Association of CYP1B1 Gene Polymorphisms with Estrogen Receptor-Positive Breast Cancer at Aga Khan University Hospital, Nairobi, Kenya

Mary Murithi[1,2 *], Steven Nyanjom[3], Victor Mobegi[4], Sayed Shahin[5], and Francis Makokha[6]

[1]Biochemistry Department, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya; [2]Pre-Clinical Department, Kabarak University, Nakuru, Kenya; [3]Biochemistry Department, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya; [4]Department of Biochemistry, University of Nairobi, Nairobi, Kenya; [5]Department of Pathology, Aga Khan University Hospital, Nairobi, Kenya, and [6]Directorate of Research and Innovation, Mount Kenya University, Thika, Kenya

***Corresponding author:** Mary Murithi. Email address: murithimary11@gmail.com

DOI: https://dx.doi.org/10.4314/ajhs.v37i2.1

Abstract

BACKGROUND

Breast cancer is a significant global health challenge, and polymorphisms in the CYP1B1 gene have been associated with its risk. Given that the effects of genetic polymorphisms on breast cancer risk vary across populations, region-specific studies are crucial. This study assessed the associations of four CYP1B1 polymorphisms (rs10012, rs1056827, rs1056836, rs1800440) with estrogen receptor-positive breast cancer (ER+BC) risk in Kenyan women.

METHODOLOGY

A retrospective hospital-based case-control study involved 64 cases and 19 controls. Oversampling adjusted the case-control imbalance, increasing the control sample size to 60. DNA was extracted from buffy coat samples, and target regions were amplified and sequenced via Sanger sequencing. Sequences were analyzed using Geneious Prime for alignment, quality trimming, and SNP identification. Statistical analysis was performed using R (R 4.3.3).

RESULTS

The study identified significant associations between CYP1B1 polymorphisms and ER+BC risk. Specifically, the variant allele C and the codominant model (CC vs. GG) of rs10012, as well as the variant allele A, dominant (CA - CC vs. AA) and log-additive models of rs1056827, demonstrated a protective effect with ORs of 0.53 (p = 0.018, 95% CI: 0.31–0.90), 0.28 (p = 0.040, 95% CI: 0.08–0.94), 0.29 (p = 0.001, 95% CI: 0.13–0.63), 0.23 (p = 0.003, 95% CI: 0.08–0.67) and 0.51 (p = 0.005, 95% CI: 0.29–0.91), respectively. In contrast, the recessive (CC vs. GG - GC) and the log-additive models of rs10012, were linked to an increased risk of ER+BC, with ORs of 2.39 (p = 0.020, 95% CI: 1.14–5.03) and 1.97 (p = 0.014, 95% CI: 1.13–3.44), respectively.

CONCLUSION

These findings reveal the complex interplay between CYP1B1 polymorphisms and ER+BC risk, with some variants protecting while others increase risk. Further research is essential to fully understand the effects of these genetic variations on breasts.

Keywords: *Estrogen Receptor-Positive Breast Cancer; CYP1B1; Single Nucleotide Polymorphisms; Sanger Sequencing*

[*Afr. J. Health Sci*. 2024 37 (2): 124-134]

Introduction

Breast cancer is the most prevalent cancer among women worldwide, with 2.3 million cases representing 23.8% of all female cancer cases. It has the highest mortality rate, causing approximately 666,000 deaths, or 15.4% of all cancer deaths in women (GLOBOCAN). In Africa, the disease poses a significant burden,

with 198,553 cases and 91,252 deaths, representing 29.2% and 21.9% of cancer cases and deaths in women, respectively. In Kenya, breast cancer is the leading malignancy and the second highest in mortality, with 7,243 cases and 3,591 deaths as of 2022, making up 25.5% and 18.5% of female cancer incidences and mortalities, respectively [1]. High mortality rates are exacerbated by inadequate diagnostic methods and clinical management, particularly in low- and middle-income countries (LMICs) [2].

Significant disparities exist in breast cancer incidence and mortality across different populations [3]. Although developed countries report higher incidence rates, most deaths occur in LMICs [2]. Nearly half of the cases in LMICs are in women under 50 [4] and many are diagnosed at advanced stages, leading to poorer outcomes [5].

Cytochrome P450 1B1, a potential breast cancer driver, is known for its polymorphic nature [6]. Research has predominantly focused on populations in developed countries [7]. Four CYP1B1 polymorphisms - rs10012, rs1056827, rs1056836, and rs1800440 - affect the enzyme's hydroxylation and catalytic activities [8]. Variants like rs10012 and rs1056827 enhance catalytic activity by upregulating gene expression and regulating substrate binding potentially increasing cancer risk, while rs1056836 and rs1800440 also enhance activity but by affecting the heme-binding domain [9]. High catalytic activity in these polymorphic variants has been linked to carcinogenic effects and potential influence on breast tumorigenesis [6]. For instance, a Nigerian study linked the Leu432Val polymorphism with breast cancer risk [10], whereas no such association was found in Egypt [11].

Ethiopian studies have identified haplotypes with altered estradiol hydroxylation kinetics [12]. Although numerous studies have explored these polymorphisms in Caucasian and Asian populations, their impact on breast cancer risk in African populations remains largely unexplored [10,13].

Given breast cancer's high burden in LMICs, exploring prevention and early diagnosis is vital. This study investigates CYP1B1 polymorphisms and ER+BC in Kenyan women to enhance understanding of disease progression.

Methodology

Study design

This retrospective case-control study at Aga Khan University Hospital (AKUHN) in Nairobi focused on estrogen receptor-positive breast cancer (ER+BC) cases diagnosed between 2019 and 2021. Cases included women aged 18 and older who were diagnosed with ER+BC and provided informed consent. Controls were healthy women aged 18 and above, with no history of malignancies or breast masses, who also provided consent. Both groups excluded women diagnosed with other cancers. All eligible ER+BC patients within the specified timeframe were included, while controls were selected from hospital staff without a history of breast cancer.

Study participants

Eighty-three participants were enrolled: 64 cases with ER+BC and 19 healthy controls from AKUHN. Data on socio-demographics, reproductive history, medical and family history, lifestyle factors, and pathological details for cases, including histological subtype, tumour grade, size, lymph node involvement, and molecular subtype, were extracted from hospital records.

Sample collection

Buffy-coat samples were obtained from the pathology department at AKUHN. Five millilitres of venous blood were drawn from each participant by skilled phlebotomists under aseptic conditions and collected in EDTA vacutainer tubes. The buffy coat fractions were stored at -20°C for further analysis.

Genomic DNA amplification

Genomic DNA was extracted using the ISOLATE II Genomic DNA Kit (Bioline, Meridian Life Science Inc., USA) per the manufacturer's guidelines. Primers were designed with Primer3Plus: a 202 bp region (containing rs1800440 and rs1056836) was amplified using Forward: 5′-ATCATCACTCTGCTGGTCA-3′ and Reverse: 5′-TGCCTGTCACTATTCCTCA-3′ primers; and a 567 bp region (containing rs1056827 and rs10012) was amplified using Forward: 5′-GAAACACACGGCACTCAT-3′ and Reverse: 5′-ACGCTCCTGCTACTCCTGT-3′ primers. PCR was performed in 30 µl reactions with 1 U MyTaq DNA polymerase (BIO-21105, Bioline), 50 ng/µL genomic DNA, 0.3 µM primers, 6 µL 5× MyTaq buffer, and nuclease-free water. Thermocycling conditions were: 95 °C for 3 min (initial denaturation), 30 cycles of 95 °C for 30 s, 55 °C for 45 s, 72 °C for 45 s, and final extension at 72 °C for 3 min, using a Thermo Fisher Scientific Applied Biosystems™ ProFlex™ 3 x 32-well PCR System (Applied Biosystems; Thermo Fischer Scientific, Waltham, UK).

Genotyping and SNP analysis

Amplicons were sequenced bidirectionally by Macrogen using Sanger sequencing. Raw data (.ab1 files) were imported into Geneious Prime for analysis, where chromatograms were reviewed and low-quality bases were trimmed. The CYP1B1 reference sequence from NCBI was aligned with the sequences using MUSCLE. SNPs were identified by comparison to the reference gene and verified against chromatograms. The identified SNPs were then filtered by quality and frequency, and the data was exported to Excel for further analysis.

Statistical analysis

Statistical analysis was conducted using R (v4.3.3). To address the case-control imbalance, the ROSE package was used to oversample controls, resulting in 60 controls and 64 cases. Descriptive statistics were calculated, and tests of independence (Wilcoxon rank-sum, Fisher's exact, Chi-squared) were applied.

Association analysis

Polymorphism-disease risk associations were analyzed with the SNPassoc package and logistic regression. Odds ratios and 95% confidence intervals were calculated.

Haplotype analysis

Haplotype analysis used the haplo. stats package to estimate probabilities and effects of CYP1B1 SNP haplotypes. Odds ratios and 95% CIs were calculated for haplotypes with ≥5% frequency.

Ethical consideration

All participants provided informed consent, and no patient identifiers were collected to maintain anonymity. The study was approved by the Aga Khan University Institutional Research Ethics Committee (Ref: 2020/IERC-26 [v4]) and received a research permit from the National Commission for Science, Technology and Innovation (License No: NACOSTI/P/21/13890).

Results

Descriptive characteristics of the study participants

The median age of ER+BC cases was 53 years (IQR 43-59) compared to 41 years (IQR 37-42) for controls ($p < 0.001$). Median BMI was 30 kg/m² (IQR 26-35) for cases and 29 kg/m² (IQR 24-30) for controls (not significant). Median age at menarche was 13 years (IQR 13-16) for cases and 14 years (IQR 12-14) for controls ($p < 0.001$). Menopause was more common in cases (55%) versus controls (3%) ($p < 0.001$). Positive parity was reported in 97% of cases and 70% of controls ($p < 0.001$). Contraceptive use was lower in cases (66%) compared to controls (93%) ($p < 0.001$). Family history of breast cancer was reported in 20% of cases and other cancers in 33%, with significant differences from controls ($p < 0.001$). Diabetes was less common among participants

but varied significantly (p = 0.006). Alcohol consumption was higher in controls (77%, p < 0.001). ER+BC cases differed significantly from controls in age, BMI, menarche age, menopause status, reproductive history, family cancer history, diabetes diagnosis, and alcohol use. Contraceptive use was less frequent in some cases.

Table 1:
Descriptive Characteristics of Study Participants

	Variables	Cases, n = 64[1]	Controls, n = 60[1]	p-value[2]	Statistic test[3]
	Age	53 (43, 59)	41 (37, 42)	<0.001	3129.5
Age categories (years)	20-29	1 (1.6)	11 (18)	<0.001	*
	30-39	9 (14)	12 (20)		
	40-49	16 (25)	35 (58)		
	50-59	24 (38)	2 (3.3)		
	>60	14 (22)	0 (0)		
	Body mass index	30 (26, 35)	29 (24, 30)	0.061	2294.5
Body mass index categories	<18.5	2 (3.1)	0 (0)	0.13	*
	18.6-24.9	10 (16)	18 (30)		
	25.0-29.9	19 (30)	18 (30)		
	>30	33 (52)	24 (40)		
	Age at menarche	14 (13, 16)	13 (12, 14)	<0.001	2636
Age at menarche categories (years)	<12	10 (16)	20 (33)	0.013	8.72
	13-14	26 (41)	27 (45)		
	>15	28 (44)	13 (22)		
Menopausal status	No	29 (45)	58 (97)	<0.001	36.59
	Yes	35 (55)	2 (3.3)		
Parity	Yes	62 (97)	42 (70)	<0.001	14.60
	No	2 (3.1)	18 (30)		
Contraceptives use	Yes	42 (66)	56 (93)	<0.001	12.72
	No	22 (34)	4 (6.7)		
Family history of breast cancer	No	51 (80)	60 (100)	<0.001	*
	Yes	13 (20)	0 (0)		
Family history of other cancers	No	43 (67)	60 (100)	<0.001	*
	Yes	21 (33)	0 (0)		
Diabetes	No	56 (88)	60 (100)	0.006	*
	Yes	8 (13)	0 (0)		
Hypertension	No	43 (67)	46 (77)	0.2	0.94
	Yes	21 (33)	14 (23)		
Radiotherapy	No	63 (98)	60 (100)	1.0	*
	Yes	1 (1.6)	0 (0)		
Tobacco	No	62 (97)	60 (100)	0.5	*
	Yes	2 (3.1)	0 (0)		
Alcohol consumption	No	40 (63)	14 (23)	<0.001	17.76
	Yes	24 (38)	46 77)		

[1] Median (Inter quartile range); n (%): Frequency and percentage
[2] P-value < 0.05: statistically significant
[3] Statistical test was calculated using the Wilcoxon rank sum test/Pearson's Chi-squared/ Fisher's exact tests.
* No test statistic

A comparison of socio-demographic, health, and lifestyle characteristics between ER+BC cases (n = 64) and controls (n = 60) at Aga Khan University Hospital, Nairobi, using Wilcoxon rank sum, Pearson's Chi-square, and Fisher's exact tests.

Pathological characteristics of estrogen receptor-positive breast cancer

The most common pathology features were invasive ductal carcinoma (77%), grade II (63%), stage II (47%), no lymph node invasion (44%), and luminal A (84%). (Table 2). Histological and molecular characteristics of 64 ER+ breast cancer cases at Aga Khan University Hospital: most common were invasive ductal carcinoma, grade II, stage II, N0, and luminal A.

Association of CYP1B1 gene polymorphisms with estrogen receptor-positive breast cancer

The study evaluated the association of CYP1B1 gene polymorphisms with ER+BC by analyzing allele and genotype frequencies. Lower frequencies of variant alleles C of rs10012, A of rs1056827, and C of rs1056836 were found in cases compared to controls. For rs10012, the variant allele C had an odds ratio (OR) of 0.53 (95% CI: 0.031–0.90, p = 0.018), indicating reduced presence in cases. The codominant model showed OR = 0.28 (95% CI: 0.08–0.94, p = 0.040), suggesting a protective CC genotype. However, the recessive model showed OR = 2.39 (95% CI: 1.14–5.03, p = 0.020), indicating increased risk with the CC genotype.

The log-additive model revealed OR = 1.97 (95% CI: 1.13–3.44, p = 0.014), suggesting each additional C allele raises risk. For rs1056827, the variant allele A showed protection with OR = 0.29 (95% CI: 0.13–0.63, p = 0.001) and OR = 0.23 (95% CI: 0.08–0.67, p = 0.003) in the dominant model, and OR = 0.51 (95% CI: 0.29–0.91, p = 0.005) in the log-additive model. The findings indicate mixed effects of SNPs on ER+BC susceptibility, with some variants showing protective effects and others suggesting potential risks (Table 3).

Association analysis of SNPs in ER+BC cases and controls assessed allele and genotype frequencies, odds ratios (OR), and 95% confidence intervals (CI) using Pearson's Chi-squared and Fisher's exact tests.

Table 2:
Pathological Characteristics of Estrogen Receptor-Positive Breast Cancer Case Participants

Characteristic		Cases, n = 64[1]
Histological type	Invasive ductal carcinoma (IDC)	49 (77)
	Invasive lobular carcinoma (ILC)	3 (4.7)
	Others	12 (18.3)
Histological grade	Grade I	3 (4.7)
	Grade II	40 (63)
	Grade III	21 (33)
Tumor stage	Stage I	18 (28)
	Stage II	30 (47)
	Stage III	16 (25)
Lymph node invasion	N0 (No lymph node invasion)	28 (44)
	N1 (Tumor invasion in 1-3 lymph nodes near breastbone)	19 (30)
	N2 (Tumor invasion in 4-9 armpit lymph nodes)	12 (19)
	N3 (Tumor invasion in 10 or more lymph nodes at the collarbone)	5 (7.8)
Molecular subtype	Luminal A	54 (84)
	Luminal B	10 (16)

n (%): Frequency and percentage

Table 3:
Association of CYP1B1 Gene Polymorphisms with Estrogen Receptor-Positive Breast Cancer

SNP ID	Variables	Allele/ genotype	Cases, n = 64[1]	Controls, n = 60[1]	P-value[2]	OR[3]	95% CI[4]
rs10012	Allele	G	57(44.5)	36(30)			
		C	71(55.4)	84(70)	0.018	0.53	0.31– 0.90
	Codominant	GG	11(17)	5(8.3)			
		GC	35(55)	26(43)	0.411	0.61	0.18 – 1.97
		CC	18(28)	29(48)	0.040	0.28	0.08 – 0.94
	Dominant	GG	11(17.2)	5(8.3)			
		GC-CC	53(82.8)	55(91.7)	0.137	2.28	0.74 – 7.01
	Recessive	GG-GC	46(71.9)	31(51.7)			
		CC	18(28.1)	29(48.3)	0.020	2.39	1.14 – 5.03
	Over-dominant	GG-CC	29(45.3)	34(56.7)			
		GC	35(54.7)	26(43.3)	0.205	0.63	0.31 – 1.29
	log-additive	0,1,2	64(51.6)	60(48.4)	0.014	1.97	1.13 – 3.44
rs1056827	Alleles	C	30(23.4)	10(8)			
		A	98(76.6)	110(92)	0.001	0.29	0.13 – 0.63
		CC	12(19)	5(8.3)			
	Codominant	CA	6(9.4)	0(0)	0.261	*	*
		AA	46(72)	55(92)	0.063	0.34	0.11 – 1.06
	Dominant	AA	46(71.9)	55(91.7)			
		CA-CC	18(28.1)	5(8.3)	0.003	0.23	0.08 – 0.67
	Recessive	AA-CA	52(81.2)	55(91.7)			
		CC	12(18.8)	5(8.3)	0.087	0.39	0.13 – 1.2
	Over-dominant	AA-CC	58(90.6)	60(100)			
		CA	6(9.4)	0	0.028	*	*
	log-additive	0,1,2	64(51.6)	60(48.4)	0.005	0.51	0.29 – 0.91
rs1056836	Alleles	G	39(30.5)	31(25.8)			
		C	89(69.5)	89(74.2)	0.418	0.79	0.45 – 1.38
	Codominant	GG	14(22)	8(13)			
		GC	11(17)	15(25)	0.143	0.41	0.13 – 1.34
		CC	39(61)	37(62)	0.309	0.60	0.22 – 1.60
	Dominant	GG	14(21.9)	8(13.3)			
		GC-CC	50(78.1)	52(86.7)	0.210	1.82	0.7 – 4.71
	Recessive	GG-GC	25(39.1)	23(38.3)			
		CC	39(60.9)	37(61.7)	0.933	1.03	0.5 – 2.13
	Over-dominant	GG-CC	53(82.8)	45(75)			
		GC	11(17.2)	15(25)	0.285	1.61	0.67 – 3.85
	log-additive	0,1,2	64(51.6)	60(48.4)	0.505	1.17	0.74 – 1.84
rs1800440	Alleles	A	95(74)	89(74)			
		G	33(26)	31(26)	0.992	0.99	0.56 – 1.77
	Codominant	AA	39(61)	33(55)			
		AG	17(27)	23(38)	0.238	0.62	0.28 – 1.36
		GG	8(13)	4(6.7)	0.422	1.69	0.47 – 6.13
	Dominant	AA	39(60.9)	33(55)			
		AG-GG	25(39.1)	27(45)	0.503	1.28	0.62 – 2.61
	Recessive	AA-AG	56(87.5)	56(93.3)			
		GG	8(12.5)	4(6.7)	0.267	0.5	0.14 – 1.76
	Over-dominant	AA-GG	47(73.4)	37(61.7)			
		AG	17(26.6)	23(38.3)	0.160	1.72	0.8 – 3.68
	log-Additive	0,1,2	64(51.6)	60(48.4)	0.993	1	0.59 – 1.7

[1] n (%): Frequency and percentage;
[2] P-value < 0.05: statistically significant;
[3] OR: Odds Ratio;
[4] 95% CI: Confidence Interval; * No OR and 95% CI

For rs10012, the variant allele C and genotype CC were protective, while the recessive model (GG - GC vs. CC) indicated increased risk. For rs1056827, the variant allele A and both the dominant (AA vs. CA – CC) and log-additive models showed a protective effect against the disease.

Factors associated with estrogen receptor-positive breast cancer risk

Multicollinearity led to the exclusion of several factors from the logistic regression analysis. Increased age was associated with a lower likelihood of ER+BC (OR 0.839, 95% CI: 0.77 – 0.899, p <0.001), whereas alcohol consumption raised the odds of ER+BC (OR 4.674, 95% CI: 1.733 – 13.666, p = 0.0031), highlighting a significant link between alcohol consumption and ER+BC risk. (Table 4).

Logistic regression analysis of ER+BC risk among cases and controls revealed that alcohol consumption increased the odds of being a case by approximately fivefold.

Haplotype analysis

Five haplotypes - C-A-C-A (44%), G-A-C-A (12%), G-C-C-A (12%), C-A-G-G (11%), and G-A-G-G (10%) - had frequencies above 5%, with C-A-C-A as the reference. None were significantly associated with ER+BC, but lower frequencies in cases suggest a potential protective effect (Table 5).

The table shows CYP1B1 haplotype frequencies, P-values, and odds ratios (OR) with 95% confidence intervals (CI) for ER+BC cases and controls. No haplotype demonstrated a statistically significant association with ER+BC compared to the reference haplotype (CACA).

Table 4:
Association of Other Factors with Estrogen Receptor Positive Breast Cancer Risk

		95% CI[2]		
Variable	**OR[1]**	**Lower**	**Upper**	**P-value[3]**
(Intercept)	770.235	7.773	128652.2	0.0069
Age of participants	0.839	0.77	0.899	<0.001
Age at menarche	0.912	0.695	1.19	0.4979
Body mass index	1.042	0.938	1.165	0.4456
Hypertension Yes	2.58	0.866	8.315	0.0973
Alcohol consumption Yes	4.674	1.733	13.666	0.0031

[1] OR: Odds Ratio
[2] 95% CI: 95% Confidence Interval
[3] P-value < 0.05 was deemed statistically significant.

Table 5:
Association Between CYP1B1 Haplotypes with ER+BC in the Study Participants

CYP1B1 Haplotypes				**n (%)[1]**			
rs10012	**rs1056827**	**rs1056836**	**rs1800440**	**Controls**	**Cases**	**P-value[2]**	**OR [3](95% CI)[4]**
C	A	C	A	34	55		1
G	A	C	A	15	7	0.087	0.17 (0.05-0.54)
G	C	C	A	15	8	0.178	0.40 (0.19-0.82)
C	A	G	G	12	8	0.461	0.25 (0.08-0.75)
G	A	G	G	7	14	0.083	1.96 (0.63-6.11)

[1] n (%): Frequency and percentage.
[2] P-value < 0.05 was deemed statistically significant.
[3] OR (95% CI): Odds Ratio
[4] 95% CI: Confidence Interval

Discussion

This study sought to investigate the relationship between polymorphisms in the CYP1B1 gene—specifically rs10012, rs1056827, rs1056836, and rs1800440—and the risk of estrogen receptor-positive breast cancer (ER+BC) in women at Aga Khan University Hospital, Nairobi. It also evaluated the combined effects of these polymorphisms on ER+BC risk.

Previous research has shown mixed results regarding the link between rs10012 and breast cancer risk. While some studies indicate an association [14], others report conflicting findings [15,16]. In our study, the rs10012 polymorphism demonstrated an increased risk for ER+BC in recessive and log-additive models. Despite this, the variant allele and genotype were notably less frequent among cases.

For rs1056827, which has been linked to breast cancer in other populations [17] our findings revealed a low association with ER+BC. The low prevalence of variant alleles in rs10012 and rs1056827 suggests a potential protective role, though the biological mechanisms remain unclear [15].

The rs1056836 polymorphism has shown varying associations with breast cancer in different studies [17]. While our results did not find a significant link, the variant genotype was highly prevalent in the cases. This finding is similar to that of De Vivo and colleagues (2020) [16]. Our finding of no significant association of rs1056836 aligns with other similar studies in Egyptians [11] and Japanese [18]. However, a Nigerian study found that the heterozygous rs1056836 genotype increased breast cancer risk by 59%, while the variant genotype showed a non-significant 51% increase [10].

Similar to our findings, studies on Caucasians found no association between rs1800440 and ER+BC [16]. A meta-analysis also reported no significant associations of rs1800440, rs10012, and rs1056827 with breast cancer risk [19]. In Africa, initial studies in Ethiopians showed a higher prevalence of rs10012, rs1056827, and rs1056836 compared to whites and Japanese, while rs1800400 was less common [12].

Breast cancer risk can be influenced by socio-demographic and pathological factors. Unlike in Europe and the US, where breast cancer is typically diagnosed at an older age, African patients often present with the disease earlier. Recent reviews indicate that most African breast cancer patients are under 50 years old [20]. Studies have shown that in Eastern Africa, a significant proportion of patients are younger [21]. For instance, the average age at diagnosis is 49.5 years in Tanzania [21] and 51 years in Nigeria [22]. In Kenya, the mean age across multiple hospitals is 49.2 years, with 50.5 years at AKUHN [4] and the median age in this study was 53 years.

Most participants in the study were obese. Obesity is associated with hormone-receptor-positive breast cancer, acting as both a risk factor (post-menopausal) and a potential protective factor (pre-menopausal) [23]. It increases aromatase activity, which converts androgens to estrogen in adipose tissue, but its protective mechanism is unclear [24].

Alcohol consumption is significantly associated with breast cancer, with the International Agency for Research on Cancer (IARC) classifying it as a group I carcinogen [25]. Alcohol and its metabolites are thought to influence estrogen levels and receptors in breast cells, contributing to cancer risk [26]. However, this study did not account for factors like consumption duration, beverage type, or amount.

Invasive ductal carcinoma was the most prevalent histological type in this study, consistent with findings from western Kenya [27]. Most tumours were grade II, similar to other Kenyan studies [4]. While a meta-analysis of African data indicated stage III breast cancer as the most common, our study observed that the majority of cases were at stages I and II [20]. This

early-stage diagnosis might reflect specific health-seeking behaviours among patients. Molecular subtypes based on hormone receptor expression identified luminal A (ER+/PR+/HER2-) as the most common subtype, aligning with results from studies of Ghanaian women [28].

This study's strength lies in combining genotype assessments with clinical and pathological factors and evaluating the impact of CYP1B1 variants on ER+BC risk. Assessing these variants across diverse populations is crucial for accurate risk prediction and prevention.

Study limitations

The study's retrospective design may have introduced recall bias. Despite efforts to recruit enough participants, the final sample size was limited by difficulties in recruiting healthy volunteers from the hospital, affecting the generalizability of the findings. Further research with larger sample sizes is needed for more definitive conclusions. Additionally, limited research on African populations results in insufficient data for comprehensive comparative analyses, impacting the understanding of SNP significance in breast cancer among Africans. Despite these limitations, this study advances our understanding of breast cancer genetics in Kenya.

Conclusion

CYP1B1 polymorphisms' association with breast cancer varies across populations. This study found mixed effects on ER+BC risk: rs10012 variants and rs1056827's dominant and log-additive models showed protective effects, while rs10012's recessive and log-additive models suggested increased risk. Further research is needed to explore these associations.

Acknowledgement

We thank Mr Erick Ouko and Dr Patrick Njage for data analysis, Ms Ann Karanu and Ms Vivian Oluoch for data collection, and Dr Michael Walekhwa and Mr Micah Lagat for manuscript review.

Author Email contacts

- SN; snyanjom@jkuat.ac.ke,
- VM; vmobegi@gmail.com,
- SS; shaheen.sayed@aku.edu,
- FM; fmakokha@mku.ac.ke

Conflicts of interests. The authors declare no conflict of interest

Source of funding. This work was supported by the Mawazo Fellowship Programme under the Mawazo Fellows Fund [Grant Number: 2022-1-06].

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024; 74:229–63. https://doi.org/10.3322/caac.21834.
2. Kashyap D, Pal D, Sharma R, Garg VK, Goel N, Koundal D, et al. Global Increase in Breast Cancer Incidence: Risk Factors and Preventive Measures. Biomed Res Int 2022:1–16. https://doi.org/10.1155/2022/9605439.
3. Omotoso O, Teibo JO, Atiba FA, Oladimeji T, Paimo OK, Ataya FS, et al. Addressing cancer care inequities in sub-Saharan Africa: current challenges and proposed solutions. Int J Equity Health 2023; 22:189. https://doi.org/10.1186/s12939-023-01962-y.
4. Sayed S, Fan S, Moloo Z, Wasike R, Bird P, Saleh M, et al. Breast cancer risk factors in relation to molecular subtypes in breast cancer patients from Kenya. Breast Cancer Research 2021;23. https://doi.org/10.1186/s13058-021-01446-3.
5. Denny L, de Sanjose S, Mutebi M, Anderson BO, Kim J, Jeronimo J, et al. Interventions to close the divide for women with breast and cervical cancer between low-income and middle-income countries and high-income countries. The Lancet 2017; 389:861–70. https://doi.org/10.1016/S0140-6736(16)31795-0.

6. Kwon YJ, Shin S, Chun YJ. Biological roles of cytochrome P450 1A1, 1A2, and 1B1 enzymes. Arch Pharm Res 2021; 44:63–83. https://doi.org/10.1007/s12272-021-01306-w.

7. Alsubait A, Aldossary W, Rashid M, Algamdi A, Alrfaei BM. CYP1B1 gene: Implications in glaucoma and cancer. J Cancer 2020; 11:4652–61. https://doi.org/10.7150/jca.42669.

8. Adab FK, Fard ZT, Akbari ME. Association between cytochrome 1B1*3 polymorphism and the breast cancer in a group of Iranian women. Int J Cancer Manag 2017;10. https://doi.org/10.17795/ijcp-6428.

9. Wen C, Wu L, Fu L, Wang B, Zhou H. Unifying mechanism in the initiation of breast cancer by metabolism of estrogen (Review). Mol Med Rep 2017; 16:1001–6. https://doi.org/10.3892/mmr.2017.6738.

10. Okobia MN, Bunker CH, Garte SJ, Zmuda JM, Ezeome ER, Anyanwu SNC, et al. Cytochrome P450 1B1 Val432Leu polymorphism and breast cancer risk in Nigerian women: A case control study. Infect Agent Cancer, vol. 4, 2009. https://doi.org/10.1186/1750-9378-4-S1-S12.

11. Ibrahim H Mona, Rashed A Reham, Hassan M Naglaa, Al-azhary M Nevin, Salama I Asmaa, Mostafa N Marwa. Ibrahim et al 2016_No Association of Cytochrome P450-1B1 Gene Polymorphisms with Risk of Breast Cancer an Egyptian Study. Asian Pacific Journal of Cancer Prevention 2016; 17:2861–6.

12. Aklillu E, Oscarson M, Hidestrand M, Leidvik B, Otter C, Ingelman-sundberg M. Functional Analysis of Six Different Polymorphic CYP1B1 Enzyme Variants Found in an Ethiopian Population. Mol Pharmacol 2002; 61:586–94.

13. Rizzolo P, Silvestri V, Valentini V, Zelli V, Bucalo A, Zanna I, et al. Evaluation of cyp17a1 and cyp1b1 polymorphisms in male breast cancer risk. Endocr Connect 2019; 8:1224–9. https://doi.org/10.1530/EC-19-0225.

14. Zimarina TS, Kristensen VN, Imyanitov EN, Berstein LM. Polymorphisms of CYP1B1 and COMT in Breast and Endometrial Cancer. Mol Biol 2004; 38:322–8.

15. Wen W, Cai Q, Shu X-O, Cheng J-R, Parl F, Pierce L, et al. Cytochrome P450 1B1 and Catechol-O-Methyltransferase Genetic Polymorphisms and Breast Cancer Risk in Chinese Women: Results from the Shanghai Breast Cancer Study and a Meta-analysis. Cancer Epidemiology, Biomarkers & Prevention 2005;14:329–35.

16. De Vivo I, Hankinson SE, Li L, Colditz GA, Hunter DJ. Association of CYP1B1 Polymorphisms and Breast Cancer Risk 1. Cancer Epidemiology, Biomarkers & Prevention 2002;11:489–92.

17. Jiao haiyan, Liu Chunlian, Guo W, Peng L, Chen Y, Martin FL. Association of CYP1B1 polymorphisms with Breast cancer: A case-control study in the Han population in Ningxia Hui Autonomous Region, P. R. China. Biomark Insights 2010; 2010:21–7.

18. Watanabe J, Shimada T, Gillam EMJ, Ikuta T, Suemasu K, Higashi Y, et al. Association of CYP1B1 genetic polymorphism with incidence to breast and lung cancer. vol. 10. Lippincott Williams & Wilkins; 2000.

19. Economopoulos KP, Sergentanis TN. Three polymorphisms in cytochrome P450 1B1 (CYP1B1) gene and breast cancer risk: a meta-analysis. Breast Cancer Res Treat 2010;122. https://doi.org/10.1007/s10549-009-0728-zï.

20. Olayide A, Isiaka A, Ganiyu R, Samuel O, Halimat A, Julius O, et al. Demographic Pattern, Tumor Size and Stage of Breast Cancer in Africa: A Meta-analysis. Asian Pacific Journal of Cancer Care 2021; 6:477–92. https://doi.org/10.31557/APJCC.2021.6.4.477.

21. Adebamowo CA, Famooto A, Ogundiran TO, Aniagwu T, Nkwodimmah C, Akang EE. Immunohistochemical and molecular subtypes of breast cancer in Nigeria. Breast Cancer Res Treat 2008; 110:183–8. https://doi.org/10.1007/s10549-007-9694-5.

22. Amadori D, Serra P, Bravaccini S, Farolfi A, Puccetti M, Carretta E, et al. Differences in biological features of breast cancer between Caucasian (Italian) and African (Tanzanian) populations. Breast Cancer Res Treat 2014; 145:177–83. https://doi.org/10.1007/s10549-014-2903-0.

23. Ritte R, Tikk K, Lukanova A, Tjønneland A, Olsen A, Overvad K, et al. Reproductive factors and risk of hormone receptor positive and

negative breast cancer: a cohort study. BMC Cancer 2013; 13:1–12.

24. Simone V, D'Avenia M, Argentiero A, Felici C, Rizzo FM, De Pergola G, et al. Obesity and Breast Cancer: Molecular Interconnections and Potential Clinical Applications. Oncologist 2016; 21:404–17. https://doi.org/10.1634/theoncologist.2015-0351.
25. Pflaum T, Hausler T, Baumung C, Ackermann S, Kuballa T, Rehm J, et al. Carcinogenic compounds in alcoholic beverages: an update. Arch Toxicol 2016; 90:2349–67. https://doi.org/10.1007/s00204-016-1770-3.
26. Choi YJ, Myung SK, Lee JH. Light alcohol drinking and risk of cancer: A meta-analysis of cohort studies. Cancer Res Treat 2018; 50:474–87. https://doi.org/10.4143/crt.2017.094.
27. Sawe RT, Kerper M, Badve S, Li J, Sandoval-Cooper M, Xie J, et al. Aggressive breast cancer in western Kenya has early onset, high proliferation, and immune cell infiltration. BMC Cancer 2016;16. https://doi.org/10.1186/s12885-016-2204-6.
28. Figueroa JD, Davis Lynn BC, Edusei L, Titiloye N, Adjei E, Clegg-Lamptey JN, et al. Reproductive factors and risk of breast cancer by tumor subtypes among Ghanaian women: A population-based case–control study. Int J Cancer 2020; 147:1535–47. https://doi.org/10.1002/ijc.32929.