RESEARCH ARTICLE

Editorial Process: Submission:12/10/2021 Acceptance:04/16/2022

Sequencing of Chromosomal Locus 6q25.1 Revealed Two Significant SNPs rs2046210 and rs2046211 Associated with Breast Cancer: A Case-Control Study in Egyptian Women

Marwa H Saied[1]*, Mohamed M Rizk[1], Nedaa Awad Abdelkader El Sayed[1], Salma Osama Ahmed Mostafa[1], Salah Marzouk[1], Haytham Fayed[1,2]

Abstract

Background: Breast cancer (BC) is one of the major health problems affecting females in Egypt. Certain chromosomal loci abnormalities were proved to be associated with BC in different populations. One of them is chromosomal locus 6q25.1, that affects estrogen receptor gene (ESR) which controls ER receptor expression. Therefore, the aim of this study was to investigate locus 6q25.1 among group of Egyptian female BC patients and compare the results to healthy matched age controls. **Methods:** Formalin fixed paraffin embedded (FFPE) samples of sixty newly diagnosed BC patients were sequenced for locus 6q25.1 using genetic analyzer with capillary electrophoresis (3500 GA). The identified single nucleotide polymorphisms (SNPs) were compared to blood samples of forty controls. Realtime PCR using TaqMan probes was used for validation. **Results:** Two SNPs rs2046210 and rs2046211 were significantly associated with BC. Frequency of rs2046210-A minor allele was 30% in controls, while the frequency of rs2046211-G minor allele was 15%. Rs2046210-A allele was associated with increased risk of BC (P=0.0001), while rs2046211-G allele was associated with reduced risk of BC (P=0.021). Combined analysis of both SNPs showed that haplotype A/C was associated with increased risk of BC (P = 0.042). No significant correlation was found between rs2046210-A allele and ER status, while positive association was observed between rs204621-C allele and ER status (p= 0.005). **Conclusion:** Our data confirmed the important association between locus 6q25.1 and risk of BC in other populations. The frequencies of minor alleles of both significant SNPs will pave the way for a wider large-scale genome study and to be investigated with other BC risk factors.

Keywords: Breast cancer- single nucleotide polymorphism- chromosomal locus- risk

Asian Pac J Cancer Prev, **23 (4)**, 1359-1364

Introduction

Breast cancer (BC) is considered the most frequent cancer among females in Egypt (Ibrahim et al., 2014). According to National Cancer Registry Program (NCRP), age-standardized incidence rate of BC per 100,000 was 32% (Ibrahim et al., 2014). The early diagnosis and subsequent management of BC assisted in improving the survival rate (Cain et al., 2019; Mahvi et al., 2018; Narod et al., 2018).

Multiple risk factors contribute to BC, for example sex, family history, hormonal factors, unhealthy lifestyle and gene mutations (Majed et al., 2014). However, the genetic risk factors are still the cornerstone in the pathogenesis of BC (Collins et al., 2011). That includes the high-penetrance genes, *BRCA1* and *BRCA2*, also H*ER2, PTEN, CHEK2, TP53* and *PALB2* (Orozoc et al., 2010) .

There are several genome wide association studies (GWAS) that contributed to discovery of new genetic defects especially novel single nucleotide polymorphisms (SNPs) (Siddiq et al., 2012; Michailidou et al., 2013). Those SNPs play important roles in the risk or prognosis of BC. The GWAS assisted also in identifying the minor allele frequency (MAF) of various SNPs among populations. Knowing the critical SNPs and investigating their MAF have a great impact on understanding the incidence and outcome of BC (Siddiq et al., 2012; Michailidou et al., 2013).

Additionally, chromosomal loci that interrogate clusters of genes gathered together may imply in the pathogenesis of BC as well (Bilal et al., 2012). For example, amplified loci on chromosomes 8 and 17 predicted early relapse of estrogen receptor (ER) positive BC (Bilal et al., 2012). The chromosomal locus 6q25.1 was also associated with higher incidence of BC (Fejerman et al., 2014; Dunning et al., 2016). SNPs at 6q25.1 such as

*[1]Department of Clinical and Chemical Pathology, Faculty of Medicine, Alexandria University, Alexandria, Egypt. [2]Department of Surgical Oncology and Breast Surgery, Faculty of Medicine, Alexandria University, Alexandria, Egypt. *For Correspondence: marwahanafi@yahoo.co.uk*

(rs204621, rs2046211, rs1219648) were studied previously in different populations (Cai et al., 2011; Dunning et al., 2016; Hein et al., 2012; Zheng et al., 2013). One of the important genes that is located at chromosome 6q25.1 is estrogen receptor (ESR) gene that encodes the expression of ER receptor and was reported to play an essential role in BC development and progression (Dahlgren et al. 2008). Although the importance of 6q25.1, yet no previous studies were performed on Egyptian patients. Therefore, the aim of this study was to investigate the chromosomal locus 6q25.1 among a cohort of Egyptian BC patients.

Materials and Methods

Patients

The study was performed on 60 female BC patients who were recruited from surgical and oncology departments, (BLINDED FOR PEER REVIEW) from December 2020 until December 2021. The selection criteria for the cases included newly diagnosed and histopathological confirmed BC, without previous radiotherapy or chemotherapy and without restriction in age or disease stage. All controls were cancer-free and matched age.

DNA extraction

Genomic DNA was extracted from formalin fixed paraffin embedded samples (FFPE) that were collected from patients after obtaining formal written consent. The patients' samples were compared to blood samples of 40 matched-age healthy females.

Genotyping

Genotyping analyses were performed using the Genetic Analyzer by capillary electrophoresis (3500 GA, Applied Biosystems, Thermofisher scientific, USA) located at Cancer Genetics Unit, Molecular biology laboratory, Clinical pathology department, (BLINDED FOR PEER REVIEW).

DNA quality was assessed using a Nanodrop Spectrophotometer ND-2000 (Thermo Scientific, Wilmington, DE, USA) and by agarose gel electrophoresis assay to ensure genomic integrity. The genomic region was amplified using primers as previously described (Nhan et al., 2018) (Figure 1). Exclusion criteria for genotyped SNPs was deviation from Hardy–Weinberg equilibrium (HWE, $P<0.05$) in the controls. Therefore, two SNPs were subjected to statistical analysis i.e. rs2046210 (Figure 2) and rs2046211 (Figure 3). Both SNPs rs2046211 and rs2046210 are located in an intergenic region 29 kb upstream of the ESR1 gene. The SNP rs2046210 was also validated using TaqMan Real Time PCR assay (assay ID, C__12034236_10, Applied Biosystems) on Stratagene (Figure 2). The concordance rate for both techniques was 98.0%.

Results

The demographic characteristics of cases and controls are presented in Table 1. No significant differences were found between cases and controls in age or governate distribution (urban or rural). More cases had a family history of any cancer than did controls ($P <0.001$).

The genotype distribution of rs2046210 was consistent with Hardy-Weinberg equilibrium in both cases and controls (supplementary Table 1). The minor A allele frequency among controls was 30% (Table 2). Comparing AA+AG genotypes versus (vs) GG genotype showed increased risk of BC (OR = 3.27, 95% CI = 1.37 to 7.83, p= 0.0078) (Table 2). Comparing to G allele, A allele was significantly associated with an increased risk of BC (chi square test, $P < 0.001$, OR=3.3810, 95% CI 1.85 to 6.17, P=0.0001) (Table 2).

Most of the patients were ER positive (50 cases, 83.3%), while ER negative were only 10 cases (10 cases, 16.7%). No significant association was found between A or G allele and ER status (supplementary Table 2). Notably, AA genotype was observed in comparable percentages in pre- and post-menopausal cases (18.3%, 21.6% respectively, supplementary Table 3), that excluded an association between AA genotype and menopausal status.

For genotype rs2046211, the SNP was consistent with Hardy-Weinberg equilibrium in both cases and controls (supplementary Table 4). The minor G allele frequency among controls was 15% (Table 3). Comparing to CC

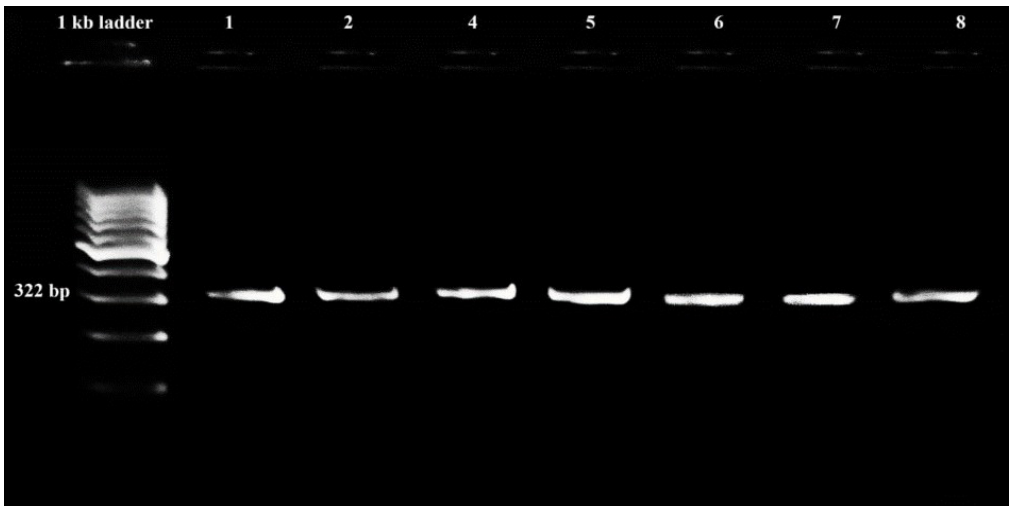

Figure 1. Gel Electrophoresis of the Amplified Chromosomal 6q25.1 Locus. The lane on the left represents 1-KB ladder, lanes (1-8) contain amplified patients' samples by PCR. The product size is 322bp

Table 1. Comparison between the Two Studied Groups According to Demographic Data

	Breast cancer (n = 60)	Control (n = 40)	Test of significance (p value)
Age (years)			
Min. – Max.	37.0 – 71.0	42.0 – 67.0	t (0.903)
Mean ± SD.	52.20 ± 8.38	52.0 ± 7.33	
Median (IQR)	51.50 (47.0 – 58.0)	50.50 (45.5 – 58.0)	
Governorate			
Urban	42 (70%)	40 (100%)	χ2 (0.585)
Rural	18 (30%)	10 (25%)	
Family History			
Negative	6	10	χ2 (<0.001*)
Positive	54	90	
Marital status			
Married	58 (96.7%)	40 (100%)	^{FE}p= 0.515
Widow	2 (3.3%)	0 (0%)	
Menstrual history			
Post-Menopausal	32 (53.3%)	26 (65%)	χ2 (0.247)
Pre-Menopausal	28 (46.7%)	14 (35%)	
Parity			
Nullipara	4	6.7	^{FE}p= 0.148
Multipara	56	93.3	
Breast feeding			
Negative	4	6.7	^{FE}p= 0.148
Positive	56	93.3	

IQR, Inter quartile range; SD, Standard deviation; t, Student t-test; χ2, Chi square test; FE, Fisher Exact; p, p value for comparing between the studied groups; *, Statistically significant at p ≤ 0.05

genotype, patients with CG had a reduced risk of BC (OR= 0.2593, CI=0.0879 to 0.7643, P=0.014) (Table 3). Using chi-square (χ2) test showed a significant C/G allelic difference (P=0.015) (Table 3). The G allele showed significant association with reduced risk of BC (OR= 0.2982, CI=0.11 to 0.83, P=0.021) (Table 3). Neither patient nor control had GG genotype in this present study.

Concerning ER status, significant association was found between rs2046211-C allele and ER positive cases (P=0.005). G allele was present in only 4 ER negative cases out of total 10 ER negative patients (Table 4).

Combined analysis of rs2046210-AA allele and the rs2046211-CC allele (haplotype AC) showed that A/C haplotype was more frequent in cases than controls (18.3%, 15% respectively) with an increased risk of BC (OR=5.5, 95% CI= 1.0645 to 28.4164, P = 0.042) (Table 5).

Statistical analysis

The chi-square (χ^2) test was used to test for deviation from Hardy-Weinberg equilibrium (HWE). Odds ratio (OR) with 95% confidence interval (CI) was used to assess the strength of association between rs2046210 or rs2046211 and BC risk. We used allele contrast AA+AG vs. GG for rs2046210 and CG+GG vs CC for rs2046211 to examine the relationship between risk allele/ genotypes and BC.

Table 2. Comparison between the Two Studied Groups According to rs2046210

	Breast cancer (n = 60) n (%)	Control (n = 40) n (%)	χ2 (p)	OR	95% CI
rs2046210					
GG®	13 (21.7)	19 (47.5)	0.001*	3.27	1.37 to 7.83
AG	23 (38.3)	18 (45.0)			
AA	24 (40.0)	3 (7.5)			
Allele					
G®	49 (40.8)	56 (70.0)	<0.001*	3.38	1.85 to 6.17
A	71 (59.2)	24 (30.0)			

®, Reference Genotype; χ2, Chi square test; p, p value for comparing between the studied groups; *, Statistically significant at p ≤ 0.05; OR, odds ratio; CI, confidence interval

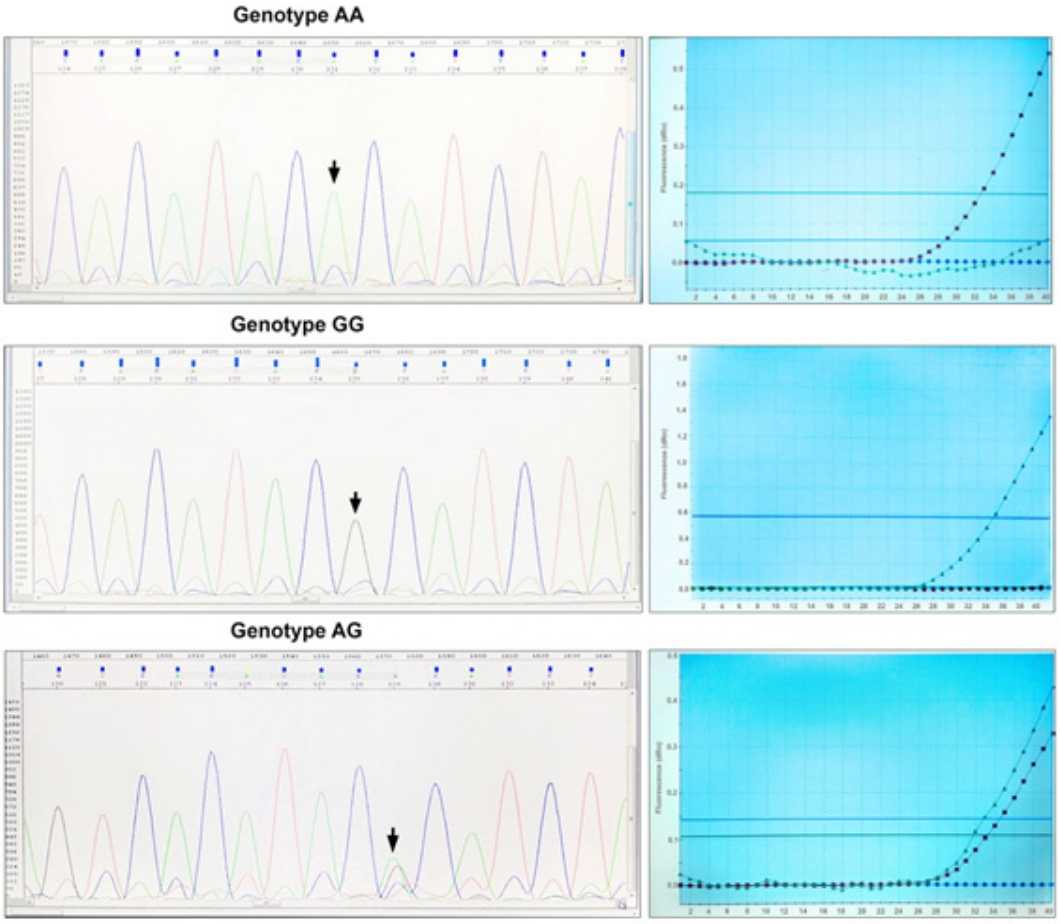

Figure 2. Sequencing Chromatogram of each rs2046210 Genotype and Realtime PCR Amplification Curves

Table 3. Comparison between the Ttwo Studied Groups According to rs2046211

	Breast cancer (n = 60) n (%)	Control (n = 40) n (%)	(χ2) p	OR	95% CI
rs2046211					
CC®	54 (90.0)	28 (70.0)	0.011*		
CG	6 (10.0)	12 (30.0)			
GG	0 (0.0)	0 (0.0)		0.26	0.088 to 0.76
Allele					
C®	114 (95.0)	68 (85.0)	0.015*	0.298	0.11 to 0.83
G	6 (5.0)	12 (15.0)			

®, Reference Genotype; χ2, Chi square test; p, p value for comparing between the studied groups; *, Statistically significant at p ≤ 0.05; OR, odds ratio; CI, confidence interval

Table 4. Relation between rs2046211 with Hormonal Receptors in Breast Cancer Group (n = 60)

	rs2046211		Test of significance (p)
	CC (n = 54) n (%)	CG (n = 6) n (%)	
ER			
Negative	6 (11.1)	4 (66.7)	FEp= 0.005*
Positive	48 (88.9)	2 (33.3)	
PR			
Negative	20 (37.0)	4 (66.7)	FEp= 0.206
Positive	34 (63.0)	2 (33.3)	
HER-2			
Negative	48 (88.9)	4 (66.7)	FEp= 0.178
Positive	6 (11.1)	2 (33.3)	
Combinations			
-ER,-PR,-HER-2	4 (7.4)	2 (33.3)	
-ER,-PR,+HER-2	2 (3.7)	2 (33.3)	
+ER,-PR,-HER-2	14 (25.9)	0 (0.0)	MCp=0.020*
+ER,+PR,-HER-2	30 (55.6)	2 (33.3)	
+ER,+PR,+HER-2	4 (7.4)	0 (0.0)	

χ2, Chi square test; MC, Monte Carlo; FE, Fisher Exact; p, p value for comparing between different genotypes

Discussion

BC is multidisciplinary disease; various risk factors contribute to its incidence including genetic factors. Chromosomal locus 6q25.1 was proved to have an association with BC risk in different ethnic groups (Wu et al., 2013). However, yet, there has not been such association studied concerning Egyptian population.

In the present study we aimed to investigate chromosomal locus 6q25.1 among a cohort of Egyptian female BC patients. The minor allele A of SNP rs2046210 had a significant association with BC. The frequency of A allele was 30% among controls. This frequency is very close to its percentage in Asian population

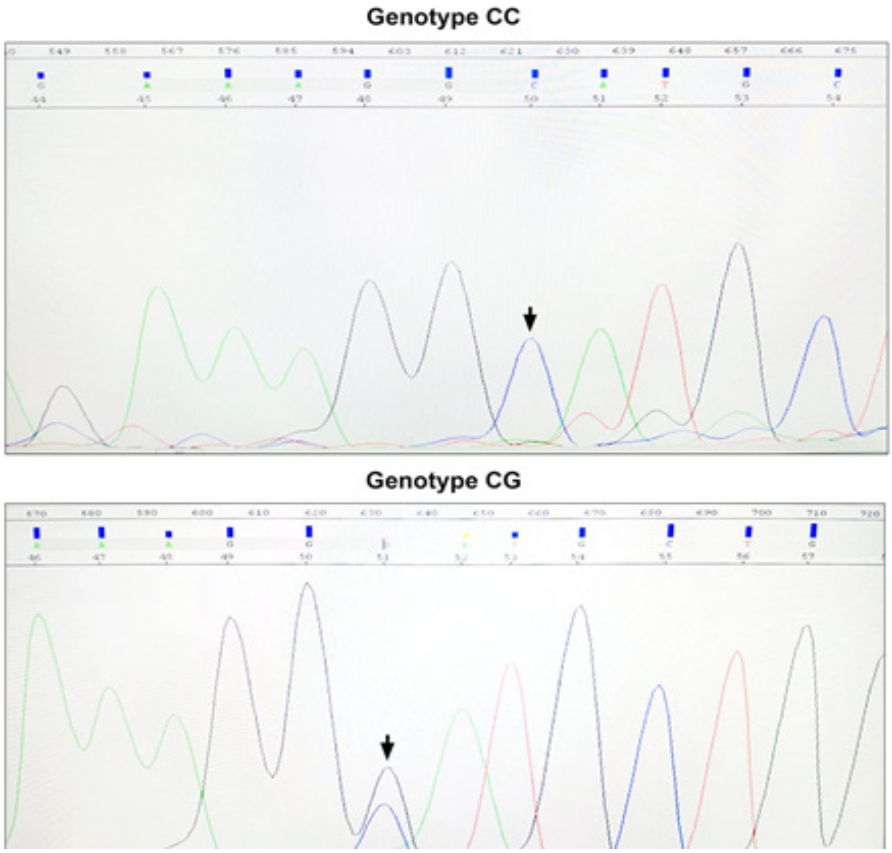

Figure 3. Sequencing Chromatogram of rs2046211 Genotypes i.e. CC and CG

Table 5. Combined Analysis of rs2046210 and rs2046211

Genotypes	AA/CC®	GG/CC®	OR	95% CI
Patients	11	6	5.5	1.07 to 28.42
Controls	3	9		

®, Reference Genotype; OR, odds ratio; CI, confidence interval

(34.8 - 39.3%) (Chan et al., 2012; Han et al., 2011), 33.7% in European population and 71.6% in African and African American (Stacey et al. 2010). This difference in allele frequencies proposes that in African and African American populations, the rs2046210-A allele may contain different haplotypes accompanied by different risk associations (Wu et al., 2013). The association between A allele and BC was proved in previous studies (Garcia-Closas et al., 2013; Han et al., 2016; Michailidou et al., 2015; Zheng et al., 2009).

Another SNP has been proved recently to be associated with BC; rs2046211 (Hamdi et al., 2018; Maxwell et al., 2013; Ruiz-Narváez et al., 2013; Ruiz-Narváez et al., 2013). The SNP showed to have a protective effect (Maxwell et al., 2013). In the current study, the minor allele G was associated with reduced risk of BC as was previously reported (Ruiz-Narváez et al., 2013). Using the human genomic variant search engine (Varsome) (Kopanos et al., 2018), the G allele frequency in our study (i.e. 15%) was most comparable to G allele frequency in African population (10.7%). However, G allele frequency in other populations was less e.g. 1.79% in East Asian, 2.04% in European (Finnish) and 2.27% % in European (non-Finnish) (Kopanos et al., 2018).

Combined analysis of both rs2046210 and rs2046211 showed that rs2046210-Allele A/ rs2046211 Allele C was associated with increased risk of BC as was previously reported {Ruiz-Narváez et al., 2013).

Limitations of the study: the random selection of cases with small number of patients might affect the expected correlation between A and ER negative status {Ruiz-Narváez et al., 2013). Similarly, that may affect potential association between A allele (rs2046210) and menopausal status (Li et al., 2011). Although the risk association between alleles and BC were significant, yet the sample size should be increased to be more confident.

Collectively, investigating chromosomal locus 6q25.1 revealed 2 SNPs that were associated with BC risk; rs2046210 and rs2046211. Our data will also assist in selection of SNPs for a larger genome-wide study in Egyptian population and in comparing the MAF with other populations (GWAS).

List of Abbreviations

BC: Breast cancer, SNP: single nucleotide, MAF: minor allele frequency, ESR: estrogen receptor gene, ER: estrogen receptor, HWE: Hardy-Weinberg equilibrium, Varsome: human genomic variant search engine

Author Contribution Statement

MHS, HF, MMR and SM conceived and designed the study. MHS, NAES and SOM did the practical work, statistical analysis of the data and writing the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We are thankful to all patients who participated in this study. The work was partially funded by Science, Technology & Innovation Funding Authority (STDF) under grant (22832). The funding bodies neither had role in the design or implementation of the experiment nor in the collection, analysis, or interpretation of the data or the writing of the manuscript.

The study is a part of theses approved by local ethics committee of Faculty of Medicine, Alexandria University in partial fulfilment of the requirements for the degree of Master of science in Clinical Pathology.

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Conflicts of interest

The authors declare no conflicts of interest.

References

Bilal E, Vassallo K, Toppmeyer D, et al (2012). Amplified loci on chromosomes 8 and 17 predict early relapse in ER-positive breast cancers. *PLoS One*, **7**, e38575.

Cai Q, Wen W, Qu S, et al (2011). Replication and functional genomic analyses of the breast cancer susceptibility locus at 6q25.1 generalize its importance in women of chinese, Japanese, and European ancestry. *Cancer Res*, **71**, 1344-55.

Cain EH, Saha A, Harowicz MR, et al (2019). Multivariate machine learning models for prediction of pathologic response to neoadjuvant therapy in breast cancer using MRI features: a study using an independent validation set. *Breast Cancer Res Treat*, **173**, 455-63.

Chan M, Ji SM, Liaw CS, et al (2012). Association of common genetic variants with breast cancer risk and clinicopathological characteristics in a Chinese population. *Breast Cancer Res Treat*, **136**, 209-20.

Collins A, Politopoulos I (2011). The genetics of breast cancer: risk factors for disease. *Appl Clin Genet*, **4**, 11-9.

Dahlgren A, Lundmark P, Axelsson T, Lind L, Syvänen AC (2008). Association of the estrogen receptor 1 (ESR1) gene with body height in adult males from two Swedish population cohorts. *PLoS One*, **3**, e1807.

Dunning AM, Michailidou K, Kuchenbaecker KB, et al (2016). Breast cancer risk variants at 6q25 display different phenotype associations and regulate ESR1, RMND1 and CCDC170. *Nat Genet*, **48**, 374-86.

Fejerman L, Ahmadiyeh N, Hu D, et al (2014). Genome-wide association study of breast cancer in Latinas identifies novel protective variants on 6q25. *Nat Commun*, **5**, 5260.

Garcia-Closas M, Couch FJ, Lindstrom S, et al (2013). Genome-wide association studies identify four ER negative-specific breast cancer risk loci. *Nat Genet*, **45**, 392-8.

Hamdi Y, Ben Rekaya M, Jingxuan S, et al (2018). A genome wide SNP genotyping study in the Tunisian population: specific reporting on a subset of common breast cancer risk loci. *BMC Cancer*, **18**, 1295.

Han MR, Long J, Choi JY, et al (2016). Genome-wide association study in East Asians identifies two novel breast cancer susceptibility loci. *Hum Mol Genet*, **25**, 3361-71.

Han W, Woo JH, Yu JH, et al (2011). Common genetic variants associated with breast cancer in Korean women and differential susceptibility according to intrinsic subtype. *Cancer Epidemiol Biomarkers Prev*, **20**, 793-8.

Hein R, Maranian M, Hopper JL, et al (2012). Comparison of 6q25 breast cancer hits from Asian and European Genome Wide Association Studies in the Breast Cancer Association Consortium (BCAC). *PLoS One*, **7**, e42380.

Ibrahim AS, Khaled HM, Mikhail NN, Baraka H, Kamel H (2014). Cancer incidence in egypt: results of the national population-based cancer registry program. *J Cancer Epidemiol*, **2014**, 437971.

Kopanos C, Tsiolkas V, Kouris A, et al (2018). VarSome: the human genomic variant search engine. *Bioinformatics*, **35**, 1978-80.

Li G, Xiang YB, Courtney R, et al (2011). Association of a single nucleotide polymorphism at 6q25.1,rs2046210, with endometrial cancer risk among Chinese women. *Chin J Cancer*, **30**, 138-43.

Mahvi DA, Liu R, Grinstaff MW, Colson YL, Raut CP (2018). Local Cancer Recurrence: The Realities, Challenges, and Opportunities for New Therapies. *CA Cancer J Clin*, **68**, 488-505.

Majeed W, Aslam B, Javed I, et al (2014). Breast cancer: major risk factors and recent developments in treatment. *Asian Pac J Cancer Prev*, **15**, 3353-58.

Maxwell KN, Nathanson KL (2013). Common breast cancer risk variants in the post-COGS era: a comprehensive review. *Breast Cancer Res*, **15**, 212.

Michailidou K, Beesley J, Lindstrom S, et al (2015). Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. *Nat Genet*, **47**, 373-80.

Narod SA (2018). Personalised medicine and population health: breast and ovarian cancer. *Hum Genet*, **137**, 769-88.

Nhan BT, Lan NTT, Thanh NTN, Thiep TV, Hue NT (2018). Primary Study of SNP rs2046210 in Vietnamese Breast Cancer Population by High-Resolution Melting Analysis (HRMA). Paper presented at: 6th International Conference on the Development of Biomedical Engineering in Vietnam (BME6); Singapore.

Orozco G, Barrett JC, Zeggini E (2010). Synthetic associations in the context of genome-wide association scan signals. *Hum Mol Genet*, **19**, 137-44.

Ruiz-Narváez EA, Rosenberg L, Yao S, et al (2013). Fine-mapping of the 6q25 locus identifies a novel SNP associated with breast cancer risk in African-American women. *Carcinogenesis*, **34**, 287-91.

Siddiq A, Couch FJ, Chen GK, et al (2012). A meta-analysis of genome-wide association studies of breast cancer identifies two novel susceptibility loci at 6q14 and 20q11. *Hum Mol Genet*, **21**, 5373-84.

Stacey SN, Sulem P, Zanon C, et al (2010). Ancestry-shift refinement mapping of the C6orf97-ESR1 breast cancer susceptibility locus. *PLoS Genet*, **6**, e1001029.

Wu X, Xu QQ, Guo L, et al (2013). Quantitative assessment of the association between rs2046210 at 6q25.1 and breast cancer risk. *PLoS One*, **8**, e65206.

Zheng W, Long J, Gao YT, et al (2009). Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. *Nat Genet*, **41**, 324-28.

Zheng W, Zhang B, Cai Q, et al (2009). Common genetic determinants of breast-cancer risk in East Asian women: a collaborative study of 23 637 breast cancer cases and 25 579 controls. *Hum Mol Genet*, **22**, 2539-50.