 1 week. Subjects were considered to have conducted regular exercise if they performed exercise at least twice a week, more than 30 min per one time [18]. Subjects were considered to have a family history of breast cancer if a mother or sister had had breast cancer.

Vital and disease status was estimated by checking the medical records from the date of the last follow-up visit. For patients lost to follow-up, vital status was confirmed from annual census registration data [19].

Genotyping and SNP selection

DNA of each subject was extracted from the buffy coat fraction using a DNA Blood Mini kit (Qiagen K.K. Tokyo, Japan). Cases and controls were genotyped using an Infinium iSelect custom array (iCOGS, Illumina Inc., San Diego, CA, USA) comprising more than 200,000 SNPs. For the purpose of our analysis, we selected all 13 SNPs at the *GANP* locus on chromosome 21: rs12482209, rs2839984, rs2839171, rs2250213, rs2839173, rs2839174, rs2839178, rs2839186, rs2839188, rs2839194, rs4438580, rs11702450, and rs9975588. We narrowed these down by considering the linkage disequilibrium (LD) using genotype information from controls. We combined neighboring SNPs with high linkage disequilibrium ($D' > 0.90$) into the same LD block (Fig. 1). Using this method, we identified two LD blocks and selected one SNP within each LD block, rs2839178 and rs11702450, for analysis.

Expression quantitative trait locus (eQTL) analysis

To speculated the functional consequences of the SNPs of GANP, we used Genotype-Tissue Expression (GTEx) [16] which is a publicity available database for eQTL.

Statistical analysis

Differences in characteristic variables between cases and controls were assessed using a Chi-squared test, excluding subjects with missing data. The ages of case and control subjects were compared using the Mann–Whitney test. The main effect of each SNP on breast cancer risk was assessed using odds ratios (ORs) and 95% confidence intervals (CIs) calculated by conditional logistic regression models (per-allele and genotype model) adjusted for potential confounder variables. Variables used in the model were age, age at menarche (≤ 12, 13–14, ≥ 15), menopausal status (premenopause and postmenopause), current BMI (< 18.5, 18.5–21.9, 22–24.9, ≥ 25), age at first live birth (no birth, 17–24, 25–29, ≥ 30), regular exercise (yes, no), family history of breast cancer (yes, no), and referral pattern to our hospital (patient discretion, recommendation by family or friends, referral from another clinic, secondary screening after primary screening, and others). Missing data for each covariate were included in the model as indicator variables.

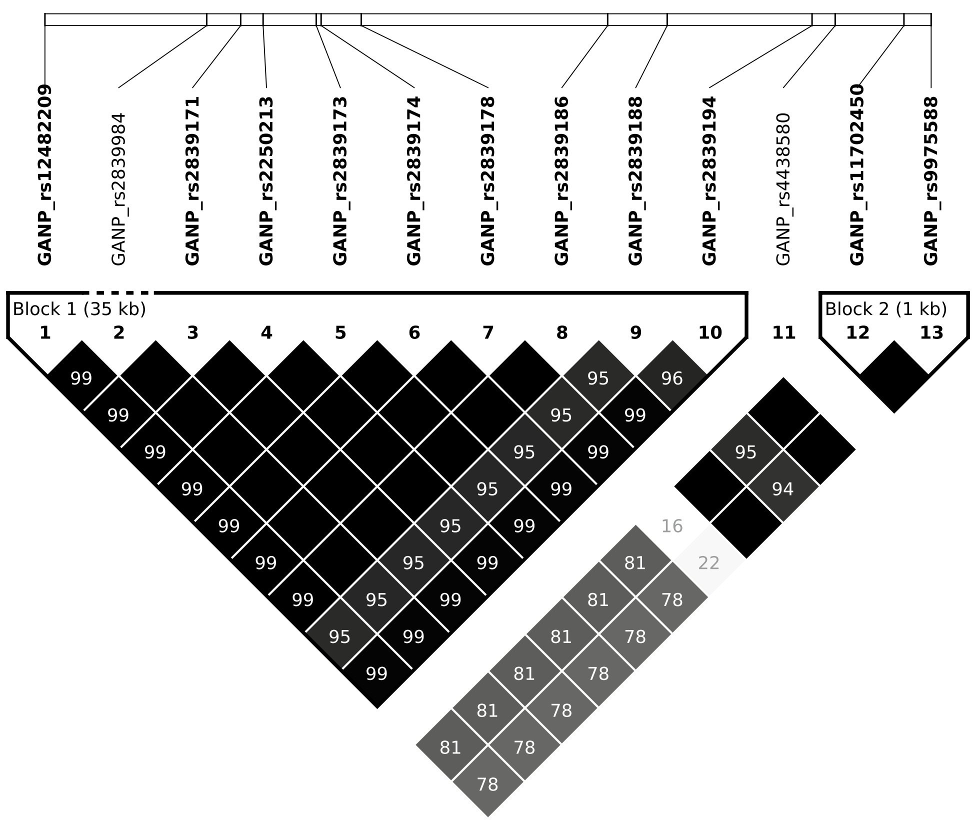

Fig. 1 Linkage-disequilibrium (LD) plot of a region covered by the 13 single-nucleotide polymorphisms (SNPs) at the germinal center-associated nuclear protein (GANP) locus. The region covered by the 13 SNPs at the GANP locus constituted two LD blocks (r^2). Two SNPs (rs2839178 and rs11702450) in each LD block were selected for evaluation

We defined the reference alleles for each SNP as the A allele of rs2839178 and G allele of rs11702450 in the model. We also stratified the analysis by menopausal status. To evaluate the potential heterogeneous impact of SNPs on breast cancer risk due to ER, PgR, Her2, and nodal status, we performed Cochran's Q test for the SNPs and each biological factor in the per-allele model.

Overall survival (OS) was defined as the interval between the date of the initiation of treatment (operation or chemotherapy) and the date of death or last follow-up. Disease-free survival (DFS) was defined as the number of days from the beginning of treatment to the date of relapse, as evaluated and recorded by the patient's physician. OS and DFS were estimated by the Kaplan–Meier method. The impact of the selected SNPs on survival was evaluated by a log-rank test. These associations were also analyzed by a multivariable Cox proportional hazards model with adjustment for potential confounders. Confounders considered in the model were age, stage, operation, recurrence, adjuvant chemotherapy, adjuvant hormonal therapy, ER, PgR, and HER2. Age and stage were treated as continuous variables, and the others were categorized into 2 values. Chemotherapy and hormonal therapy were considered whether they were done or not regardless of type or duration. ER and PgR positivity were defined as more than 3 points using the Allred score. HER2 status was determined using immunohistochemistry (IHC) and/or fluorescence in-situ hybridization (FISH); positive HER2 was defined as 3+ by IHC or gene amplification by FISH. HER2 gene amplification was considered when the ratio of HER2 gene copy number relative to CEP17 was greater than 2.0. Hazard ratio (HR) and 95% CI were used as the measures of association. We also evaluated the impact of the SNPs on clinical outcome (OS and DFS) stratified by characteristic breast cancer variables (ER, PgR, Her2, and nodal status). In addition, we used a multiplicative assumption to assess the heterogeneity of the impact of genotype across breast cancer characteristic variables (ER, PgR, Her2, and node positivity) using an interaction term to represent the interaction between each genotype and variable. For example, the product of the scores for each genotype (0, homozygous for risk allele; 1, heterozygous for risk allele; 2, homozygous for non-risk allele) and variant status (0, ER negative; 1, ER positive) was used as an interaction term in the Cox's proportional hazard model for the heterogeneity of HRs across the variants.

Concordance of the two SNPs with the Hardy–Weinberg law was evaluated using the Chi-squared test with one degree of freedom.

All statistical analyses were carried out using STATA ver.12 (StataCorp, College Station, TX). All tests were two-sided and P values of <0.05 were considered statistically significant.

Results

Characteristics of patients and controls

A total of 697 breast cancer cases and 1394 control subjects were included in the analysis. Table 1 summarizes the characteristics of the subjects. Age and menopausal status were similar between the groups. Breast cancer patients were more likely to have a family history of the disease ($P=0.03$) and a later first live birth among parous women (mean age 26.2 ± 3.4 vs 25.6 ± 3.6, $P=0.02$) compared to control subjects. Similarly, case subjects were more likely to have been referred through a family recommendation or by another clinic, and less likely to have attended at patient discretion and by secondary screening after a primary screening. The other factors did not significantly differ between the two groups.

Association between GANP SNPs and breast cancer risk

The genotype frequencies of the two SNPs at the *GANP* locus among controls were in accordance with the Hardy–Weinberg law (rs2839178: $P=0.51$; rs11702450: $P=0.36$). Table 2 shows the association between the two *GANP* SNPs and breast cancer risk. The minor allele (G) of rs2839178 showed a statistically significant inverse association with breast cancer risk. The OR for the GG genotype was 0.48 (95% CI 0.30–0.76, $P=0.002$), and the OR in the per-allele model was 0.84 (95% CI 0.72–0.99, $P=0.038$). The rs11702450 SNP was not statistically significantly associated with breast cancer risk.

Stratified analysis by menopausal status showed a similar association between the GG genotype of rs2839178 and reduced breast cancer risk for premenopausal (OR 0.43, 95% CI 0.22–0.85, $P=0.015$) and postmenopausal cases (OR 0.47, 95%CI 0.24–0.93, $P=0.029$).

We evaluated the potential heterogeneous impact of the SNPs on breast cancer risk according to ER, PgR, Her2, and nodal status using a per-allele model (Table 3). No significant heterogeneous impact was observed between either of the SNPs and breast cancer risk for any of the characteristic variables.

Impact of SNPs on clinical outcomes (Fig. 2; Table 4 and supplementary table 1)

Breast cancer patients with the G allele of rs2839178 had longer DFS (HR 0.62, 95% CI 0.42–0.92, $P=0.02$) and a tendency for longer OS, although this was not statistically significant (HR 0.65, 95% CI 0.40–1.09, $P=0.11$).

Table 1 Characteristics of breast cancer cases and cancer-free controls

	Case		Control		*P*
	(*n* = 697)	(%)	(*n* = 1394)	(%)	
Age (years)					
≤ 29	10	(1.4)	27	(1.9)	
30–39	95	(13.6)	177	(12.7)	
40–49	197	(28.3)	416	(29.8)	
50–59	222	(31.9)	429	(30.8)	
60–69	132	(18.9)	271	(19.4)	
70–79	41	(5.9)	74	(5.3)	
Mean age ± SD	51.6 ± 11.0		51.4 ± 11.0		0.65
Menopausal status					
Premenopausal	354	(50.8)	708	(50.8)	
Postmenopausal	343	(49.2)	686	(49.2)	1.00
Family history of breast cancer					
No	643	(92.3)	1320	(94.7)	
Yes	54	(7.8)	74	(5.3)	0.03
Parity					
0	108	(15.5)	206	(14.8)	
1–2	452	(64.9)	870	(62.4)	
≥ 3	137	(19.7)	316	(22.7)	0.28
Unknown	0	(0.0)	2	(0.1)	
Age at first live birth (years)					
≤ 24	201	(28.8)	494	(35.4)	
25–29	301	(43.2)	532	(38.2)	
≥ 30	84	(12.1)	149	(10.7)	
Nonparous	108	(15.5)	206	(14.8)	0.02
Unknown	3	(0.4)	13	(0.9)	
Mean age ± SD	26.2 ± 3.4		25.6 ± 3.6		0.02
Age at menarche (years)					
≤ 12	216	(31.0)	439	(31.5)	
13–14	340	(48.8)	648	(46.5)	
≥ 15	133	(19.1)	277	(19.9)	0.72
Unknown	8	(1.2)	30	(2.2)	
Mean age ± SD	13.3 ± 1.5		13.4 ± 1.6		0.44
Age at menopause (postmenopausal women only) (years)					
< 50	118	(34.4)	261	(38.1)	
≥ 50	223	(65.0)	417	(60.8)	0.23
Unknown	2	(0.6)	8	(1.2)	
Mean age ± SD	49.7 ± 4.9		49.4 ± 4.7		0.32
Lactation					
No	131	(18.8)	260	(18.7)	
Yes	558	(80.1)	1117	(80.1)	0.94
Unknown	8	(1.2)	17	(1.2)	
BMI					
< 18.5	56	(8.0)	116	(8.3)	
18.5–21.9	274	(39.3)	598	(42.9)	
22–24.9	224	(32.1)	423	(30.3)	
≥ 25	143	(20.5)	245	(17.6)	0.25
Unknown	0	(0.0)	12	(0.9)	
Drinking habit					
Never	416	(59.7)	779	(55.9)	

Table 1 (continued)

	Case		Control		*P*
	(*n* = 697)	(%)	(*n* = 1394)	(%)	
Former[a]	12	(1.7)	37	(2.7)	
Current-light[b]	131	(18.8)	309	(22.2)	
Current-moderate[c]	81	(11.6)	173	(12.4)	
Current-heavy[d]	49	(7.0)	77	(5.5)	0.12
Unknown	8	(1.2)	19	(1.4)	
Smoking habit					
Never	585	(83.9)	1113	(79.8)	
Former[a]	33	(4.7)	79	(5.7)	
Current: 0–19 (pack-years)	49	(7.0)	136	(9.8)	
Current: ≥ 20 (pack-years)	28	(4.0)	63	(4.5)	0.12
Unknown	2	(0.3)	3	(0.2)	
Regular exercise					
No	351	(50.4)	699	(50.1)	
Yes	338	(48.5)	685	(49.1)	0.59
Unknown	8	(1.2)	10	(0.7)	
Referral pattern to our hospital					
Patient's discretion	186	(26.7)	432	(31.0)	
Family recommendation	159	(23.0)	214	(15.4)	
Referral from another clinic	205	(29.4)	286	(20.5)	
Secondary screening after primary screening	138	(19.8)	445	(31.9)	
Other	6	(0.9)	11	(0.8)	< 0.01
Unknown	3	(0.4)	6	(0.4)	
Breast cancer characters					
ER positive	391	(56.1)			
ER negative	142	(20.4)			
Unknown	164	(23.5)			
PgR positive	327	(46.9)			
PgR negative	202	(29.0)			
Unknown	168	(24.1)			
Her2 positive	98	(14.1)			
Her2 negative	331	(47.5)			
Unknown	268	(38.5)			
Node positive	190	(27.3)			
Node negative	359	(51.5)			
Unknown	148	(21.2)			

BMI body mass index, *ER* estrogen receptor, *PR* progesterone receptor, *Her2* human epidermal growth factor receptor 2, *Node* lymph-node metastasis

[a]Former drinkers and smokers were defined as those who had quit drinking or smoking at least 1 year ago

[b]Light drinker refers to those consuming < 5 g ethanol/day

[c]Moderate drinker refers to those consuming between 5 and 23 g ethanol/day

[d]Heavy drinker refers to those consuming > 23 g/day

Similar to the results of our risk evaluation, the genotypes of rs11702450 showed no association with the clinical outcomes of breast cancer patients. No significant interaction was observed between genotype and clinical outcome on the stratified analysis by breast cancer characteristic variables for either rs2839178 or rs11702450 (data not shown).

Expression quantitative trait locus (eQTL) analysis

We found that rs2839173, which is tightly linked with rs2839178 (D'=1.00 and $R^2 = 0.825$), was significantly associated with the gene expression of *GAMP*/*MCM3AP* ($P = 1.9 \times 10^{-7}$) in normal breast tissue (Supplementary

Table 2 Logistic regression analysis of 2 SNPs at the GANP locus for breast cancer risk and stratified analysis by menopausal status

SNP	MAF		Case	Control	Genotype			Premenopausal				Postmenopausal			
					OR[a]	(95% CI)	*P*	Case/control	OR[a]	(95% CI)	*P*	Case/control	OR[a]	(95% CI)	*P*
rs2839178	0.265				0.84	(0.72–0.99)	0.038	353/700	0.87	(0.69–1.09)	0.217	341/676	0.827	(0.65–1.05)	0.119
		AA	394	749	1.00	(Reference)		201/384	1.00	(Reference)		193/365	1.00	(Reference)	
		AG	273	526	0.99	(0.81–1.21)	0.903	139/266	1.08	(0.81–1.44)	0.6	134/260	0.96	(0.71–1.29)	0.774
		GG	27	101	0.48	(0.30–0.76)	0.002	13/50	0.43	(0.22–0.85)	0.015	14/51	0.47	(0.24–0.93)	0.029
rs11702450	0.113				0.89	(0.71–1.10)	0.277	353/700	0.83	(0.61–1.13)	0.244	341/676	0.94	(0.67–1.30)	0.702
		GG	553	1086	1.00	(Reference)		283/546	1.00	(Reference)		270/540	1.00	(Reference)	
		AG	136	269	0.95	(0.75–1.22)	0.702	67/139	0.94	(0.66–1.33)	0.725	69/130	0.981	(0.68–1.41)	0.918
		AA	5	21	0.44	(0.16–1.21)	0.112	3/15	0.34	(0.09–1.25)	0.105	2/6	0.559	(0.11–2.89)	0.488

MAF minor allele frequency, *CI* confidence interval

[a]Odds ratios (ORs) were adjusted for age, family history of breast cancer, age at first live birth, body mass index, drinking habit, smoking habit, regular exercise, and referral pattern to our hospital

Fig. 1). In addition, all the other SNPs which belong to the same LD block with rs2839178 were significantly associated with the gene expression of *GANP*/*MCM3AP* (rs12482209, rs2839984, rs2839171, rs2250213, rs2839174, rs2839186, rs2839188, and rs2839194; $P = 9.7 \times 10^{-7}$, 1.0×10^{-5}, 1.1×10^{-5}, 4.4×10^{-7}, 1.9×10^{-7}, 1.9×10^{-7}, 2.5×10^{-7}, 2.1×10^{-5}, and 9.2×10^{-8}, respectively). Rs11702450 were also found to be eQTL ($P = 1.3 \times 10^{-8}$) for the *GANP*/*MCM3AP* expression in breast tissue.

Discussion

In this analysis of the association between GANP genotype and breast cancer risk, we found that the G allele of rs2839178 at the *GANP* locus was significantly associated with reduced breast cancer risk. In addition, the G allele of rs2839178 also increased breast cancer DFS and decreased breast cancer susceptibility in the same direction, albeit that the statistical significance of the latter was attenuated after Bonferroni correction. To our knowledge, this is the first study to report an association between polymorphisms at the *GANP* locus and breast cancer risk and prognosis. In addition, eQTL analysis indicated that the associations which we found in this study might be due to functional significance of the locus on expression of GANP/MCM3AP.

DNA recombination and RNA export, in which GANP plays a role [7], are important for the maintenance of genome integrity [8, 20]. GANP is a member of the Transcription-Export-2 (TREX-2) complex [21], which is involved in mRNA export and prevents genome instability [20]. Interestingly, TREX-2 binds to BRCA2 repair factor and affects genome integrity by preventing R-loop accumulation [22], which may be a chief source of replication stress and cancer-associated instability.

A previous study showed that the expression of GANP was significantly decreased in human breast cancer tissue, and that lower expression of GANP was an independent risk factor for poor prognosis of the disease [12]. Together with our finding of a consistent direction of association for low-risk susceptibility and prognosis with rs2839178 in this study, one might hypothesize that *GANP*/*MCM3AP* rs2839178 is associated with GANP/MCM3AP expression via truly functional variants around rs2839178. Interestingly, rs2839173 belongs to the same LD block with rs2839178 and is a variant in exon, the A allele of this locus which correspond to G allele of rs2839178 was significantly associated with the expression of GANP/MCM3AP in breast tissue in the eQTL analysis (Supplementary Fig. 1). These might indicate that G allele of rs2839178 may indirectly enhance *GANP*/*MCM3AP* expression in breast tissue, leading to modify susceptibility and prognosis. Moreover, a recent whole-exome association study conducted in Chinese

Table 3 Stratified analysis assessing the heterogeneous impact of SNPs on breast cancer risk according to ER, PgR, Her2, and nodal status

SNP	OR[a]	(95% CI)	*P*	OR[a]	(95% CI)	*P*	*P* for interaction
rs2839178							
	ER positive			ER negative			
	0.82	(0.67–0.99)	0.039	0.79	(0.59–1.08)	0.138	0.887
	PgR positive			PgR negative			
	0.84	(0.69–1.04)	0.127	0.77	(0.60–1.00)	0.051	0.604
	Her2 positive			Her2 negative			
	0.80	(0.56–1.13)	0.207	0.76	(0.62–0.94)	0.013	0.837
	Node positive			Node negative			
	0.67	(0.51–0.88)	0.004	0.88	(0.72–1.07)	0.174	0.122
rs11702450							
	ER positive			ER negative			
	0.94	(0.72–1.22)	0.636	0.80	(0.52–1.24)	0.322	0.547
	PgR positive			PgR negative			
	0.94	(0.71–1.24)	0.641	0.88	(0.61–1.25)	0.470	0.777
	Her2 positive			Her2 negative			
	0.96	(0.59–1.54)	0.86	0.92	(0.69–1.22)	0.551	0.881
	Node positive			Node negative			
	0.82	(0.57–1.19)	0.299	0.97	(0.74–1.27)	0.83	0.476

ER estrogen receptor, *PR* progesterone receptor, *Her2* human epidermal growth factor receptor 2, *Node* lymph-node metastasis

[a]Odds ratios (ORs) were adjusted for age, family history of breast cancer, age at first live birth, body mass index, drinking habit, smoking habit, regular exercise, and referral pattern to our hospital

women provided supporting evidence for above our speculation [23]. In the study, rs13047478 on *C21or58* which is a nearby gene of *GANP*/*MCM3AP* was identified as one of the novel missense variants for breast cancer risk. They confirmed that G allele of rs13047478 decreased the mRNA levels of *MCM3AP* in MCF7 cells by luciferase reporter assay. According to HapMap JPT data, rs13047478 and rs2839178 are in strong LD ($D'=1$ and $R^2=0.8573$). This also supports our idea that rs2839178 affects expression of *GANP*/*MCM3AP* and leads to risk of breast cancer as well as prognosis as a consequence. Further studies evaluating functional significance of loci which were related to the *GANP*/*MCM3AP* expression on breast cancer risk and prognosis are warranted.

This study has several methodological strengths. First, the size of our study was sufficient to assess the association between *GANP* polymorphisms and risk/prognosis of breast cancer. Second, controls were selected from the same hospital and almost all the participants lived in the same area, making it likely that all patients would be referred to Aichi Cancer Center Hospital for cancer diagnosis and treatment. Therefore, case and control subjects were assumed to be from the same population base, indicating the internal validity of this case-control study. Third, we previously confirmed that the questionnaire-based lifestyle characteristics in this population were consistent with those of the general population in Nagoya City [14], indicating the external generalizability of our findings to the Japanese population. Fourth, potential confounding, including age and menopausal status, was considered by individual matching when selecting subjects and by statistical adjustment in the analyses. Fifth, allele frequencies were comparable to those previously reported as Japanese population in public databases, such as 1000 Genome project [24], indicating that bias in the distribution of selected polymorphisms was negligible.

Potential limitations of our study also warrant mention. First, we collected information about confounders via a self-reported questionnaire, making it difficult to rule out the potential for some information bias. If present, however, the effect of any misclassification due to potential under-adjustment is expected to be negligible. Second, control participants were selected among non-cancer patients at our hospital. To dilute any bias that might have resulted from the inclusion of a specific diagnostic group associated with the exposure, we did not set eligibility criteria for control diseases. Third, the retroactive nature of this study raises the possibility of recall bias. However, as the *GANP* genotype does not change throughout life, this bias is unlikely. Fourth, we may need to consider about multiple comparisons in our analysis. For rs2839178, although the statistical significance in the per-allele model is attenuated ($P=0.038>0.025$) after Bonferroni correction, the genotype model kept statistical significance ($P=0.002<0.025$), warranting replication in the other studies. Finally, although we

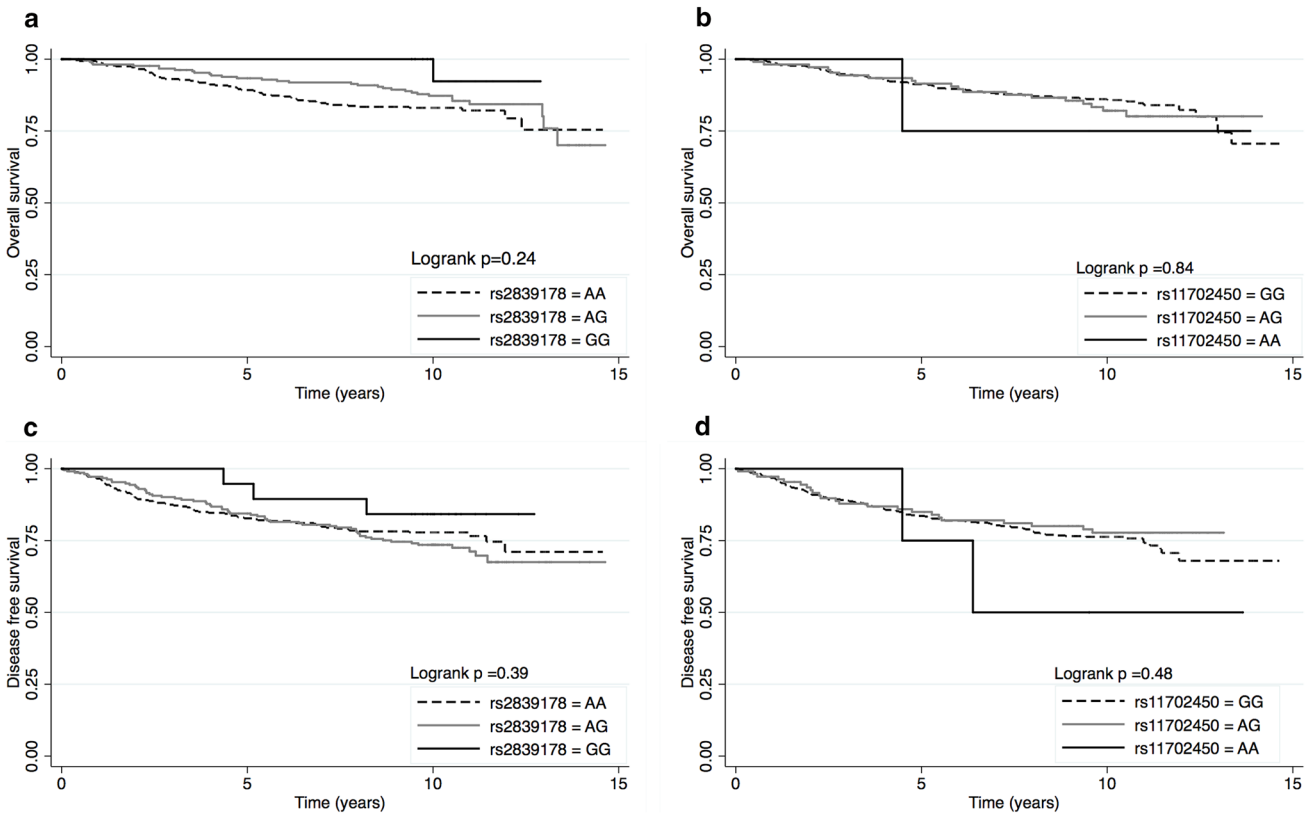

Fig. 2 Kaplan–Meier estimates of disease-free survival and overall survival according to the genotypes of each SNP. **a, b** Overall survival according to the genotypes of rs2839178 and rs11702450, respectively. **c, d** Disease-free survival according to the genotypes of rs2839178 and rs11702450, respectively. Values are based on the Kaplan–Meier method. Median observation period was 10.1 years

Table 4 Cox hazard models for the association between SNPs and overall survival (OS) and disease-free survival (DFS)

	DFS				OS			
	HR[a]	95% CI	*P*	Median F/U	HR[b]	95% CI	*P*	Median F/U
rs2839178	0.62	(0.42–0.92)	0.02		0.66	(0.40–1.09)	0.11	
AA	1.00	(Reference)		10.00	1.00	(Reference)		10.03
AG	0.60	(0.39–0.93)	0.02	10.00	0.64	(0.37–1.09)	0.10	10.14
GG	0.42	(0.12–1.66)	0.23	10.02	0.63	(0.08–4.93)	0.66	10.32
rs11702450	1.25	(0.79–1.98)	0.34		1.48	(0.85–2.59)	0.17	
GG	1.00	(Reference)		10.01	1.00	(Reference)		10.08
AG	1.11	(0.65–1.90)	0.60	9.97	1.33	(0.72–2.44)	0.36	10.00
AA	2.88	(0.64–12.89)	0.17	7.95	6.06	(0.72–51.18)	0.10	11.59

MedianF/U median follow-up period (years)

[a]Hazard ratios (HRs) were adjusted for age, stage, operation, adjuvant chemotherapy, adjuvant hormonal therapy, ER, PgR, and Her2

[b]Hazard ratios (HRs) were adjusted for age, stage, operation, adjuvant chemotherapy, adjuvant hormonal therapy, ER, PgR, Her2, and recurrence

minimized confounding by individual matching and statistical adjustment, we cannot completely rule out the effect of residual confounding from unevaluated factors. For example, in survival analysis, we could not evaluate the performance status because of the lack of information on this variable.

In conclusion, we evaluated the association between *GANP*/*MCM3AP* SNPs and breast cancer risk and prognosis in a Japanese population. We found that the G allele of rs2839178 at the *GANP*/*MCM3AP* locus was significantly associated with the reduced breast cancer risk. Furthermore,

it was associated with increased DFS in breast cancer patients, showing consistency in the direction of association with susceptibility and clinical outcome. eQTL analysis suggested that an impact of rs2839178 on risk and prognosis of sporadic breast cancer is via the expression of GANP/MCM3AP on breast tissue. Our results prompt the need for further basic studies to clarify the biological mechanism of GANP/MCM3AP in breast cancer.

Acknowledgements We thank the many doctors, nurses, and technical and administrative staff of Aichi Cancer Center Hospital for their daily administration of the HERPACC study.

Funding This study was funded by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports, Culture and Technology of Japan (17015018, 15K08792 and 24590776) and by a Grant-in-Aid for Third-Term Comprehensive 10-Year Strategy for Cancer Control from Ministry of Health, Labour and Welfare of Japan, and Japan Agency for Medical Research and Development (15655291).

Conflict of interest Haruru Kotani declares that she has no conflict of interest. Hidemi Ito declares that she has no conflict of interest. Kazuhiko Kuwahara declares that he has no conflict of interest. Kiyotaka Kuzushima declares that he has no conflict of interest. Hiroji Iwata declares that he has no conflict of interest. Nobuyuki Tsunoda declares that he has no conflict of interest. Masato Nagino declares that he has no conflict of interest. Keitaro Matsuo declares that he has no conflict of interest.

References

1. International Agency for Research on Cancer. GLOBOCAN, World Health Organization, vol. 2017. http://globocan.iarc.fr/old/FactSheets/cancers/breast-new.asp. Accessed 6 Aug 2018.
2. Kelsey JL, Gammon MD, John EM. Reproductive factors and breast cancer. Epidemiol Rev. 1993;15(1):36–47.
3. Taira N, Arai M, Ikeda M, Iwasaki M, Okamura H, Takamatsu K, Nomura T, Yamamoto S, Ito Y, Mukai H. The Japanese Breast Cancer Society clinical practice guidelines for epidemiology and prevention of breast cancer, 2015 edition. Breast Cancer. 2016;23(3):343–56. https://doi.org/10.1007/s12282-016-0673-8.
4. Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q, Cochran C, Bennett LM, Ding W, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science. 1994;266(5182):66–71.
5. Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, Collins N, Gregory S, Gumbs C, Micklem G. Identification of the breast cancer susceptibility gene BRCA2. Nature. 1995;378(6559):789–92. https://doi.org/10.1038/378789a0.
6. Kuwahara K, Yoshida M, Kondo E, Sakata A, Watanabe Y, Abe E, Kouno Y, Tomiyasu S, Fujimura S, Tokuhisa T, Kimura H, Ezaki T, Sakaguchi N. A novel nuclear phosphoprotein, GANP, is up-regulated in centrocytes of the germinal center and associated with MCM3, a protein essential for DNA replication. Blood. 2000;95(7):2321–8.
7. Okamoto N, Kuwahara K, Ohta K, Kitabatake M, Takagi K, Mizuta H, Kondo E, Sakaguchi N. Germinal center-associated nuclear protein (GANP) is involved in mRNA export of Shugoshin-1 required for centromere cohesion and in sister-chromatid exchange. Genes Cells. 2010;15(5):471–84. https://doi.org/10.1111/j.1365-2443.2010.01396.x.
8. Ohta K, Kuwahara K, Zhang Z, Makino K, Komohara Y, Nakamura H, Kuratsu J, Sakaguchi N. Decreased expression of germinal center-associated nuclear protein is involved in chromosomal instability in malignant gliomas. Cancer Sci. 2009;100(11):2069–76. https://doi.org/10.1111/j.1349-7006.2009.01293.x.
9. Fujimura S, Xing Y, Takeya M, Yamashita Y, Ohshima K, Kuwahara K, Sakaguchi N. Increased expression of germinal center-associated nuclear protein RNA-primase is associated with lymphomagenesis. Cancer Res. 2005;65(13):5925–34. https://doi.org/10.1158/0008-5472.CAN-04-3259.
10. Kageshita T, Kuwahara K, Oka M, Ma D, Ono T, Sakaguchi N. Increased expression of germinal center-associated nuclear protein (GANP) is associated with malignant transformation of melanocytes. J Dermatol Sci. 2006;42(1):55–63. https://doi.org/10.1016/j.jdermsci.2005.12.007.
11. Chan-On W, Kuwahara K, Kobayashi N, Ohta K, Shimasaki T, Sripa B, Leelayuwat C, Sakaguchi N. Cholangiocarcinomas associated with long-term inflammation express the activation-induced cytidine deaminase and germinal center-associated nuclear protein involved in immunoglobulin V-region diversification. Int J Oncol. 2009;35(2):287–95.
12. Kuwahara K, Yamamoto-Ibusuki M, Zhang Z, Phimsen S, Gondo N, Yamashita H, Takeo T, Nakagata N, Yamashita D, Fukushima Y, Yamamoto Y, Iwata H, Saya H, Kondo E, Matsuo K, Takeya M, Iwase H, Sakaguchi N. GANP protein encoded on human chromosome 21/mouse chromosome 10 is associated with resistance to mammary tumor development. Cancer Sci. 2016. https://doi.org/10.1111/cas.12883.
13. Yoshida M, Kuwahara K, Shimasaki T, Nakagata N, Matsuoka M, Sakaguchi N. GANP suppresses DNA recombination, measured by direct-repeat beta-galactosidase gene construct, but does not suppress the type of recombination applying to immunoglobulin genes in mammalian cells. Genes Cells. 2007;12(10):1205–13. https://doi.org/10.1111/j.1365-2443.2007.01119.x.
14. Sueta A, Ito H, Kawase T, Hirose K, Hosono S, Yatabe Y, Tajima K, Tanaka H, Iwata H, Iwase H, Matsuo K. A genetic risk predictor for breast cancer using a combination of low-penetrance polymorphisms in a Japanese population. Breast Cancer Res Treat. 2012;132(2):711–21. https://doi.org/10.1007/s10549-011-1904-5.
15. Hamajima N, Matsuo K, Saito T, Hirose K, Inoue M, Takezaki T, Kuroishi T, Tajima K. Gene-environment Interactions and Polymorphism Studies of Cancer Risk in the Hospital-based Epidemiologic Research Program at Aichi Cancer Center II (HERPACC-II). Asian Pac J Cancer Prev. 2001;2(2):99–107.
16. Tajima K, Hirose K, Inoue M, Takezaki T, Hamajima N, Kuroishi T. A Model of Practical Cancer Prevention for Out-patients Visiting a Hospital: the Hospital-based Epidemiologic Research Program at Aichi Cancer Center (HERPACC). Asian Pac J Cancer Prev. 2000;1(1):35–47.
17. Inoue M, Tajima K, Hirose K, Hamajima N, Takezaki T, Kuroishi T, Tominaga S. Epidemiological features of first-visit outpatients in Japan: comparison with general population and variation by sex, age, and season. J Clin Epidemiol. 1997;50(1):69–77.
18. Ministry of Health, Labour and Welfare, Japan. Physical activity for promoting health. https://www.mhlw.go.jp/www1/topics/kenko21_11/b2.html. Accessed 2 Nov 2018.
19. Kawakita D, Oze I, Hosono S, Ito H, Watanabe M, Yatabe Y, Hasegawa Y, Murakami S, Tanaka H, Matsuo K. Prognostic value of drinking status and aldehyde dehydrogenase 2 polymorphism in patients with head and neck squamous cell carcinoma. J Epidemiol. 2016;26(6):292–9. https://doi.org/10.2188/jea.JE20140240.
20. Gonzalez-Aguilera C, Tous C, Gomez-Gonzalez B, Huertas P, Luna R, Aguilera A. The THP1-SAC3-SUS1-CDC31 complex works in transcription elongation-mRNA export

preventing RNA-mediated genome instability. Mol Biol Cell. 2008;19(10):4310–8. https://doi.org/10.1091/mbc.E08-04-0355.
21. Jani D, Lutz S, Hurt E, Laskey RA, Stewart M, Wickramasinghe VO. Functional and structural characterization of the mammalian TREX-2 complex that links transcription with nuclear messenger RNA export. Nucleic Acids Res. 2012;40(10):4562–73. https://doi.org/10.1093/nar/gks059.
22. Bhatia V, Barroso SI, Garcia-Rubio ML, Tumini E, Herrera-Moyano E, Aguilera A. BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. Nature. 2014;511(7509):362–5. https://doi.org/10.1038/nature13374.
23. Zhang B, Chen MY, Sheng YJ, Zhuo XB, Gao P, Zhou FS, Liang B, Zu J, Zhang Q, Suleman S, Xu YH, Xu MG, Xu JK, Liu CC, Giannareas N, Xia JH, Zhao Y, Huang ZL, Yang Z, Cheng H, Li N, Hong YY, Li W, Zhang MJ, Yu KD, Li G, Sun MH, Chen ZD, Wei GH, Shao ZM. A large-scale, exome-wide association study of Han Chinese women identifies three novel loci predisposing to breast cancer. Cancer Res. 2018. https://doi.org/10.1158/0008-5472.CAN-17-1721.
24. 1000 Genomes Project, vol. 2018. https://www.ncbi.nlm.nih.gov/variation/tools/1000genomes/. Accessed 6 Aug 2018.