

Whole-Exome Sequencing Identifies Novel Germline Risk Variants for Breast Cancer in Ghanaian Women

Lily Paemka

lpaemka@ug.edu.gh

University of Ghana

Claudia Adzo Anyigba

University of Ghana

Saikou Y. Bah

University of Glasgow

Vincent Appiah

University of Ghana

Enock K. Amoako

University of Ghana

Derrick Antwi

University of Ghana

Bismark Tetteh

University of Ghana

Josephine Nsaful

University of Ghana

Joe-Natt Clegg-Lampitey

University of Ghana

Florence Dedey

University of Ghana

Lawrence Edusei

University of Ghana

Ralph Armah

Greater Accra Regional Hospital

Alfred Twumasi

Greater Accra Regional Hospital

Nelson Affram

Ho Teaching Hospital

Lucas N. Amenga Etego

Article

Keywords:

Posted Date: June 16th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9972157/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

This study explored germline risk variants in a Ghanaian breast cancer cohort (BCOPGh-Ghana Breast Cancer Omics Project). Whole-exome sequencing of 113 unrelated breast cancer cases, analysed alongside 43 controls, revealed 26 pathogenic variants in 27.4% (31/113) of the cases. Variants in DNA-damage-response genes, including 3 canonical breast cancer susceptibility genes, were found in 4.4% (6/113) of the patients. This included a known *BRCA1* variant in 1.8% (2/113) of participants and two novel protein-truncating *BRCA2* variants— p.(R2112fs*15) and p.(N1066fs*10) — in 1.8% (2/113) of participants. Additionally, 11 candidate breast cancer genes were identified. Notably, approximately 96.1% (25/26) of the variants have no prior association with breast cancer. The discovery of 2 novel *BRCA2* variants and 11 candidate genes highlights the unique germline landscape of breast cancer in Ghanaian women and the value of African populations in identifying novel breast cancer susceptibility genes and variants.

Introduction

Breast cancer is a complex and heterogeneous genetic disease that affects mostly women ¹. In 2022, GLOBOCAN recorded approximately 2.3 million new cases with 666,103 related deaths, placing breast cancer as the second most common cancer and fourth leading cause of malignancy-associated deaths in all sexes and ages ². In females, however, it is the leading diagnosed cancer and cause of cancer-related mortalities, accounting for 24% of cancer incidence and 15.4% of all related mortality globally ³. In the coming decade, the global breast cancer incidence is projected to increase by 48.9% and mortality by 60.6% ⁴. Breast cancer was the second most diagnosed cancer in 2022, and it remains the most prevalent cancer and a significant public health challenge ⁵.

Approximately 5–10% of breast cancer cases are hereditary and often associated with early-onset disease ⁶. Since the detection of *BRCA1* and *BRCA2* and their implication in familial breast cancer, several studies have investigated other hereditary breast cancer-associated genes with varying penetrance ^{7,8}. However, the identified genes only explain approximately 25% of the hereditary in cases ⁹. Notably, indigenous African populations, despite being the most genetically diverse, are woefully underrepresented in genetic studies ¹⁰. Furthermore, breast cancer genetic studies in African populations are often restricted to investigating *BRCA1* and *BRCA2*. The racial variations in breast cancer-associated risk factors prompt urgent population-specific genetic risk factor profiling.

In Ghana, breast cancer is increasingly affecting younger women ¹¹. Despite this trend, genetic studies in this population have so far relied mainly on genotyping ^{12,13} and targeted sequencing ¹⁴, with risk assessment confined to specific molecular subtypes using largely Eurocentric risk genes and variants. Essentially, these studies confirmed known and candidate risk variants in the canonical breast cancer genes ¹⁴ and reaffirmed the high burden of aggressive subtypes in Ghanaian women ^{12,13}. However, the West African ancestry-associated germline risk variants for breast cancer/triple-negative breast cancer

(TNBC) in these studies were not enriched in the Ghanaian cohorts. This further emphasises the substantial heterogeneity and unshared genetic predispositions across indigenous African populations, reinforcing the need for population-specific germline risk assessment. While targeted sequencing and genotyping are cost-effective and efficient for detecting variants in established genes, they are inherently unable to identify novel risk genes or their variants. Hence, a targeting sequencing and genotyping approach is a critical limitation in African populations where breast cancer susceptibility may be driven by population-specific and previously uncharacterized genetic risk factors. In this study, we explored the germline landscape of the Ghana Breast Cancer Omics Project cohort (BCOPGh) using whole-exome sequencing. This approach offers a more comprehensive opportunity to interrogate the breast cancer genetic risk architecture and improve population-specific variant detection and risk assessment in Ghanaian women.

Results

Patient's demographics and breast cancer phenotypes

Figure 1 provides an overview of the participants' demographic and clinicopathologic characteristics. The cohort consisted of 113 women recruited from the Greater Accra region, who self-reported ethnicities from 10 of Ghana's 16 regions (Fig. 1a). A majority of participants, 52.2% (59/113), were aged 35–60 years, while 38.9% (44/113) were over 60 years and 8.9% (10/113) were younger than 35 years, with the youngest being 18 years old (Fig. 1b). Clinical data availability varied, with complete datasets for tumour type ($n = 81$), tumour grade ($n = 57$), initial symptom ($n = 69$), and tumour stage ($n = 15$). Three major breast cancer types were diagnosed in this cohort. The most prevalent was invasive breast carcinoma of no special type (80.2%, 65/81), and the least prevalent was invasive lobular carcinoma (2.5%, 2/81) (Fig. 1c). Other malignant neoplasms, including mucinous carcinoma and squamous cell carcinoma, accounted for 6.2% (5/81) of cases. A breast lump was the most common initial symptom, reported in 72.5% (50/69) of cases (Fig. 1d). Less common presentations included breast pain (15.9%, 11/69), nipple discharge (5.8%, 4/69), distorted breast shape (4.3%, 3/69), and breast ulcer (1.4%, 1/69). Among participants with available staging data ($n = 15$), stage II disease was most frequent (46.7%, 7/15), and stage IV was the least frequent (6.7%, 1/15). For tumour grade ($n = 57$), grade 2 was predominant (50.9%, 29/57), and grade 1 was the least prevalent 21.1% (12/57) (Fig. 1e).

Variant Validation and Interpretation

Genetic Variant Classification and Prioritisation

Initial variant calling identified 433,423 variants in cases and 200,324 in controls (Fig. 2a). Sequential exclusion of synonymous variants, variants with Minor Allele Frequency (MAF) $\geq 1\%$ in gnomAD and ExAC, and those present in COSMIC, yielded 266,289 variants in cases and 51,792 in controls. Subsequent exclusion of variants classified as benign or “unknown” in ClinVar retained 9,401 variants in

the cases and 2,521 in the controls. Based on ClinVar annotations, the filtered variants were categorised as pathogenic, likely pathogenic, likely benign, variant of uncertain significance (VUS) and other variants (Fig. 2b). The vast majority of variants in both cases (66.4%) and controls (66.3%) were likely benign. VUS represented 32.5% in both groups, while pathogenic (0.4% in cases and 0.5% in controls), likely pathogenic (0.3% in cases and 0.4% in controls) and “Other” variants (0.5% in both cases and controls) were rare. In this study, pathogenic variants encompassed those classified in ClinVar as “Pathogenic,” “Association,” or “risk factor. The “Other” category included ClinVar annotations such as “protective,” “risk factor,” “affect,” and “other (no appropriate term)”.

From an initial set of 33 pathogenic variants in cases and 13 in controls, subsequent validation and prioritisation steps (Fig. 3a) yielded 26 high-confidence variants (in 25 genes) and 9 high-confidence variants (in 9 genes) for the case and control cohorts, respectively (Fig. 3b; Supplementary Tables S1 and S2). Both cases and controls shared one variant in the *GJB2* gene; thus, 25 of the finalised variants were exclusive to the case group, while 8 were exclusive to the control group.

Pathogenic Variant in Known and Putative Breast Cancer Genes within the Cases

Overall, pathogenic variants were detected in 27.3% (31/113) of the cases (Supplementary Table S1). Fifteen of the 25 variants identified in the cases occurred in 14 genes previously implicated in breast cancer (Table 1). These included the *BRCA1* pathogenic variant rs80357708, detected in two Ghanaian participants and previously associated with breast cancer¹⁵. Additionally, 2 novel *BRCA2* frameshift variants—c.6331_6335del; p.(R2112fs*15) and c.3192_3195del; p.(N1066fs*10)—were each identified in 2 participants and absent from gnomAD, ExAC and TOPMed databases. The c.6331_6335del; p.(R2112fs*15) variant is a 5-bp deletion that begins 2-bp upstream of and overlaps the first 3-base deletions of the previously associated breast-ovarian cancer susceptibility variant rs397507369 described in a Chinese population¹⁶. Similarly, the c.3192_3195del; p.(N1066fs*10) variant, a 4-bp deletion, starts 3 bases upstream of and overlaps the first base of the breast cancer-associated variant rs80359375 previously reported in a Korean population¹⁷. The p.(R2112fs*15) mutant results in loss of all domains downstream of the RAD51-interacting domain (Fig. 4a, Supplementary Figures S2a and S2b), while the p.(N1066fs*10) mutant truncates all domains downstream of the first BRC repeat (Fig. 4b, Supplementary Figures S2c and S2d). *BRCA1* and *BRCA2* variants were identified in 12.9% (4/31) of cases carrying pathogenic alterations. Pathogenic variants in DNA repair genes (*BRCA1*, *BRCA2*, *FANCA* and *FUS*) were found in 19.4% (6/31) of cases with any pathogenic variant, representing 5.3% (6/113) of all cases. These variants accounted for 23.1% (6/26) of all pathogenic variants detected in the cases. Two participants each carried variants in *MAN1B1*, *SERPINB8* and *RDH12*, while variants in the remaining eight genes were found in single individuals. The remaining 11 variants were found in genes that had not previously been studied in the context of breast cancer (Table 2). However, experimental evidence links 5 of these genes (*ARSA*, *COLEC11*, *DYSF*, *ECEL1* and *PGAM*) to various cancers (Supplementary Table S2), including melanoma, hepatocellular carcinoma (HCC), and lung, bladder, prostate, and colorectal

cancers. Despite the absence of loss-of-heterozygosity data, copy-number alterations in the 11 genes have been reported in several breast cancer cohorts (Supplementary Table S2). The most recurrent variant was the stop-gain *PGAM2* (rs10250779), present in 16.1% (5/31) of participants carrying pathogenic variants. Additionally, a rare frameshift and protein-truncating variant in *ARSA*, c.418dupC; p.(H140fs*36), was detected in one participant. This variant introduces a premature stop codon, resulting in a truncated protein that lacks two-thirds of the catalytic domain and downstream structural regions (Fig. 4c and Supplementary Figures S2e and S2f). Generally, 61.5% (16/26) of the variants were single-nucleotide variants (SNVs), with 81% (13/16) being transitions; INDELs, including duplications, constituted 38.5% (10/26) of the variants.

Table 1
Pathogenic Variants in known breast cancer-associated genes

Genes	SNP	Nucleotide Change	Amino acid change	Number of Participants
ABHD12	rs758316679	c. 193C > T	p.(R65X)	1
BRCA1	rs80357708	c.344_345del	p.(V115Efs*19)	2
BRCA2	-	c.6331_6335del	p.(R2112fs*15)	1
BRCA2	-	c.3192_3195del	p.(N1066fs*10)	1
FANCA	rs2040094645	c.1085T > C	p.(L362P)	1
FUS	rs387906628	c.604G > A	p.(G202S)	1
GALC	rs200532368	c.310C > T	p.(R104X)	1
GIPC3	rs387907002	c.565C > T	p.(R189C)	1
GJB2	rs80338948	c.427C > T	p.(R143W)	1
HBB	rs34704828	c.-29G > A	-	1
HNF1A	rs1593058932	c.865dupC	p.(G292Rfs*25)	1
MAN1B1	rs1382373816	c.1833_1834del	p.(D613Pfs*40)	2
RDH12	rs144148976	c.698T > A	p.(V233D)	2
SERPINB8	rs762923677	c.946delA	p.(K316Sfs*90)	2
WDR45	rs781998959	c.292G > T	p.(E98X)	1
X – Termination codon; fs* - frameshift variant introducing a stop codon after a few to several amino acids. Rs number (-) – previously unreported SNP. SNP – Single Nucleotide Polymorphism				

Table 2
Variants in Candidate Breast Cancer Genes

Genes	SNP	Nucleotide Change	Amino acid change	Expression in breast (level*)	Number of Participants
ARSA	-	c.418dupC	p.(H140fs*36)	Yes (Medium)	1
CAPN15	rs762523863	c.2398C > T	p.(R800W)	Yes (Medium)	1
CHRNA1	rs767503038	c.753_754del	p.(V253Afs*44)	Below cutoff	1
COLEC11	rs762031914	c.231_232insGA	p.(R31Dfs*21)	Yes (Low)	1
CYP4F22	rs755885838	c.847C > T	(p.R283W)	Yes (Low)	1
DYSF	rs778065845	c.4643dupT	p.(M1548Ifs*39)	Yes (Medium)	1
ECEL1	rs370167241	c.997C > T	p.(R333X)	Yes (Low)	1
GUCY2D	rs281865409	c.3G > C	p.(M1?)	Below cutoff	1
KASH5	rs188572864	c.1146 + 5G > A	-	Below cutoff	1
PGAM2	rs10250779	c.233G > A	p.(W78X)	Below cutoff	5
SLC6A19	rs147837686	c.532C > T	p.(R178X)	Below cutoff	1
Breast tissue gene expression data were obtained from Ensembl. Expression levels*: Low (< 1TPM), Medium (> 10TPM), High (in PPB) and below cutoff (< 1TPM). fs* (frameshift event resulting in a premature stop codon). TPM – Transcript per million, PPB, parts per billion					

Variants found in the Controls

Pathogenic variants were detected in 40.9% (18/43) of the controls (Supplementary Table S3). Five variants—*CDH23* (rs1230303971), *GJB2* (rs80338948), *MYO7A* (rs121965085), *OTOF* (rs1558492758) and *SLC26A4* (rs727503430)—along with their corresponding genes (Table 3), were consistent with those previously reported in the larger hearing impairment cohort¹⁸. The predominant variant in the controls was *GJB2* rs80338948, present in 27.8% (5/18) of controls and 3.2% (1/31) of cases harbouring pathogenic variants. Variants in *MPDZ* and *PALB2* each occurred in approximately 7% (3/43), and a *MYO7A* variant was observed in 4.7% (2/43) of the control participants. Notably, the *PALB2* variant rs995629797, found exclusively in the controls, has previously been linked to familial breast cancer¹⁹. A novel frameshift variant, c.273dupC: p.(H93Pf*73), was identified in *B4GALT7*. This variant, along with the *CDH23* and *SLC26A4* variants, was detected in a single participant. Generally, 77.8% (7/9) of the variants were single-nucleotide variants (SNVs), with 85.7% (6/7) being transitions; INDELs, including duplications, constituted 22.2% (2/9) of the variants (Table 3).

Table 3
Pathogenic Variants within the Controls

Gene	SNP	Nucleotide change	Amino acid change	Number of Participants	Variant-associated Diseases
AGPAT2	rs116807569	c.589-2A > G (UTR5)	-	1	Congenital generalized lipodystrophy type 1
B4GALT7	-	c.273dupC	p.(H93Pf*73)	1	Ehlers-Danlos syndrome progeroid type
CDH23	rs1230303971	c.220C > T	p.(R736X)	1	Usher syndrome type 1D
GJB2	rs80338948	c.427C > T	p.(R143W)	5	Hearing impairment, Keratitis-ichthyosis-deafness syndrome
MPDZ	rs781251438	c.4429del	p. (Gln1490fs*3)	3	Hydrocephalus congenital 2 with or without brain or eye anomalies
MYO7A	rs121965085	c.1996C > T	p.(R666X)	2	Usher syndrome type 1, Deafness, Usher syndrome, type 1B
OTOF	rs1558492758	c.1051C > T	p.(Q351X)	1	Hearing impairment, Deafness
PALB2	rs995629797	c.2730T > A	p.(Y910X)	3	Familial breast cancer, Hereditary cancer predisposing syndrome
SLC26A4	rs727503430	c.2089 + 1G > A (splicing)	-	1	Enlarged vestibular aqueduct, Pendred syndrome, deafness

The variant-associated conditions were obtained from ClinVar and OMIM. UTR5–5-prime untranslated region. X – Termination codon, fs* - frameshift variant introducing a stop codon after a few to several (*) amino acids

Variant Consequences

Overall, approximately 22.2% (2/9) of the pathogenic variants in the controls (Supplementary Table S2) and 26.9% (7/26) in the cases (Supplementary Table S1) were absent from ExAC, gnomAD²⁰ and TOPMed. Among the cases, pathogenic variants were predominantly frameshift insertions/deletions (INDELs; 38.5%), followed by nonsynonymous variants (26.9%) and stop-gain variants (23.1%), with the remaining comprising non-coding (7.7%) and start-loss (3.8%) variants (Supplementary Fig. S1A). In the controls, stop-gain variants predominated (44.4%), whereas frameshift INDELs and non-coding variants each accounted for 22.2%, and nonsynonymous variants constituted 11.1% (Supplementary Fig. S1B).

Discussion

Previous breast cancer genetic studies in the Ghanaian population primarily utilised targeted sequencing panels and genotyping arrays to assess germline susceptibility to the disease^{12–14}. While these approaches have provided valuable insights into the circulation of known and novel variants in known breast cancer genes, their limited genomic coverage has constrained the discovery of novel genes and variants that may be unique to Ghanaians. In this study, we address this gap by generating the first Ghanaian breast cancer whole-exome sequence data, enabling a more comprehensive investigation of germline predisposition in this population. Our exome-based approach broadened the range of detectable variants and genes, while offering a rich resource to facilitate the discovery of Ghanaian-specific breast cancer genetic risk factors.

We detected pathogenic variants in 27.4% of the BCOPGh (Ghana Breast Cancer Omics Project) cohort, including breast cancer 1 early-onset (*BRCA1*) and breast cancer 2 early-onset *BRCA2* mutations in 3.5%. The previously observed autosomal dominant *BRCA1* variant rs80357708 has been associated with hereditary breast cancer risk in Lebanese Arab women²¹. To the best of our knowledge, this is the first observation of this variant in the Ghanaian population. Two novel *BRCA2* protein-truncating variants were identified in our cohort: a 5-base frameshift deletion c.6331_6335del;p.(R2112fs*15) and a 4-bp deletion c.3192_3195del; p.(N1066fs*10) detected in the *BRCA2* mutational hotspot (exon 11), harbouring the most significant pathogenic and frequent variants²². The former variant begins 2 base pairs upstream of rs397507369, a variant previously associated with breast cancer susceptibility in women from Shanghai, China¹⁶, and lies at the start of the *RAD51*-binding domain²³. Although its precise functional impact remains to be experimentally validated, this frameshift deletion is predicted to produce a truncated *BRCA2* protein approximately one-third the normal length, retaining only the *PALB2*-interaction domain. The second variant begins 3-bp upstream of rs80359375, which has been reported in Korean breast cancer patients¹⁷, and does not fall within any annotated functional domain. Nonetheless, this frameshift deletion is expected to truncate the *BRCA2* protein to roughly half its original size, losing all regions downstream the *RAD51*-interaction domain²³. Notably, our study is the first to report these novel *BRCA2* variants. Contrary to the report by Ahearn *et al.*¹⁴ no *BRCA1* and *BRCA2* variants were observed in the controls in this study. Additionally, the lower *BRCA1* and *BRCA2* mutation frequency observed in our study, compared with the 6.1% reported by Ahearn *et al.* in Ghana¹⁴, may be attributable to our moderate sample size. Nonetheless, our findings support prior evidence that Ghanaian women are more

likely to carry deleterious *BRCA1* and *BRCA2* mutations¹⁴ and the observation of higher *BRCA2* sequence diversity in Africans²⁴.

Additionally, the Fanconi Anaemia complementation group A (*FANCA*) variant rs2040094645, previously associated with Fanconi anaemia (FA)²⁵, was detected in one breast cancer patient. Although this variant has not been directly implicated in breast cancer, the *FANCA* gene is located within a loss-of-heterozygosity hotspot in breast tumours²⁶. Moreover, a rare *FANCA* duplication has been associated with breast cancer in an Iranian cohort²⁷, and the intronic variant rs1061646 has been linked to an 8% increase in breast cancer risk²⁸. *FANCA* is a core component of the Fanconi anaemia/BRCA DNA repair pathway and its associated protein network²⁹. Functionally, *FANCA* facilitates the activation of Fanconi anaemia complementation group D1 (*FANCD1*), better known as *BRCA2*³⁰. Notably, relatives of FA-D1 (homozygous *FANCD1* mutant) patients of Ashkenazi Jewish ancestry exhibit an elevated lifetime risk of breast and ovarian cancers³¹. Emerging evidence supports a broader role for Fanconi anaemia pathway genes in breast cancer predisposition³². Although biallelic *BRCA2/FANCD1* loss is usually lethal in early childhood, thereby preventing adult-onset solid tumours in most FA patients³³, heterozygous *BRCA2/FANCD1* carriers typically develop breast and ovarian cancers in the fourth decade of life³⁴. Other established Fanconi anaemia genes implicated in breast cancer susceptibility include *FANCI* and *FANCF* (also known as *PALB2*)³². Notably, the partner and localizer of *BRCA2* (*PALB2*) variant rs995629797, previously linked to hereditary breast cancer^{19,35}, was found in approximately 7% of the controls but was absent in the cases. Its frequency in the control group was markedly higher than that reported by Ahearn *et al.*¹⁴, who found the variant in 0.06% of controls and 1.03% of cases, potentially suggesting an increasing circulation of *PALB2* variants in the Ghanaian population. Although the control group in this study was drawn from a hearing impairment cohort without prior breast cancer status and family history assessment, the observed prevalence is consistent with the characterisation of *PALB2* variants as moderate-penetrance breast cancer susceptibility alleles³⁶. This further underscores the complex and cumulative interplay of genetic and environmental factors in breast carcinogenesis³⁷. Nevertheless, as high-penetrant breast cancer genes, the *FANCA*, *BRCA1* and *BRCA2* variants in this study's cohort may contribute to the rising early-onset breast cancer in Ghanaian women¹¹, underscoring the need for larger studies with well-characterised control groups to explore potential co-occurrence.

Overall, the pathogenic variants were detected in 25 genes that are typically expressed in human breast tissues at varying levels³⁸. The gap junction protein beta 2 (*GJB2*) gene has been studied in relation to breast cancer; experimental evidence suggests that it mediates ductal carcinoma stemness and progression and may act as a prognostic marker for early-stage breast cancer progression and metastasis³⁹. The *GJB2* variant (rs80338948) is a known variant implicated in autosomal recessive nonsyndromic hearing loss in populations, including the Ghanaian population¹⁸. This variant was observed in one breast cancer patient in this study. Data on hearing was not captured during patient

recruitment; therefore, further studies may explore the possible link between the variant and breast cancer.

HBB (haemoglobin beta), a key oxygen transporter, has been implicated in modulating immune cell infiltration within the breast cancer microenvironment ⁴⁰. Although the detected variant has not been observed in breast cancer cases, the gene has emerged as a potential biomarker for breast cancer progression and tumour aggressiveness ⁴¹. *FUS* has been identified as a driver of metastasis in breast cancer and an early marker of therapy resistance in breast cancer ⁴². Predictive scores suggest that the *FUS* (Fused in sarcoma) variant (rs387906628) in our cohort is likely benign (Supplementary Table S1). Nevertheless, it has been reported as a pathogenic germline variant associated with amyotrophic lateral sclerosis in a South Korean family ⁴³. Given the absence of functional studies and the fact that two independent studies – ⁴³ and ours – have detected this variant, it is crucial to evaluate its true biological effect and potential role in breast tumorigenesis. *ABHD12* (Abhydrolase Domain Containing 12) has been implicated in breast cancer cell migration ⁴⁴. While the functional role of *SERPINB8* (Serpin Family B Member 8) in breast cancer remains unclear, evidence suggests a null expression in certain breast cancer cell lines ⁴⁵. Although population-based studies assessing associations of *HBB*, *FUS*, and *ABHD12* with breast cancer are lacking, existing experimental data and observed variation in population-specific risk profiles highlight the need for well-powered, population-focused investigations as well as functional studies to clarify their potential contributions to the disease susceptibility.

GIPC3 (GIPC PDZ domain containing family member 3) has been identified as a novel functional genetic interactor of *BRCA2*, suggesting a role in modulating *BRCA2*-associated pathways ⁴⁶. Additionally, missense variants in *GIPC3* and *GALC* (galactosylceramidase) have been jointly reported as high-risk breast cancer variants in Swedish familial breast cancer cases ⁴⁷. Despite this, *GALC* has not been functionally investigated in breast cancer, although there are reports of its functional role in lung ⁴⁸ and colorectal ⁴⁹ cancers. These observations necessitate functional characterisation of these genes and their variants in breast cancer, as well as evaluation of their potential as Ghanaian population-specific risk factors in a larger cohort.

Currently, there is no experimental or population-level evidence linking the 11 genes [*ARSA* (arylsulfatase A), *CAPN15* (calpain 15), *CHRNA7* (cholinergic receptor nicotinic gamma subunit), *COLEC11* (collectin subfamily member 11), *CYP4F22*, *DYSF* (dysferlin), *ECEL1*, *GUCY2D*, *KASH5*, *PGAM*, or *SLC6A19*] specifically to breast cancer. Of these genes, *CAPN15*, *CHRNA7* and *CYP4F22*, have not been investigated in cancers. Available expression data ³⁸ suggest that these genes are expressed in breast tissues—*ARSA*, *CAPN15* and *DYSF* at moderate levels, and the remaining seven genes at low levels. Furthermore, homozygous deletions in six of these genes (*ARSA*, *CAPN15*, *ECEL1*, *GUCY2D*, *KASH5* and *SLC6A19*) have been observed in multiple breast cancer cohorts. Given the characteristic dysregulated gene expression observed in cancers, owing to genomic instability ⁵⁰, future studies must explore their potential functional roles in breast tumorigenesis and their relevance in the Ghanaian population. Overall, the

functional roles of the mutated genes in this cohort suggest their involvement in DNA damage repair, tumour immune infiltration, and tumour progression.

To the best of our knowledge, this study represents the first whole-exome sequencing analysis of germline variants in a Ghanaian breast cancer cohort. We identified 26 pathogenic variants across 25 genes, including 14 previously implicated in breast cancer. Variants were detected in four DNA damage-repair genes, three of which are well-established high-penetrance breast cancer genes. Notably, we uncovered 11 putative breast cancer susceptibility genes and three novel variants, including two protein-truncating *BRCA2* alterations. Fully elucidating the functional significance of these variants and genes will require population-specific experimental models, beginning with cell lines that capture the genetic background and phenotypic diversity of Ghanaian breast cancer.

Methods

Ethical approval

The study was conducted in accordance with the Declaration of Helsinki and ethics approval was obtained from the Ghana Health Service Ethics Review Committee (GHS-ERC 006/12/20), the Noguchi Memorial Institute for Medical Research Institutional Review Board (NMIMR-IRB 004/18-19), the Korle-Bu Teaching Hospital Institutional Review Board (KBTH-IRB/00009/2018), and the Ethics committee of the College of Basic and Applied Sciences (ECBAS 045/18-19), University of Ghana. Additional permission to conduct the study was sought from the Ho Teaching Hospital Administration. The study objectives were clearly explained to the participants in English or their chosen Ghanaian languages.

Patient recruitment

After written informed consent documentation, women aged 18 years and older diagnosed with malignant tumours and scheduled for surgery were recruited from three hospitals in the Greater Accra region in Ghana. Each participant was interviewed using a structured questionnaire to obtain information on previous breast cancer and familial cancer history. Subsequently, 2 mL of saliva was collected from all patients using the OGR-500 kit (DNA Genotek, Ottawa, Canada), per the manufacturer's protocol. Approximately 100 mg each of breast tumour and adjacent non-neoplastic breast tissue (tissue estimated to be 5cm away from the tumour) were obtained only from neoadjuvant-naïve patients by expert pathologists.

Whole-exome sequencing and data analyses

Genomic DNA was extracted from 113 participant samples (15 saliva, 77 tumours and 21 non-neoplastic tissues). Initial DNA concentration and quality checks were assessed using the QuantiFluor dsDNA System and Qubit 4 Fluorometer (Invitrogen, Waltham, MA, USA); approximately 220 ng of gDNA per patient sample was submitted for whole-exome sequencing at the Genomics Division – Iowa Institute of Human Genetics, Iowa, USA. The exome libraries were prepared using the Agilent SureSelect XT V6+UTR

human exome baits (Agilent, California, USA) and sequenced on the Illumina NovaSeq 6000 (Illumina, CA, USA) sequencer to generate 150 bp paired-end reads at 100X coverage. The quality of the generated exome sequence data was determined using FastQC (v0.11, <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>, RRID: SCR_014583). The reads were then trimmed with TRIMMOMATIC (v0.39, <https://trimmomatic.com/>, RRID: SCR_011848), and sequences with per-base Phred quality score (Q, RRID: SCR_001017) ≥ 30 and a minimum read length of 30 bp were subsequently mapped to the human reference genome (build hg38) using Burrows-Wheeler Aligner (BWA v0.7, <https://github.com/lh3/bwa/releases>, RRID: SCR_010910). The aligned base quality was recalibrated using Base Quality Score Recalibration (BQSR) and ApplyBQSR (<https://gatk.broadinstitute.org/hc/en-us/articles/360035890531-Base-Quality-Score-Recalibration-BQSR>) as stipulated by the Genome Analysis Toolkit (GATK v4.1.8, RRID: SCR_001876). Germline variants in the samples were jointly called using HaplotypeCaller, following the GATK best-practice guidelines (<https://gatk.broadinstitute.org/hc/en-us/articles/360037225632-HaplotypeCaller>). The variant quality was recalibrated using the Variant Quality Score Recalibration (VQSR) and ApplyVQSR functions of GATK. A hard-filtering criterion prioritised variants, with SNP (-g) and indel (-G) gaps of 10, allelic depth (AD) ≥ 20 QualByDepth (QD) ≥ 2 , mean mapping quality (MMQ) ≥ 60 , mean base quality (MBQ) ≥ 30 and FS ≤ 60 . Existing whole-exome sequencing data (BAM format) for 43 participants of a Ghanaian Genetics of Hearing Impairment cohort ¹⁸ served as controls.

Variant Annotation and Filtering Strategy

Single-nucleotide and insertion/deletion variants were annotated using SnpEff (RRID: SCR_005191), ANNOVAR (RRID: SCR_012821) and dbSNP (<http://www.ncbi.nlm.nih.gov/SNP/>, RRID: SCR_002338) and filtered using the Genome Aggregation Database (gnomAD, <https://gnomad.broadinstitute.org/>) and Exome Aggregation Consortium (ExAC) African-specific MAF $< 0.01\%$. Variants present in the catalogue of somatic mutations in cancer (COSMIC; <http://cancer.sanger.ac.uk/cancergenome/projects/cosmic/>, RRID: SCR_002260) were excluded, and the retained variants were further evaluated using bioinformatic prediction scores from Sorting Intolerant From Tolerant (SIFT), Polymorphism Phenotyping version 2 (Polyphen2), Combined Annotation Dependent Depletion (CADD), MutationAssessor (RRID: SCR_005762), MutationTaster (RRID: SCR_010777), Functional Analysis through Hidden Markov Models (FATHMM), and MetaLR. The pathogenic variants were further prioritised using the American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG-AMP) classification guidelines ⁵¹ and evidence from ClinVar (RRID: SCR_006169), Online Mendelian Inheritance in Man (OMIM, <https://www.omim.org/>), Trans-Omics for Precision Medicine (TOPMed, <https://topmed.nhlbi.nih.gov/>) and Ensembl (<http://www.ensembl.org/>, RRID: SCR_002344) databases. Pathogenicity of the variants was further supported by at least four concordant predictive scores.

Structural Modelling of Variant Consequence

The coding sequences of the MANE select transcripts were retrieved from the National Centre for Biotechnology Information (NCBI; <https://www.ncbi.nlm.nih.gov/>), and variants were manually introduced

into the sequences. The “wildtype” and mutant sequences were translated using ExPASy (<https://web.expasy.org/translate/>, RRID: SCR_024703), the resulting protein structures were modelled with AlphaFold3 (<https://alphafoldserver.com/>), and the structures were visualised and overlapped using PyMol (version 3.18, RRID: SCR_000305). Details on protein features and domains were obtained from UniProtKB (<https://www.uniprot.org/>, RRID: SCR_004426).

Statistical Analysis

Descriptive statistical analysis was performed using the open-source statistical analysis software R (version 4.5.0, <https://cran.r-project.org/>). Proportions of variants detected between disease groups were calculated. The sample size was not based on any established sample-size estimation formula because this manuscript reports on participants recruited to date for an ongoing breast cancer genetic project.

Declarations

Data Availability

The data in this study are not publicly available because they are part of a larger, unfinished project, but are available upon reasonable request to the corresponding author.

Code availability

No custom code was generated in this study. Analyses were performed using publicly available bioinformatic pipelines, with any modifications described in the Sequence Analysis section of the Methods.

Acknowledgements

We are grateful to Ambroise Wonkam for providing the BAM files for the control, and to Michelle Abena Buckman, Elizabeth Mbirbah, Naa Dedei Quarshie, Samuel Degraft Mensah, Francis Asiedu Ampaw, Precious Oporum, and Queenster Owusu for their assistance with participant recruitment and data collection. This study was funded by the World Bank African Centres of Excellence (grant number WACCBIP+NCDs: Awandare) and the DELTAS African Initiative (grant number DEL-22-014: Awandare). The funders played no role in the study design, data collection, analysis and interpretation of data, or the writing of this manuscript.

Author contributions

LP conceptualised the study, secured funding, supervised the project, and contributed substantially to writing the original draft, reviewing and editing the manuscript. CAA recruited patients, prepared samples for sequencing, analysed, visualised, and interpreted the data, performed structural prediction, and contributed significantly to the original draft and manuscript revision. VP managed the data, refined analysis codes, analysed the data, and contributed to writing the original draft. EKA analysed the data

and contributed to the review and editing of the manuscript. DA and BT performed structural modelling of novel variants and contributed to writing the original draft. JN, JNC, FD, LE, RA, AT, and NA recruited patients, confirmed cases, and contributed significantly to the review and editing of the manuscript. SYB and LNA supervised data analysis and interpretation and contributed substantially to the manuscript's review and editing. RJW supervised the work, performed case classification, and was a major contributor to reviewing and editing. All authors read and approved the final manuscript.

Competing interests

All authors declare no financial or non-financial competing interests.

References

1. Xiong, X. *et al.* Breast cancer: pathogenesis and treatments. *Signal Transduction and Targeted Therapy* 10, 49–49, doi:<https://doi.org/10.1038/s41392-024-02108-4> (2025).
2. Bray, F. *et al.* Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians* 74, 229–263, doi:<https://doi.org/10.3322/caac.21834> (2024).
3. GLOBOCAN. *Cancer Today*, <https://gco.iarc.fr/today/en/dataviz/pie?mode=cancer&group_populations=1&populations=288&sexes=2&types=0&age_end=9&cancers=20> (2024).
4. GLOBOCAN. *Cancer Tomorrow*, <https://gco.iarc.fr/tomorrow/en/dataviz/bubbles?types=1&sexes=1_2&mode=cancer&group_populations=1&multiple_populations=1&multiple_cancers=1&cancers=20&populations=903_904_905_908_909_935&apc=cat_ca20v1.5_ca23v-1.5&group_cancers=1&bar_mode=stacked&years=2035> (2024).
5. International Agency for Research on, C. (2024).
6. Claus, E. B., Risch, N. J. & Thompson, W. D. Age at onset as an indicator of familial risk of breast cancer. *American journal of epidemiology* 131, 961–972, doi:<https://doi.org/10.1093/oxfordjournals.aje.a115616> (1990).
7. Pal, M., Das, D. & Pandey, M. Understanding genetic variations associated with familial breast cancer. *World Journal of Surgical Oncology* 22, 271–271, doi:10.1186/S12957-024-03553-9 (2024).
8. Girard, E. *et al.* Familial breast cancer and DNA repair genes: Insights into known and novel susceptibility genes from the GENESIS study, and implications for multigene panel testing. *International Journal of Cancer* 144, 1974–1974, doi:<https://doi.org/10.1002/ijc.31921> (2018).
9. Wendt, C. & Margolin, S. Identifying breast cancer susceptibility genes – a review of the genetic background in familial breast cancer. *Acta Oncologica* 58, 135–146, doi:<https://doi.org/10.1080/0284186X.2018.1529428> (2019).
10. Pfennig, A., Petersen, L. N., Kachambwa, P. & Lachance, J. Evolutionary Genetics and Admixture in African Populations. *Genome Biology and Evolution* 15, evad054-evad054,

doi:<https://dx.doi.org/10.1093/gbe/evad054> (2023).

11. Akakpo, P. K. et al. Clinicopathologic characteristics of early-onset breast cancer: a comparative analysis of cases from across Ghana. *BMC Women's Health* 23, 5–5, doi:<https://doi.org/10.1186/s12905-022-02142-w> (2023).
12. Newman, L. A. et al. Hereditary Susceptibility for Triple Negative Breast Cancer Associated with Western Sub-Saharan African Ancestry: Results from an International Surgical Breast Cancer Collaborative. *Annals of Surgery* 270, 484–492, doi:10.1097/SLA.0000000000003459 (2019).
13. Martini, R. et al. Investigation of triple-negative breast cancer risk alleles in an International African-enriched cohort. *Scientific Reports* 2021 11:1 11, 1–12, doi:<https://doi.org/10.1038/s41598-021-88613-w> (2021).
14. Ahearn, T. U. et al. Breast Cancer Risk in Women from Ghana Carrying Rare Germline Pathogenic Mutations. *Cancer Epidemiology Biomarkers and Prevention* 31, 1593–1601, doi:<https://doi.org/10.1158/1055-9965.epi-21-1397> (2022).
15. Borg, A. et al. Characterization of BRCA1 and BRCA2 deleterious mutations and variants of unknown clinical significance in unilateral and bilateral breast cancer: the WECARE study. *Hum Mutat* 31, E1200-1240, doi:<https://doi.org/10.1002/humu.21202> (2010).
16. Suter, N. M. et al. BRCA1 and BRCA2 mutations in women from Shanghai China. *Cancer Epidemiol Biomarkers Prev* 13, 181–189, doi:<https://doi.org/10.1158/1055-9965.EPI-03-0196> (2004).
17. Kim, B. Y. et al. Identification of BRCA1 and BRCA2 mutations from Korean breast cancer patients using denaturing HPLC. *Biochem Biophys Res Commun* 349, 604–610, doi:<https://doi.org/10.1016/j.bbrc.2006.08.086> (2006).
18. Wonkam, A. et al. Exome sequencing of families from Ghana reveals known and candidate hearing impairment genes. *Communications Biology* 5, 369–369, doi:<https://doi.org/10.1038/s42003-022-03326-8> (2022).
19. Antoniou, A. C. et al. Breast-Cancer Risk in Families with Mutations in PALB2. *New England Journal of Medicine* 371, 497–506, doi:<http://www.nejm.org/doi/10.1056/NEJMoa1400382> (2014).
20. Chen, S. et al. A genomic mutational constraint map using variation in 76,156 human genomes. *Nature* 625, 92–100, doi:<https://doi.org/10.1038/s41586-023-06045-0> (2024).
21. El Saghir, N. S. et al. BRCA1 and BRCA2 Mutations in Ethnic Lebanese Arab Women With High Hereditary Risk Breast Cancer. *The Oncologist* 20, 357–364, doi:<https://doi.org/10.1634/theoncologist.2014-0364> (2015).
22. Karami, F. & Mehdipour, P. A comprehensive focus on global spectrum of BRCA1 and BRCA2 mutations in breast cancer. *Biomed Res Int* 2013, 928562, doi:<https://doi.org/10.1155/2013/928562> (2013).
23. Ahmad, S. et al. The UniProt website API: facilitating programmatic access to protein knowledge. *Nucleic Acids Research* 53, W547-W553, doi:<https://doi.org/10.1093/nar/gkaf394> (2025).
24. Wagner, T. M. U. et al. Global Sequence Diversity of BRCA2: Analysis of 71 Breast Cancer Families and 95 Control Individuals of Worldwide Populations. *Human molecular genetics* 8, 413–423,

- doi:<https://doi.org/10.1093/hmg/8.3.413> (1999).
25. Gille, J. J. et al. Diagnosis of Fanconi Anemia: Mutation Analysis by Multiplex Ligation-Dependent Probe Amplification and PCR-Based Sanger Sequencing. *Anemia* 2012, 603253, doi:<https://doi.org/10.1155/2012/603253> (2012).
 26. Cleton-Jansen, A. M. et al. Loss of heterozygosity mapping at chromosome arm 16q in 712 breast tumors reveals factors that influence delineation of candidate regions. *Cancer Res* 61, 1171–1177 (2001).
 27. Abbasi, S. & Rasouli, M. A rare FANCA gene variation as a breast cancer susceptibility allele in an Iranian population. *Mol Med Rep* 15, 3983–3988, doi:<https://doi.org/10.3892/mmr.2017.6489> (2017).
 28. Haiman, C. A. et al. Comprehensive association testing of common genetic variation in DNA repair pathway genes in relationship with breast cancer risk in multiple populations. *Hum Mol Genet* 17, 825–834, doi:<https://doi.org/10.1093/hmg/ddm354> (2008).
 29. Venkitaraman, A. R. Tracing the network connecting BRCA and Fanconi anaemia proteins. *Nat Rev Cancer* 4, 266–276, doi:<https://doi.org/10.1038/nrc1321> (2004).
 30. Howlett, N. G. et al. Biallelic Inactivation of *BRCA2* in Fanconi Anemia. *Science* 297, 606–609, doi:<https://www.science.org/doi/abs/10.1126/science.1073834> (2002).
 31. Offit, K. et al. Shared Genetic Susceptibility to Breast Cancer, Brain Tumors, and Fanconi Anemia. *JNCI: Journal of the National Cancer Institute* 95, 1548–1551, doi:<https://doi.org/10.1093/jnci/djg072> (2003).
 32. D'Andrea, A. D. Susceptibility pathways in Fanconi's anemia and breast cancer. *N Engl J Med* 362, 1909–1919, doi:<https://www.nejm.org/doi/full/10.1056/NEJMra0809889> (2010).
 33. Berwick, M. et al. Genetic heterogeneity among Fanconi anemia heterozygotes and risk of cancer. *Cancer Res* 67, 9591–9596, doi:<https://doi.org/10.1158/0008-5472.CAN-07-1501> (2007).
 34. Ford, D., Easton, D. F., Bishop, D. T., Narod, S. A. & Goldgar, D. E. Risks of cancer in BRCA1-mutation carriers. Breast Cancer Linkage Consortium. *Lancet* 343, 692–695, doi:[https://doi.org/10.1016/s0140-6736\(94\)91578-4](https://doi.org/10.1016/s0140-6736(94)91578-4) (1994).
 35. Janatova, M. et al. The PALB2 gene is a strong candidate for clinical testing in BRCA1- and BRCA2-negative hereditary breast cancer. *Cancer Epidemiol Biomarkers Prev* 22, 2323–2332, doi:<https://doi.org/10.1158/1055-9965.EPI-13-0745-T> (2013).
 36. Shiovitz, S. & Korde, L. A. Genetics of breast cancer: a topic in evolution. *Annals of oncology: official journal of the European Society for Medical Oncology* 26, 1291–1299, doi:<https://doi.org/10.1093/annonc/mdv022> (2015).
 37. Anothaisintawee, T. et al. Risk factors of breast cancer: a systematic review and meta-analysis. *Asia Pac. J. Public Health* 25, 368–387, doi:<https://doi.org/10.1177/1010539513488795> (2013).
 38. Dyer, S. C. et al. Ensembl 2025. *Nucleic Acids Research* 53, D948–D957, doi:<https://doi.org/10.1093/nar/gkae1071> (2024).

39. Liu, Y. *et al.* ID2 and GJB2 promote early-stage breast cancer progression by regulating cancer stemness. *Breast Cancer Research and Treatment* 175, 77–90, doi:<https://doi.org/10.1007/s10549-018-05126-3> (2019).
40. Yin, B., Luo, X., Yan, X., Shen, H. & Jiang, J. PDE3B and HBB are key prognostic biomarkers driving cell proliferation and regulating immune microenvironment in breast cancer. *Hereditas* 162, 97–97, doi:<https://doi.org/10.1186/s41065-025-00470-z> (2025).
41. Ponzetti, M. *et al.* Non-conventional role of haemoglobin beta in breast malignancy. *British Journal of Cancer* 117, 994–1006, doi:<https://doi.org/10.1038/bjc.2017.247> (2017).
42. Yang, L. *et al.* LINC00460-FUS-MYC feedback loop drives breast cancer metastasis and doxorubicin resistance. *Oncogene* 43, 1249–1262, doi:<https://doi.org/10.1038/s41388-024-02972-y> (2024).
43. Yan, J. *et al.* Frameshift and novel mutations in FUS in familial amyotrophic lateral sclerosis and ALS/dementia. *Neurology* 75, 807–814, doi:<https://doi.org/10.1212/wnl.0b013e3181f07e0c> (2010).
44. Jun, S., Kim, S. W., Lim, J. Y. & Park, S. J. ABHD12 Knockdown Suppresses Breast Cancer Cell Proliferation, Migration and Invasion. *Anticancer research* 40, 2601–2611, doi:<https://doi.org/10.21873/anticancer.14231> (2020).
45. Yamamoto, F. & Yamamoto, M. Scanning copy number and gene expression of the 18q21-qter chromosomal region by the systematic multiplex PCR and reverse transcription-PCR methods. *Electrophoresis* 28, 1882–1895, doi:<https://doi.org/10.1002/elps.200700093> (2007).
46. Ding, X. *et al.* Survival of BRCA2-deficient cells is promoted by GIPC3, a novel genetic interactor of BRCA2. *Genetics* 207, 1335–1345, doi:<https://doi.org/10.1534/genetics.117.300357> (2017).
47. Helgadottir, H. T., Thutkawkorapin, J., Lagerstedt-Robinson, K. & Lindblom, A. Sequencing for germline mutations in Swedish breast cancer families reveals novel breast cancer risk genes. *Scientific Reports* 11, 1–7, doi:<https://doi.org/10.1038/s41598-021-94316-z> (2021).
48. Liu, D. G., Xue, L., Li, J., Yang, Q. & Peng, J. Z. Epithelial-mesenchymal transition and GALC expression of circulating tumor cells indicate metastasis and poor prognosis in non-small cell lung cancer. *Cancer Biomarkers* 22, 417–426, doi:<https://doi.org/10.3233/cbm-170995> (2018).
49. Yang, M. *et al.* GALC Triggers Tumorigenicity of Colorectal Cancer via Senescent Fibroblasts. *Frontiers in Oncology* 10, 380–380, doi:<https://doi.org/10.3389/fonc.2020.00380> (2020).
50. Li, F. *et al.* R-Loops in Genome Instability and Cancer. *Cancers (Basel)* 15, doi:<https://doi.org/10.3390/cancers15204986> (2023).
51. Biesecker, L. G. & Harrison, S. M. The ACMG/AMP reputable source criteria for the interpretation of sequence variants. *Genet Med* 20, 1687–1688, doi:[10.1038/gim.2018.42](https://doi.org/10.1038/gim.2018.42) (2018).

Figures

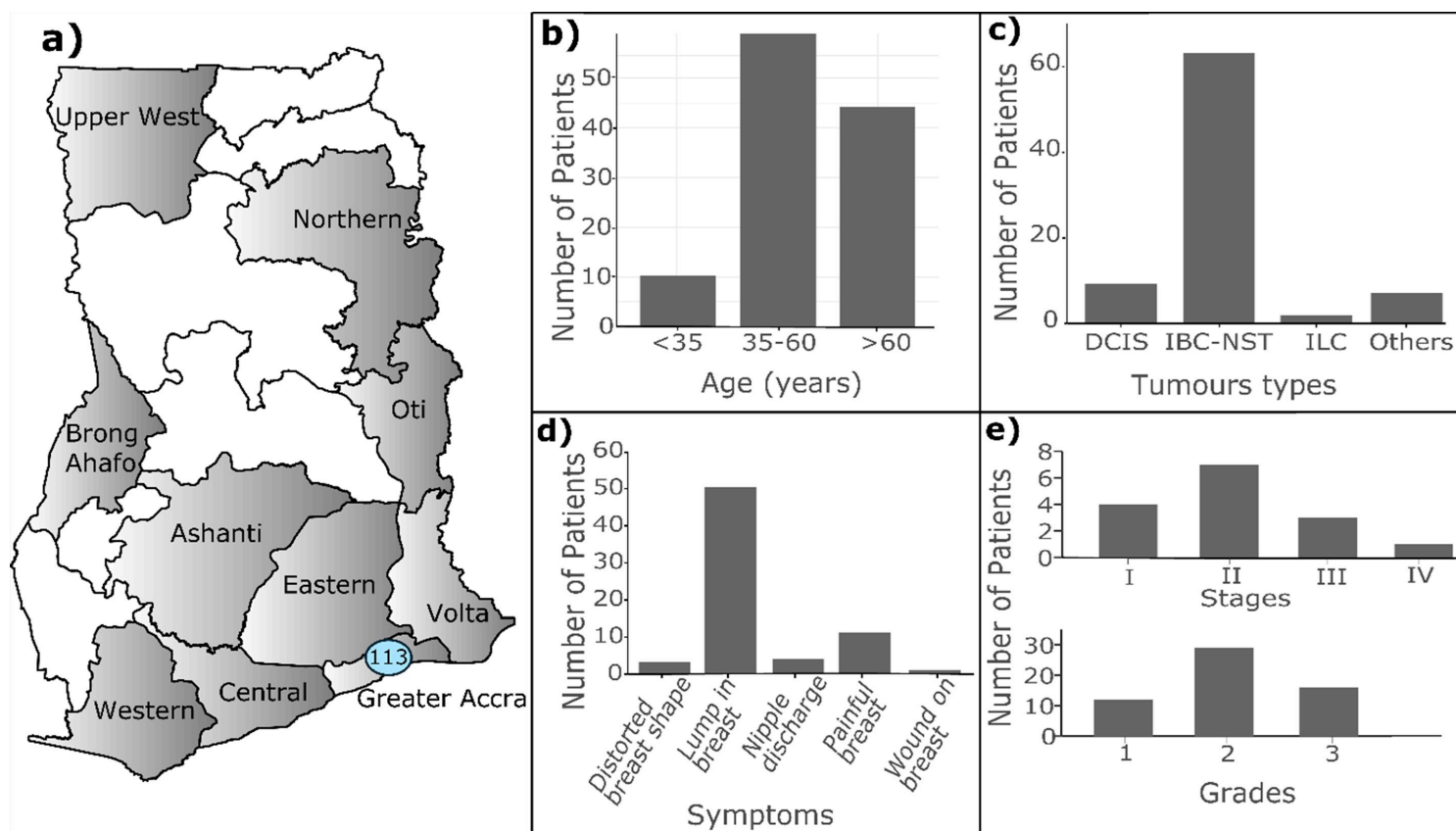


Figure 1

summary of patient demographics and clinicopathological features. The geographical origin of participants by self-reported ethnicity (a), the age ranges of participants (b), tumour types (c), disease symptoms (d) and tumour stages and grades (e). The participants originated from 10 regions (shaded regions) in Ghana; the recruitment sites were in the Greater Accra region and the number of participants recruited is indicated in the blue circle (a). DCIS – Ductal carcinoma in situ; IBC-NST – Invasive breast carcinoma No Special Type (NST); IDC – Invasive ductal Carcinoma; ILC – Invasive lobular carcinoma; Others – Intracystic papillary carcinoma, mucinous carcinoma, squamous cell carcinoma and other (malignant neoplasm of central portion without histology).

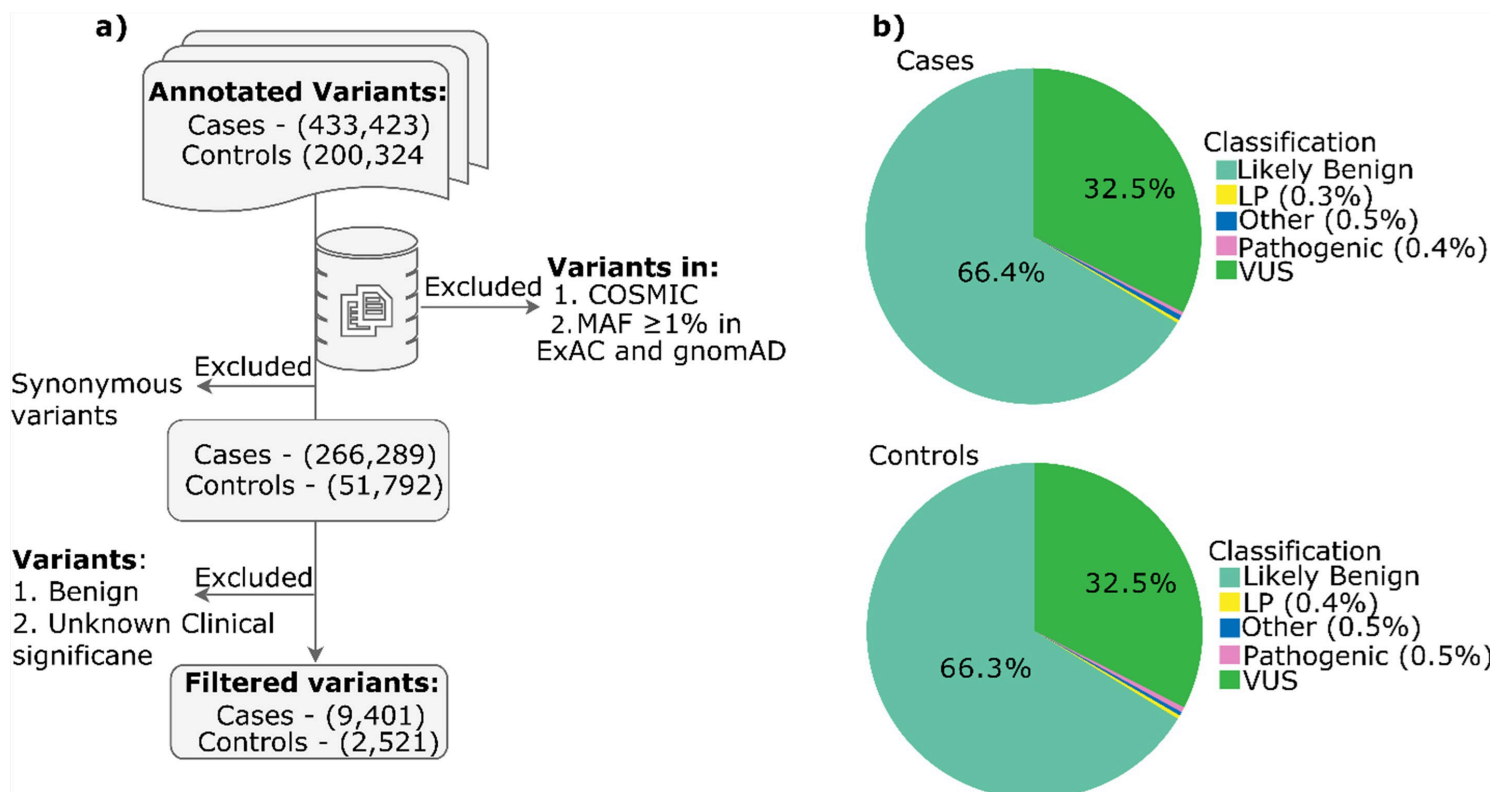


Figure 2

summary of the variant filtering and classification. Panel a: Variant filtration steps panel b: and classification. Most of the variants were likely benign in both cases and controls. ExAC – Exome Aggregation Consortium; gnomAD – Genome Aggregation Consortium; COSMIC – Catalogue of somatic mutations in cancer; MAF – Minor Allele Frequency; LP – Likely Pathogenic.

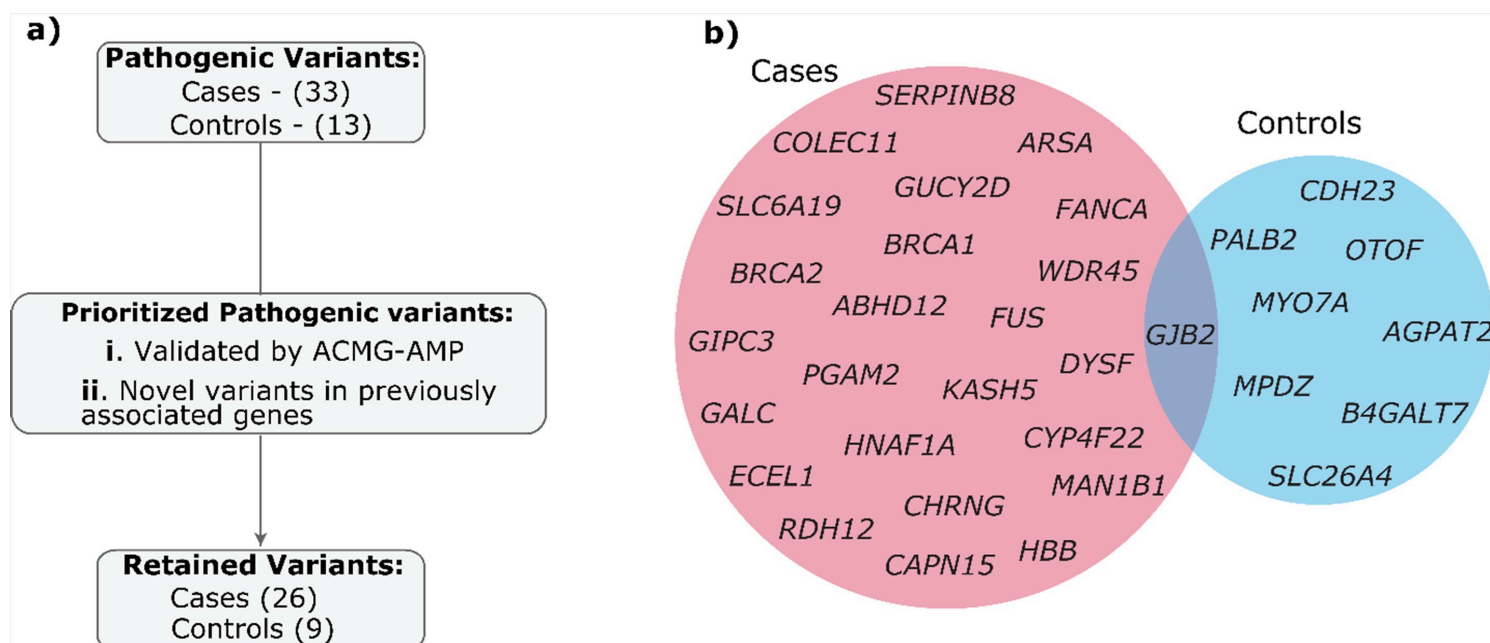


Figure 3

Schematic representation of pathogenic variant prioritisation and validation. Panel a: variant identification and prioritisation by study. Panel b: genes containing the prioritised variants post ACMP-AMP validation. Twenty-six pathogenic variants were detected in 25 genes in the cases and 9 variants in 9 genes in the controls. The same *GJB2* variant was detected in both cases and controls. ACMP-AMP is the American College of Medical Genetics and Genomics and the Association for Molecular Pathology

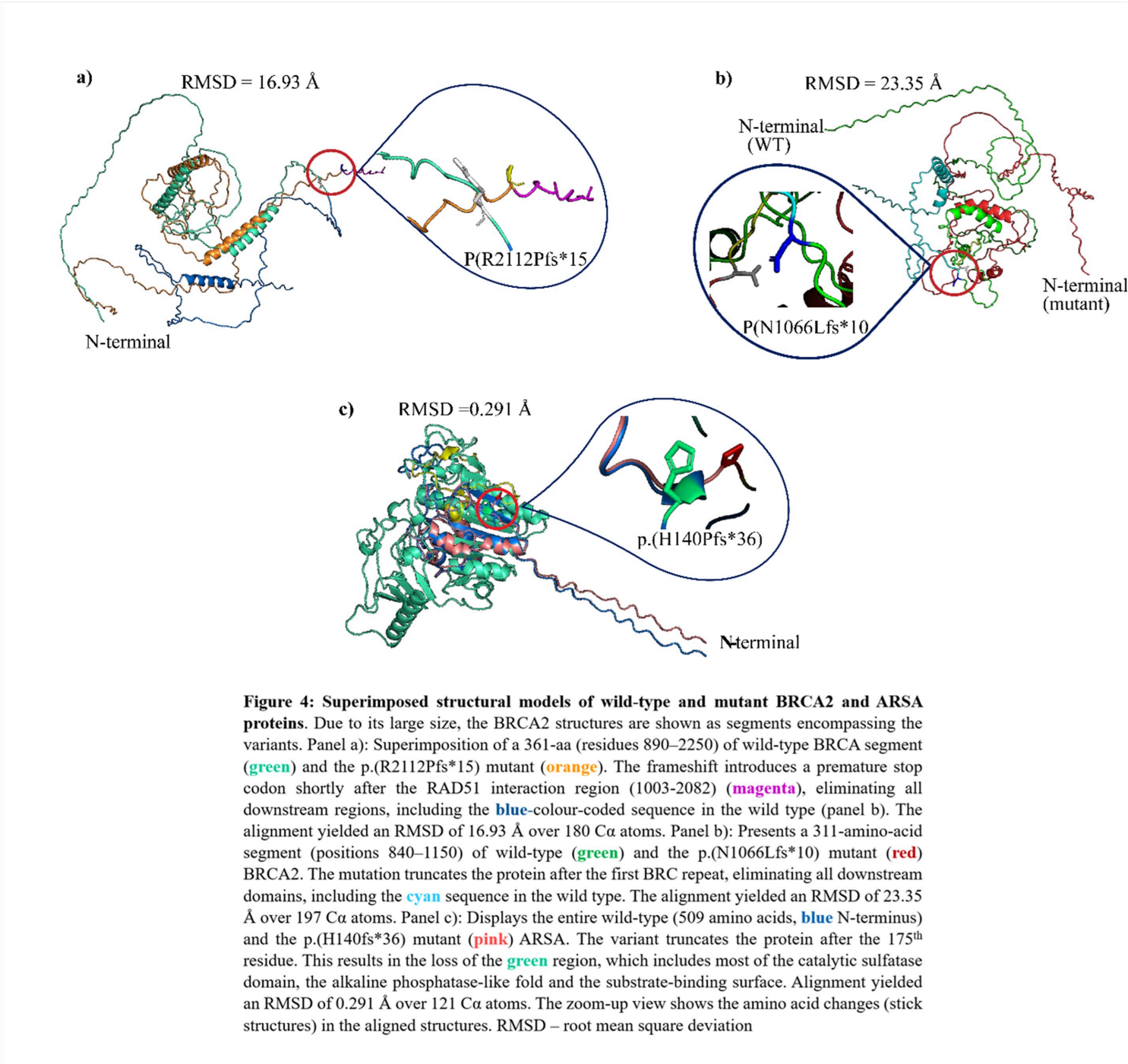


Figure 4

See image above for figure legend

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryData.docx](#)