



Colorectal cancer susceptibility variants alter risk of breast cancer in a Chinese Han population

W. Wei^{1*}, M. Jiang^{2*}, L. Luo³, Z. Li¹, P. Wang¹ and W.Q. Dong¹

¹School of Biotechnology, Southern Medical University, Guangzhou, Guangdong Province, China

²Central Laboratory, The Second Affiliated Hospital of Fujian University of Traditional Chinese Medicine, Fujian, China

³Department of Medical Intensive Care Unit, the First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China

*These authors contributed equally to this study.

Corresponding author: W.Q. Dong

E-mail: dongwq63@263.net

Genet. Mol. Res. 12 (4): 6268-6274 (2013)

Received February 18, 2013

Accepted July 31, 2013

Published December 4, 2013

DOI <http://dx.doi.org/10.4238/2013.December.4.14>

ABSTRACT. Recent genome wide association studies (GWAS) and candidate gene studies have revealed many novel loci associated with colorectal cancer susceptibility. We evaluated the effect of these colorectal cancer-associated variants on the risk of breast cancer in a Chinese Han population. Seven single nucleotide polymorphisms (SNPs) (rs3856806 in *PPARG*, rs7014346 in *POU5F1P1*, rs989902 in *PTPN13*, rs1801278 in *IRSI*, rs7003146 in *TCF7L2*, rs1503185 in *PTPRJ*, and rs63750447 in *MLH1*) were genotyped in Han Chinese subjects, including 216 patients with breast cancer and 216 matched controls, using the Sequenom MassARRAY platform. The association of genotypes with susceptibility to breast cancer was analyzed using the odds ratio (OR), with 95% confidence interval (CI) and logistic regression. Three SNPs (rs7014346, rs989902, and rs7003146) were found to be significantly associated

with the susceptibility of breast cancer. The GA and AA genotypes of rs7003146 in *TCF7L2*, and the CA and CC genotype of rs989902 in *PTPN13* were associated with reduced breast cancer risk in the Chinese Han population based on the best-fit dominant model. The GG genotype of rs7014346 in *POU5F1P1* was also significantly associated with decreased breast cancer risk under the best-fit additive model. Our results confirmed the association of rs7014346 in *POU5F1P1*, rs989902 in *PTPN13*, and rs7003146 in *TCF7L2* with variations in the risk of breast cancer in a Chinese Han population.

Key words: Breast cancer; *PTPN13*; *TCF7L2*; *POU5F1P1*

INTRODUCTION

Breast cancer is one of the most common fatal and malignant diseases of women, and its incidence was ranked first in 2011 (Jemal et al., 2011). Although compared with western countries, the breast cancer incidence is currently lower in China, the rate of breast cancer in China is expected to increase substantially, from 10-60 cases per 100,000 women in 2008 to more than 100 cases per 100,000 women aged 55-69 years by 2021 (Linos et al., 2008; Ziegler et al., 2008). Therefore, there is an urgent need to identify the key factors related to breast cancer in Chinese women, and to take effective measures to control this projected increase in breast cancer incidence.

Breast cancer is a complex disease that is caused by multiple genetic and environmental factors. Many candidate genes that may cause breast cancer have been identified in recent years including the high penetrance genes *BRCA1/2* (Ford et al., 1998) and *TP53* (Hemel and Domchek, 2010), the medium penetrance genes *CHEK2* (Meijers-Heijboer et al., 2002), *BRIP1*, and *PALB2* (Vargas et al., 2011), and the low penetrance genes *FGFR2*, *ESR1*, and *TOX3* (Easton et al., 2007). Inherited mutations in these genes predispose individuals to high risks of breast cancer. Lifetime risks of breast cancer can reach up to 80% in individuals with mutations in *BRCA1/2*. Inherited mutations of *CHEK2*, *BRIP1* (Seal et al., 2006), and *PALB2* (Rahman et al., 2007) show 2.2-, 2.0- and 2.3-fold increased risks of breast cancer compared to the normal population, respectively. Numerous genome-wide association studies (GWAS) have also identified approximately 50 new alleles among low penetrance genes such as *FGFR2*, *ESR1*, and *TOX3*, with each locus conferring no more than an odds ratio (OR) of 1.4 for breast cancer. Therefore, it is likely that there are several low penetrance genetic variants contributing to the remaining unexplained excess breast cancer risks, which have not yet been identified.

Some loci regulate carcinogenic pathways that are common to multiple cancers. A study investigating the association between *BRAC1* and the development of colorectal cancer suggested that *BRAC1* was a susceptibility gene for breast cancer as well as colorectal cancer (Garcia-Patino et al., 1998). *SMAD7* was found to be associated with colorectal cancer risk (Pittman et al., 2009), and also acts as an important factor in the development of breast cancer (Scollen et al., 2011; Kim et al., 2012). Candidate gene studies and GWAS have identified a number of novel genetic variants and loci including rs3856806 in *PPARG*, rs7014346 in *POU5F1P1*, rs989902 in *PTPN13*, rs1801278 in *IRS1*, rs7003146 in *TCF7L2*, rs1503185 in *PTPRJ*, and rs63750447 in *MLH1* (Zhang et al., 2004; Pechlivanis et al., 2007; Folsom et al.,

2008; Kury et al., 2008; Tenesa et al., 2008; Mita et al., 2010; Pancione et al., 2010) that are associated with the susceptibility of colorectal cancer. Due to the similarities between the etiology of colorectal and breast cancer, we performed the present study to evaluate the association between these colorectal susceptibility variants and breast cancer risk.

MATERIAL AND METHODS

Study population

Following pathology-based diagnoses, 216 patients with breast cancer were enrolled in the study at the Nanfang Hospital, Southern Medical University, Guangzhou, Guangdong Province, China. A further 216 subjects were randomly selected among cancer-free, healthy individuals from the same geographic region to comprise the control group. All of the subjects involved were Han Chinese females. The mean ages of the patients and control subjects were 47.62 years and 47.46 years, respectively.

Genotyping

Five milliliter peripheral blood samples were collected after informed consent was obtained, and were delivered and stored in a frozen state. Genomic DNA was extracted from 200 μ L peripheral blood samples using a Genomic DNA Purification Kit (Omega, China) according to manufacturer instructions, and was stored at -70°C until use. All of the single nucleotide polymorphisms (SNPs) were genotyped using the SEQUENOM MassARRAY matrix-assisted laser desorption ionization-time of flight mass spectrometry platform (Sequenom, USA). Primers were designed using a semi-automated method (Assay Design3.1, Sequenom). The call rate for each assay was set at $>90\%$.

Statistical analysis

Hardy-Weinberg equilibrium (HWE) was examined using Haploview 4.1. Codominant, dominant, overdominant, and additive genetic models of inheritance were chosen to systematically evaluate the associations between each SNP and breast cancer. Association analysis based on unconditional logistic regression was carried out by calculating the OR and 95% confidence interval (CI) for each SNP in codominant, dominant, overdominant, and additive genetic models; the significance level was set to $P < 0.05$. The statistical tests were implemented in the web-based tool SNPstats (<http://bioinfo.iconcologia.net/SNPstats>) (Sole et al., 2006).

RESULTS

The gene symbol, position, and alleles for each SNP are presented in Table 1. The genotype distributions and allele frequencies of seven SNPs in healthy control subjects and breast cancer patients are summarized in Table 2. The results were adjusted with respect to age and gender, and are shown for each genetic model. All SNPs in the present study conformed to HWE in the controls ($P > 0.1$).

Table 1. Markers genotyped in the current study.

Gene	Polymorphism (SNP)	Location	Alleles
<i>PPARG</i>	rs3856806	3p25	C/T
<i>POU5F1P1</i>	rs7014346	8q24	G/A
<i>PTPN13</i>	rs989902	4q21.3	C/A
<i>IRS1</i>	rs1801278	2q36.3	G/A
<i>TCF7L2</i>	rs7003146	10q25.2	G/A
<i>PTPRJ</i>	rs1503185	11p11.2	C/T
<i>MLH1</i>	rs63750447	3p22.2	A/T

Table 2. The results of 7 SNPs in breast cancer case-control group.

Nearest Gene	SNP_ID	Model	Genotype	Control (N, %)	Case (N, %)	OR ^a (95%CI)	P ^b
<i>PPARG</i>	rs3856806	-	C/C	122 (61%)	115 (73.7%)	1	0.41
			C/T	69 (34.5%)	69 (34.7%)	1.07 (0.70-1.63)	
			T/T	9 (4.5%)	15 (7.5%)	1.79 (0.75-4.26)	
<i>POU5F1P1</i>	rs7014346	-	G/G	122 (81.9%)	175 (95.1%)	1	1.00E-04
			G/A	27 (18.1%)	9 (4.9%)	0.23 (0.11-0.51)	
<i>PTPN13</i>	rs989902	Codominant	A/A	52 (26.8%)	74 (36.8%)	1	0.074
			C/A	107 (55.1%)	90 (44.8%)	0.59 (0.38-0.93)	
			C/C	35 (18%)	37 (18.4%)	0.74 (0.42-1.33)	
		Dominant	A/A	52 (26.8%)	74 (36.8%)	1	0.033
			C/A-C/C	142 (73.2%)	127 (63.2%)	0.63 (0.41-0.97)	
		Overdominant	A/A-C/C	87 (44.9%)	111 (55.2%)	1	0.04
C/A	107 (55.1%)		90 (44.8%)	0.66 (0.44-0.98)			
<i>IRS1</i>	rs1801278	-	G/G	31 (20.9%)	41 (26.3%)	1	0.26
			G/A	117 (79%)	115 (73.7%)	0.74 (0.43-1.26)	
			C/A	107 (55.1%)	90 (44.8%)	0.66 (0.44-0.98)	
<i>TCF7L2</i>	rs7003146	Codominant	G/G	75 (37.5%)	102 (51%)	1	0.021
			G/A	95 (47.5%)	78 (39%)	0.60 (0.40-0.92)	
			A/A	30 (15%)	20 (10%)	0.49 (0.26-0.94)	
		Dominant	G/G	75 (37.5%)	102 (51%)	1	0.0066
			G/A-A/A	125 (62.5%)	98 (49%)	0.58 (0.39-0.86)	
		Log-additive	-	-	-	-	0.67 (0.50-0.90)
<i>PTPRJ</i>	rs1503185		Codominant	C/C	87 (43.9%)	104 (52.3%)	1
C/T		92 (46.5%)		75 (37.7%)	0.68 (0.45-1.04)		
T/T		19 (9.6%)		20 (10.1%)	0.89 (0.44-1.77)		
<i>MLH1</i>	rs63750447	-	T/T	185 (93%)	193 (95.5%)	1	0.27
			T/T	185 (93%)	193 (95.5%)	1	
			A/T	14 (7%)	9 (4.5%)	0.62 (0.26-1.47)	

^aThe corresponding OR is counted by age and gender adjustment. ^bThe P value is counted by the web-based tool SNPstats.

The rs7014346 SNP was found to differ significantly between patients and healthy controls, with OR values of 0.23 (95%CI = 0.11-0.51, P = 0.0001). Compared with the GG genotype, the GA genotype showed a remarkable reduced risk of developing breast cancer. Furthermore, a significant difference was also found in the genotype distributions of *TCF7L2* rs7003146 in three genetic models (codominant, dominant, and log-addictive). The GA and AA genotypes of rs7003146 were significantly associated with a decreased breast cancer risk (OR = 0.60, 95%CI = 0.40-0.92, P = 0.021; OR = 0.58, 95%CI = 0.39-0.86, P = 0.0066; OR = 0.67, 95%CI = 0.50-0.90, P = 0.0069, for the three models respectively). Additionally, rs989902 of *PTPN13* was observed to differ significantly between patients and controls in dominant and overdominant genetic models. Generally, the CA genotype of rs989902 was associated with increased breast cancer risk (OR = 0.63, 95%CI = 0.41-0.97, P = 0.033; OR = 0.66, 95%CI = 0.44-0.98, P = 0.04).

No association of rs3856806 in *PPARG*, rs1801278 in *IRS1*, rs1503185 in *PTPRJ*, or

rs63750447 in *MLHI* and breast cancer risk was observed in the present population ($P = 0.41$, $P = 0.26$, $P = 0.2$, and $P = 0.27$, respectively).

DISCUSSION

Breast and colorectal cancer have similar disease pathogenesis, and it appears that these two malignancies might share some common genetic risks. Therefore, colorectal cancer susceptibility loci might also be associated with the risk of breast cancer. In this study, we investigated the association between seven SNPs, which have been most strongly associated with colorectal cancer risk in colorectal cancer GWAS, and breast cancer risk in a Chinese population.

The results showed a strong association between the rs7014346 polymorphism with susceptibilities to breast cancer in the Chinese Han population, which is consistent with results of a previous study in human breast cancer cells (Wang et al., 2003). The SNP rs7014346, located on chromosome 8q24, is on the POU class 5 homeobox 1 pseudogene 1 gene (*POU5F1P1*). Studies related to the POU homeodomain protein OCT3 in human breast cancer cells have suggested that deregulated expression of POU genes in breast cancer cells could repress the expression of a tumor suppressor and activate the expression of an oncogenic growth factor (Wang et al., 2003). In the present study, the GA genotype of rs7014346 was significantly associated with decreased breast cancer risk, which demonstrated the similar role of *POU5F1P1* in the development of breast cancer. In addition, rs7014346 in *POU5F1P1* was found to be associated with colorectal cancer susceptibility in a GWAS (Tenesa et al., 2008). Therefore, genetic variants of *POU5F1P1* in colorectal cancer are also associated with breast cancer susceptibility.

Furthermore, the SNPs rs989902 and rs7003146 were also found to be associated with the risk of breast cancer. The non-receptor type protein-tyrosine phosphatase 13 gene (*PTPN13*) of rs989902 is located on chromosome 4q21.3. *PTPN13* belongs to the protein tyrosine phosphatase (PTP) gene family, which has been linked to the regulation of cell growth and apoptosis (Andersen et al., 2001), and is considered as an enhancer of cell proliferation that is mediated by several oncogenic protein-tyrosine kinases. Although *PTPN13* has been linked with the development of cancer, association studies between *PTPN13* and breast cancer remain limited. Our present study provides the first evidence demonstrating this relationship. The third SNP that was associated with breast cancer risk, rs7003146, is located on the transcription factor 7-like 2 (*TCF7L2*) gene, which plays a key role in the Wnt signaling pathway, and its variants are reported to be likely involved in many types of cancers (Tang et al., 2008). The results of this SNP with respect to breast cancer are consistent with previous reports.

However, the remaining SNPs evaluated in the *PPARG*, *IRS1*, *PTPRJ*, and *MLHI* genes showed no significant associations with the risk of breast cancer. Among these genes, many studies have already confirmed their association with a variety of human cancers. For example, variations in *PPARG* expression or gene mutations have been associated with tumorigenesis (Sabatino et al., 2005). Transgenic mice overexpressing *IRS1* developed breast cancer (Dearth et al., 2006). Missense mutations of *PTPRJ* were identified in human colon, lung, and breast cancers (Ruivenkamp et al., 2002; 2003). Similarly, Mita et al. (2010) demonstrated that missense polymorphisms of *PTPRJ* genes affected susceptibility to a variety of human cancers. The *MLHI* gene is a member of the mismatch repair gene family. People with a mutation in this gene appear to have an increased chance of developing several types of cancer, such as colorectal, ovary, and stomach, among others. Nevertheless, the present

study indicated no significant association between these genetic variants and breast cancer in a Chinese population.

In summary, our study provides the first evidence that rs7014346 in *POU5F1P1*, rs989902 in *PTPN13*, and rs7003146 in *TCF7L2* appear to be associated with breast cancer susceptibility in Chinese women, whereas another four tested susceptibility variants confirmed in colorectal cancer were not found association with breast cancer. Further work will be needed to determine the functional role that rs7014346 in *POU5F1P1*, rs989902 in *PTPN13*, and rs7003146 in *TCF7L2*, or any correlated variants, play in breast cancer development. Furthermore, due to the relatively small scale of our study, more in-depth studies with larger samples are needed to confirm our results.

REFERENCES

- Andersen JN, Mortensen OH, Peters GH, Drake PG, et al. (2001). Structural and evolutionary relationships among protein tyrosine phosphatase domains. *Mol. Cell. Biol.* 21: 7117-7136.
- Dearth RK, Cui X, Kim HJ, Kuitatse I, et al. (2006). Mammary tumorigenesis and metastasis caused by overexpression of insulin receptor substrate 1 (IRS-1) or IRS-2. *Mol. Cell Biol.* 26: 9302-9314.
- Easton DF, Pooley KA, Dunning AM, Pharoah PD, et al. (2007). Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 447: 1087-1093.
- Folsom AR, Pankow JS, Peacock JM, Bielinski SJ, et al. (2008). Variation in TCF7L2 and increased risk of colon cancer: the Atherosclerosis Risk in Communities (ARIC) Study. *Diabetes Care* 31: 905-909.
- Ford D, Easton DF, Stratton M, Narod S, et al. (1998). Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. *Am. J. Hum. Genet.* 62: 676-689.
- Garcia-Patino E, Gomendio B, Lleonart M, Silva JM, et al. (1998). Loss of heterozygosity in the region including the BRCA1 gene on 17q in colon cancer. *Cancer Genet. Cytogenet.* 104: 119-123.
- Hemel D and Domchek SM (2010). Breast cancer predisposition syndromes. *Hematol. Oncol. Clin. North Am.* 24: 799-814.
- Jemal A, Bray F, Center MM, Ferlay J, et al. (2011). Global cancer statistics. *CA Cancer J. Clin.* 61: 69-90.
- Kim S, Han J, Lee SK, Koo M, et al. (2012). Smad7 acts as a negative regulator of the epidermal growth factor (EGF) signaling pathway in breast cancer cells. *Cancer Lett.* 314: 147-154.
- Kury S, Buecher B, Robiou-du-Pont S, Scoul C, et al. (2008). Low-penetrance alleles predisposing to sporadic colorectal cancers: a French case-controlled genetic association study. *BMC Cancer* 8: 326.
- Linos E, Spanos D, Rosner BA, Linos K, et al. (2008). Effects of reproductive and demographic changes on breast cancer incidence in China: a modeling analysis. *J. Natl. Cancer Inst.* 100: 1352-1360.
- Meijers-Heijboer H, van den Ouweland A, Klijn J, Wasielewski M, et al. (2002). Low-penetrance susceptibility to breast cancer due to CHEK2(*)110delC in noncarriers of BRCA1 or BRCA2 mutations. *Nat. Genet.* 31: 55-59.
- Mita Y, Yasuda Y, Sakai A, Yamamoto H, et al. (2010). Missense polymorphisms of PTPRJ and PTPN13 genes affect susceptibility to a variety of human cancers. *J. Cancer Res. Clin. Oncol.* 136: 249-259.
- Pancione M, Sabatino L, Fucci A, Carafa V, et al. (2010). Epigenetic silencing of peroxisome proliferator-activated receptor gamma is a biomarker for colorectal cancer progression and adverse patients' outcome. *PLoS One* 5: e14229.
- Pechlivanis S, Pardini B, Bermejo JL, Wagner K, et al. (2007). Insulin pathway related genes and risk of colorectal cancer: INSR promoter polymorphism shows a protective effect. *Endocr. Relat. Cancer* 14: 733-740.
- Pittman AM, Naranjo S, Webb E, Broderick P, et al. (2009). The colorectal cancer risk at 18q21 is caused by a novel variant altering SMAD7 expression. *Genome Res.* 19: 987-993.
- Rahman N, Seal S, Thompson D, Kelly P, et al. (2007). PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nat. Genet.* 39: 165-167.
- Ruivenkamp CA, van WT, Zanon C, Stassen AP, et al. (2002). Ptpnj is a candidate for the mouse colon-cancer susceptibility locus Sccl and is frequently deleted in human cancers. *Nat. Genet.* 31: 295-300.
- Ruivenkamp C, Hermsen M, Postma C, Klous A, et al. (2003). LOH of PTPRJ occurs early in colorectal cancer and is associated with chromosomal loss of 18q12-21. *Oncogene* 22: 3472-3474.
- Sabatino L, Casamassimi A, Peluso G, Barone MV, et al. (2005). A novel peroxisome proliferator-activated receptor gamma isoform with dominant negative activity generated by alternative splicing. *J. Biol. Chem.* 280: 26517-26525.
- Scollen S, Luccarini C, Baynes C, Driver K, et al. (2011). TGF-beta signaling pathway and breast cancer susceptibility. *Cancer Epidemiol. Biomarkers Prev.* 20: 1112-1119.

- Seal S, Thompson D, Renwick A, Elliott A, et al. (2006). Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. *Nat. Genet.* 38: 1239-1241.
- Sole X, Guino E, Valls J, Iniesta R, et al. (2006). SNPStats: a web tool for the analysis of association studies. *Bioinformatics* 22: 1928-1929.
- Tang W, Dodge M, Gundapaneni D, Michnoff C, et al. (2008). A genome-wide RNAi screen for Wnt/beta-catenin pathway components identifies unexpected roles for TCF transcription factors in cancer. *Proc. Natl. Acad. Sci. U. S. A.* 105: 9697-9702.
- Tenesa A, Farrington SM, Prendergast JG, Porteous ME, et al. (2008). Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. *Nat. Genet.* 40: 631-637.
- Vargas AC, Reis-Filho JS and Lakhani SR (2011). Phenotype-genotype correlation in familial breast cancer. *J. Mammary. Gland. Biol. Neoplasia.* 16: 27-40.
- Wang P, Branch DR, Bali M, Schultz GA, et al. (2003). The POU homeodomain protein OCT3 as a potential transcriptional activator for fibroblast growth factor-4 (FGF-4) in human breast cancer cells. *Biochem. J.* 375: 199-205.
- Zhang XM, Li JT, Zhu M, Wu XL, et al. (2004). Study on the relationship between genetic polymorphism Val384Asp in hMLH1 gene and the risk of four different carcinomas. *Zhonghua Liu Xing Bing Xue Za Zhi.* 25: 978-981.
- Ziegler RG, Anderson WF and Gail MH (2008). Increasing breast cancer incidence in China: the numbers add up. *J. Natl. Cancer Inst.* 100: 1339-1341.