PHYTON-ANNALES REI BOTANICAE
ISSN: 0079-2047
Volume 66 No. 1 (2026)
https://phyton-annales.com

Genetic polymorphisms Of EGFR and CYP19 genes among Iraqi breast cancer patients

Running title EGFR and CYP19 genes in breast cancer

Hasan Alhayali[1], Mustafa Isin 2*
Department of Bioengineering, Cyprus International University, Cyprus.
[1]E mail: hasanalialhayali@gmail.com ; ORCID: https://orcid.org/0000-0001-7909-087
E mail: misin@ciu.edu.tr ; ORCID: https://orcid.org/0000-0002-8728-7571

Abstract

Objective: The study looked at the relationship between breast cancer and CYP19A1 and EGFR single nucleotide polymorphisms (SNPs).

Methods: The sequences of the CYP19A1 and EGFR genes were examined in ten breast cancer patient samples, corresponding to two genetic loci measuring 600 bp and 658 bp, respectively. We used a direct Sanger sequencing method on the amplified locus to look at the pattern of SNP variations we saw in the population we studied.

Results: The sequencing experiments revealed one variation in the intron-3 region, specifically 175T>A, identified as rs7175531, and two variations in the exon-13 region, namely 84A>G and 324G>A, designated as rs11536635 and rs2227983, respectively. The statistical analyses indicated that the identified polymorphism in both rs11536635 and rs2227983 SNPs demonstrated a significant correlation with the advancement of breast cancer.

Conclusion: rs7175531 SNP in the CYP19A1 gene and the rs11536635 and rs2227983 SNPs in the EGFR gene exhibited a notable association as potential markers linked to the development of breast cancer in the examined Iraqi samples. It showed a strong link between these SNPs and the progression of this disease.

Keywords: *Breast cancer, CYP19A1 gene, EGFR gene, DNA sequencing.*

Introduction

Breast cancer is the predominant cancer affecting women globally and is a major public health issue worldwide [1]. The etiology of breast cancer arises from the interplay of several variables, integrating environmental conditions with genetic predispositions [2].

The scientific community has demonstrated increasing interest in investigating gene polymorphisms to comprehend their implications in breast cancer development and disease progression. The precise etiology of breast cancer is uncertain; nevertheless, experts concur that genetic alterations significantly influence susceptibility to the illness [3].

Mutations in BRCA1 and BRCA2 are significantly associated with increased risks of hereditary breast and ovarian cancers. The comprehensive knowledge of genetic determinants in sporadic breast cancer remains ambiguous, highlighting the need for more research on genetic variants in cancer risk [4].

The scientific community regards CYP19A1 as noteworthy due to its role in producing the aromatase enzyme, which is a member of the cytochrome P450 superfamily.

Aromatase enzyme expression occurs in the ovaries, testes, placenta, bone tissue, adipose cells, skin tissue, and brain cells, governed by unique regulatory systems. In premenopausal women, ovarian granulosa and luteal cells synthesize
the majority of estrogens, whereas postmenopausal women derive their estrogen from peripheral tissues, such as adipose and bone. Aromatase is present in both normal and cancerous cells of the mammary epithelium and stroma [5].

The CYP19A1 gene resides on chromosome 15 band q21.1 and spans 123 kb while having 30 kb of coding sequence and 93 kb of untranslated regions. The CYP19A1 gene consists of nine coding exons together with ten untranslated exons that include I.1, I.2, 2a, I.3, I.4, I.5, I.6, I.7, If, and PII. The distal I.4 promoter serves as the main transcription site in normal breast tissue. The I.3 and PII promoters become primary transcription sites during the development of breast cancer. Scientific studies show that CYP19A1 transcripts exist at higher levels in cancerous tissue than in normal tissue. Research findings indicate elevated CYP19A1 mRNA and estrogen receptor alpha (ERα) and circulating estrogen levels in breast cancer patients and ovarian cancer patients and endometrial cancer patients [6].

The CYP19A1 rs4646 polymorphism shows clinical significance regarding letrozole treatment outcomes for patients with advanced cancer stages. Research suggests that disease development and treatment outcomes might be linked to gene variants of CYP19A1 because this gene produces aromatase which generates estrogen [7].

The EGFR gene which produces the epidermal growth factor receptor (also known as HER-1 or c-erbB-1) functions as one of four transmembrane receptor tyrosine kinases that show structural and functional similarities [8]. A ligand that binds to EGFR causes the receptor to form dimers which then activate its intracellular tyrosine kinase domain through autophosphorylation. The receptor activation triggers signaling pathways which manage basic cellular operations including proliferation and angiogenesis and programmed cell death. [9].

The genetic amplification of EGFR remains poorly understood despite its notable research interest regarding its impact on tumor behavior and response to tyrosine kinase inhibitors. Research shows that 16–36% of breast tumors express EGFR protein yet no study has measured gene amplification and mRNA levels and protein expression across the same patient population. Research has investigated EGFR amplification together with HER-2 protein levels and EGFR expression in numerous breast tumor samples [10–12].

The research investigated genetic variations between polymorphisms and mutations of CYP19A1 and EGFR genes to determine their potential role in breast cancer initiation and progression.

Methodology

Study Design

The Oncology Clinic at Merjan Hospital in Babylon Province, Iraq, took seventy-five samples of blood from cancerous breasts patients and 40 samples from healthy subjects. The study participants spanned different age groups with patients averaging 51.74 years old and controls averaging 34.5 years old. The collected samples were preserved at −20 °C until researchers conducted additional molecular tests.

Extraction of DNA for CYP19A1 and EGFR genes using conventional PCR

The SYNCTM DNA Extraction Kit extracted genomic DNA from whole blood samples. The extracted DNA received storage at −80 °C until additional procedures were performed. The Applied Biosystems thermal cycler performed polymerase chain reaction (PCR) in 25 µL reaction volumes. The reaction mixture contained 5 µL of Master Mix together with 1 µL of each primer pair:

(CYP19A1:rs10046 Forward: 5′-GCTCTGAGCCTCCCTTCTCT-3′ and Reverse: 5′-TGGGAATGTGGTGATAGTGG-3′ and EGFR Forward: 5′-CTCCCACCAGGAGGAAGAC-3′ and Reverse: 5′-AAGATACTTCCAGGAAAAGAGATTC-3) and 16.5 µL of nuclease-free water and 1.5 µL of template DNA.

The PCR amplification methodology started with a 5-minute denaturation phase at 95 °C in the initial cycle, succeeded by thirty-five cycles including denatured state at 95 °C for 30 seconds, heating at 61 °C for 30 seconds, and elongation at 72 °C for 1 minute. The procedure concluded with a 5-minute elongation at 72 °C. The PCR results were subjected to 2% agarose gel electrophoresis at 70 volts for 30 minutes, followed by inspection with a UV transilluminator to identify DNA bands.

Sequencing the PCR amplicons of CYP19A1 and EGFR genes

The sequencing business issued directives to commercially sequence the resolved PCR amplicons with both sides and reverse orientations (Macrogen Inc., Geumchen, Seoul, South Korea). The ABI (Applied Biosystem) sequence files produced clear chromatographs which were analyzed to display annotations and variations through bioinformatics analysis. The PCR fragments' virtual locations and additional information were determined by comparing local DNA sequences with the retrieved DNA sequences.

Bioinformatics analysis of the CYP19A1 and EGFR genes sequencing

The following was utilized by the application that was used to analyses the sequencing results:

a) DNASTAR laser gene Edit Seq, a program developed by DNASTAR and distributed by the company in Madison, Wisconsin, USA, was utilized to align and analyze sequencing findings of PCR products from various samples in cases where reference database sequences were accessible.

b) Using Snap Gene Viewer ver. 4.0.4 (https://www.snapgene.com), the program allotted numerical values to observed variations in sequenced samples' PCR amplicons and reference genome locations.

c) The Single Nucleotide Polymorphism database (dbSNP) was used to verify novelty (https://www.ncbi.nlm.nih.gov/snp/): The software checked the SNP position against the reference genome and evaluated its presence in the dbSNP server.

d) NCBI BLASTn engine (https://blast.ncbi.nlm.nih.gov/Blast.cgi): It used to find the similarities between the sequenced samples and the reference target sequences that were intended.
e) UniProtKB (https://www.uniprot.org/uniprot/P09211): It used to identified the function of glutathione S-transferase pi 1 proteins that encoded by CYP19A1 and EGFR genes.

Ethical approval

All procedures followed the guidelines of the sake of scientific integrity, the ethical, human and scientific aspects was taken into account in the process of collecting samples, especially since it was approved by the official authorities

Statistical analysis

Statistical analysis serves to analyze quantitative data through methods which describe data and perform basic inference for both continuous and categorical data. The process starts with data collection before testing the connection between two statistical data sets. The research presents all data through mean ± standard deviation values. The statistical analysis was conducted through SPSS version 26 by implementing dependent t-tests (two-tailed) and independent t-tests (two-tailed) for normally distributed variables while using the Mann-Whitney and Wilcoxon tests for variables that did not follow a normal distribution. The study considered any statistical significance at $P < 0.05$.

Results

The amplicon of DNA at CYP19A1 and EGFR genes

The results of multiplex PCR for detection the amplicon of DNA at CYP19A1 and EGFR genes: Gel electrophoresis revealed bands of extracted nucleic acid from CYP19A1 and EGFR genes from patients with breast cancer, Figure 1 A and B.

The sequencing of CYP19A1 gene

The CYP19A1 genetic sequences on chromosome 15 were analyzed in ten patient samples (S1 to S10) to study CYP19A1 gene genetic polymorphism. The CYP19A1 gene encodes the cytochrome P450 family 19 subfamily A member 1 protein. The protein functions as a key component in multiple metabolic processes that affect endogenous substrates (https://www.uniprot.org/uniprotkb/P10632/entry). The sequencing reactions revealed the precise genetic fragment identity following NCBI blastn (https://blast.ncbi.nlm.nih.gov/Blast.cgi) execution.

PHYTON-ANNALES REI BOTANICAE
ISSN: 0079-2047
Volume 66 No. 1 (2026)
https://phyton-annales.com

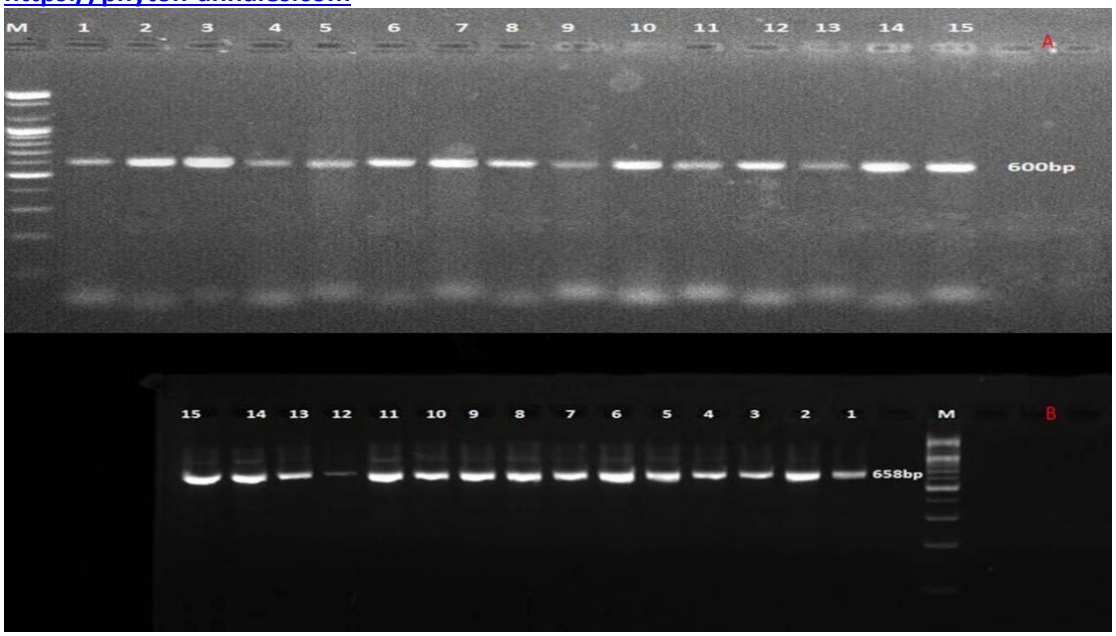

Figure 1. The conventional PCR, electrophoresis gel, M = Molecular weight marker of 100 bp; 1-15 lines of **A**. CYP19A1 gene at 600bp. **B.** EGFR gene at 658bp.

The sequencing chromatogram from the detected variation region contained both the chromatogram of the observed variation and its placements within the PCR amplicons as shown in Figure 2A. The examined samples contained an unusual SNP that replaced Thymine with Adenine at position 75 of the amplified PCR products which was referred to as T175A or 175T>A. The SNP was found in two polymorphic configurations: TT and TA. The T/T homozygous genotype was the most common genotype observed in the samples but the T/A heterozygous genotype was detected in S2, S6 and S8.

The polymorphic locus showed a high frequency in the examined samples. It is important to investigate the more details of this SNP at its specific locations within the genomic DNA sequences of the CYP19A1 gene as recorded in the dbSNP database. The positions of the targeted SNPs in relation to their recorded SNP database for the sequenced 600 bp segments were obtained from the dbSNP service for the CYP19A1 gene. https://www.ncbi.nlm.nih.gov/projects/SNP/

The distinctiveness of the identified SNP concerning the CYP19A1 dbSNP database on chromosome 15 was evaluated by graphical representations utilizing GenBank Acc. No. NC_000015.10. The dbSNP database indicated that the 175T>A SNP is already present in the genome as rs7175531, as seen in Figure 2B.

Table 1. The position and length of the 600 bp PCR amplicons used to entirely amplify the intron-3 of the *CYP19A1* gene within chromosome 15 (GenBank acc. no. NG_007982.1). And 658 bp of the exon-13 of the *EGFR* gene within chromosome 7 (GenBank acc. no. NG_00726.3).

Amplicon	locus sequences (5′ - 3′) Reference	length
sequences of DNA within the *CYP19A1* gene	*GCTCTGAGCCTCCCTTCTCTAGAGTTGACCCTGTCAGGGA GCAGAGAACAAAATTAACTAAATGTTCCTTTAAACTGGGC ATTAAAGTAATCTAAAATATCAGGCTGAATCAGGCAGAG AGACAGGCAATTATCTGGAATGTTAAAAAGAATCCAGAC ACAGAGCATTTTGTGCTAGCCCCATGCTCCCACTTCTGAT CCCACTGAGACTAGGCCCACCACACTTACCAGTGTGTAAC TATGCTCTTAAAGTCTCATTGGATATCTTATGGAATTGGG AACAGATGTGTCTGCAATGAAGCATTTTCACTGAAACCTA AAGAAGCAGTCTGAAATTTCAAGTCATTCTGAATTAATCT GATATCTGATCATTGCGGCTTTAAGGGTGGAAGGGTGGGG TTTTCTATTGGGATTATCTGAATAATTATATTATTCTCTGC TTTTGATTGAACCACAGATCTTGTTGAAGATTCTCCCGGG GTATGATGTTCCAAGAGCACTGGAGTGGTCCCCATTTTTC TCCCACTACCAGGCTCCCATAAGGGTCGAATGGGATCCAG ACAGTGGGAGAAAAATGGGGACCACTATCACCACATTCC CA**	600 bp
DNA sequences within the *EGFR* gene	*CTCCCACCAGGAGGAAGACCTGTCCTCCACTGTCAGGCA CATTTCAGTCTTCCCAGCAGCCAGCACAACTACTTTGTCCT TCCAGTCACGGTCGGCCTCTGGGAAGCCCAGTCTGTGTCC TCCTCCTTCAGGGGTAGCCAGCATGTCTGTGTCACCCAAG GTCATGGAGCACAGGGCCCCTCCCGGGAAGGTGCCGTCTC CTCCGGCCCCTCGGGTCCCTGCTCTGTCACTGACTGCTGTG ACCCACTCTGTCTCCGCAGAGGCCACAGGCCAGGTCTGCC ATGCCTTGTGCTCCCCCGAGGGCTGCTGGGGCCCGGAGCC CAGGGACTGCGTCTCTTGCCGGAATGTCAGCCGAGGCAG GGAATGCGTGGACAAGTGCAACCTTCTGGAGGGGTAGGA GGTTATTTCTTTAATCCCCTTGCGTTGATCAAAAATAAGG CTCCAGGTTGTTGTTATAGCTTTACAGGCATTCTGTTTGAT TTTCTCTTCCTTTTATTCTTTGCCCTTGGCTTTTGGAGGTTT TGGGTTTTCTGTGGGGAGACGGGAAGTTGTTTGATTGCGT TATTTTTGGCAAATTTAAGCACAATAGGAAATAAGCAAGT ATTATTGCCTAATATAATCCAATAATTTATAGAATCTCTTT TCCTGGAAGTATCTT**	658 bp

* The forward primer sequences are placed in a forward direction; ** The reverse primer sequences are placed in a reverse complement direction I

A large amount of data indicates that cytochrome P450 family 19 subgroup Member 1 is crucial in the etiology of breast cancer. Therefore, understanding the genetic variety of the CYP19A1 gene is essential for clarifying the growth of breast carcinoma. Consequently, it is essential to enhance our comprehension of the function of the CYP19A1 gene in this context. This growth has resulted from our capacity to comprehend the pattern of genetic variation within this genomic segment. The rs7175531 single nucleotide polymorphism is situated inside the CYP19A1 gene, as previously documented in our database. It was, however, deposited at a high frequency (0.311, 0.30, and 0.32, as indicated in the TOPMED, Genom AD, and ALFA programmers, respectively).
(https://www.ncbi.nlm.nih.gov/snp/rs7175531).

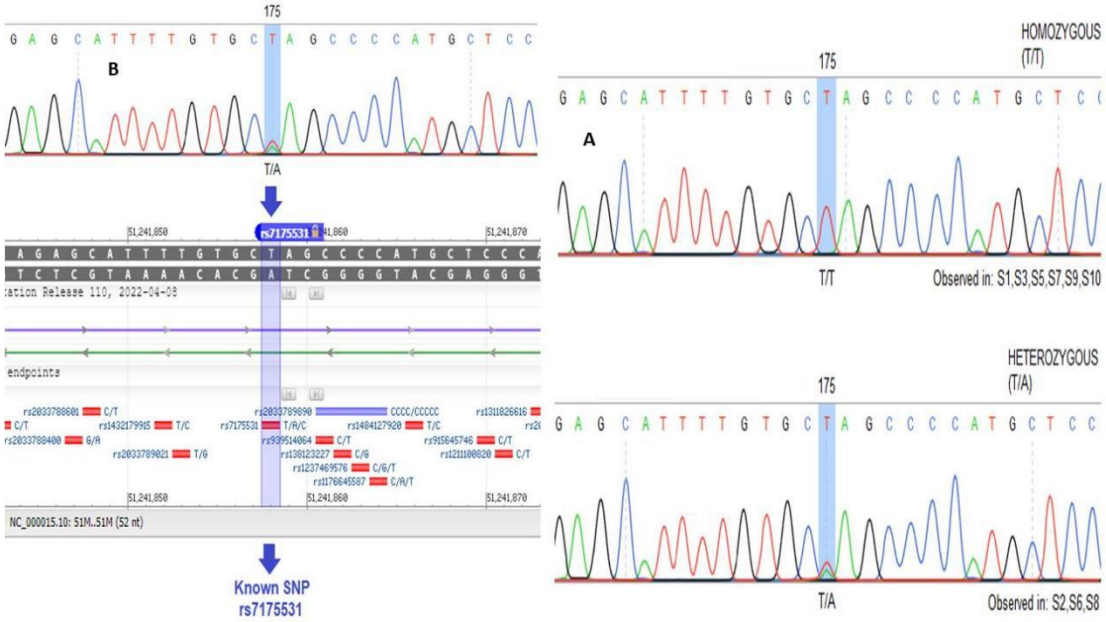

Figure 2. A. The DNA chromatogram shows the identified single nucleotide polymorphism (SNP) in the CYP19A1 gene's specific 600 bp amplicons. The detected SNP was emphasized based on its location within the PCR amplicons. B. The CYP19A1 genetic single nucleotide polymorphisms were evaluated using the dbSNP service. The designated SNP was highlighted in blue. The GenBank accession number NC_000011.10 was used to locate the emphasized substitution SNP. The relevant sequences were located in the positive strand.

Three publications documented the progression of endometrial cancer[13] and cardiovascular disease[14] and gastric cancer[15] through this SNP. Scientists have not previously linked the rs7175531 SNP with the disease development process. The examined 600 bp of the CYP19A1 gene contains more than one SNP besides rs7175531. The targeted PCR amplicons contained numerous SNPs besides the ones found in this study. Several intronic SNPs found throughout intron-3 produce different impacts on human health.

The SNPs found throughout intron-3 were documented at minimal frequencies by the most complete allele frequency data from dbSNP including Gnom AD exome, Gnom AD genome, TopMed, ALFA projects and ExAC. The extremely low occurrence of these SNPs

resulted in no genetic variation across all studied population members. These findings help explain why no polymorphism exists for these SNPs. A research investigation detected high rs7175531 SNP prevalence throughout the Iraqi population under study. The study results validate the documented high frequency of deposition in the dbSNP database .

The rs7175531 SNP third variant TT was not detected in the analyzed population through the research. The rs7175531 SNP displayed A/T heterozygous status in patient samples S2, S6 and S8 while other samples lacked this heterozygous form. The research data suggested that individuals who carry this heterozygous variation of the SNP have a small chance of developing breast cancer.

The rs7175531 SNP showed medical relevance in this study yet it remained unrecorded in the ClinVar database because no researchers previously submitted it to the server. The CYP19A1 gene contains the detected rs7175531 SNP within its targeted region. Two other genes were discovered to include this SNP within their sequences, namely MIR4713HG and MIR4713 genes. The SNP appears in three independent gene sequences at three different locations throughout the genome. These three genes function together because they share overlapping positions with the rs7175531 SNP.

The research outcome emphasized how this SNP in the CYP19A1 gene outshines other SNPs found in the amplified locus because of its breast cancer connection. The change would result in breast cancer patients producing different modified substances. Research has not established any link between this SNP and breast cancer diagnosis or prognosis before this study. This research represents the first investigation into genetic variations of CYP19A1 gene that might link to breast cancer development in Iraqi individuals.

Small sample size was a significant obstacle to establishing any large-scale statistical correlation with breast cancer development. The limited sample size in the studied population may account for the lack of the homozygous A/A pattern among the few samples examined. The diminished size of the studied group may be the primary barrier to adequately recognizing the significance of this SNP in relation to the illness. Nevertheless, the diseased condition of the examined region concerning the documented phenotypic criteria cannot be disregarded in our analysis.

Consequently, an increased population size is strongly advised in these instances to provide an accurate evaluation of the rs7175531 SNP in the diagnosis and prognosis of breast cancer. Additionally, amplifying and sequencing other genomic segments within the CYP19A1 gene is crucial to evaluate the extent of their association with this pathological condition. These segments may signify coding and noncoding sequences, which may be examined and subjected to extensive sequencing procedures to provide a more thorough understanding of the genotype-phenotype relationship in the studied population.

In all racial and cultural groupings, breast cancer continues to be the most common type. Breast cells experience changes in their development and growth patterns when exposed to estrogens. Scientific research identifies genetic polymorphisms responsible for estrogen synthesis and breakdown as potential factors that contribute to breast cancer development. Research since the early 1970s shows that breast cancer develops mainly from higher exposure to estrogens and progesterone [16].

The discovery by Siiteri revealed that adipose tissues produce a major portion of endogenous estrogens while obesity emerged as a risk factor[17]. The study by O'Neill and Miller proved that breast tissue estrogen production creates high hormone levels which drive cell expansion and accelerate breast cell proliferation and cancerous transformations[18].

A transition mutation exists in the 3′ UTR of the CYP19A1 gene which generates some degree of polymorphism. Scientists have recently classified this harmless genetic variation as the most recent known form of aromatase deficiency. Recent research by Farzaneh et al. revealed that the rs10046 polymorphism demonstrates notable genetic and allelic variations among Iranian women with and without breast cancer[19].

Countless scientific research has examined the association between the rs10046 polymorphism and breast cancer risk; nevertheless, their results have been conflicting among various demographic groups [20]. Zhang et al. investigated Chinese women with estrogen receptor-positive (ER+) breast cancer and discovered that the T allele heightened risk for breast cancer, particularly in before menopause individuals [21].

Yoshimoto et al. performed research on Japanese women that indicated the C allele might function as an indicator of breast cancer risk [22]. The rs10046 polymorphism failed to demonstrate any association with breast cancer susceptibility according to research conducted by Ghisari et al. among Greenlandic Inuit women. Pineda et al. found potential evidence that the C allele linked to higher breast cancer risk in Spanish women but the connection vanished during their subsequent meta-analysis [23].

The researchers proposed that rs10046 could produce different effects based on the combination with other genetic variations [24]. Zins et al. discovered that patients carrying the TT genotype typically develop breast cancer before age 50 which indicates a possible link to early age of onset rather than cancer risk [25]. Research studies have shown that rs10046 genotypes regulate different biochemical markers through their influence on estrogen and estrogen to testosterone ratios [26].

The CYP19A1 gene polymorphism rs4646, situated in its 3' untranslated region (3' UTR), has a more pronounced association with the onset of breast cancer. The A allele of rs4646 was shown to be linked to increased breast cancer risk, reduced tumor sizes, and poorer histological grades in Swedish patients. Shao et al. discovered in their research that Chinese patients with the AA genotype experienced better disease-free survival outcomes than those with CC or CA genotypes [27].

Fasching et al. also supported the notion that rs4646 may influence disease-free survival duration. Moreover, several single nucleotide polymorphisms (SNPs)—including rs4646, rs10046, rs700518, rs749292, rs2289106, rs3759811, and rs4775936—have been associated with changes in lipid profiles, such as reduced triglyceride levels and variable effects on HDL-C, particularly in patients treated with letrozole [28]

The sequencing of EGFR gene

This study examined the genetic sequences of the *EGFR* gene located on chromosome 7 across ten samples, labeled S1 through S10. These included five patient samples (S1–S5) and five control samples (S6–S10), all analyzed to investigate genetic polymorphisms within the *EGFR* gene. This gene encodes the epidermal growth factor receptor, a protein essential

for cellular functions through its interaction with several ligands to facilitate cellular responses to external stimuli.

The NCBI BLASTn program (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was employed to analyze sequencing findings, revealing up to 99% sequence similarity with reference sequences. The aligned sequences encompassed the complete exon 13 region of the EGFR gene. The obtained sequences were aligned with the GenBank reference sequence (NG_00726.3) to determine the exact positions and attributes of the PCR-amplified segments, as illustrated in Figure 3.

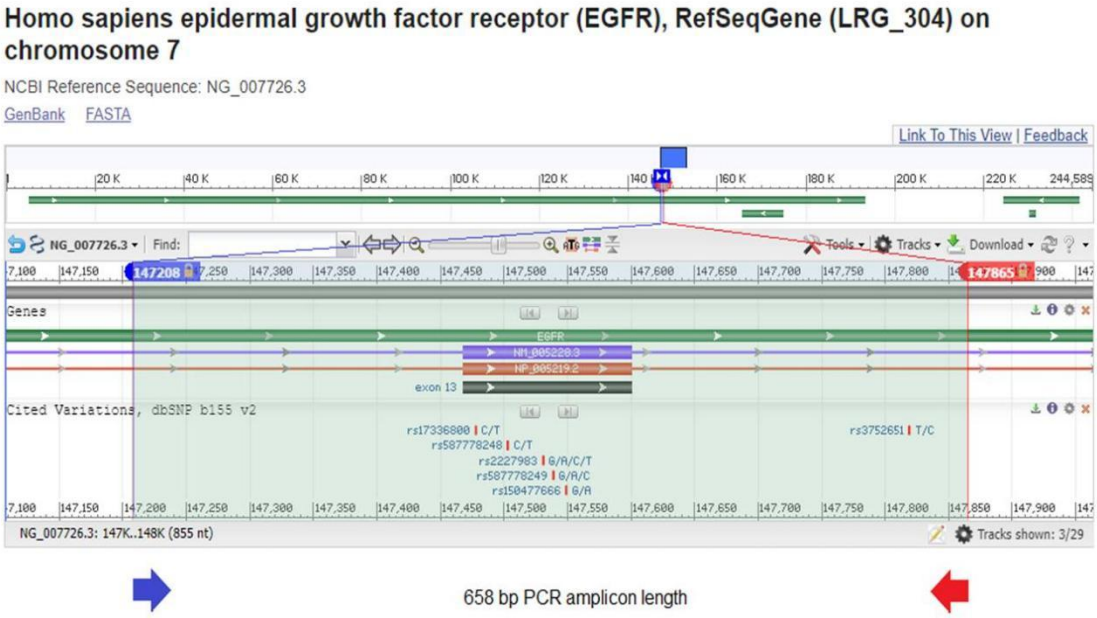

Figure 3. The precise location of the obtained 658 bp amplicons that completely included exon-13 of the EGFR gene on chromosome 7 (GenBank accession number NG_00726.3). The greenish-blue arrow indicates the starting place of this amplicon, while the red arrow denotes its termination.

The forward and reverse primer locations within the amplified regions were mapped after the 658 bp amplicon sequences within chromosome 7 were aligned, and the results are listed in Table 1. The variations and their precise locations within the PCR amplicons were highlighted in detailed sequencing chromatograms that corresponded to the identified variant regions. Among the examined samples, two notable single nucleotide polymorphisms (SNPs) were found. As seen in Figure 4A, the first SNP featured an adenine-to-guanine substitution at position 84 of the PCR product (84A>G), resulting in three genotypic forms: homozygous wild-type (AA), heterozygous (AG), and homozygous mutant (GG). The AG and GG genotypes were only found in samples S3, S4, and S5, whereas the AA genotype predominated.

A guanine-to-adenine substitution at position 324 (324G>A) was the second SNP found. As seen in Figure 4B, it displayed three genotypes, similar to the first variant: homozygous wild-type (GG), heterozygous (GA), and homozygous mutant (AA). Once more,

the most common genotype was GG, though samples S3, S4, and S5 also contained the GA and AA genotypes.

Interestingly, both SNPs showed the same zygosity patterns in all of the examined samples. The study group's comparatively high frequency of these polymorphisms highlights the significance of additional research into their genomic context. The identified SNPs were cross-referenced with the dbSNP database using the sequenced 658 bp regions in order to pinpoint their exact locations within the EGFR gene (https://www.ncbi.nlm.nih.gov/projects/SNP/).

The discovered SNPs were compared to entries in the dbSNP database for chromosome 7 using GenBank accession number NC_000007.14 to confirm whether they represent distinct variations within the EGFR gene by graphical analysis. As shown in Figures 5A and 5B, this analysis confirmed that the two variants—84A>G and 324G>A—correspond to previously identified SNPs, specifically rs11536635 and rs2227983, respectively.

The occurrence and advancement of breast cancer have been significantly associated with the epidermal growth factor receptor (EGFR). Identifying the genetic variability of the EGFR gene is essential for elucidating its role in the progression of diseases. It is becoming more and more obvious that EGFR should be the main focus of studies on breast cancer as knowledge about the genetic variability within this gene grows. The validity of the discovered polymorphisms is supported by the identification of rs11536635 and rs2227983 within the examined EGFR region, which is in accordance with their current records in the dbSNP repository.

According to information from the dbSNP database, rs11536635 is classified as an intronic variant. Interestingly, the ALFA (Allele Frequency Aggregator) project reports that it has a relatively high allele frequency of 0.24, indicating that it might be a common variant with possible biological importance (https://www.ncbi.nlm.nih.gov/snp/rs11536635).

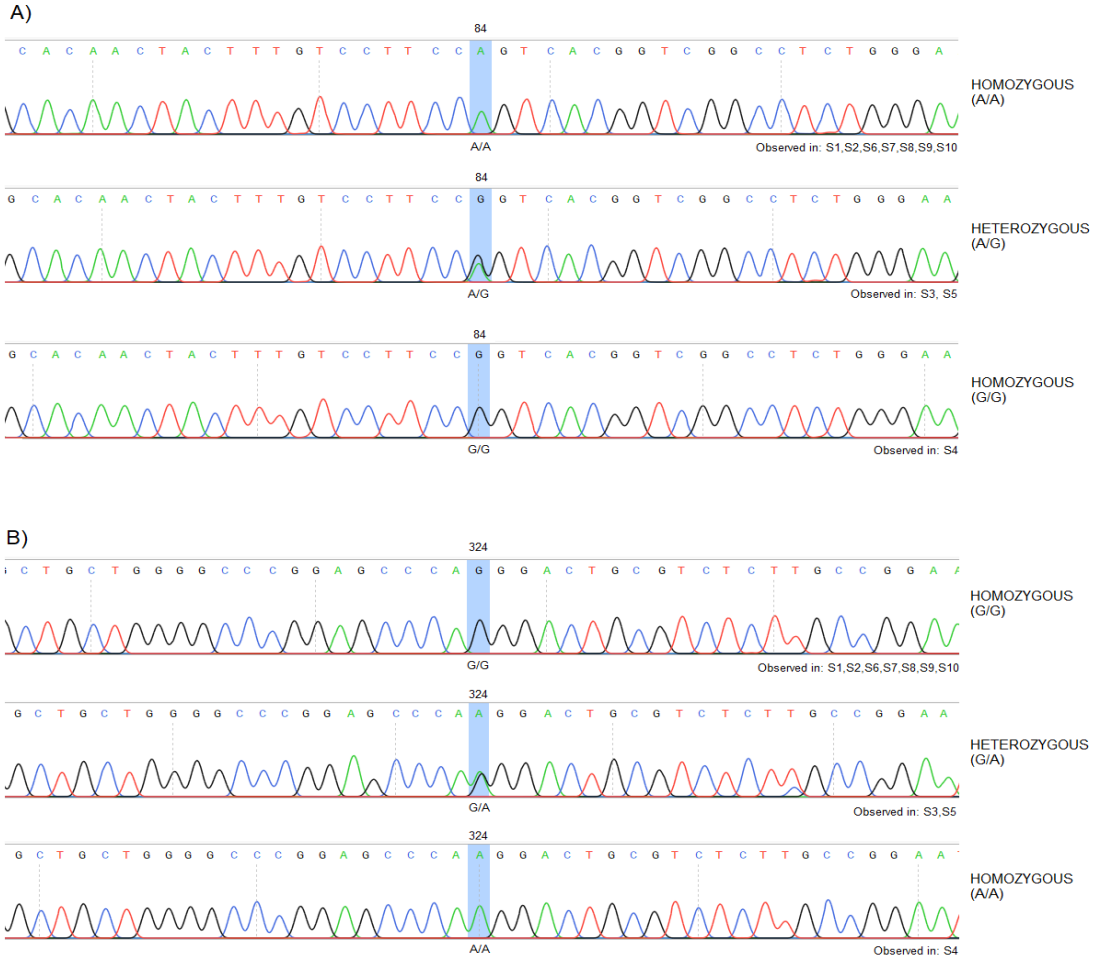

Figure 4. A and B. The pattern of the detected single nucleotide polymorphism (SNP) within the DNA chromatogram of the targeted 658 bp amplicons of the *EGFR* gene. The identified SNPs were highlighted according to their positions in the PCR amplicons.

The research findings demonstrate that rs11536635 SNP leads to breast cancer development despite previous studies showing no link between this genetic variant and disease progression. The rs11536635 variant appears in patient samples at high rates but remains absent from control samples which supports the observed correlation. The ClinVar database lacks any previous clinical submissions of rs11536635 because this variant shows clinical significance in this particular study. The rs2227983 SNP exists at a 0.26 allele frequency according to ALFA project and dbSNP (https://www.ncbi.nlm.nih.gov/snp/rs2227983) data which classifies it as a missense mutation. The SNP results in a substitution of Arg to Lys at position 254 (NP_001333870.1:p.Arg254Lys) [29].

PHYTON-ANNALES REI BOTANICAE
ISSN: 0079-2047
Volume 66 No. 1 (2026)
https://phyton-annales.com

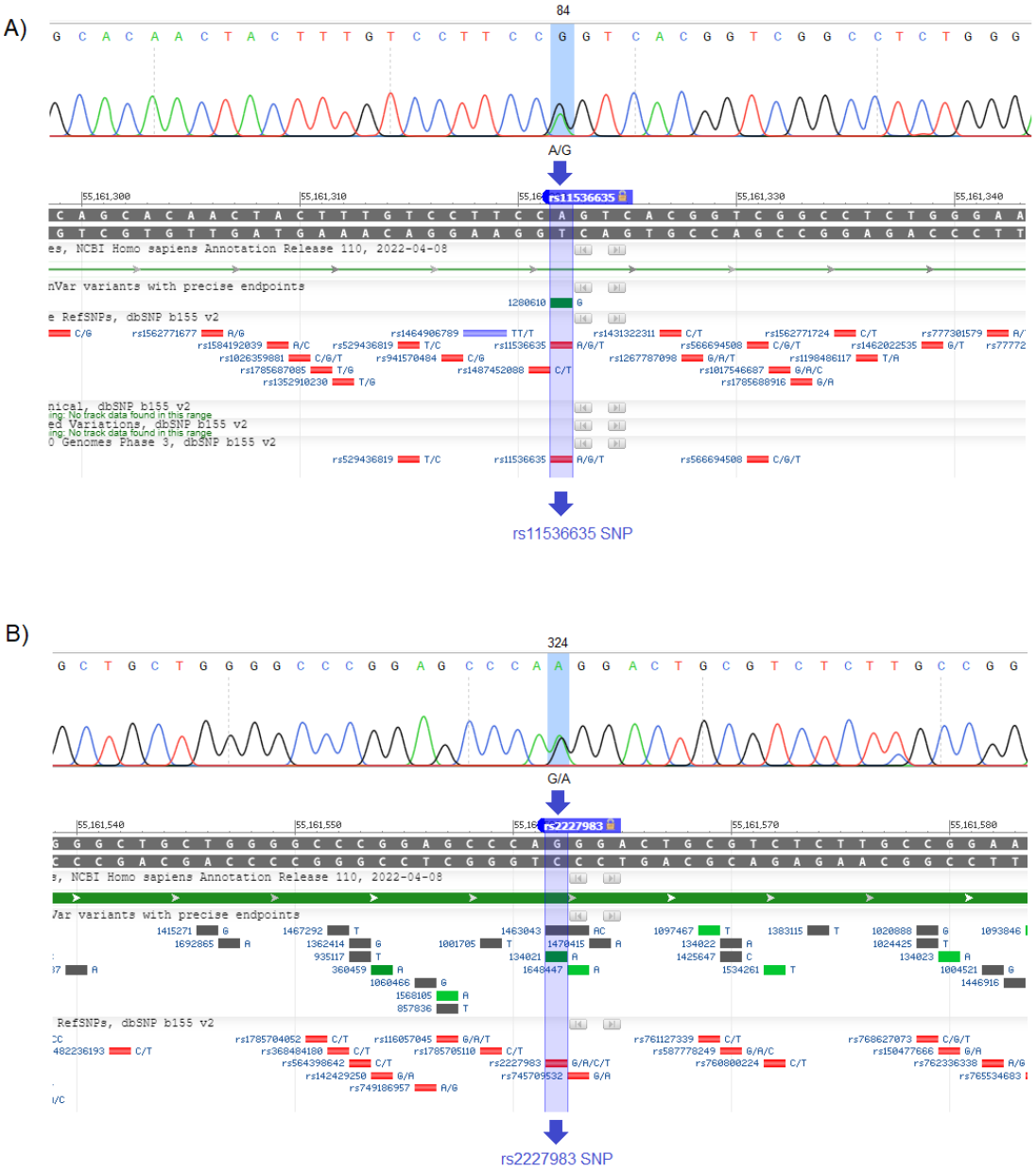

Figure 5 Both A and B. The SNP's assessment of the originality of EGFR genetic single nucleotide polymorphisms utilising the dbSNP website. The designated SNPs were highlighted in blue. The marked substitution SNPs were positioned using GenBank accession number NC_000007.14. The locations of the targeted sequences were identified on the positive strand of the referenced genomic DNA.

Numerous studies have linked this SNP to a range of metabolic disorders and malignancies, including cervical cancer, oral squamous cell carcinoma, lung cancer, and colorectal carcinoma (https://www.ncbi.nlm.nih.gov/snp/rs2227983 publications)[30,31]. The rs2227983 SNP has a variety of reported clinical implications and has been recorded in the ClinVar database due to its medical significance. In this study, the rs2227983 variant also showed a consistent and noteworthy association with the onset and progression of breast

cancer, much like rs11536635. Crucially, the 658 bp region of the EGFR gene that was examined contained more variants than just these two SNPs.

The same PCR amplicons also contained a number of other SNPs, mainly in exon 13 and the upstream and downstream regions that surround it. The majority of these SNPs occur at very low allele frequencies, according to data from large-scale population databases, include the GnomAD genome, GnomAD exome, ALFA, TopMed, and ExAC, even though they may have different biological effects.

This is supported by the fact that no polymorphisms were found at these extra SNP loci in the current study population. The lack of these variants in the examined samples may be due to their rarity. Additionally, it's possible that the detection of less prevalent polymorphisms was hampered by the study's comparatively small sample size. Therefore, more research with bigger sample sizes is required to draw firmer conclusions about the distribution and clinical significance of rs11536635 and rs2227983.

Notwithstanding this drawback, both SNPs were completely absent in the control group and primarily found in patient samples (S3, S4, and S5). The possible connection between these variations and the onset of breast cancer is supported by this clear pattern. As far as we are aware, this is the first study to raise the possibility of a link between rs11536635 and rs2227983 and the development and risk of breast cancer in an Iraqi population.

The research establishes new correlations between rs11536635 and rs2227983 SNPs and breast cancer disease in the Iraqi population being studied. The pathological condition of the examined region concerning the documented phenotypic characteristics remains ambiguous. The evaluation of rs11536635 and rs2227983 SNP in breast cancer diagnosis and prognosis requires a bigger population size to achieve accurate results.

Essential amplification and sequencing procedures are needed to evaluate clinical condition correlations with other genetic fragments located within the same gene. The genotype-phenotype relationship in the studied population needs further investigation and extensive sequencing procedures to establish these segments contain coding and noncoding sequences.

The Nomura study indicated that the EGFR-rs1050171G/A and EGFR-rs2227983G/G SNPs independently elevate the risk of non-small cell lung cancer and function as diagnostic indicators for its early identification **32**. The study by Connor et al. examined EGFR SNPs in connection to breast cancer risk and identified associations between these SNPs and the risk of breast cancer. The data did not demonstrate any statistical significance after doing multiple comparison analysis [33]. The research conducted by Abdraboh et al. indicated that rs1136201 (HER2 I655V) and rs11543848 (EGFR R497K) were associated with an elevated risk of breast cancer [34].

Escórcio-Dourado et al. identified that the HER2 gene polymorphism Rs1136201 (Ile 655 Val) serves as a possible risk factor for breast cancer in Brazilian individuals. Research has identified genetic polymorphisms in the EGFR family and alterations in the EGF signaling pathway as risk factors for cancer, including breast cancer. EGF is essential in carcinogenesis, proliferation, metastasis, and angiogenesis. The EGF signaling pathway may be significant in forecasting cancer onset and progression. The EGFR promoter contains identified SNPs while polymorphic microsatellite sequences within the first intron control

gene expression. The transcription factor binding site in the promoter region experienced enhanced EGFR production due to specific SNPs[35].

Conclusions

The rs7175531 SNP located in the CYP19A1 gene demonstrated a substantial link to breast cancer development in the studied population. The rs11536635 and 2227983 SNPs in the EGFR gene were found to contribute to the development of breast cancer. The genetic mutations and polymorphisms in CYP19A1 and EGFR genes show potential links to breast cancer development in female patients.

Acknowledgements

The study is funded by the Department of Bioengineering at Cyprus International University, Cyprus.

References

`1. Hammadi AH, Ali SH. Trastuzumab Response in Iraqi Women with Her2 Positive Breast Cancer: The Role of PIK3CA Oncogene Mutations at E542k and E545k. Academic J Health Sciences 2024; 39 (2):46-52.
2. Siegel R, DeSantis C, Virgo K, Stein K, Mariotto A, Smith T, et al. Cancer treatment and survivorship statistics, 2012. CA Cancer J Clin. 2012; 62:220–41.
3. Movchan OV, Smolanka II, Lyashenko AO, Loboda AD, Dosenko IV, Ivankova OM. Tumor infiltrating lymphocytes density as the chemoresistance and relapse factor for advanced breast cancer patients. Academic J Health Sciences 2024;39 (4): 9-16.
4. Go´mez FL, Escoto-De DA, Puebla-Pe´rez AM, Figuera-Villanueva LE, Ramos-Silva A, Ramı´rez-Patiño R, et al. Association of the tumor necrosis factor-alpha -308G > A polymorphism with breast cancer in Mexican women. Genet Mol Res. 2013; 12:5680–93.
5. Mar N, Antoni A, Dora R, Alice C, Javier CB. Evaluation of dietary habits and nutritional needs in women treated for breast cancer in the Balearic Islands. Academic J Health Sciences 2023;38 (4):125-8.
6. Gallegos-Arreola MP, Figuera-Villanueva LE, Ramos-Silva A, Salas-Gonza´lez E, Puebla-Pe´rez AM, Peralta-Leal V, et al. The association between the 844ins68 polymorphism in the CBS gene and breast cancer. Arch Med Sci. 2014; 6:1214–24.
7. de Jong MM, Nolte IM, te Meerman GJ, van der Graaf WT, Oosterwijk JC, Kleibeuker JH, et al. Genes other than BRCA1 and BRCA2 involved in breast cancer susceptibility. J Med Genet. 2002; 39:225–42.
8. Dunning AM, Healey CS, Pharoah PD, Teare MD, Ponder BA, Easton DF. A systematic review of genetic polymorphisms and breast cancer risk. Cancer Epidemiol Biomarkers Prev. 1999; 8:843–54.
9. Colomer R, Monzo M, Tusquets I, Rifa J, Baena JM, Barnadas A, et al. A single-nucleotide polymorphism in the aromatase gene is associated with the efficacy of the aromatase inhibitor letrozole in advanced breast carcinoma. Clinical Cancer Research, 2008;14(3):811-816.

10. Costa-Silva DR, da Conceição BM, Borges RS, Campos-Verdes LM, da Silva-Sampaio JP, Escorcio-Dourado CS, et al. Insulin-like growth factor 1 gene polymorphism in women with breast cancer. Medical Oncology, 2017; 34(4):1-5.
11. Behavioral studies of mice with breast cancer after treatment with new anticancer agent, Rh2-containing arginine-graphene Academic Journal of Health Sciences 2023;38 (1):63-65.
12. Hong Y, Li H, Yuan YC, Chen S. Molecular characterization of aromatase. Annals of the New York Academy of Sciences,2009; 1155(1):112-120.
13. Holloway KR, Barbieri A, Malyarchuk S, Saxena M, Nedeljkovic-Kurepa A, Cameron MM, et al. SIRT1 positively regulates breast cancer associated human aromatase (CYP19A1) expression. Molecular endocrinology, 2013;27(3):480-490.
14. Chan HJ, Petrossian K, Chen S. Structural and functional characterization of aromatase, estrogen receptor, and their genes in endocrine-responsive and–resistant breast cancer cells. J steroid biochemistry and molecular biology, 2016;161:73-83.
15. Chand AL, Herridge KA, Howard TL, Simpson ER, Clyne CD. Tissue-specific regulation of aromatase promoter II by the orphan nuclear receptor LRH-1 in breast adipose stromal fibroblasts. Steroids, 2011;76(8):741-744.
16. Feigelson HS, Henderson BE. Estrogens and breast cancer. Carcinogenesis. 1996;17(11):2279-2284
17. Siiteri PK. Adipose tissue as a source of hormones. American J Clinical Nutrition. 1987;45(1 Suppl):277-282.
18. O'Neill JS, Miller WR. Aromatase activity in breast adipose tissue from women with benign and malignant breast diseases. British J Cancer. 1987;56(5):601-604
19. Farzaneh F, Noghabaei G, Barouti E, Pouresmaili F, Jamshidi J, Fazeli A, et al. Analysis of CYP17, CYP19 and CYP1A1 gene polymorphisms in Iranian women with breast cancer. Asian Pacific J Cancer Prevention. 2016;17:(Spec No.):23-6.
20. Zhang L, Gu L, Qian B, Hao X, Zhang W, Wei Q, et al. Association of genetic polymorphisms of ER-alpha and the estradiol-synthesizing enzyme genes CYP17 and CYP19 with breast cancer risk in Chinese women. Breast Cancer Research and Treatment. 2009;114(2):327-338.
21. Yoshimoto N, Nishiyama T, Toyama T, Takahashi S, Shiraki N, Sugiura H, et al. Genetic and environmental predictors, endogenous hormones and growth factors, and risk of estrogen receptor-positive breast cancer in Japanese women. Cancer Science. 2011;102(11):2065-2072
22. Ghisari M, Eiberg H, Long M, Bonefeld-Jørgensen EC. Polymorphisms in phase I and phase II genes and breast cancer risk and relations to persistent organic pollutant exposure: A case-control study in Inuit women. Environmental Health. 2014;13(1):19
23. Pineda B, García-Pérez MÁ, Cano A, Lluch A, Eroles P. Associations between aromatase CYP19 rs10046 polymorphism and breast cancer risk: From a case-control to a metaanalysis of 20,098 subjects. PLoS One. 2013;8(1):e53902

24. Zins K, Mogg M, Schneeberger C, Abraham D, Schreiber M. Analysis of the rs10046 polymorphism of aromatase (CYP19) in premenopausal onset of human breast cancer. International Journal of Molecular Sciences. 2014;15(1):712-724
25. Darabi H, Czene K, Wedrén S, Li Y, Liu J, Hall P, et al. Genetic variation in the androgen estrogen conversion pathway in relation to breast cancer prognosticators. Breast Cancer Research and Treatment. 2011;127(2):503-509
26. Shao X, Guo Y, Xu X, Zheng Y, Wang J, Chen Z, et al. The CYP19 RS4646 polymorphism IS related to the prognosis of stage I-II and operable stage III breast cancer. PLoS One. 2015;10(3):e0121535.
27. Fasching PA, Loehberg CR, Strissel PL, Lux MP, Bani MR, Schrauder M, et al. Single nucleotide polymorphisms of the aromatase gene (CYP19A1), HER2/neu status, and prognosis in breast cancer patients. Breast Cancer Research and Treatment. 2008;112(1):89-98.
28. Santa-Maria CA, Blackford A, Nguyen AT, Skaar TC, Philips S, Oesterreich S, et al. Association of variants in candidate genes with lipid profiles in women with early breast cancer on adjuvant aromatase inhibitor therapy. Clinical Cancer Research. 2016;22(6):1395-1402
29. Herbst RS. Review of epidermal growth factor receptor biology. International J Radiation Oncology Biology Physics, 2004;59(2):S21-S26.
30. Lo HW, Xia W, Wei Y, Ali-Seyed M, Huang SF, Hung MC. Novel prognostic value of nuclear epidermal growth factor receptor in breast cancer. Cancer research, 2005;65(1):338-348.
31. Castilla LH, Couch FJ, Erdos MR, Hoskins KF, Calzone K, Garber JE, et al. Mutations in the BRCA1 gene in families with early-onset breast and ovarian cancer. Nature Genetics. 1994;8(4):387-391
32. Nomura M, Shigematsu H, Li L, Suzuki M, Takahashi T, Estess P, et al. Polymorphisms, mutations, and amplification of the EGFR gene in non-small cell lung cancers. PLoS Med. 2007;4(4):e125.
33. Connor AE, Baumgartner RN, Baumgartner KB, Pinkston CM, John EM, Torres-Mejía G, et al. Epidermal growth factor receptor (EGFR) polymorphisms and breast cancer among Hispanic and non-Hispanic white women: the Breast Cancer Health Disparities Study. Int J Mol Epidemiol Genet. 2013;4(4):235-49.
34. AbdRaboh NR, Shehata HH, Ahmed MB, Bayoumi FA. HER1 R497K and HER2 I655V polymorphisms are linked to development of breast cancer. Dis Markers. 2013;34(6):407-17.
35. Escórcio-Dourado CS, Alves-Ribeiro FA, Lima-Dourado JC, Dos Santos AR, de Oliveira Pereira R, Tavares CB, et al. Human Epidermal Growth Factor Receptor-2 gene polymorphism and breast cancer risk in women from the Northeastern region of Brazil. Clinics (Sao Paulo). 2020;75: e2360.