International Journal of Clinical Oncology (2019) 24:179–188
https://doi.org/10.1007/s10147-018-1346-8

ORIGINAL ARTICLE

Correlations between *CYP3A4* polymorphism and susceptibility to breast cancer in Chinese Han population

Xu Liu[1] · Xi Huang[2] · Shanshan Zhang[3] · Fanglin Niu[3] · Yongri Ouyang[3] · Zhexing Shou[4] · Jikui Liu[1]

Received: 9 July 2018 / Accepted: 6 September 2018 / Published online: 14 September 2018

Abstract

Background CYP3A4 is a major enzyme catalyzing the metabolism of endogenous steroids that play an important role in the etiology of carcinogenesis. This study was designed to investigate the contribution of *CYP3A4* polymorphism to breast cancer in Chinese Han female population.

Methods To examine whether variants of *CYP3A4* contribute to breast cancer, 5 single-nucleotide polymorphisms (SNPs) of *CYP3A4* were genotyped by Sequenom MassARRAY in 267 breast cancer patients and 302 healthy controls. Odds ratio (OR) and 95% confidence intervals (CIs) were calculated by unconditional logistic regression adjusted for age.

Results We found that the TT genotype of *CYP3A4*1G* (rs2242480) polymorphism was associated with increased risk of breast cancer using the fixed effects model (recessive model: OR = 2.34, $p = 0.018$). Stratified according to age, CYP3A4*1G increased the risk of breast cancer especially in less than 50-year-old group (codominant model OR = 3.68, $p = 0.041$; recessive model: OR = 3.55, $p = 0.012$). Furthermore, TT genotype of rs2242480 was associated with Cerb-B2 positive (recessive model: OR = 2.47, $p = 0.025$) and stage I/II (recessive model: OR = 2.32, $p = 0.041$). However, no statistically significant associations in other polymorphisms and haploview analysis were observed.

Conclusions This study provides an evidence for polymorphism of *CYP3A4* gene associated with the development of breast cancer, also a new insight into etiology of breast cancer. However, the underlying mechanism of the *CYP3A4* gene in breast cancer is necessary for further study.

Keywords *CYP3A4* gene · Polymorphism · Breast cancer · Susceptibility

Introduction

Breast cancer is the most frequently diagnosed cancer and the leading cause of cancer death among females worldwide. According to 2012 Global Cancer Statistics, there were about 1.68 million new cases of breast cancer and 522 thousand deaths worldwide annually, which accounted for 25.2% and 14.7% of all female malignant tumor incidences and mortality, respectively [1]. In China, breast cancer is the most common cancer and is the sixth main cause of death among Chinese women [2]. It is known that breast cancer is a disease with unknown etiology which involves multiple risk factors, including the sex hormone, environment, lifestyle, and the genetic background [3]. There is a high degree of inter-individual variance in susceptibility, clinical outcome, and therapeutic response to breast cancer, highlighting the importance of genetic alterations [4]. Clearly, mutations in breast cancer susceptibility genes (*BRCA* genes 1 and 2) may increase the risk for developing hereditary breast

Xu Liu and Xi Huang cofirst authors.

✉ Zhexing Shou
36747601@qq.com

✉ Jikui Liu
liu8929@126.com

1 Department of Hepato-Pancreato-Biliary Surgery, Peking University Shenzhen Hospital, #1120, Lianhua Road, Futian District, Shenzhen 518036, Guangdong, China

2 Department of Thyroid and Breast Surgery, Peking University Shenzhen Hospital, Shenzhen 518036, Guangdong, China

3 Key Laboratory of Resource Biology and Biotechnology in Western China (Northwest University), Ministry of Education, Xi′an 710069, Shaanxi, China

4 Department of Integrated Traditional Chinese and Western Medicine, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, #1277, Jiefang Avenue, Jianghan District, Wuhan 430022, Hubei, China

cancer during the lifetime [5]. Besides, there are a number of common variants (SNPs) associated with the development of breast cancer [6–8].

Cytochrome P450s (CYPs) are a group of complex and structurally related enzymes, which are essential role for metabolizing many physiologic compounds including steroids such as estrogen and progesterone [9, 10]. One of the physiological functions of CYP epoxygenases is metabolizing arachidonic acid (AA) to biologically active epoxyeicosatrienoic acids (EETs), which has been reported to play a critical role in the tumorigenesis of diverse cancers [11]. The *CYP3A4* gene is part of a cluster of cytochrome P450 genes on chromosome 7q21.1, which has an important role in the development and progression of breast cancer [12, 13]. As known, CYP3A4 participates in synthesizing EETs, which promotes the proliferation of estrogen receptor (ER)-positive breast cancer cells, in part, through EET-mediated Stat3 activation [14]. In addition, it has been reported that CYP3A4 is involved in the oxidative metabolites of estrone, which is considered a crucial factor in the development and evolution of breast cancer [15, 16]. Many polymorphisms have been identified in the *CYP3A4* gene, which impact the expression and activity of CYP3A4 and may be associated with the carcinogenic process [17, 18].

In the present study, we devised a case–control study to investigate the frequencies distribution of *CYP3A4* SNPs in healthy population and breast cancer cohort and the potential association with breast cancer susceptibility in Han Chinese population.

Materials and methods

Study participants

Using a case–control design, a total of 569 female participants including 267 patients with newly diagnosed breast cancer and 302 healthy controls were recruited. All of blood samples of patients were collected by the Second Affiliated Hospital of Xi′an Jiaotong University. None of the patients had received radiotherapy, chemotherapy, or endocrine therapy before blood collection. All included patients had recently diagnosed and histopathologically confirmed as breast cancer. Patients with breast cancer were selected from the Department of Oncology. The exclusion criteria for patients were: patients who had any history of cancer or with complicated blood diseases, metabolic, cardiovascular disease or other systemic inflammatory diseases were excluded from this study. The controls were healthy volunteers from the medical examination of the Second Affiliated Hospital of Xi′an Jiaotong University during the same period. All of the healthy controls were ascertained to be no evidence of mammary tumors proved by recent physical and mammograph examination (within past 2 years from sample collection), no self-reported history of cancer, family history of breast cancer or sexually transmitted diseases. All subjects were genetically unrelated Chinese Han women.

Data collection

Information on demographic data of subjects and the clinical characteristics of patients was collected by trained personnel using a structured questionnaire. The clinical information of patients included estrogen receptor (ER), progesterone receptor (PR) and Cerb-B2 status, menopausal status, tumor status, stage, and lymph nodes metastasis. Subsequently, 5 mL of peripheral blood from each participant was collected by a specialized technician and stored into tubes containing ethylenediamine tetraacetic acid (EDTA) for anticoagulation in a freezer, at − 20 °C. Genomic DNA was extracted from blood samples of subjects using the Whole Blood Genomic DNA Extraction Kit (Tiangen Biotech, Beijing, China) following the manufacturer′s protocol, and then stored at − 80 °C until analysis. All of the subjects were well informed of the purpose of our study and all of them signed a written informed consent form prior to biological material collection in this study. All procedures in the present study were carried out with the approval of the Research Ethics Committee of the Second Affiliated Hospital of Xi′an Jiaotong University and the Peking University Shenzhen Hospital, and in compliance with the Declaration of Helsinki.

SNPs genotyping

This study focused on the relationship between *CYP3A4* SNPs and breast cancer in Chinese Han population. The selection of candidate SNPs in the *CYP3A4* gene was based on the International HapMap Project (http://www.hapmap.org), dbSNP (http://www.ncbi.nlm.nih.gov/projects/SNP/) and UCSC (http://genome.ucsc.edu/) databases. The minor allele frequency (MAF) value with > 0.05 in Chinese Han population was also applied to select candidate SNPs. Five SNPs (rs12333983, rs3735451, rs2242480, rs4646437, and rs2246709) in *CYP3A4* were ultimately identified for the present case–control study (Table 1). Genotyping of variants was performed by two laboratory personnel in a double-blinded fashion using the Agena MassARRAY system (Agena, San Diego, CA, U.S.A.), which have been reported in the previous study [19, 20]. To verify genotyping accuracy, marker and sample genotyping efficiencies were examined along with performance of positive and negative controls. In addition, about 10% of the total samples were randomly selected to repeat genotyping and the reproducibility was 100%.

Table 1 Comparison of basic characteristics between cases and controls

Characteristics	Cases	Controls
Number	267	302
Age (mean ± SD)	50.64 ± 11.71	49.60 ± 8.72
Menopausal status		
Premenopausal	105 (39.3%)	
Postmenopausal	156 (58.4%)	
Unavailable	6 (2.2%)	
ER status		
Negative	75 (28.1%)	
Positive	157 (58.8%)	
Unavailable	35 (13.1%)	
PR status		
Negative	104 (39.0%)	
Positive	127 (47.6%)	
Unavailable	36 (13.5%)	
Cerb-B2 status		
Negative	64 (24.0%)	
Positive	162 (60.7%)	
Unavailable	41 (15.4%)	
Lymph nodes metastasis		
Negative	115 (43.1%)	
Positive	114 (42.7%)	
Unavailable	38 (14.2%)	
T stage		
T1	80 (30%)	
T2 + T3 + T4	130 (56.1%)	
Unavailable	37 (13.9)	
TNM stage		
I + II	162 (60.7%)	
III + IV	69 (25.8%)	
Unavailable	36 (13.5%)	

ER estrogen receptor, *RP* progesterone receptor

Data analyses

The differences in demographic and clinical characteristics of study participants were evaluated using the Pearson's χ^2 test for categorical variables and independent sample Student's t test for continuous variables. Deviation from Hardy–Weinberg equilibrium (HWE) was assessed using the Chi-square test to compare the observed and expected genotype frequencies among the control subjects. The polymorphisms were excluded if they deviated from the HWE or if missing data comprised more than 10% of the total data. The allelic and genotype frequencies were compared between patients with breast cancer and controls using Pearson Chi-squared test or Fisher's exact test. Odds ratios (ORs) and 95% confidence intervals (CIs) were calculated to estimate the association between polymorphisms in the *CYP3A4* gene and the risk of breast cancer using logistic regression analysis by adjusting covariates such as age. The wild-type allele was used as a reference. The association between SNPs in *CYP3A4* and breast cancer risk was estimated using multiple inheritance model analyses (codominant, dominant, recessive and log-additive) adjusted for age by SNPstats software (http://bioinfo.iconcologia.net/snpstats/start.htm). Furthermore, we calculated stratification factors using age (≤ 50 and > 50 years) to adjust for possible cofounders. Association between *CYP3A4* polymorphisms and clinical parameters in patients with breast cancer was investigated. Finally, the pairwise linkage disequilibrium (LD), haplotype construction, and genetic association of polymorphism loci were assessed using the Haploview software package (version 4.2) and the SHEsis software (http://analysis.bio-x.cn/myAnalysis.php). To avoid false-positive results, permutation testing was performed ($n = 10{,}000$) for multiple testing correction of the haplotype analysis. Statistical analyses were performed in IBM® SPSS Statistics software version 22 (IBM®, Armonk, New York City, NY, USA). All p values of statistical tests were two-sided, and $p < 0.05$ was regarded as statistically significant.

Results

This case–control study included 267 patients with breast cancer and 302 healthy controls. The mean age and standard deviation were 50.64 ± 11.71 years for cases and 49.60 ± 8.72 years for controls. The demographic and clinical information about the cases and the controls enrolled in the present study is summarized in Table 1. The basic information on the five SNPs in *CYP3A4* regarding chromosomal position, role, MAF of cases and controls, HWE test results and call rate, were demonstrated in Table 1. Five SNPs in *CYP3A4* gene were successfully genotyped for further analysis, and the call rate for all SNPs was > 99% among the breast cancer cases and the controls. Distribution of genotype frequencies of all SNPs in the control groups was no deviation from the HWE ($p > 0.05$). The differences in allele frequency between cases and controls were compared by Chi-squared test and ORs to evaluate the associations with the risk of developing breast cancer. The minor allele of each SNP as a risk factor was compared to the wild-type (major) allele. However, no significant differences in allele frequencies were observed for patients compare with healthy controls, and these SNPs were not significantly associated with the risk of breast cancer in the allele model (Table 2).

Furthermore, multiple inheritance models (dominant, recessive, additive, and codominant models) were applied for analyzing the association between each SNP and breast cancer risk by unconditional logistic regression analysis adjusted for age (Table 3). The "TT" genotype distribution

Table 2 Basic Information about *CYP3A4* candidate SNPs and association with risk of breast cancer in allele model

SNP ID	Chr	Position	Role	Alleles (minor/major)	Case[a]	Control[a]	Call rate (%)	MAF (case/controls)	HWE[b]	OR (95% CI)	χ^2	*p* value[c]
rs12333983	7q22.1	99354114	Downstream	A/T	171/363	189/415	100	0.320	0.313	0.2837	1.034 (0.805–1.329)	0.070	0.791
rs3735451	7q22.1	99355975	Intron	C/T	172/362	189/413	99.82	0.322	0.314	0.2314	1.038 (0.808–1.333)	0.087	0.769
rs2242480	7q22.1	99361466	Intron	T/C	148/384	141/457	99.3	0.278	0.236	0.1976	1.249 (0.956–1.632)	2.660	0.103
rs4646437	7q22.1	99365083	Intron	A/G	80/454	84/518	99.82	0.150	0.140	0.8135	1.087 (0.780–1.513)	0.242	0.623
rs2246709	7q22.1	99365719	Intron	G/A	193/341	230/374	100	0.361	0.381	0.2222	0.920 (0.723–1.171)	0.455	0.500

SNP single-nucleotide polymorphism, *MAF* minor allele frequency, *HWE* Hardy–Weinberg equilibrium, *OR* odds ratio, *95% CI* 95% confidence interval

[a]The frequencies of minor/major allele between cases and controls

[b]*p* values for the Hardy–Weinberger equilibrium (HWE) test

[c]*p* values were calculated with Pearson's χ^2 tests. $p < 0.05$ indicates statistical significance

of rs2242480 in controls was 4%, while in patients was 8.7%, respectively. The frequency of mutant homozygote was found to be significantly high among the breast cancer group compared to controls. The result indicated that TT homozygote in rs2242480 had a risk-increasing effects compared with CC + CT genotype (OR = 2.34, 95% CI: 1.14–4.81, $p = 0.018$) in recessive model. No significant associations between breast cancer risk and the remaining four SNPs (rs12333983, rs3735451, rs4646437 and rs2246709) were found.

Furthermore, interaction analysis between age and *CYP3A4* gene polymorphisms was carried out, and the results (Table 4) showed that *CYP3A4* rs2242480 polymorphisms were associated with the increased risk of breast cancer at age ≤ 50 years under a codominant model (TT vs CC: OR = 3.68, 95% CI: 10.25–10.78, $p = 0.041$) and a recessive model (TT vs CC + CT: OR = 3.55, 95% CI: 1.23–10.22, $p = 0.012$). However, we still did not find any other SNPs were associated with breast cancer in age stratification (data not shown).

We also investigated the relationship of *CYP3A4* SNPs with clinical and histological features of breast cancer, including tumor status, stage, lymph node metastasis, and the statuses of ER, PR, and Cerb-B2. Multiple inheritance models analysis showed that clinical parameter such as tumor status, stage, and the statuses of ER was significantly associated with rs2242480 polymorphism. As for rs2242480, we found that the TT genotype might be related to a higher incidence of ER-positive breast cancer (TT vs. CC: OR = 2.36, 95% CI: 1.02–5.46), but this difference was not statistically significant ($p = 0.13$). This finding indicates a possible association of rs2242480 with the risk of ER-positive breast cancer. The TT genotype of rs2242480 associates with a higher Cerb-B2 positive rate compared with CC + CT genotypes (recessive model: OR = 2.47, 95% CI: 1.12–5.48, $p = 0.025$) (Table 5). Moreover, the homozygote mutant-type TT genotype was predominant in Stages I–II (TT vs. CC–CT: OR = 2.32, 95% CI: 1.04–5.18, $p = 0.025$). No significant association exists between the other four polymorphisms (rs12333983, rs3735451, rs4646437 and rs2246709) and any of the clinical parameters (data not shown).

Finally, linkage disequilibrium (LD) and haplotype analyses of the SNPs were further studied. The LD block in *CYP3A4* gene was comprised of four SNPs including rs12333983, rs3735451, rs2242480, and rs4646437, as shown in Fig. 1. The frequencies' distribution of haplotypes in case and control group is presented in Table 6. The most represented haplotype in the whole cohort of controls and cases was TTCG, followed by ACTA and ACTGB, ACCG, and TTTG. To examine the effect of haplotypes on the risk of breast cancer, the haplotype-based logistic regression method adjusted by age was carried out within the case–control cohort. Out of the haplotypes obtained, the TTCG was

Table 3 Relationship between *CYP3A4* gene polymorphisms and risk of breast cancer under multiple models of inheritance

SNP ID	Model	Genotype	Control	Case	Crude analysis		Adjusted by age and gender	
					OR (95% CI)	*p* value	OR (95% CI)	*p* value
rs12333983	Codominant	T/T	138 (45.7%)	125 (46.8%)	1.00	0.48	1.00	0.44
		A/T	139 (46%)	113 (42.3%)	0.90 (0.63–1.27)		0.89 (0.63-1.26)	
		A/A	25 (8.3%)	29 (10.9%)	1.28 (0.71–2.30)		1.29 (0.72–2.33)	
	Dominant	T/T	138 (45.7%)	125 (46.8%)	1.00	0.79	1.00	0.76
		A/T–A/A	164 (54.3%)	142 (53.2%)	0.96 (0.69–1.33)		0.95 (0.68–1.32)	
	Recessive	T/T–A/T	277 (91.7%)	238 (89.1%)	1.00	0.29	1.00	0.27
		A/A	25 (8.3%)	29 (10.9%)	1.35 (0.77–2.37)		1.37 (0.78–2.41)	
	Log-additive	–	–	–	1.04 (0.80–1.33)	0.79	1.03 (0.80–1.33)	0.79
rs3735451	Codominant	T/T	137 (45.5%)	126 (47.2%)	1.00	0.29	1.00	0.27
		C/T	139 (46.2%)	110 (41.2%)	0.86 (0.61–1.22)		0.85 (0.60–1.21)	
		C/C	25 (8.3%)	31 (11.6%)	1.35 (0.76–2.41)		1.35 (0.76–2.42)	
	Dominant	T/T	137 (45.5%)	126 (47.2%)	1.00	0.69	1.00	0.66
		C/T–C/C	164 (54.5%)	141 (52.8%)	0.93 (0.67–1.30)		0.93 (0.67–1.29)	
	Recessive	T/T–C/T	276 (91.7%)	236 (88.4%)	1.00	0.19	1.00	0.18
		C/C	25 (8.3%)	31 (11.6%)	1.45 (0.83–2.53)		1.46 (0.84–2.54)	
	Log-additive	–	–	–	1.04 (0.81–1.34)	0.77	1.04 (0.81–1.33)	0.78
rs2242480	Codominant	C/C	170 (56.9%)	141 (53%)	1.00	0.07	1.00	0.058
		C/T	117 (39.1%)	102 (38.4%)	1.05 (0.74–1.49)		1.05 (0.74–1.48)	
		T/T	12 (4%)	23 (8.7%)	**2.31 (1.11–4.81)**		**2.38 (1.14–4.97)**	
	Dominant	C/C	170 (56.9%)	141 (53%)	1.00	0.36	1.00	0.35
		C/T–T/T	129 (43.1%)	125 (47%)	1.17 (0.84–1.63)		1.17 (0.84–1.63)	
	Recessive	C/C–C/T	287 (96%)	243 (91.3%)	1.00	**0.022***	1.00	**0.018***
		T/T	12 (4%)	23 (8.7%)	**2.26 (1.10–4.64)**		**2.34 (1.14–4.81)**	
	Log-additive	–	–	–	1.26 (0.96–1.65)	0.1	1.26 (0.96–1.66)	0.093
rs4646437	Codominant	G/G	222 (73.8%)	192 (71.9%)	1.00	0.88	1.00	0.86
		G/A	74 (24.6%)	70 (26.2%)	1.09 (0.75–1.60)		1.10 (0.75–1.61)	
		A/A	5 (1.7%)	5 (1.9%)	1.16 (0.33–4.05)		1.22 (0.35–4.29)	
	Dominant	G/G	222 (73.8%)	192 (71.9%)	1.00	0.62	1.00	0.6
		G/A–A/A	79 (26.2%)	75 (28.1%)	1.10 (0.76–1.59)		1.11 (0.76–1.60)	
	Recessive	G/G–G/A	296 (98.3%)	262 (98.1%)	1.00	0.85	1.00	0.79
		A/A	5 (1.7%)	5 (1.9%)	1.13 (0.32–3.95)		1.19 (0.34–4.17)	
	Log-additive	–	–	–	1.09 (0.78–1.52)	0.62	1.10 (0.79–1.54)	0.58
rs2246709	Codominant	A/A	121 (40.1%)	107 (40.1%)	1.00	0.38	1.00	0.39
		G/A	132 (43.7%)	127 (47.6%)	1.09 (0.76–1.55)		1.10 (0.77–1.58)	
		G/G	49 (16.2%)	33 (12.4%)	0.76 (0.46–1.27)		0.78 (0.46–1.30)	
	Dominant	A/A	121 (40.1%)	107 (40.1%)	1.00	0.99	1.00	0.93
		G/A–G/G	181 (59.9%)	160 (59.9%)	1.00 (0.71–1.40)		1.01 (0.72–1.42)	
	Recessive	A/A–G/A	253 (83.8%)	234 (87.6%)	1.00	0.19	1.00	0.2
		G/G	49 (16.2%)	33 (12.4%)	0.73 (0.45–1.17)		0.74 (0.46–1.18)	
	Log-additive	–	–	–	0.92 (0.73–1.17)	0.5	0.93 (0.73–1.18)	0.56

Bold values are statistically significant

p values were calculated by unconditional logistic regression analysis with adjustments for age

OR odd ratio, *95% CI* 95% confidence interval

* indicates statistical significance ($p < 0.05$)

considered as a reference according to which is the most commonly found haplotype. The results did not reveal a significant association between common haplotypes and the risk of breast cancer ($p > 0.05$, Table 5). Interestingly, we found that rare haplotype in the block was associated with an increased risk for breast cancer under before and

Table 4 Relationship of *CYP3A4* rs2242480 polymorphisms with breast cancer stratified by age

Model	Genotype	> 50 years				≤ 50 years			
		Control	Case	OR (95% CI)	*p*	Control	Case	OR (95% CI)	*p*
Codominant	C/C	73 (52.9%)	74 (53.6%)	1.00	0.98	97 (60.2%)	67 (52.3%)	1.00	**0.041***
	C/T	58 (42%)	56 (40.6%)	0.97 (0.59–1.61)		59 (36.6%)	46 (35.9%)	1.09 (0.66–1.81)	
	T/T	7 (5.1%)	8 (5.8%)	1.09 (0.36–3.28)		5 (3.1%)	15 (11.7%)	**3.68 (1.25–10.78)**	
Dominant	C/C	73 (52.9%)	74 (53.6%)	1.00	0.95	97 (60.2%)	67 (52.3%)	1.00	0.28
	C/T–T/T	65 (47.1%)	64 (46.4%)	0.98 (0.60–1.60)		64 (39.8%)	61 (47.7%)	1.30 (0.81–2.10)	
Recessive	C/C–C/T	131 (94.9%)	130 (94.2%)	1.00	0.86	156 (96.9%)	113 (88.3%)	1.00	**0.012***
	T/T	7 (5.1%)	8 (5.8%)	1.10 (0.37–3.24)		5 (3.1%)	15 (11.7%)	**3.55 (1.23–10.22)**	
Log-additive	–	–	–	1.00 (0.67–1.51)	0.99	–	–	1.44 (0.98–2.12)	0.062

Bold values are statistically significant

p values were calculated by unconditional logistic regression analysis with adjustments for age

OR odds ratio, *95% CI* 95% confidence interval

* indicates statistical significance ($p < 0.05$)

after the adjustment for age (OR = 2.93, 95% CI: 1.08–7.95, $p = 0.036$).

Discussion

With the general understanding that breast cancer occurs by the interaction of gene with gene or environment, gene variants are likely to be involved in the etiology of breast cancer. Some studies have focused on *CYPs* gene polymorphisms which participate in the metabolism of steroids such as estrogen and progesterone. There is a paucity of studies on *CYP3A4* gene polymorphism associated with breast cancer susceptibility in Chinese Han population. This study was designed to investigate the contribution of genetic variation in the *CYP3A4* gene to breast cancer risk in a Chinese Han population. Allele, genotype, and haplotype frequencies of five SNPs (rs12333983, rs3735451, rs2242480, rs4646437, and rs2246709) in the *CYP3A4* gene between breast cancer patients and healthy controls were compared and stratification analyses by age were conducted. To our knowledge, this is the first study to evaluate the association these SNPs with the risk of breast cancer in Chinese Han population. Our study found that the genotype frequency distribution of rs2242480 was significantly different in patients and controls, and this variant was a susceptibility locus of breast cancer, which increased the risk in Chinese women population. Stratified analysis indicated that this association was especially remarkable in case subjects who were diagnosed at or before age 50 years. We also found that clinical parameter such as stage, and the status of ER and Cerb-B2 was significantly associated with rs2242480 polymorphism. These findings provide some support for our hypothesis that genetic variation in *CYP3A4* may contribute to the etiology of breast cancer.

The previous studies suggest that genetic variation in CYPs enzyme activities such as CYP3A4 may alter circulating estradiol and progesterone levels, and/or rate of formation of reactive compounds and/or rates of detoxification of steroid hormone metabolites, thus leading to differing risks of developing breast cancer [21, 22]. Besides, it has been reported that CYP3A4 is overexpressed in breast cancer cells and plays an important role in cell proliferation, angiogenesis, and migration, in part through 11,12-EET biosynthesis [23]. Given the central role of CYP3A4 enzyme in the metabolism of endogenous hormones and the progression of breast cancer, polymorphisms in *CYP3A4* gene, especially those that may potentially alter the activity or inducibility and relative to the common alleles, may influence breast cancer risk [24]. Genetic variation found in the flanking, intronic, and exonic regions of the *CYP3A4* gene may influence the level or function of the CYP3A4 protein [25]. The CYP3A4*1G (g.20230G>A, rs2242480), a single-nucleotide polymorphism present in intron 10 of CYP3A4, is the most frequent SNP of *CYP3A4* in Chinese population. The mutant allele T frequency in the study was 23.6%, similar to Healthy Han Chinese and Asian [26], but significantly different from Caucasian (7.34%) and African (85.71%). Polymorphisms in noncoding sequences may influence gene function by altering the level, location, stability, or timing of gene expression [27]. *CYP3A4* rs2242480 located in the noncoding regions of *CYP3A4* could be as a possible enhancer and promoter activity to increase the activity of CYP3A4 [28, 29]. The CYP3A4*1G polymorphism has been identified as being associated with specific types of cancers, including prostate cancer, breast cancer, and ovarian

Table 5 rs2242480 polymorphisms that were associated with clinicopathological characteristics

ER status			ER positive			ER negative		
Model	Genotype	Controls, *N* (%)	Case, *N* (%)	OR (95%)	*p*	Case, *N* (%)	OR (95%)	*p*
Codominant	C/C	170 (56.9%)	80 (51.3%)	1.00	0.13	43 (57.3%)	1.00	0.63
	C/T	117 (39.1%)	63 (40.4%)	1.14 (0.75–1.71)		27 (36%)	0.91 (0.53–1.55)	
	T/T	12 (4%)	13 (8.3%)	**2.36 (1.02–5.46)**		5 (6.7%)	1.59 (0.53–4.78)	
Dominant	C/C	170 (56.9%)	80 (51.3%)	1.00	0.27	43 (57.3%)	1.00	0.91
	C/T–T/T	129 (43.1%)	76 (48.7%)	1.25 (0.84–1.85)		32 (42.7%)	0.97 (0.58–1.62)	
Recessive	C/C–C/T	287 (96%)	143 (91.7%)	1.00	0.056	70 (93.3%)	1.00	0.38
	T/T	12 (4%)	13 (8.3%)	2.23 (0.98–5.08)		5 (6.7%)	1.66 (0.56–4.87)	
Log-additive	–	–	–	1.31 (0.95–1.82)	0.098	–	1.05 (0.68–1.62)	0.81
Cerb-B status			Cerb-B2 positive			Cerb-B2 negative		
Model	Genotype	Controls, *N* (%)	Case, *N* (%)	OR (95%)	*p*	Case, *N* (%)	OR (95%)	*p*
Codominant	C/C	170 (56.9%)	80 (49.7%)	1.00	0.062	39 (60.9%)	1.00	0.78
	C/T	117 (39.1%)	66 (41%)	1.17 (0.78–1.76)		22 (34.4%)	0.82 (0.46–1.46)	
	T/T	12 (4%)	15 (9.3%)	**2.65 (1.17–5.97)**		3 (4.7%)	1.08 (0.29–4.02)	
Dominant	C/C	170 (56.9%)	80 (49.7%)	1.00	0.18	39 (60.9%)	1.00	0.55
	C/T–T/T	129 (43.1%)	81 (50.3%)	1.31 (0.89–1.93)		25 (39.1%)	0.85 (0.49–1.47)	
Recessive	C/C–C/T	287 (96%)	146 (90.7%)	1.00	**0.025***	61 (95.3%)	1.00	0.82
	T/T	12 (4%)	15 (9.3%)	**2.47 (1.12–5.48)**		3 (4.7%)	1.17 (0.32–4.26)	
Log-additive	–	–	–	1.38 (1.00–1.89)	0.047	–	0.90 (0.56–1.45)	0.67
Stage			Stages I + II			Stages III + IV		
Model	Genotype	Controls, *N* (%)	Case, *N* (%)	OR (95%)	*p*	Case, *N* (%)	OR (95%)	*p*
Codominant	C/C	170 (56.9%)	86 (53.1%)	1.00	0.12	37 (54.4%)	1.00	0.84
	C/T	117 (39.1%)	62 (38.3%)	1.04 (0.69–1.55)		27 (39.7%)	1.06 (0.61–1.84)	
	T/T	12 (4%)	14 (8.6%)	**2.35 (1.03–5.34)**		4 (5.9%)	1.45 (0.44–4.79)	
Dominant	C/C	170 (56.9%)	86 (53.1%)	1.00	0.46	37 (54.4%)	1.00	0.73
	C/T–T/T	129 (43.1%)	76 (46.9%)	1.16 (0.78–1.70)		31 (45.6%)	1.10 (0.64–1.87)	
Recessive	C/C–C/T	287 (96%)	148 (91.4%)	1.00	**0.041***	64 (94.1%)	1.00	0.57
	T/T	12 (4%)	14 (8.6%)	**2.32 (1.04–5.18)**		4 (5.9%)	1.41 (0.44–4.57)	
Log-additive	–	–	–	1.26 (0.91–1.72)	0.16	–	1.12 (0.72–1.75)	0.62

Bold values are statistically significant

p values were calculated by unconditional logistic regression analysis with adjustments for age

OR odds ratio, *95% CI* 95% confidence interval

* indicates statistical significance ($p < 0.05$)

cancer [30–32]. Our findings are consistent with at least one other case–control study of American women in which the CT–TT genotype was found to be associated with increased breast cancer incidence compared with the CC genotype (OR = 1.25, 95% CI: 1.00–1.56) [33]. However, another study did not supported an association between breast cancer risk and *CYP3A4* rs2242480 variants in Western Irish population [34]. The inconsistencies in these reports may result from different ethnic, environment, or insufficient sample size. However, there was no finding that rs2242480 was associated with the risk of breast cancer among Chinese Han population in previous studies. Therefore, these results have to be taken with caution, and replication of the studies with a larger sample size is required to confirm our findings.

Moreover, it is interesting that the *CYP3A4* rs2242480 was associated with the increased risk of breast cancer, especially at age younger than 50 years. Notably, the median age at diagnosis of breast cancer is between 60 and 70 years in Western countries, whereas the peak age is between 40 and 50 years in Asian countries [35]. This finding partly reflects the age–gene interactions among women with breast cancer, and also implies that breast

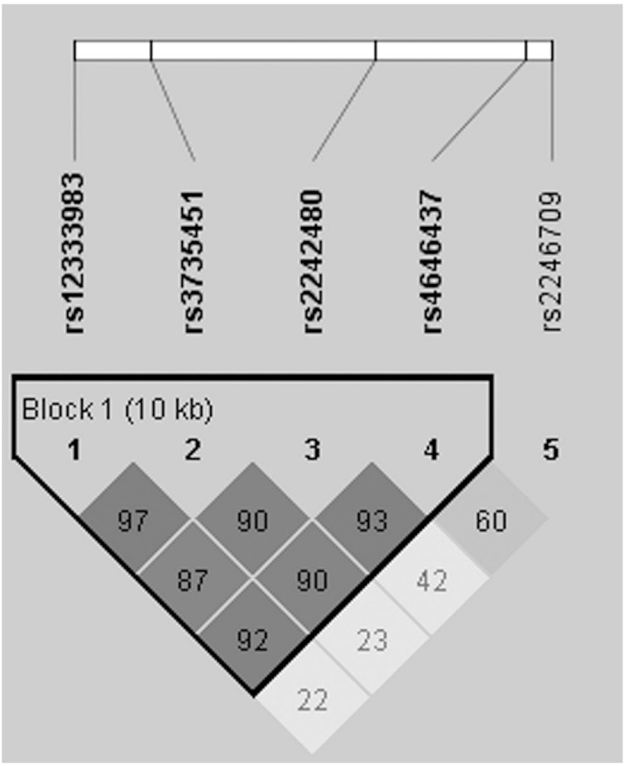

Fig. 1 Haplotype block map for SNPs in the *CYP3A4* gene

cancers among younger women tend to maybe have specific risk factors and molecular characteristics in Chinese Han population. In addition, we further explored the association of *CYP3A4* polymorphisms and clinical features of breast cancer. Our results suggested that rs2242480 polymorphism was associated with stage and the status of ER and Cerb-B2. Mitra et al. demonstrated that CYP3A4 epoxygenase promotes the growth of estrogen receptor (ER)-positive breast cancer cells [14]. Our results showed that rs2242480-TT was potentially associated with the increased risk of breast cancer in the ER-positive patients. This result implied that CYP3A4*1G might play an important role in the pathogenesis of ER-positive breast cancer. Cerb-B2 (also named HER2) protooncogene is an indicator of clinical tumor aggressiveness and poor prognosis [36]. CYP3A4*1G TT genotype had a risk-increasing effects compared with CC + CT genotype in the CerbB2-positive patients, which indicates that CYP3A4 may be associated with prognosis of breast cancer. Our results also displayed that rs2242480 increased the risk of breast cancer in the in Stage I–II patients under the recessive model. The underlying mechanisms for this finding are unknown by now, which can be a promising field for future studies. We also performed stratified analyses by PR status, tumor status, and lymph nodes metastasis, though there was no significant difference in the subgroups.

Inevitably, our study has potential limitations that should be considered during the interpretation of our results. First, the inherent selecting bias and information bias were the unavoidable problems, because all participants were recruited from the identical hospitals. Second, the number of cases in our study was not large and our study population was all Chinese Han people, which cannot preclude false-negative results and cannot be extrapolated to other populations. Finally, there is the unavailability of complete clinical information (including menopausal status of healthy controls) and the absence of some environmental exposures factors (such as high-dose radiation exposure, alcohol consumption). Accordingly, explicit mechanisms of *CYP3A4* SNPs on development of breast cancer are still bewildered and further research is needed. Despite the limitations mentioned above, the results of our present study provided scientific evidence of *CYP3A4* gene with breast cancer in the future studies.

Table 6 Haplotype frequencies of *CYP3A4* gene and the association with breast cancer risk

Haplotype	Frequency		χ^2	*p* value	Crude analysis		Adjusted by age	
	Case	Control			OR (95% CI)	*p*	OR (95% CI)	*p*
TTCG	0.645	0.669	0.72	0.3961	1.00	–	1.00	–
ACTA	0.138	0.131	0.124	0.7246	1.09 (0.76–1.55)	0.65	1.10 (0.77–1.57)	0.62
ACTG	0.113	0.096	0.859	0.354	1.26 (0.85–1.86)	0.26	1.25 (0.85–1.86)	0.26
ACCG	0.058	0.086	3.428	0.0641	0.68 (0.42–1.10)	0.12	0.67 (0.41–1.08)	0.1
TTTG	0.017	0.009	1.574	0.2096	1.92 (0.65–5.70)	0.24	1.97 (0.66–5.90)	0.22
Rare					**2.98 (1.10–8.10)**	**0.033***	**2.93 (1.08–7.95)**	**0.036***

Bold values are statistically significant

Block comprised of the three closely linked SNPs rs12333983, rs3735451, rs2242480, and rs4646437

OR odds ratio, *95% CI* 95% confidence interval

p values were calculated with Pearson's χ^2 tests

* indicates statistical significance ($p < 0.05$)

In summary, our study provides evidence for a Chinese population that the potential role of the *CYP3A4* variant in breast cancer risk. Combined with the previous studies, this association may be a promising starting point for a functional profile of the *CYP3A4* gene and increased understanding of the biological processes associated with breast cancer formation and progression. Larger well-designed epidemiological studies with ethnically diverse populations and functional evaluations are warranted.

Acknowledgements This study was funded by Sanming Project of Medicine in Shenzhen (No. SZSM201612021) and Science-Technology Developing Project of Guangdong Province (No. 2017B090904010). We are grateful to the individuals for their participation in this study. We also thank the clinicians and hospital staff who contributed to the sample and data collection for this study. We would also like to thank all participants for this manuscript. Special thanks go to Zhijun Dai (Department of Oncology, The Second Affiliated Hospital, Xi'an Jiaotong University) for the sample and data collection and Bo Xu (Department of Bioinformatics, College of Life Sciences, Northwest Agriculture and Forestry University) for helpful checking our statistical methods.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

References

1. Torre LA, Bray F, Siegel RL et al (2015) Global cancer statistic, 2012. CA Cancer J Clin 65(2):87–108
2. Chen W, Zheng R, Baade PD et al (2016) Cancer statistics in China, 2015. CA Cancer J Clin 66(2):115–132
3. Fasching PA, Ekici AB, Adamietz BR et al (2011) Breast cancer risk—genes, environment and clinics. Geburtsh Frauenheilk 71(12):1056–1066
4. Sparano JA, Gray RJ, Makower DF et al (2015) Prospective validation of a 21-gene expression assay in breast cancer. N Engl J Med 373(21):2005–2014
5. Easton DF, Bishop DT, Ford D et al (1993) Genetic linkage analysis in familial breast and ovarian cancer: results from 214 families. The Breast Cancer Linkage Consortium. Am J Hum Genet 52(4):678–701
6. Dai ZJ, Liu XH, Ma YF et al (2016) Association between single nucleotide polymorphisms in DNA polymerase kappa gene and breast cancer risk in Chinese Han population: a STROBE-compliant observational study. Medicine 95(2):e2466
7. Xia P, Li B, Geng T et al (2014) FGFR2 gene polymorphisms are associated with breast cancer risk in the Han Chinese population. Am J Cancer Res 5(5):1854–1861
8. Zhou L, He N, Feng T et al (2014) Association of five single nucleotide polymorphisms at 6q25.1 with breast cancer risk in northwestern China. Am J Cancer Res 5(8):2467–2475
9. Modugno F, Knoll C, Kanbourshakir A et al (2003) A potential role for the estrogen-metabolizing cytochrome P450 enzymes in human breast carcinogenesis. Breast Cancer Res Treat 82(3):191–197
10. Lee AJ, Cai MX, Thomas PE et al (2003) Characterization of the oxidative metabolites of 17beta-estradiol and estrone formed by 15 selectively expressed human cytochrome p450 isoforms. Endocrinology 144(8):3382–3398
11. Capdevila J, Falck J, Harris RC (2000) Cytochrome P450 and arachidonic acid bioactivation: molecular and functional properties of the arachidonate monooxygenase. J Lipid Res 41(2):163–181
12. Yang X, Zhang B, Molony C et al (2010) Systematic genetic and genomic analysis of cytochrome P450 enzyme activities in human liver. Genome Res 20(8):1020–1036
13. Murray GI, Patimalla S, Stewart KN et al (2010) Profiling the expression of cytochrome P450 in breast cancer. Histopathology 57(2):202–211
14. Mitra R, Guo Z, Milani M et al (2011) CYP3A4 mediates growth of estrogen receptor-positive breast cancer cells in part by inducing nuclear translocation of phospho-Stat3 through biosynthesis of (±)-14,15-epoxyeicosatrienoic acid (EET). J Biol Chem 286(20):17543–17559
15. Rebbeck TR, Troxel AB, Walker AH et al (2007) Pairwise combinations of estrogen metabolism genotypes in postmenopausal breast cancer etiology. Cancer Epidemiol Biomark Prev 16(3):444–450
16. Russo J, Hasan LM, Balogh G et al (2003) Estrogen and its metabolites are carcinogenic agents in human breast epithelial cells. J Steroid Biochem Mol Biol 87(1):1–25
17. Bozina N, Bradamante V, Lovrić M (2009) Genetic polymorphism of metabolic enzymes P450 (CYP) as a susceptibility factor for drug response, toxicity, and cancer risk. Arhiv Za Higijenu Rada I Toksikologiju 60(2):217–242
18. Eshragh J, Dhruva A, Paul SM et al (2017) Associations between neurotransmitter genes and fatigue and energy levels in women after breast cancer surgery. J Pain Symptom Manag 53(1):67–84 e67
19. Gabriel S, Ziaugra L, Tabbaa D (2009) SNP genotyping using the Sequenom MassARRAY iPLEX platform. Curr Protoc Hum Genet. https://doi.org/10.1002/0471142905.hg0212s60
20. Thomas RK, Baker AC, Debiasi RM et al (2007) High-throughput oncogene mutation profiling in human cancer. Nat Genet 39(3):347–351
21. Brockstedt U, Krajinovic M, Richer C et al (2002) Analyses of bulky DNA adduct levels in human breast tissue and genetic polymorphisms of cytochromes P450 (CYPs), myeloperoxidase (MPO), quinone oxidoreductase (NQO1), and glutathione *S*-transferases (GSTs). Mutat Res Genet Toxicol Environ Mutagen 516(1–2):41–47
22. Hormones TE, Group BCC (2011) Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies. J Natl Cancer Inst 105(5):709–722
23. Nithipatikom K, Brody DM, Tang AT et al (2010) Inhibition of carcinoma cell motility by epoxyeicosatrienoic acid (EET) antagonists. Cancer Sci 101(12):2629–2636
24. Keshava C, Mccanlies EC, Weston A (2004) CYP3A4 polymorphisms—potential risk factors for breast and prostate cancer: a huge review. J Urol 160(9):825–841
25. Lamba JK, Lin YS, Schuetz EG et al (2002) Genetic contribution to variable human CYP3A-mediated metabolism. Adv Drug Deliv Rev 54(10):1271–1294
26. Du J, Xing Q, Xu L et al (2006) Systematic screening for polymorphisms in the CYP3A4 gene in the Chinese population. Pharmacogenomics 7(6):831–841
27. Tabor HK, Risch NJ, Myers RM (2002) Candidate-gene approaches for studying complex genetic traits: practical considerations. Nat Rev Genet 3(5):391–397
28. He BX, Shi L, Qiu J et al (2011) A functional polymorphism in the CYP3A4 gene is associated with increased risk of coronary heart disease in the Chinese Han population. Basic Clin Pharmacol 108(3):208–213

29. Yang W, Zhao D, Han S et al (2015) CYP3A4*1G regulates CYP3A4 intron 10 enhancer and promoter activity in an allelic-dependent manner. Int J Clin Pharmacol Ther 53(8):647–657
30. Chau CH, Price DK, Till C et al (2015) Finasteride concentrations and prostate cancer risk: results from the Prostate Cancer Prevention Trial. PLoS One 10(5):e0126672
31. Pascual T, Apellaniz-Ruiz M, Pernaut C et al (2017) Polymorphisms associated with everolimus pharmacokinetics, toxicity and survival in metastatic breast cancer. PLoS One 12(7):e0180192
32. Fiszer-Maliszewska L, Laczmanski L, Dolinska A et al (2018) Polymorphisms of ABCB1, CYP3A4 and CYP3A5 genes in ovarian cancer and treatment response in poles. Anticancer Res 38(3):1455–1459
33. Parada H Jr, Steck SE, Cleveland RJ et al (2017) Genetic polymorphisms of phase I metabolizing enzyme genes, their interaction with lifetime grilled and smoked meat intake, and breast cancer incidence. Ann Epidemiol 27(3):208–214
34. Johnson N, Walker K, Gibson LJ et al (2012) CYP3A variation, premenopausal estrone levels, and breast cancer risk. J Natl Cancer Inst 104(9):657–669
35. Leong SPL, Shen ZZ, Liu TJ et al (2010) Is breast cancer the same disease in asian and western countries? World J Surg 34(10):2308–2324
36. Gown AM (2008) Current issues in ER and HER2 testing by IHC in breast cancer. Mod Pathol 21(Suppl 2):S8–S15