Genetics and Molecular Research

Validation of associations between *ESR1* variants and breast cancer risk in Chinese cohorts

X. Li[1]*, G.Y. Yao[2]*, F.X. Li[3], Y.R. Qiu[1] and X.X. Yang[3]

[1]Clinical Laboratory of Nanfang Hospital, Southern Medical University, Guangzhou, China
[2]Breast Cancer Center, Nanfang Hospital, Southern Medical University, Guangzhou, China
[3]School of Biotechnology, Southern Medical University, Guangzhou, China

*These authors contributed equally to this study.
Corresponding authors: X.X. Yang / Y.R. Qiu
E-mail: yxxzb@sohu.com / qiuyuronggz@126.com

Genet. Mol. Res. 15 (3): gmr.15036683
Received July 21, 2015
Accepted April 23, 2016
Published July 15, 2016
DOI http://dx.doi.org/10.4238/gmr.15036683

ABSTRACT. Estrogen receptor-α (ER) protein plays a key role in breast carcinogenesis, and common genetic variants in the corresponding gene locus have been associated with breast cancer risk in different populations. Here, we analyzed estrogen receptor 1 (*ESR1*) associations in two hospital-based studies of patients from the south of China. Three single-nucleotide polymorphisms (SNPs; rs3757318, rs2046210, and rs3734805) in *ESR1* were selected from previous genome-wide association study results and were genotyped using the Sequenom MassARRAY® iPLEX System in 845 breast cancer patients and 882 healthy controls. Association analysis based on unconditional logistic regression was carried out to determine the odds ratio (OR) and 95%

confidence interval (95%CI) for each SNP. Stratified analyses according to the status of ER and progesterone receptor (PR) were also performed. Of the three SNPs, rs3757318 did not pass the Hardy-Weinberg equilibrium test and was excluded from the subsequent analysis. The other two SNPs (rs2046210 and rs3734805) were strongly associated with susceptibility to breast cancer. Allele T of rs2046210 and allele C of rs3734805 were risk alleles and the adjusted ORs were 1.348 (95%CI = 1.172-1.550, P = 0.0001) and 1.319 (95%CI = 1.144-1.522, P = 0.0001), respectively. Furthermore, the risk allele of rs2046210 gave negative results for ER and PR expression in an immunohistochemical test, with ORs of 0.602 (95%CI = 0.384-0.944, P = 0.027) and 0.532 (95%CI = 0.338-0.837, P = 0.006), respectively. Our study further supports associations between rs2046210 and rs3734805 and breast cancer risk in Chinese women.

Key words: Breast cancer; Estrogen receptor 1; Susceptibility; Single-nucleotide polymorphism

INTRODUCTION

Breast cancer is one of the most common malignant tumors among women worldwide. The incidence of breast cancer is low but increasing in China (Jemal et al., 2011); its pathogenesis is multifactorial and remains unclear. However, there is increasing evidence that genetic factors play an important role in the development and progression of breast cancer.

Estrogen receptor-α (ERα), which is encoded by the *ESR1* gene, stimulates the proliferation and differentiation of mammary epithelial tissue by combining with estrogen. Previous studies have demonstrated that ERα is associated with the development of breast cancer (Clemons and Goss, 2001). In this paper, we use "ER" to indicate "ERα," if not otherwise specified. Currently, endocrine therapy treatment occurs mainly under the influence of ER and progesterone receptor (PR) immunochemistry. ER-positive (ER+) breast cancer is typically associated with lower proliferation rate and grade, and positive PR expression (Wenger et al., 1993). Owing to the importance of ER in breast cancer, many researchers have focused on the ER gene because its variants may change the normal expression of ER or its functions. Recently, three genome-wide association studies (GWASs) have revealed that three single-nucleotide polymorphisms (SNPs; rs3757318, rs2046210, and rs3734805) near or in the *ESR1* gene are associated with breast cancer (Zheng et al., 2009; Turnbull et al., 2010; Fletcher et al., 2011). However, the results of replication studies were not consistent owing to the sample sizes used and the diverse genetic backgrounds of the ethnic groups involved (Stacey et al., 2010; Zhong and Prentice, 2010; Campa et al., 2011).

Therefore, we carried out a hospital-based case-control study of moderate size on subjects selected from northern and southern Chinese populations. Stratification analyses based on the status of ER and PR among breast cancer patients were also carried out.

MATERIAL AND METHODS

Study population

In order to eliminate regional discrepancies between southern and northern Chinese

populations, a total of 845 female patients were recruited between 2010 and 2012 from the outpatient and inpatient clinics of: Nanfang Hospital, Southern Medical University, Guangzhou, Guangdong Province; the First Affiliated Hospital of Nanchang University, Nanchang, Jiangxi Province; the Affiliated Hospital of the Medical College of Qingdao University, Qingdao; and the Chongqing Cancer Hospital, Chongqing Province. Patients selected for the study were those with histologically confirmed breast cancer. Healthy women with no history of cancer (882) were recruited randomly from the same geographic regions as the patients. The mean ages of the patients and controls were 48.42 years (SD: 9.75 years) and 45.22 years (SD: 10.65 years), respectively. All participants gave their written informed consent before participating in the trial.

DNA extraction

Peripheral blood samples were drawn from all participants. The sample were delivered in a frozen state by express mail and stored at -70°C until required for DNA extraction. Genomic DNA was extracted using a commercial blood DNA kit (TIANamp Genomic DNA Purification Kit; Tiangen Biotech, Beijing, China), following the manufacturer instructions, and stored at -80°C before testing.

SNP selection, genotyping, and quality control

Three SNPs (rs3757318, rs2046210, and rs3734805) were selected from previous GWASs on *ESR1* (Zheng et al., 2009; Turnbull et al., 2010; Fletcher et al., 2011). The following primer sets, which included a pair of amplicons and an extension primer for each SNP, were designed using the Assay Design 3.1 software (Sequenom, San Diego, CA, USA): for rs3757318, forward 5'-ACGTTGGATGGCAGGGTGGTTCAGGAATTT-3', reverse 5'-ACGTTGGATGCTTTGCAGAGAGCATGGAAC-3', and extension 5'-CACATATGGGTCAGAGTCC-3'; for rs2046210, forward 5'-ACGTTGGATGTGAAACCATCAGGGTGCCTC-3', reverse 5'-ACGTTGGATGCCTCACACATACATACAGTC-3', and extension 5'-GAATCTTTTATTTCAGGTAGATG-3'; for rs3734805, forward 5'-ACGTTGGATGTTGAGAGAGGTGGCCATAAG-3', reverse 5'-ACGTTGGATGCACAAACTGATCAAATCTC-3', and extension 5'-ATTCCCAATTTCACAAATTCCTC-3'. Three SNPs were genotyped using the MassARRAY® matrix-assisted laser desorption ionization-time of flight mass spectrometry platform (Sequenom, San Diego, CA, USA), according to the manufacturer instructions (www.sequenom.com). The DNA sample quality control threshold was set at 90%.

Statistical analysis

Deviations from the Hardy-Weinberg equilibrium (HWE) were assessed by Fisher's exact and χ^2 tests. Generally, dominant, recessive, overdominant, and additive genetic models were carried out based on unconditional logistic regression with calculation of the odds ratio (OR) and 95% confidence interval (95%CI) for each SNP after adjustment for age. We conducted our statistical analyses not only on the association between SNPs and susceptibility to breast cancer, but also on the association between SNPs and status of ER and PR. P values of < 0.05 were considered statistically significant. All statistical tests were carried out using the SPSS 13.0 software (SPSS, Chicago, IL, USA) and repeated with the web-based tool SNPstats (http://bioinfo.iconcologia.net/SNPstats) (Solé et al., 2006).

RESULTS

Of the three SNPs, rs3757318 did not conform to the HWE in the controls ($P < 0.01$), and was excluded from further analysis. The DNA sample quality control reached 98.09% in the genotyping of rs2046210 and 99.07% in that of rs3734805.

Polymorphisms rs2046210 and rs3734805 were both found to be significantly associated with susceptibility to breast cancer, as shown in Table 1. For rs2046210, compared with allele C, subjects with allele T had an increased risk of developing breast cancer, with an OR of 1.348 (95%CI = 1.172-1.550, P = 0.0001). In the analysis of rs3734805, subjects with allele C were prone to developing breast cancer, with an OR of 1.319 (95%CI = 1.144-1.522, P = 0.0001).

Table 1. rs2046210 and rs3734805 genotype distributions (%) and association with breast cancer risk.

SNP	Model	Genotype	Control [N (%)]	Case [N (%)]	P[a]	Adjusted[b] OR (95%CI)	P[c]
rs2046210	Co-dominant	C/C	355 (41.5%)	254 (30.3%)	0.0001	1	
		C/T	383 (44.8%)	442 (52.7%)		1.616 (1.305-2.001)	0.0001
		T/T	117 (13.7%)	143 (17%)		1.676 (1.247-2.254)	0.001
	Dominant	C/C	355 (41.5%)	254 (30.3%)	0.0001	1	
		C/T-T/T	500 (58.5%)	585 (69.7%)		1.630 (1.331-1.997)	0.0001
	Recessive	C/C-C/T	738 (86.3%)	696 (83%)	0.055	1	
		T/T	117 (13.7%)	143 (17%)		1.270 (0.971-1.661)	0.081
	Overdominant	C/C-T/T	472 (55.2%)	397 (47.3%)	0.001	1	
		C/T	383 (44.8%)	442 (52.7%)		1.383 (1.139-1.678)	0.001
	Log-additive	C	1093 (63.9%)	950 (56.6%)	0.0001	1	
		T	617 (36.1%)	728 (43.4%)		1.348 (1.172-1.550)	0.0001
rs3734805	Co-dominant	A/A	419 (47.9%)	314 (37.5%)	0.0001	1	
		C/A	360 (41.2%)	414 (49.5%)		1.558 (1.268-1.914)	0.001
		C/C	95 (10.9%)	109 (13%)		1.505 (1.098-2.063)	0.011
	Dominant	A/A	419 (47.9%)	314 (37.5%)	0.0001	1	
		C/A-C/C	455 (52.1%)	523 (62.5%)		1.547 (1.272-1.880)	0.0001
	Recessive	A/A-C/A	779 (89.1%)	728 (87%)	0.169	1	
		C/C	95 (10.9%)	109 (13%)		1.198 (0.891-1.612)	0.232
	Overdominant	A/A-C/C	514 (58.8%)	423 (50.5%)	0.001	1	
		C/A	360 (41.2%)	414 (49.5%)		1.424 (1.173-1.728)	0.0001
	Log-additive	A	1198 (68.5%)	1042 (62.2%)	0.0001	1	
		C	550 (31.5%)	632 (37.8%)		1.319 (1.144-1.522)	0.0001

SNP = single-nucleotide polymorphism. [a]P values for differences between controls and patients by χ^2 test; [b]adjusted for age; [c]P values for differences between controls and patients by logistic regression analysis.

In addition, we conducted further analysis on the breast cancer group to investigate whether there was any relationship between the two SNPs and the status of ER and PR. Interestingly, rs2046210 distribution showed significant differences between the ER and PR groups (Table 2). In the ER and PR groups, patients with the TT genotype were prone to be ER-negative, with an OR of 0.602 (95%CI = 0.384-0.944, P = 0.027), and to be PR-negative, with an OR of 0.532 (95%CI = 0.338-0.837, P = 0.006). No significant associations were found between rs3734805 and the status of ER and PR in the present study.

DISCUSSION

The aim of this study was to evaluate the association between three SNPs (rs3757318, rs2046210, and rs3734805) and susceptibility to breast cancer in a Chinese population. We collected samples from a total of 1727 cases (845 breast cancer cases and 882 controls) from northern and southern China for a hospital-based case-control study in order to overcome regional differences. Besides replicating the association study, we also conducted stratification

Table 2. Stratified analysis of rs2046210 based on the status of estrogen receptor-α (ER) and progesterone receptor (PR) among breast cancer patients.

SNP	Genotype	ER^- N (%)	ER^+ N (%)	Adjusted[b] OR (95%CI)	P[c]
rs2046210	C/C	80 (27.5%)	137 (32%)	1	
	C/T	150 (51.5%)	229 (53.5%)	0.892 (0.632-1.259)	0.515
	T/T	61 (21%)	62 (14.5%)	0.602 (0.384-0.944)	0.027
	C	310 (53.3%)	503(58.8%)	1	
	T	272 (46.7%)	353(41.2%)	0.805 (0.651-0.996)	0.046
		PR^- N (%)	PR^+ N (%)		
	C/C	85 (25.9%)	131 (33.6%)	1	
	C/T	174 (53%)	205 (52.6%)	0.763 (0.542-1.074)	0.121
	T/T	69 (21%)	54 (13.8%)	0.532 (0.338-0.837)	0.006
	C	344 (52.4%)	467 (59.9%)	1	
	T	312 (47.6%)	313 (40.1%)	0.753 (0.609-0.930)	0.008

SNP = single nucleotide polymorphism.

analysis to determine whether there was any link between SNPs and the status of ER and PR, two important clinical characteristics, among breast cancer cases.

Among the three SNPs, rs3757318 was unfortunately excluded because it did not comply with HWE. The other two SNPs (rs2046210 and rs3734805) showed a significant association with susceptibility to breast cancer. For rs2046210, compared with allele C, subjects with allele T had an increased risk of developing breast cancer, with an OR of 1.348 (95%CI = 1.172-1.550, P = 0.0001), which is consistent with data from one of the GWASs (Zheng et al., 2009). In the analysis of rs3734805, subjects with allele C were prone to developing breast cancer, with an OR of 1.319 (95%CI = 1.144-1.522, P = 0.0001), which is also consistent with the result of the relevant GWAS (Fletcher et al., 2011). We also found that rs2046210 had a statistical association with the expression of ER and PR. It seems that women with breast cancer who carry the TT genotype are more likely to lose the expression of ER and PR, with ORs of 0.602 (95%CI = 0.384-0.944, P = 0.027) and 0.523 (95%CI = 0.338-0.837, P = 0.006), respectively.

Breast cancer, a hormone-depended cancer, is the most common malignant tumor among women worldwide. To date, many papers have demonstrated that the estrogen receptor has a strong connection with the pathogenesis of breast cancer (Clemons and Goss, 2001). There are two kinds of ER that are encoded by distinct genes: ERα, which play a vital role in breast cancer initiation and progression, and ERβ, whose biological function is still subject to debate (Osborne et al., 2001). Related studies have revealed that ER acts as a hormone-activated transcription factor, which regulates the expression of various genes that promote breast cancer cell proliferation and survival after binding with the hormone estrogen (Clemons and Goss, 2001; Osborne et al., 2001). Moreover, many investigations have been conducted on risks associated with sequence variants in *ESR1*, generally with equivocal results (Figtree et al., 2009). In the present study, rs2046210 and rs3734805, which are located upstream of the transcription initiation site and at the 5' untranslated region of the *ESR1* gene, respectively, (Zheng et al., 2009; Fletcher et al., 2011) showed statistical association with the risk of breast cancer in the Chinese Han population. A change of allele in the gene may result in a change in the working mechanisms of ER (Herynk and Fuqua, 2004), but our data did not provide evidence to support this hypothesis.

In clinical situations, ER status is an important biomarker to predict response to endocrine treatment, which is the first and most efficacious target treatment for breast cancer in ER-positive patients (Giuliano et al., 2011). Moreover, the loss of ER expression is one of the reasons for resistance to endocrine treatment (Musgrove and Sutherland, 2009). Interestingly,

in the further analysis of rs2046210 in ER-positive and ER-negative breast cancer subjects, the susceptibility genotype TT seemed to correspond to loss of expression. Similar results appeared in the PR-positive and PR-negative breast cancer subjects. It is possible that ER is usually positive when PR expressed (Wenger et al., 1993). Unfortunately, we did not have more information about therapeutic effectiveness in the subjects. However, the data revealed that rs2046210 may not only be a susceptibility SNP for breast cancer, but may also be a functional SNP that effects ER expression levels. Further mechanism studies are needed to confirm this assumption.

In conclusion, our results suggest that rs2046210 and rs3734805 represent suitable markers for assessing susceptibility to breast cancer among the Chinese Han population. However, breast cancer patients with the risk genotype of rs2046210 were prone to lose the expression of ER and PR.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

Research supported by the National Natural Science Foundation of China (Grant #81302327).

REFERENCES

Campa D, Kaaks R, Le Marchand L, Haiman CA, et al. (2011). Interactions between genetic variants and breast cancer risk factors in the breast and prostate cancer cohort consortium. *J. Natl. Cancer Inst.* 103: 1252-1263. http://dx.doi.org/10.1093/jnci/djr265

Clemons M and Goss P (2001). Estrogen and the risk of breast cancer. *N. Engl. J. Med.* 344: 276-285. http://dx.doi.org/10.1056/NEJM200101253440407

Figtree GA, Noonan JE, Bhindi R and Collins P (2009). Estrogen receptor polymorphisms: significance to human physiology, disease and therapy. *Recent Pat. DNA Gene Seq.* 3: 164-171. http://dx.doi.org/10.2174/187221509789318397

Fletcher O, Johnson N, Orr N, Hosking FJ, et al. (2011). Novel breast cancer susceptibility locus at 9q31.2: results of a genome-wide association study. *J. Natl. Cancer Inst.* 103: 425-435. http://dx.doi.org/10.1093/jnci/djq563

Giuliano M, Schifp R, Osborne CK and Trivedi MV (2011). Biological mechanisms and clinical implications of endocrine resistance in breast cancer. *Breast* 20 (Suppl 3): S42-S49. http://dx.doi.org/10.1016/S0960-9776(11)70293-4

Herynk MH and Fuqua SA (2004). Estrogen receptor mutations in human disease. *Endocr. Rev.* 25: 869-898. http://dx.doi.org/10.1210/er.2003-0010

Jemal A, Bray F, Center MM, Ferlay J, et al. (2011). Global cancer statistics. *CA Cancer J. Clin.* 61: 69-90. http://dx.doi.org/10.3322/caac.20107

Musgrove EA and Sutherland RL (2009). Biological determinants of endocrine resistance in breast cancer. *Nat. Rev. Cancer* 9: 631-643. http://dx.doi.org/10.1038/nrc2713

Osborne CK, Schiff R, Fuqua SA and Shou J (2001). Estrogen receptor: current understanding of its activation and modulation. *Clin. Cancer Res.* 7 (Suppl): 4338s-4342s, discussion 4411s-4412s.

Solé X, Guinó E, Valls J, Iniesta R, et al. (2006). SNPStats: a web tool for the analysis of association studies. *Bioinformatics* 22: 1928-1929. http://dx.doi.org/10.1093/bioinformatics/btl268

Stacey SN, Sulem P, Zanon C, Gudjonsson SA, et al. (2010). Ancestry-shift refinement mapping of the C6orf97-ESR1 breast cancer susceptibility locus. *PLoS Genet.* 6: e1001029. http://dx.doi.org/10.1371/journal.pgen.1001029

Turnbull C, Ahmed S, Morrison J, Pernet D, et al.; Breast Cancer Susceptibility Collaboration (UK) (2010). Genome-wide association study identifies five new breast cancer susceptibility loci. *Nat. Genet.* 42: 504-507. http://dx.doi.org/10.1038/ng.586

Wenger CR, Beardslee S, Owens MA, Pounds G, et al. (1993). DNA ploidy, S-phase, and steroid receptors in more than 127,000 breast cancer patients. *Breast Cancer Res. Treat.* 28: 9-20. http://dx.doi.org/10.1007/BF00666351

Zheng W, Long J, Gao YT, Li C, et al. (2009). Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. *Nat. Genet.* 41: 324-328. http://dx.doi.org/10.1038/ng.318

Zhong H and Prentice RL (2010). Correcting "winner's curse" in odds ratios from genomewide association findings for major complex human diseases. *Genet. Epidemiol.* 34: 78-91.