

## **Genome-Wide Association Studies (GWAS) breast cancer susceptibility loci in Arabs: susceptibility and prognostic implications in Tunisians.**

Jingxuan Shan, Wijden Mahfoudh, Shoba P Dsouza, Elham Hassen, Nouredine Bouaouina, Sonia Abdelhak, Ahlem Benhadjayed, Hager Memmi, Rebecca Ann Mathew, Idil I Aigha, et al.

► **To cite this version:**

Jingxuan Shan, Wijden Mahfoudh, Shoba P Dsouza, Elham Hassen, Nouredine Bouaouina, et al.. Genome-Wide Association Studies (GWAS) breast cancer susceptibility loci in Arabs: susceptibility and prognostic implications in Tunisians.. Breast Cancer Research and Treatment, Springer Verlag, 2012, 135 (3), pp.715-24. 10.1007/s10549-012-2202-6 . pasteur-01375153

**HAL Id: pasteur-01375153**

**<https://hal-riip.archives-ouvertes.fr/pasteur-01375153>**

Submitted on 2 Oct 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Genome-Wide Association Studies (GWAS) breast cancer susceptibility loci in Arabs: susceptibility and prognostic implications in Tunisians

Jingxuan Shan · Wijden Mahfoudh · Shoba P. Dsouza · Elham Hassen ·  
Noureddine Bouaouina · Sonia Abdelhak · Ahlem Benhadjayed · Hager Memmi ·  
Rebecca Ann Mathew · Idil I. Aigha · Sallouha Gabbouj · Yasmine Remadi ·  
Lotfi Chouchane

Received: 21 June 2012 / Accepted: 7 August 2012 / Published online: 22 August 2012  
© The Author(s) 2012. This article is published with open access at Springerlink.com

**Abstract** Genome-wide Association Studies (GWAS) revealed novel genetic markers for breast cancer susceptibility. But little is known about the risk factors and molecular events associated with breast cancer in Arab Population. Therefore, we designed a broad study to investigate the susceptibility and prognostic implications of the GWAS breast cancer loci in the Tunisian population. In a cohort of 640 unrelated patients with breast cancer and 371 healthy control subjects, we characterized the variation of 9 single nucleotide polymorphisms (SNPs), namely rs1219648, rs2981582; rs8051542, rs12443621, and rs3803662; rs889312; rs3817198; rs13387042 and rs13281615. Only 5 out of 9 GWAS breast cancer loci were found to be significantly associated with breast cancer in Tunisians: The rs1219648 (G vs. A allele: OR = 1.36,  $P = 1 \times 10^{-3}$ ) and rs2981582 (A vs. G allele: OR = 1.55,  $P = 3 \times 10^{-6}$ ) of

*FGFR2* gene; the rs8051542 of the *TNRC9* gene (T vs. C allele: OR = 1.40,  $P = 4 \times 10^{-4}$ ); the rs889312 of the *MAP3K1* gene (C vs. A allele: OR = 1.33,  $P = 3 \times 10^{-3}$ ) and the rs13281615 located on 8q24 (G vs. A allele: OR = 1.21,  $P = 0.03$ ). Homozygous variant genotypes of rs2981582 were strongly related to lymph node negative breast cancer (OR = 3.33,  $P = 6 \times 10^{-7}$ ) and the minor allele of rs2981582 was associated with increased risk of ER+ tumors (OR = 1.57,  $P = 0.02$ ; OR = 2.15,  $P = 0.001$ , for heterozygous and homozygous variant genotypes, respectively) and increased risk of distant metastasis development (OR = 2.30,  $P = 4 \times 10^{-3}$ ; OR = 3.57,  $P = 6 \times 10^{-5}$ , for heterozygous and homozygous variant genotypes, respectively) in a dose dependent manner. The association for rs8051542 was stronger for high-grade SBR tumors (OR = 2.54,  $P = 2 \times 10^{-4}$ ). GG genotype of rs13387042 on 2q35 showed a significant association with the risk of developing distant metastasis (OR = 1.94,  $P = 0.02$ ). The G allele of rs1219648 in *FGFR2* and the A allele of rs13387042 on 2q35 indicated a better prognosis by showing a significantly higher overall survival rates ( $P = 0.013$  and  $P = 0.005$ , respectively). In conclusion, GWAS breast cancer *FGFR2*, *TNRC9*, *MAP3K1*, and 8q24 loci are associated with an increased risk of breast cancer and genetic variation in *FGFR2* gene may predict the aggressiveness of breast cancer in Tunisians.

Jingxuan Shan and Wijden Mahfoudh are contributed equally to the study.

**Electronic supplementary material** The online version of this article (doi:10.1007/s10549-012-2202-6) contains supplementary material, which is available to authorized users.

J. Shan · S. P. Dsouza · R. A. Mathew ·  
I. I. Aigha · L. Chouchane (✉)  
Genetic Medicine and Immunology Laboratory, Weill Cornell  
Medical College in Qatar, Qatar Foundation, Education City,  
P.O. Box 24144, Doha, Qatar  
e-mail: loc2008@med.cornell.edu

W. Mahfoudh · E. Hassen · N. Bouaouina · A. Benhadjayed ·  
H. Memmi · S. Gabbouj · Y. Remadi  
Laboratory of Immuno-Oncology, Faculty of Medicine of  
Monastir, University of Monastir, Monastir, Tunisia

S. Abdelhak  
Research Unit on Orphan Genetic Diseases, Institut Pasteur de  
Tunis, 13 Place Pasteur, 1002 Tunis, Tunisia

**Keywords** Breast cancer · Tunisians · Arabs · GWAS · Prognosis · Survival

## Introduction

Breast cancer is the most common form of malignancy affecting women worldwide [1, 2]. Breast cancer

incidence rates have increased progressively in Arab populations over the last 10 years, probably due to more reliable data being collected from cancer registries and to easier access by patients screening and diagnostic program [3]. In Arab populations, breast cancer represents ~13–30 % of newly diagnosed malignancies in women and occurs at a median age of 49–52 years as compared to 63 in industrialized nations [3]. It is characterized by younger age at onset, advanced stage and poor prognosis [3]. In Tunisia, breast cancer remains the most common cancer among women, and it is considered to be a real problem of public health. During the period from 1993 to 2007, the Cancer Registry of the Center of Tunisia counted 2,404 new cases of breast cancer. The median age at diagnosis was 48 years and the age-standardized incidence rate (ASR) was 29.2 per 100,000 s during the study period [4].

The etiology of breast cancer is extremely complex and, while not yet elucidated, appears to involve numerous genetic, endocrine, and external environmental factors. Family history represents the most prominent risk factor for the development of the disease. It is estimated that about 5–10 % of all breast cancers may arise from the inheritance of germline mutations in dominant highly penetrant susceptibility genes such as *BRCA1* and *BRCA2*. Mutations in these genes are rare and explain only a small fraction of the familial risk for the disease [5, 6]. This leads to the suggestion that the remaining breast cancer susceptibility is likely to be explained by a polygenic model involving a combination of low-penetrance alleles, each conferring a small increase in risk [7]. Recently, genome-wide association studies (GWAS) have provided a systematic way to search for genetic variants and have successfully identified several low-penetrance susceptibility loci for breast cancer [8–19]. Most of breast cancer GWAS and replications published today has been conducted in Northern European populations [8–10, 12, 14–16, 18] and to a lesser extent in Asians [13, 17, 19] and Ashkenazi Jews [11]. Thus, it is important to assess whether these variants confer risk across different populations with diverse ancestry backgrounds, including women of Arab ancestry. Moreover, little is known about risk factors and molecular events associated with breast cancer in Arab populations, which differ strongly from the other populations by ethnicity, lifestyle, reproductive behavior, and environmental exposure. This prompts us to analyze the previously GWAS-identified breast cancer risk variants in Tunisians using a case–control study. In this report, we focus on nine polymorphisms in the following genes/regions: *FGFR2*, *TNRC9* (also known as *TOX3*), *MAP3K1*, *LSP1*, 2q35, and 8q24. We further explored other potential effects of these risk loci on disease characteristics and survival.

## Materials and methods

### Study population

A total of 1,011 individuals, comprising 640 breast cancer patients and 371 healthy controls, were included in this study. Controls and patients were selected from the same population living in the middle coast of Tunisia and including only unrelated subjects. The sporadic breast cancer patients were recruited from the department of Radiation Oncology of Sousse Hospital (Sousse, Tunisia) between 1996 and 2011. Their disease information was obtained from their hospital medical records. All patients included in this study had primary breast cancer, with unilateral breast tumors and with no family history of the disease. The diagnosis of cancer was confirmed by histopathological analyses. The patients had a mean age of  $47.9 \pm 10$  years. The median follow-up was 65 months (range, 1–276 months). At the time of analysis, 118 patients relapsed (local or distant recurrence). Among them, 20 (17 %) patients died from breast cancer. A detailed description of the clinic-pathological characteristics of this cohort was summarized in Table 1.

Controls were healthy women having a mean age of  $55 \pm 14$  years. They were blood donors with no evidence of any personal or family history of cancer (or other chronic illness). Samples from healthy controls were collected consecutively between 2004 and 2010 and were age matched to the cases.

Both patients and controls gave their written consent to participate in the study and to allow their biological samples to be genetically analyzed. Approval for the study was given by the Tunisian National Ethical Committee and by Weill Cornell Medical College in Qatar IRB committee.

### Genomic DNA extraction

Genomic DNA was extracted from peripheral blood leukocytes by a “salting out” procedure [20]. Briefly, 10 ml of blood was mixed with Triton lysis buffer (0.32 M sucrose, 1 % Triton X-100, 5 mM  $MgCl_2$ ,  $H_2O$ , 10 mM Tris-HCl, pH 7.5). Leukocytes were spun down and washed with  $H_2O$ . The pellet was incubated with proteinase K at 56 °C and subsequently salted out at 4 °C using a saturated NaCl solution. Precipitated proteins were removed by centrifugation. The DNA in supernatant fluid was precipitated with ethanol. The DNA pellet was dissolved in 400  $\mu$ l of sterile distilled water. DNA concentration and quality were analyzed by the nanodrop 2000.

### SNP selection and genotyping

We selected and genotyped 9 SNPs that had been associated with increased risk of breast cancer in GWAS studies. This includes rs1219648 and rs2981582 in *FGFR2*.

**Table 1** Description of the study population

Characteristics	Cases (%)
Sample size	640
Age range (year)	23–91
Age at diagnosis (mean $\pm$ SD)	47.9 $\pm$ 10
Menopausal status	
Pre-menopausal	304 (53.1)
Post-menopausal	269 (46.9)
Tumor size	
T1–T2	421 (71)
T3–T4	172 (29)
Lymph node involvement	
Positive	367 (61)
Negative	235 (39)
Histological grade	
SBR1-2	354 (62.9)
SBR3	209 (37.1)
Estrogen receptor (ER) status	
Positive	202 (51.1)
Negative	193 (48.9)
Progesterone receptor (PR) status	
Positive	167 (42.1)
Negative	230 (57.9)
ER/PR status	
ER+/PR+	143 (36.2)
ER+/PR–	59 (14.9)
ER–/PR+	22 (5.6)
ER–/PR–	171 (43.3)

rs8051542, rs12443621 and rs3803662 in *TNRC9*, rs889312 in *MAP3K1*, rs3817198 in *LSPI*, rs13387042 in 2q35 and rs13281615 in 8q24.

Genotyping was performed using the TaqMan<sup>®</sup> SNP genotyping assays. PCR mixture was as follows: 12.5  $\mu$ l of TaqMan<sup>®</sup> 2 $\times$  Universal PCR Master Mix, 5–25 ng of DNA, 0.625  $\mu$ l of predesigned TaqMan<sup>®</sup> SNP Genotyping Assay mix (40 $\times$ ) and water to bring the final reaction volume to 25  $\mu$ l. The PCR thermal cycling was as follows: initial denaturing at 95  $^{\circ}$ C for 10 min; 40 cycles of 92  $^{\circ}$ C for 15 s and 60  $^{\circ}$ C for 1 min. Thermal cycling was performed using the Applied Biosystems 7500 Real-Time PCR System. All reactions were carried out with no template as negative controls. Genotype call success rate for cases and for controls was 97.6 and 97.8 %, respectively. Randomly selected DNA sequencing and PCR replication with a coincidence rate greater than 99 % verified genotype reproducibility.

#### Statistical analyses

The genotype and allele frequencies of the 9 SNPs that had been associated with increased risk of breast cancer in

GWAS were tested for the Hardy–Weinberg equilibrium for both patient and control groups using the Chi-square test. According to the general genotype model, risk association between the genotypes and breast cancer susceptibility and tumors characteristics was estimated by crude odds ratio (OR) and 95 % confidence intervals (95 % CI) using the unconditional logistic regression analysis with the low-risk genotype as a reference category [21, 22]. A *P* value of less than 0.05 was required for statistical significance.

Clinical pathological parameters were dichotomised as follows: nodal status ( $\geq 1$  vs. no positive lymph node), SBR (Scarff, Bloom and Richardson) tumor grade (1–2 vs. 3), clinical tumor size (T<sub>1</sub>–T<sub>2</sub> vs. T<sub>3</sub>–T<sub>4</sub>) and estrogen/progesterone receptor status (positive vs. negative).

The statistical analysis was performed using the Epi-Info statistical program (version 5.01; Centers for disease Control and Epidemiology Program office, Atlanta Georgia, USA). Breast cancer-specific overall survival (OVS) was defined as the time from the date of diagnosis to death if the patient died from breast cancer or to last contact. Six-year survival rates were estimated, and survival curves were plotted according to Kaplan and Meier [23]. The differences between groups were calculated by the log-rank test [24]. Univariate analyses for each SNP were carried out by estimating Kaplan–Meier survival curves stratified by genotypes using SEM-STATISTIQUES software (Centre Jean Perrin, Clermont-Ferrand, France).

#### Results

GWAS-identified loci as risk factors for breast cancer in Tunisians

Minor allele frequencies and estimates for the association between the nine SNPs and overall breast cancer risk are shown in Table 2. Genotype frequencies in cases and controls appear in Supplementary Table 1. Genotype distribution of all SNPs in both patients and controls did not deviate from Hardy–Weinberg equilibrium (*P* > 0.05) (Supplementary Table 1).

In this study, 5 out of the 9 breast cancer-associated SNPs discovered in GWAS were replicated in the Tunisian population. Both polymorphisms of *FGFR2* (rs1219648, rs2981582), *TNRC9* rs8051542, *MAP3K1* rs889312, and the SNP located on 8q24 (rs13281615) were statistically significantly associated with breast cancer risk at *P* less than 0.05 (Table 2). A suggestive association was observed between rs3817198 in *LSPI* gene and breast cancer risk with an increased risk for GG genotype and G allele (OR = 1.45, *P* = 0.1; OR = 1.19, *P* = 0.08, respectively). However, no significant evidence was observed for associations between the two other SNPs in *TNRC9* and the SNP located on 2q35

**Table 2** Association of nine SNPs identified from previous GWAS with breast cancer risk in Tunisian women

SNP	Gene/ locus	Allele <sup>a</sup>	Risk allele Frequency (%)		HetOR (95 % CI)	P value	HomOR (95 % CI)	P value	Per risk allele (95 % CI)	P value
			Case	Control						
rs1219648	FGFR2	A/G	45.7	38.2	1.60 (1.19–2.16)	0.001	1.70 (1.15–2.51)	0.004	1.36 (1.13–1.65)	0.001
rs2981582	FGFR2	G/A	50.2	39.4	2.00 (1.47–2.73)	0.000004	2.23 (1.51–3.29)	0.00001	1.55 (1.28–1.88)	0.000003
rs8051542	TNRC9 <sup>b</sup>	C/T	44.5	36.4	1.15 (0.86–1.54)	0.32	2.11 (1.39–3.19)	0.0001	1.40 (1.16–1.69)	0.0004
rs12443621	TNRC9	G/A	46.6	48.2	0.86 (0.63–1.18)	0.34	0.89 (0.61–1.30)	0.53	0.94 (0.78–1.13)	0.49
rs3803662	TNRC9	C/T	45.9	43.5	1.12 (0.83–1.52)	0.45	1.19 (0.82–1.72)	0.34	1.10 (0.91–1.33)	0.3
rs889312	MAP3K1	A/C	35.1	28.9	1.37 (1.03–1.81)	0.02	1.74 (1.08–2.82)	0.01	1.33 (1.09–1.63)	0.003
rs3817198	LSP1	T/C	32.1	28.4	1.15 (0.87–1.52)	0.3	1.45 (0.90–2.34)	0.1	1.19 (0.97–1.46)	0.08
rs13387042	2q35	A/G	41.6	44.8	0.92 (0.68–1.25)	0.5	0.76 (0.52–1.12)	0.15	0.88 (0.73–1.06)	0.16
rs13281615	8q24	A/G	55.2	50.4	1.19 (0.85–1.66)	0.3	1.44 (0.99–2.10)	0.04	1.21 (1.00–1.46)	0.03

CI confidence interval, OR odds ratio, HetOR odds ratio in heterozygote, HomOR odds ratio in homozygote for risk allele (relative to homozygote for non-risk allele)

<sup>a</sup> Reference allele/risk allele

<sup>b</sup> Also known as TOX3

and breast cancer risk ( $P = 0.49$ ,  $0.3$  and  $0.16$  for rs12443621, rs3803662, and rs13387042, respectively).

Most significant associations were of high magnitude. The strongest associations were found for rs2981582 in the *FGFR2* gene and rs8051542 in the *TNRC9* gene. Homozygous variant genotypes of rs2981582 and rs8051542 were associated with over a two-fold increased risk of breast cancer (OR = 2.23,  $P = 0.00001$ ; OR = 2.11,  $P = 0.0001$ , respectively).

#### Effects of GWAS risk loci on disease characteristics of breast cancer

In this study, we also analyzed the effects of the 9 GWAS-identified SNPs on a series of disease clinico-pathological characteristics, including clinical tumor size, SBR tumor grade, lymph node metastasis, distant metastasis, and estrogen/progesterone receptor status.

Both polymorphisms of the *FGFR2* gene and rs8051542 in the *TNRC9* gene were found to be associated with either small or large clinical tumor size. However, slightly stronger associations were found for *FGFR2* rs2981582 and *TNRC9* rs8051542 with T1–T2 tumor size ( $P = 0.0001$ ,  $P = 0.0002$ , respectively) (Supplementary Table 2). Variant allele of rs889312 in the *MAP3K1* gene seems to be associated with increased risk of both small and large tumors (OR = 1.70,  $P = 0.03$ ; OR = 1.89,  $P = 0.04$ ) (Supplementary Table 2). However, homozygous variant genotype of rs13281615 on 8q24 was found to be associated with increased risk of small tumor size (OR = 1.63,  $P = 0.01$ ).

For homozygous variant genotypes of both polymorphisms in the *FGFR2* gene, the associations were also

stronger with low than with high-grade SBR (OR = 2.01,  $P = 0.0008$  for rs1219648; OR = 2.53,  $P = 0.00001$  for rs2981582) (Supplementary Table 3). However, stronger association with high-grade SBR was found for rs8051542 in the *TNRC9* gene (OR = 2.54,  $P = 0.0002$ ) (Supplementary Table 3).

Associations of both polymorphisms in the *FGFR2* gene (rs1219648 and rs2981582) with breast cancer risk were stronger for patients with negative than with positive nodal involvement. The strongest association was found with rs2981582. Homozygous variant genotype was associated with over threefold increased risk of lymph node negative breast cancer (OR = 3.33,  $P = 0.000006$ ) (Table 3). For both polymorphisms, increased breast cancer risk with negative nodal involvement was associated with the minor allele in a dose-dependant manner. Moreover, among cases, rs2981582 AA genotype was more frequent in lymph node negative compared to lymph node positive breast cancers (28.5 vs. 21.9 %; OR = 0.56,  $P = 0.018$ ) (Table 3). The association of rs8051542 in the *TNRC9* gene with breast cancer risk tended to be slightly stronger for patients with positive nodal involvement (OR = 2.15,  $P = 0.0004$ ) (Table 3). However, rs889312 in the *MAP3K1* gene was equally associated with lymph node negative and positive breast cancer (Table 3).

Associations by ER and PR tumor status revealed also some findings. Stratification of tumors by ER status indicated that rs2981582 *FGFR2* polymorphism increased risk of both ER+ and ER– tumors. Slightly stronger association was observed with ER+ ( $P = 0.001$ ,  $P = 0.01$  for ER+ and ER–, respectively) (Supplementary Table 4). Moreover, the minor allele was associated with increased risk of ER+ tumors in a dose-dependant manner (OR = 1.57,  $P = 0.02$ ;

**Table 3** Associations between GWAS breast cancer loci and lymph node involvement

SNP	Controls	No cases	Yes cases	No vs. Yes OR (95 %CI)	<i>P</i> value	Controls vs. No OR (95 %CI)	<i>P</i> value	Controls vs. Yes OR (95 %CI)	<i>P</i> value
rs1219648									
AA	146	62	111	1		1		1	
AG	153	113	183	0.90 (0.60–1.36)	0.6	1.74 (1.16–2.60)	0.004	1.57 (1.12–2.21)	0.006
GG	61	57	70	0.69 (0.42–1.13)	0.1	2.20 (1.34–3.61)	0.0008	1.51 (0.97–2.36)	0.05
G allele				0.83 (0.65–1.06)	0.1	1.55 (1.22–1.98)	0.0002	1.29 (1.04–1.60)	0.01
rs2981582									
GG	140	44	94	1		1		1	
AG	154	124	191	0.72 (0.46–1.12)	0.1	2.56 (1.66–3.96)	0.000006*	1.85 (1.30–2.62)	0.0003
AA	64	67	80	0.56 (0.33–0.93)	0.01	3.33 (2.00–5.56)	0.0000006*	1.86 (1.20–2.90)	0.003
A allele				0.76 (0.60–0.97)	0.02	1.87 (1.47–2.39)	0.0000002*	1.43 (1.15–1.77)	0.0008
rs8051542									
CC	146	73	127	1		1		1	
CT	176	115	150	0.75 (0.51–1.11)	0.1	1.31 (0.89–1.92)	0.1	0.98 (0.70–1.37)	0.9
TT	46	46	86	1.07 (0.66–1.75)	0.75	2.00 (1.18–3.39)	0.005	2.15 (1.37–3.38)	0.0004
T allele				1.00 (0.79–1.28)	0.9	1.38 (1.09–1.77)	0.006	1.39 (1.12–1.73)	0.001
rs12443621									
GG	98	69	110	1		1		1	
AG	180	117	164	0.88 (0.59–1.31)	0.5	0.92 (0.62–1.38)	0.6	0.81 (0.57–1.16)	0.2
AA	85	48	92	1.20 (0.74–1.96)	0.43	0.80 (0.49–1.32)	0.3	0.96 (0.63–1.47)	0.8
rs3803662									
CC	126	76	114	1		1		1	
TC	165	102	169	1.10 (0.74–1.65)	0.6	1.02 (0.69–1.52)	0.8	1.13 (0.80–1.60)	0.4
TT	78	57	84	0.98 (0.61–1.57)	0.9	1.21 (0.76–1.94)	0.3	1.19 (0.78–1.81)	0.3
rs889312									
AA	187	106	146	1		1		1	
AC	151	97	178	1.33 (0.92–1.92)	0.1	1.13 (0.79–1.63)	0.4	1.51 (1.10–2.08)	0.008
CC	31	32	42	0.95 (0.55–1.66)	0.8	1.82 (1.02–3.27)	0.03	1.74 (1.01–2.99)	0.03
C allele				1.07 (0.83–1.37)	0.5	1.28 (0.99–1.66)	0.04	1.37 (1.10–1.72)	0.004
rs3817198									
TT	192	111	172	1		1		1	
CT	147	96	157	1.06 (0.73–1.52)	0.7	1.13 (0.79–1.62)	0.4	1.19 (0.87–1.64)	0.2
CC	32	28	38	0.88 (0.49–1.56)	0.6	1.51 (0.83–2.74)	0.1	1.33 (0.77–2.29)	0.2
rs13387042									
AA	115	84	118	1		1		1	
AG	175	107	183	1.22 (0.83–1.79)	0.2	0.84 (0.57–1.23)	0.3	1.02 (0.72–1.44)	0.9
GG	77	44	65	1.05 (0.64–1.74)	0.8	0.78 (0.48–1.28)	0.3	0.82 (0.53–1.28)	0.3
rs13281615									
AA	93	45	81	1		1		1	
AG	176	116	165	0.79 (0.50–1.25)	0.2	1.36 (0.87–2.13)	0.1	1.08 (0.73–1.58)	0.6
GG	96	74	120	0.90 (0.55–1.47)	0.6	1.59 (0.97–2.61)	0.05	1.44 (0.94–2.19)	0.07

\**P* value < 0.0001 (cut-off for correction)

OR = 2.15, *P* = 0.001, for heterozygous and homozygous variant genotypes, respectively) (Supplementary Table 4). On the other hand, homozygous variant genotype of rs8051542 in the *TNRC9* gene was found to be associated with an increased risk of both ER+ and ER– tumors with the same extent (OR = 2.38, *P* = 0.0005, OR = 2.35,

*P* = 0.0007 for ER+ and ER–, respectively) (Supplementary Table 4). *FGFR2* rs2981582 and *TNRC9* rs8051542 polymorphisms were also associated with both PR+ and PR– tumors, while rs889312 in the *MAP3K1* gene was only associated with PR+ tumors (OR = 2.07, *P* = 0.01) (Supplementary Table 5). The rs13281615 SNP on 8q24 was

found to be associated with ER+ and PR– tumors (Supplementary Table 4 and 5).

Regarding disease progression, it was noted that *FGFR2* rs2981582 and rs1219648, *TNRC9* rs8051542, and *MAP3K1* rs889312 polymorphisms were associated with increased

risk of distant metastasis development (Table 4). The strongest association was found with the minor allele of *FGFR2* rs2981582 in a dose-dependant manner (OR = 2.30,  $P = 0.004$ , OR = 3.57,  $P = 0.00006$ ) (Table 4). For rs13387042 on 2q35, GG genotype was more frequent in

**Table 4** Associations between GWAS breast cancer loci and distant metastasis

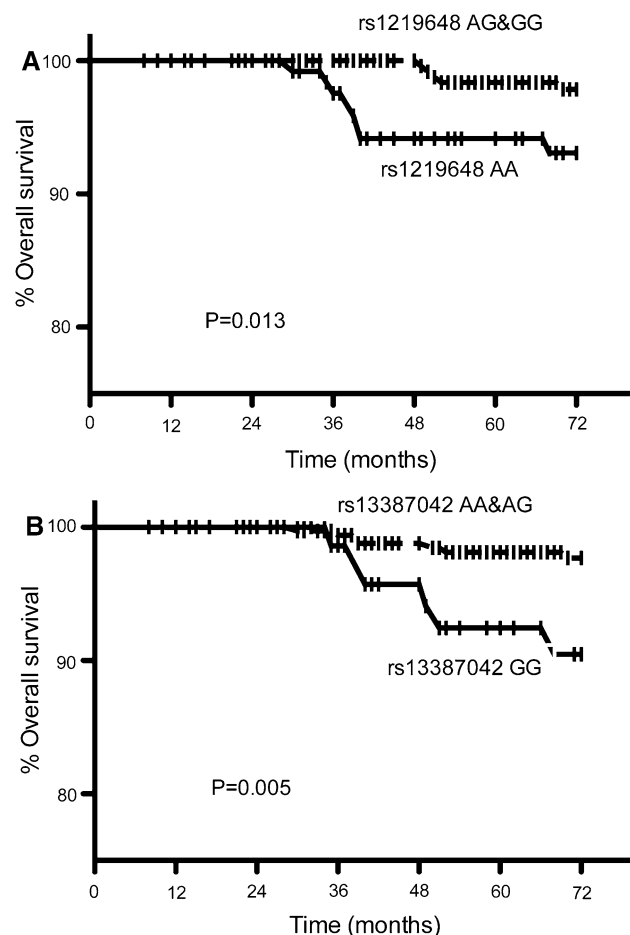
SNP	Controls	No cases	Yes cases	No vs. Yes OR (95 %CI)	<i>P</i> value	Controls vs. No OR (95 %CI)	<i>P</i> value	Controls vs. Yes OR (95 %CI)	<i>P</i> value
rs1219648									
AA	146	154	30	1		1		1	
AG	153	266	44	0.85 (0.50–1.45)	0.5	1.65 (1.21–2.25)	0.001	1.40 (0.81–2.42)	0.2
GG	61	106	24	1.16 (0.62–2.18)	0.6	1.65 (1.10–2.47)	0.01	1.91 (0.99–3.70)	0.03
G allele				1.06 (0.77–1.46)	0.6	1.35 (1.11–1.64)	0.002	1.43 (1.03–1.99)	0.02
rs2981582									
GG	140	129	19	1		1		1	
AG	154	281	48	1.16 (0.63–2.14)	0.6	1.98 (1.44–2.73)	0.00001*	2.30 (1.24–4.27)	0.004
AA	64	120	31	1.75 (0.90–3.43)	0.07	2.03 (1.36–3.05)	0.0002	3.57 (1.79–7.15)	0.00006*
A allele				1.32 (0.96–1.82)	0.07	1.49 (1.22–1.81)	0.00004*	1.97 (1.41–2.74)	0.00002*
rs8051542									
CC	146	169	37	1		1		1	
CT	176	251	32	0.58 (0.34–1.00)	0.03	1.23 (0.91–1.67)	0.1	0.72 (0.41–1.25)	0.2
TT	46	108	28	1.18 (0.66–2.12)	0.5	2.03 (1.32–3.12)	0.0006	2.40 (1.27–4.53)	0.003
T allele				1.05 (0.76–1.44)	0.7	1.38 (1.14–1.69)	0.0009	1.45 (1.04–2.02)	0.02
rs12443621									
GG	98	150	37	1		1		1	
AG	180	256	39	0.62 (0.37–1.04)	0.05	0.93 (0.67–1.29)	0.6	0.57 (0.33–0.99)	0.03
AA	85	123	23	0.76 (0.41–1.39)	0.3	0.95 (0.64–1.40)	0.7	0.72 (0.38–1.35)	0.2
rs3803662									
CC	126	172	27	1		1		1	
TC	165	238	48	1.28 (0.75–2.21)	0.3	1.06 (0.77–1.45)	0.7	1.36 (0.78–2.38)	0.2
TT	78	121	24	1.26 (0.67–2.39)	0.4	1.14 (0.78–1.67)	0.4	1.44 (0.74–2.79)	0.2
rs889312									
AA	187	232	32	1		1		1	
AC	151	237	54	1.65 (1.00–2.73)	0.03	1.27 (0.95–1.69)	0.1	2.09 (1.25–3.50)	0.002
CC	31	61	13	1.55 (0.72–3.28)	0.2	1.59 (0.96–2.62)	0.05	2.45 (1.08–5.50)	0.01
C allele				1.32 (0.96–1.83)	0.07	1.26 (1.02–1.56)	0.02	1.67 (1.19–2.34)	0.001
rs3817198									
TT	192	253	45	1		1		1	
CT	147	220	45	1.15 (0.72–1.85)	0.5	1.14 (0.85–1.52)	0.3	1.31 (0.80–2.14)	0.2
CC	32	58	9	0.87 (0.37–1.98)	0.7	1.38 (0.84–2.26)	0.1	1.20 (0.49–2.85)	0.6
rs13387042									
AA	115	185	28	1		1		1	
AG	175	260	46	1.17 (0.68–2.00)	0.5	0.92 (0.68–1.26)	0.6	1.08 (0.62–1.89)	0.7
GG	77	85	25	1.94 (1.03–3.68)	0.02	0.69 (0.46–1.03)	0.05	1.33 (0.69–2.57)	0.3
G allele				1.38 (1.01–1.89)	0.03	0.84 (0.69–1.02)	0.07	1.16 (0.84–1.61)	0.3
rs13281615									
AA	93	106	27	1		1		1	
AG	176	253	44	0.68 (0.39–1.20)	0.1	1.26 (0.89–1.79)	0.1	0.86 (0.48–1.53)	0.5
GG	96	171	28	0.64 (0.35–1.20)	0.1	1.56 (1.06–2.31)	0.01	1.00 (0.53–1.91)	0.9

\*  $P$  value < 0.0001 (cut-off for correction)

patients developing distant metastasis compared to patients without distant metastasis (25.2 vs. 16 %) (OR = 1.94,  $P = 0.02$ ).

#### Effects of GWAS risk loci on survival from breast cancer

In this study, we also assessed the effect of GWAS risk loci on the OVS of patients. A significant difference was observed between the OVS Kaplan–Meier survival curves for rs1219648 in the *FGFR2* gene. As shown in Fig. 1a, the breast cancer-specific OVS rate was significantly higher among patients carrying the G variant allele. The OVS rates in the group of patients with or without the rs1219648 G allele were 98.1 versus 92 %, respectively (log-rank test,  $P = 0.013$ ). However, no significant difference between the OVS Kaplan–Meier survival curves was observed for rs2981582. In addition, significant difference between OVS Kaplan–Meier curves was observed for rs13387042 on



**Fig. 1** The 6-year breast cancer-specific overall survival of 640 breast cancer patients stratified by genotype. Overall survival of 640 patients according to the presence or absence of **a** rs1219648-G allele and **b** rs13387042-A allele ( $P$  denotes the log-rank test value)

2q35. The OVS rate was significantly lower in the group of patients without rs13387042 A allele compared to patients with rs13387042 A allele (89.9 vs. 97.7 % respectively; log-rank test,  $P = 0.005$ ) (Fig. 1b).

#### Discussion

GWAS have led to the identification of multiple new genetic variants associated with breast cancer risk. Most of these breast cancer GWAS and replication studies have been conducted in European populations [8–10, 12, 14–16, 18] and to a lesser extent in Asians [13, 17, 19]. However, there are significant differences in allele frequencies and the prevalence of breast cancer among different populations. It is, therefore, important to explore the effects of the GWAS-identified markers in other ethnic populations, including women of Arab ancestry. Thus, we carried out this study to estimate the allele frequencies of 9 GWAS-identified loci in the Tunisian population and to investigate, with a case–control study, the potential association of these loci with the risk of breast cancer among Tunisian women.

The 10q26 (*FGFR2*) locus was discovered in two GWAS among women of European descent [8, 9] and the index SNPs rs1219648 and rs2981582 have since been consistently replicated in European [12, 25] and Chinese populations [26] as well as in several ethnic groups including Hispanic and non-Hispanic white women [27] and African American women [28, 29]. In this study, it was also confirmed that *FGFR2* rs1219648 and rs2981582 were significantly associated with increased breast cancer risk in the Tunisian population, which strengthens the conclusion that this locus plays an important role in the development of breast cancer. *FGFR2* is a member of the receptor tyrosine kinase family, involved in mammary gland proliferation and development [30, 31]. It has been shown that *FGFR2* can transform normal human mammary epithelial cells and is over-expressed in breast tumors [32]. The two polymorphisms in the *FGFR2* gene were originally identified by Hunter et al. [9] and were associated with risk of sporadic post-menopausal breast cancer Slattery et al. [27] reported similar findings for post-menopausal Hispanic women. In this study, we have not evaluated the risk of breast cancer according to the menopausal status. However, although 53.1 % of our cases were premenopausal, we found strong associations between both *FGFR2* polymorphisms and breast cancer risk. This may suggest that in the Tunisian population, SNPs rs1219648 and rs2981582 may confer similar effects in both pre and post-menopausal women.

Of the 3 SNPs evaluated in the *TNRC9* locus, only rs8051542 replicated breast cancer risk among Tunisian women. No associations were found with rs3803662 and rs12443621. The SNP rs3803662 was identified as breast



cancer susceptibility variant in two GWAS, both conducted in European populations [8, 10]. This SNP remained the strongest signal for the 16q12 region in further studies of European ancestry and has also been confirmed from deep sequencing study as a key *TNRC9* SNP associated with breast cancer [16, 33]. Zheng et al. [34] found that rs3803662 was associated with breast cancer risk in a Southern Chinese population. However, the association with *TNRC9* rs3803662 was not confirmed in other ethnic groups including Hispanic [27] and African American women [28, 29, 33, 35]. In this study, we also showed the lack of association of rs12443621 with risk of breast cancer among Tunisian women. SNP rs12443621 was identified to increase breast cancer risk by Easton et al. [8] and was found to be in strong linkage disequilibrium with SNP rs3803662 of the *TNRC9* gene. In addition, at the 16q12 locus, the LD pattern between rs3803662 and rs3104793 was also different across populations [36]. Taken together, LD pattern difference across populations may explain the discrepancy between these studies. Thus, a fine-mapping study might be an effective approach to identify the causal variant(s) in the 16q12 locus in Arab women.

In this study, we also found the SNP rs889312 in the *MAP3K1* gene and rs13281615 on 8q24 to be associated with an increased risk of breast cancer in the Tunisian population. The rs889312 SNP in the *MAP3K1* gene was identified by Easton et al. and has been shown to be involved in a potential key pathway for breast cancer [8, 9, 14]. The rs13281615 variant lies in a non-genetic region of chromosome 8q24. Other independent variants in the region 8q24 have been associated with the risk of prostate, colorectal and ovarian cancer [37–39].

Associations with most of the susceptibility loci identified to date are evidently stronger for ER+ than for ER– disease. The strongest evidence is for a variant in *FGFR2* that was primarily associated with ER+ disease [12]. Garcia-Closas et al. [40] also confirmed a stronger association between *FGFR2* rs2981582 and ER+ tumors. Similarly, *FGFR2* rs1219648 and rs2981582 genotypes were significantly associated with breast cancer in European-American only in ER+ and PR+ tumors [25]. Findings of Slattery et al. [27] suggest that *FGFR2* polymorphisms decrease the likelihood of ER–/PR– tumors among non-Hispanic white women, while increasing the likelihood of ER+/PR+ among Hispanic women. Our data showed that *FGFR2* rs2981582 and *TNRC9* rs8051542 were strongly associated with both positive and negative tumor status of ER and PR receptors.

*MAP3K1* variants were found to be relevant in ER+ and PR+ tumors to greater degree than in ER negative or PR negative tumors [40]. Moreover, *MAP3K1* rs889312 variant genotype was associated with larger tumors in Asians but not in European populations, and less likely to be

associated with lymph node positive at breast cancer diagnosis in a Dutch population [40, 41]. Rebbeck et al. [25] showed that the same SNP was associated with breast cancer in African-American women, but again limited only to ER+, PR+ tumors. Our data showed that homozygous variant genotype of *MAP3K1* rs889312 was associated with PR+ but not with PR– tumors, and while not reaching significance, was more likely to be associated with ER– tumors but not with ER+ tumors. Moreover, rs889312 homozygous variant genotype seems to be associated with an increased risk of both small and large tumors and was equally associated with negative and positive nodal involvement. Taken together, associations between *MAP3K1* rs889312 and breast cancer characteristics need to be further explored in other ethnic groups.

Regarding disease progression, we found that rs13387042 homozygous variant genotype was associated with distant metastasis. This finding suggested that rs13387042 variant allele might affect the progression of breast cancer. The rs13387042 variant lies in a non-genic region of chromosome 2q35. Thus, functional studies in this region are likely to lead to a better understanding of mechanisms of carcinogenesis and progression of breast cancer.

Two out of the nine SNPs included in this study had a significant association with the OVS. A higher rate of survival was observed in patients carrying the variant allele of rs1219648 in the *FGFR2* gene. Conversely, a lower OVS rate was found in patients carrying homozygous variant genotype of rs13387042 on 2q35. In the Tunisian population, homozygous variant genotype of rs1219648 was associated with ER+ but not with ER– tumors. The good prognosis known for ER+ tumors could explain the association found between rs1219648 variant of the *FGFR2* gene and the high rate of survival.

The present association study in the Tunisian population highlighted genetic susceptibility patterns different from that reported for other populations. These differences could stem from disparities in genetic background, including differences in allele frequencies and LD pattern, and gene–environment interaction. Moreover, clinical and biological differences in breast cancer have been found in Arab women compared to Europeans [42]. Early disease onset and aggressive forms of breast cancer are seen more frequently in Arab populations [42]. Breast cancer tumors in Arab populations are frequently characterized by large tumor size, high histo-pronostic SBR grade and molecular luminal subtype B [42]. In addition, apart differences in biological characteristics and genetic background, Arab populations greatly differ from European and Asian populations by lifestyle, reproductive behavior, family history status, and environmental exposure, suggesting that risk factors associated with breast cancer development and progression might be different in these different

populations. The high proportion of young-onset and poor prognosis of breast cancer in women of Arab ancestry is probably due to a correspondingly high prevalence of pertinent genetic risk factors that may be uniquely associated with these populations. Thus, new GWAS in women of Arab ancestry may promise to reveal new causal variants and are needed to fully uncover the genetic basis for breast cancer susceptibility in Arab population.

In conclusion, the present association study in the Tunisian population revealed several implications of the 9 SNPs that had been associated with increased risk of breast cancer in GWAS. It reinforces the need to replicate the GWAS discovered variants across different populations and ethnicities.

**Acknowledgments** This publication was made possible by a grant from the Qatar National Research Fund (NPRP08-083-3-031). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Qatar National Research Fund. This work was also supported by the Ministry of Higher Education and Scientific Research and the Ministry of Public Health of the Republic of Tunisia.

**Conflict of interest** The authors declare no conflict of interest.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

- Curado MP, Edwards B, Shin HR (2007) Cancer incidence in five continents, vol IX. IARC Scientific Publications, Lyon
- Ferlay J, Parkin DM, Steliarova-Foucher E (2010) GLOBOCAN 2008. Cancer incidence and mortality worldwide IARC Cancer Base No. 10. International Agency for Research on Cancer, Lyon. <http://globocan.iarc.fr>
- El Saghir NS, Khalil MK, Eid T, El Kinge AR, Charafeddine M, Geara F, Seoud M, Shamseddine AI (2007) Trends in epidemiology and management of breast cancer in developing Arab countries: a literature and registry analysis. *Int J Surg* 5:225–233
- Missaoui N, Jaidene L, Abdelkrim SB, Abdelkader AB, Beizig N, Yaacoub LB, Yaacoubi MT, Hmissa S (2011) Breast cancer in Tunisia: clinical and pathological findings. *Asian Pac J Cancer Prev* 12:169–172
- Antoniou A, Pharoah PD, Narod S, Risch HA, Eyfjord JE, Hopper JL, Loman N, Olsson H, Johannsson O, Borg A et al (2003) Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies. *Am J Hum Genet* 72:1117–1130
- Mahfoudh W, Bouaouina N, Ahmed SB, Gabbouj S, Shan J, Mathew R, Uhrhammer N, Bignon YJ, Troudi W, Elgaaid AB, Hassen E, Chouchane L (2012) Hereditary breast cancer in Middle Eastern and North African (MENA) populations: identification of novel, recurrent and founder BRCA1 mutations in the Tunisian population. *Mol Biol Rep* 39:1037–1046
- Pharoah PD, Antoniou A, Bobrow M, Zimmern RL, Easton DF, Ponder BA (2002) Polygenic susceptibility to breast cancer and implications for prevention. *Nat Genet* 31:33–36
- Easton DF, Pooley KA, Dunning AM, Pharoah PD, Thompson D, Ballinger DG, Struwing JP, Morrison J, Field H, Luben R et al (2007) Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 447:1087–1093
- Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, Hankinson SE, Wacholder S, Wang Z, Welch R, Hutchinson A (2007) A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet* 39:870–874
- Stacey SN, Manolescu A, Sulem P, Rafnar T, Gudmundsson J, Gudjonsson SA, Masson G, Jakobsdottir M, Thorlacius S, Helgason A et al (2007) Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat Genet* 39:865–869
- Gold B, Kirchhoff T, Stefanov S, Lautenberger J, Viale A, Garber J, Friedman E, Narod S, Olshen AB, Gregersen P et al (2008) Genome-wide association study provides evidence for a breast cancer risk locus at 6q22.33. *Proc Natl Acad Sci USA* 105:4340–4345
- Stacey SN, Manolescu A, Sulem P, Thorlacius S, Gudjonsson SA, Jonsson GF, Jakobsdottir M, Bergthorsson JT, Gudmundsson J, Aben KK et al (2008) Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat Genet* 40:703–706
- Zheng W, Long J, Gao YT, Li C, Zheng Y, Xiang YB, Wen W, Levy S, Deming SL, Haines JL (2009) Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. *Nat Genet* 41:324–328
- Thomas G, Jacobs KB, Kraft P, Yeager M, Wacholder S, Cox DG, Hankinson SE, Hutchinson A, Wang Z, Yu K (2009) A multistage genome-wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (RAD51L1). *Nat Genet* 41:579–584
- Ahmed S, Thomas G, Ghousaini M, Healey CS, Humphreys MK, Platte R, Morrison J, Maranian M, Pooley KA, Luben R et al (2009) Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2. *Nat Genet* 41:585–590
- Turnbull C, Ahmed S, Morrison J, Pernet D, Renwick A (2010) Genome-wide association study identifies five new breast cancer susceptibility loci. *Nat Genet* 42:504–507
- Long J, Cai Q, Shu XO, Qu S, Li C, Zheng Y, Gu K, Wang W, Xiang YB, Cheng J (2010) Identification of a functional genetic variant at 16q12.1 for breast cancer risk: results from the Asia Breast Cancer Consortium. *PLoS Genet* 6:e1001002
- Fletcher O, Johnson N, Orr N, Hosking FJ, Gibson LJ, Walker K, Zelenika D, Gut I, Heath S (2011) Novel breast cancer susceptibility locus at 9q31.2: results of a genome-wide association study. *J Natl Cancer Inst* 103:425–435
- Long J, Cai Q, Sung H, Shi J, Zhang B, Choi JY, Wen W, Delahanty RJ, Lu W (2012) Genome-wide association study in East Asians identifies novel susceptibility loci for breast cancer. *PLoS Genet* 8:e1002532
- Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucl Acids Res* 16:1215
- Lunetta KL (2008) Genetic association studies. *Circulation* 118:96–101
- Lewis CM (2002) Genetic association studies: design, analysis and interpretation. *Brief Bioinf* 3:146–153
- Kaplan EL, Meir PO (1958) Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 53:457–461
- Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, Mantel N, McPherson K, Peto J, Smith PG (1977) Design

- and analysis of randomized clinical trials requiring prolonged observation of each patient. *Br J Cancer* 35:1–39
25. Rebbeck TR, DeMichele A, Tran TV, Panossian S, Bunin GR, Troxel AB, Strom BL (2009) Hormone-dependent effects of FGFR2 and MAP3K1 in breast cancer susceptibility in a population-based sample of post-menopausal African-American and European-American women. *Carcinogenesis* 30:269–274
  26. Liang J, Chen P, Hu Z, Zhou X, Chen L, Li M, Wang Y, Tang J, Wang H, Shen H (2008) Genetic variants in fibroblast growth factor receptor 2 (FGFR2) contribute to susceptibility of breast cancer in Chinese women. *Carcinogenesis* 29:2341–2346
  27. Slattery ML, Baumgartner KB, Giuliano AR, Byers T, Herrick JS, Wolff RK (2011) Replication of five GWAS-identified loci and breast cancer risk among Hispanic and non-Hispanic white women living in the Southwestern United States. *Breast Cancer Res Treat* 129:531–539
  28. Barnholtz-Sloan JS, Shetty PB, Guan X, Nyante SJ, Luo J, Brennan DJ, Millikan RC (2010) FGFR2 and other loci identified in genome-wide association studies are associated with breast cancer in African-American and younger women. *Carcinogenesis* 31:1417–1423
  29. Zheng W, Cai Q, Signorello LB, Long J, Hargreaves MK, Deming SL, Li G, Li C, Cui Y, Blot WJ (2009) Evaluation of 11 breast cancer susceptibility loci in African-American women. *Cancer Epidemiol Biomarkers Prev* 18:2761–2764
  30. Lu P, Ewald AJ, Martin GR, Werb Z (2008) Genetic mosaic analysis reveals FGF receptor 2 function in terminal end buds during mammary gland branching morphogenesis. *Dev Biol* 321:77–87
  31. Koziczak M, Holbro T, Hynes NE (2004) Blocking of FGFR signaling inhibits breast cancer cell proliferation through down-regulation of D-type cyclins. *Oncogene* 23:3501–3508
  32. Moffa AB, Ethier SP (2007) Differential signal transduction of alternatively spliced FGFR2 variants expressed in human mammary epithelial cells. *J Cell Physiol* 210:720–731
  33. Udler MS, Ahmed S, Healey CS, Meyer K, Struewing J, Maranian M, Kwon EM, Zhang J, Tyrer J (2010) Fine scale mapping of the breast cancer 16q12 locus. *Hum Mol Genet* 19:2507–2515
  34. Zheng W, Wen W, Gao YT, Shyr Y, Zheng Y, Long J, Li G, Li C, Gu K, Cai Q, Shu XO, Lu W (2010) Genetic and clinical predictors for breast cancer risk assessment and stratification among Chinese women. *J Natl Cancer Inst* 102:972–981
  35. Ruiz-Narváez EA, Rosenberg L, Cozier YC, Cupples LA, Adams-Campbell LL, Palmer JR (2010) Polymorphisms in the TOX3/LOC643714 locus and risk of breast cancer in African-American women. *Cancer Epidemiol Biomarkers Prev* 19:1320–1327
  36. Huo D, Zheng Y, Ogundiran TO, Adebamowo C, Nathanson KL, Domchek SM, Rebbeck TR, Simon MS, John EM, Hennis A, Nemesure B, Wu SY, Leske MC, Amb S, Niu Q, Zhang J, Cox NJ, Olopade OI (2012) Evaluation of 19 susceptibility loci of breast cancer in women of African ancestry. *Carcinogenesis* 33:835–840
  37. Gudmundsson J, Sulem P, Manolescu A, Amundadottir LT, Gudbjartsson D, Helgason A, Rafnar T et al (2007) Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. *Nat Genet* 39:631–637
  38. Haiman CA, Le Marchand L, Yamamoto J, Stram DO, Sheng X, Kolonel LN, Wu AH, Reich D, Henderson BE (2007) A common genetic risk factor for colorectal and prostate cancer. *Nat Genet* 39:954–956
  39. Ghousaini M, Song H, Koessler T, Al Olama AA, Kote-Jarai Z, Driver KE, Pooley KA, Ramus SJ et al (2008) Multiple loci with different cancer specificities within the 8q24 gene desert. *J Natl Cancer Inst* 100:962–966
  40. Garcia-Closas M, Hall P, Nevanlinna H, Pooley K, Morrison J, Richesson DA, Bojesen SE, Nordestgaard BG, Axelsson CK, Arias JI, Milne RL, Ribas G (2008) Heterogeneity of breast cancer associations with five susceptibility loci by clinical and pathological characteristics. *PLoS Genet* 4:e1000054
  41. Huijts PE, Vreeswijk MP, Kroeze-Jansema KH, Jacobi CE, Seynaeve C, Krol-Warmerdam EM, Wijers-Koster PM, Blom JC, Pooley KA, Klijn JG, Tollenaar RA, Devilee P, van Asperen CJ (2007) Clinical correlates of low-risk variants in FGFR2, TNRC9, MAP3K1, LSP1 and 8q24 in a Dutch cohort of incident breast cancer cases. *Breast Cancer Res* 9:R78
  42. Chalabi N, Bernard-Gallon DJ, Bignon YJ et al (2008) Comparative clinical and transcriptomal profiles of breast cancer between French and South Mediterranean patients show minor but significant biological differences. *Cancer Genomics Proteomics* 5:253–261