

Mutational and Expression Profile of ZNF217, ZNF750, ZNF703 Zinc Finger Genes in Kenyan Women Diagnosed with Breast Cancer

Michael Kitoi

michael.o.kitoi@gmail.com

Jomo Kenyatta University of Agriculture and Technology

John Gitau

Université du Québec à Montréal

Godfrey Wagutu

Mount Kenya University

Kennedy Mwangi

International Livestock Research Institute

Florence Ngonga

Jomo Kenyatta University of Agriculture and Technology

Francis Makokha

Mount Kenya University

Research Article

Keywords:

Posted Date: September 25th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10598686/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

**Mutational and Expression Profile of ZNF217, ZNF750, ZNF703 Zinc Finger Genes in
Kenyan Women Diagnosed with Breast Cancer**

**Michael Kitoi^{13*}, John Gitau², Godfrey Wagutu³, Kennedy Mwangi⁴, Florence Ngonga¹,
Francis Makokha³**

¹ Jomo Kenyatta University of Agriculture and Technology (JKUAT)

² Département de Chimie, Université du Québec à Montréal, Montréal, Québec, Canada

³ Directorate of Research and Innovation, Mount Kenya University, Thika, Kenya

⁴ International Livestock Research Institute, Nairobi, Kenya

*** Correspondence:**

Michael Kitoi

Michael.o.kitoi@gmail.com

Abstract

Background: Breast cancer remains a major public health challenge globally and is the leading cancer among women in Kenya. Despite remarkable progress in molecular developments, the genomic and transcriptional features of breast cancer in African populations are still not fully elucidated. While zinc finger proteins (ZNFs) are central regulators of gene expression and have implications for cancer, their involvement in development has not yet been elucidated in breast cancer patients in Kenya. The current study profiled the mutational characteristics of ZNF217, ZNF703, and ZNF750, compared their expression patterns in tumor and matched adjacent non-tumor tissues, and analyzed associations with clinical aspects.

Methods: Whole-exome and RNA sequencing analyses were performed on 23 pair-matched tumor-normal samples. The somatic mutations were extracted using MuTect2, following standard preprocessing and annotation steps, and identified at protein regions. Gene expression was quantified using STAR, featureCounts, and DESeq2. Paired statistical methods were used to assess differential expression between tumor and matched normal tissues using false discovery rate correction. Associations between gene expression and clinical variables were evaluated using correlation analysis and multivariable regression models.

Results: ZNF217 (170 mutations) and ZNF750 (164 mutations) had high mutation load consisting mostly of single-nucleotide variants, and that of ZNF703 (24 mutations) was reported less frequently. However, mutations occurred across protein sequences in general, with limited domain enrichment; a recurrent ZNF217 mutation overlapped with a zinc finger domain. Comparison with TCGA data showed a much lower mutation frequency, indicating population specificity. Tumors were notable for upregulation of all three genes relative to normal tissues (ZNF217: $p = 0.00068$; ZNF703: $p = 0.00475$; ZNF750: $p = 0.00366$). In most samples, tumor expression exceeded matched normal expression, and a small number of samples showed the opposite expression trend, indicating molecular heterogeneity. PC1 and PC2 accounted for 75.41% and 13.41% of the total variance for tumor and normal samples, respectively, according to principal component analysis. Clinical results were limited; however, ZNF703 expression was associated with body mass index ($\beta = 0.194$, $p = 0.025$), and ZNF750 expression was elevated in estrogen receptor-positive tumors ($\beta = 1.050$, $p = 0.005$).

Conclusions: These results show that ZNF217, ZNF703, and ZNF750 are functionally both GM (genetically and transcriptionally dysregulated) in breast cancer and play different functions in tumor biology in breast cancer. Findings of this study underscore the necessity of community-derived genomic research and offer novel perspectives on zinc finger gene expression in overlooked groups.

Background

Breast cancer is the most frequently diagnosed cancer among women worldwide and plays a prominent role in cancer-related death. Latest global estimates show that there were around 2.3 million new breast cancer cases and 670,000 deaths in 2022, which accounts for around one in four cases diagnosed in women [1, 2]. Despite the progress achieved by early identification of cancers as well as molecular diagnosis and targeted therapies, disparities remain between high-income and low-income nations in terms of incidence, mortality, and survival in general [3]. Such inequalities are prevalent in LMICs, where limited screening, diagnosis, and treatment lead to unsatisfactory clinical outcomes [4].

Breast cancer is the most common cancer diagnosed in women in Kenya and a leading cause of cancer-related death [5]. Specifically, Kenya recorded approximately 7,243 new breast cancer cases and 3,398 deaths in 2022, reflecting the extensive public health burden the disease contributes to [1]. Breast cancer is among the top five cancers amongst women population in sub-Saharan Africa, which is higher (29.2%) compared to the global statistics (23.8%) [6]. Yet, African populations are systematically underrepresented in global or genomic studies that have provided further insights into breast cancer biology [7]. As a result, crucial, population-specific molecular mutations might remain unexplored, limiting the translation of precision oncology methodologies developed on predominantly European and North American cohorts.

Breast cancer is a heterogeneous disease with multiple molecular and genetic factors that can cause changes in the initiation, progression, treatment, and prognosis of the tumor [8]. Extensive genomic and transcriptomic profiling has revealed genetic alterations, copy number variations and recurrent genetic fusion, ergo, improving molecular classification and enabling targeted treatments [9]. Despite many of these successes, most of these observations have been made in

populations that do not reflect the high genetic diversity of African patients, indicating the need for a population-specific investigation.

Zinc finger proteins (ZNFs) are among the largest families of transcription factors in the human genome and are essential regulators of gene expression, cell differentiation, expansion, proliferation, apoptosis, chromatin remodeling, and DNA repair [10, 11]. Several zinc finger proteins have been identified as playing roles in cancer development and progression through their effects on epithelial-mesenchymal transition, metastasis, stemness, and therapeutic resistance [12]. Recently, it was found that ZNF217, ZNF703, and ZNF750 are three of the specific genes of interest in breast cancer biology.

ZNF217, which resides on chromosome 20q13, is often overexpressed and expanded in breast tumors, further contributing to enhanced tumor aggressiveness, metastatic potential, endocrine resistance, and adverse clinical outcome [13, 14]. ZNF703, an oncogenic transcription factor found on chromosome 8p12 amplicon, is frequently overexpressed in luminal B breast cancer and is associated with increased proliferation, tumor progression, and resistance against endocrine therapy [15, 16]. Whereas ZNF750 is classified as a tumor suppressor, and interacts with chromatin regulators CTBP1/2, KFM1A and transcriptional factors KLF4 [17]. ZNF750 has been implicated in epithelial differentiation and also in maintaining cellular homeostasis [18]. Alteration of expression or functional disruption of ZNF750 has been implicated in aggressive tumor phenotypes and pathogenesis in various epithelial cancers, including breast cancer [19].

Despite the biological significance of ZNF217, ZNF703, and ZNF750 and their roles in breast cancer, their mutational and transcriptional phenotypes have not been widely studied in African populations. Thus far, there is no comprehensive study characterizing these genes in a Kenyan female population with breast cancer using integrative genomic and transcriptomic approaches. Accordingly, the current study also aimed to reveal the mutational and transcriptional profiles of ZNF217, ZNF703, and ZNF750 in Kenyan women with breast cancer and to assess their associations with clinical and pathological features of interest. This study contributes to the growing body of evidence needed to advance precision oncology in underrepresented African populations by providing population-specific molecular data.

Materials and Methods

Design Study and Sample Collection

This study was a secondary genomic analysis of the Kenyan cohort from a multicenter breast cancer sequencing study [20], studying population-specific mutational patterns in African American, European American, Asian American, and Kenyan patients using whole-exome sequencing (WES) techniques. To characterize the mutational and transcriptional signatures of the breast cancer mutations of ZNF217, ZNF703, and ZNF750, the current study was conducted in a Kenyan cohort only. Patients were recruited after submitting written informed consent at AIC Kijabe Hospital and Aga Khan University Hospital in Nairobi, Kenya for histologically confirmed breast cancer between 2019 and 2021. Ethical approval was obtained from the Mount Kenya University Institutional Ethics Review Committee (IERC Ref: 2018/REC-80 v7) as well as National Commission for Science, Technology and Innovation (NACOSTI License No. NACOSTI/P/22/21374).

A total of 80 fresh-frozen tissue specimens were collected from 40 Kenyan patients who were subjected to pathological review at the Laboratory of Human Carcinogenesis, National Cancer Institute (USA). Out of these individuals, 23 showed features according to inclusion criteria of tumor specimens having >10% tumor cellularity and matched adjacent normal tissues without tumor cells, yielding 46 specimens for WES and RNA-seq analysis. The corresponding WES and RNA-seq collections are hosted in the NCBI Sequence Read Archive (SRA; <https://www.ncbi.nlm.nih.gov/sra>; accession PRJNA913947) and Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>; accession GSE225846) respectively.

Quality Control of Sequencing Information and Pre-processing of Sequencing Evidence

Quality assessment of raw WES and RNA-seq reads was done at FastQC v0.11.9. The quality metrics assessed included per-base sequence quality, sequence duplication rates, GC content, sequence length distribution, adapter contamination, overrepresented sequences, and k-mer enrichment. Reads with poor-quality bases or adapter contamination were processed through Trimmomatic v0.39. After trimming, FastQC was repeated to ensure that sequence quality

improved and technical artifacts were eliminated. Subsequent analyses retained only high-quality reads.

Whole-Exome Sequencing Analysis and Somatic Variant Detection

Sequence reads were matched to the human reference genome (GRCh38/hg38), and the maps extracted from the alignment files were converted to BAM format, sorted, indexed, and annotated with read groups using Picard. Base Quality Score Recalibration (BQSR) was implemented in Genome Analysis Toolkit (GATK v4.2), using known variant datasets from dbSNP and Mills, as well as 1000G gold-standard indels. The results from this benchmark will be applied to other studies. All downstream analyses used recalibrated BAM files. Somatic variants were identified using GATK Mutect2, following a tumour–normal paired analysis strategy. Normal samples were initially processed to generate single-variant call files, which were then blended into a Panel of Normals (PoN) using CreateSomaticPanelOfNormals. Tumor samples were then assessed in comparison with matched normal controls and the PoN to enhance discrimination between somatic variants and technical artifacts. Variant calls were set to normalize using vt normalize, and filtered to maintain high-confidence single-nucleotide variants and insertions/deletions. Functional annotation was obtained using Ensembl Variant Effect Predictor (VEP).

Characterization of ZNF Gene Mutations

Genomic Coordinates of ZNF217, ZNF703, and ZNF750 were extracted from the UCSC Genome Browser (GRCh38). Differences in genes were extracted and visually assessed using Integrative Genomics Viewer (IGV) to assess read support and alignment quality. Variants were grouped by structural type (single nucleotide modifications, insertions, and deletions) and predicted functional effect (synonymous variants, missense, frameshift, and in-frame deletion variants). The protein domain annotations were obtained from UniProt to represent the mapped mutations in both functional domains of each protein. The recurrence of mutations was defined as identical amino acid changes present in more than one patient sample. Mutation profiles were also compared with publicly available breast cancer mutation data from The Cancer Genome Atlas (TCGA).

RNA Sequencing and Gene Expression Analysis

STAR-aligned RNA-seq reads were mapped to the human reference genome, and gene-level counts were generated using featureCounts. Raw counts were normalized, and the median of ratios was used as designed in DESeq2. A variance-stabilizing transformation was used for visualization and multivariate analyses, and normalized count data were employed for statistical analyses. Results on changes in ZNF217, ZNF703, and ZNF750 were analyzed using differential expression. The paired statistical method was applied to matched normal tissue samples paired with tumor-sized counterparts. Normality of paired expression differences was assessed using Shapiro–Wilk's paired-samples test. Differences in expression between tumor and normal tissues were then assessed using paired t-tests. Effect sizes were estimated using Cohen's d_z , and p-values for multiple testing were adjusted using the Benjamini–Hochberg false discovery rate (FDR) procedure. $FDR < 0.05$ was considered statistically significant.

Visualization and Multivariate Analysis

Gene expression distributions were visualized with violin plots, boxplots, paired expression plots, histograms, density plots, and ranked difference plots. Heatmaps with Z-score-transformed expression values were generated to compare samples. PCA was performed on variance-stabilized expression data to assess clustering patterns and overall variance between tumor and normal samples.

Association of Gene Expression with Clinical Characteristics

Tumor expression, normalized, was combined with clinical and pathologic characteristics, including age, BMI, tumor grade, cancer stage, HER2 status, estrogen receptor (ER) status, and progesterone receptor (PR) status. The correlation between the gene expressions and the continuous variables was calculated using Spearman's rank correlation. The comparison of tumor grade and stage type was performed using the Kruskal–Wallis test, and receptor status analysis was conducted using the Wilcoxon rank-sum test. Multiple linear regression models [ZNF217, ZNF703, ZNF750] were generated separately to reveal independent associations between clinical variables and gene expression. Residual diagnostics, variance inflation factors, and Shapiro–

Wilk tests of residual normality were utilized to evaluate model assumptions. All analyses were performed in R version 4.5.0.

Results

Patient Characteristics

A total of 23 women diagnosed with breast cancer were included in the study. The mean age of the participants was 48.7 ± 10.8 years (range, 30–72 years), while the mean body mass index (BMI) was 27.8 ± 4.4 kg/m² (range, 21.1–37.8 kg/m²). Most patients were aged 40–49 years (30.4%), followed by those aged 50–59 years (26.1%), whereas patients younger than 40 years and those aged 60 years or older each accounted for 21.7% of the cohort. The largest proportion of patients were overweight (39.1%), with 34.8% having normal BMI and 26.1% classified as obese.

Regarding tumour characteristics, 65.2% of patients had Grade 3 tumours, while 34.8% had Grade 2 tumours; no Grade 1 tumours were observed. Most patients presented with Stage II disease (47.8%), followed by Stage III (30.4%) and Stage I (21.7%), indicating that the majority were diagnosed with intermediate to advanced disease. Analysis of receptor status showed that HER2 positivity was present in 52.2% of patients, while 52.2% were also ER-positive. PR positivity was more common, occurring in 69.6% of the cohort. Based on combined receptor profiles, 34.8% of patients were classified as triple-positive (ER+/PR+/HER2+), 21.7% were triple-negative, and the remaining 43.5% exhibited other receptor combinations. These demographics and clinicopathological characteristics are summarized in Table 1.

Mutation burden, prevalence, and spectrum

We measured mutation burden across ZNF217, ZNF703, and ZNF750. ZNF217 has the highest mutation burden (~195), followed by ZNF750 (~160), whereas ZNF703 has a much lower burden (~25) (Figure 1A; Table 4, Table S1). Across the genes, single-nucleotide polymorphisms (SNPs) dominated, with fewer insertions and deletions (indels). The coding analysis showed that missense and synonymous mutations were the predominant classes (Figure 1B). In addition, characterization of the nucleotide substitution spectrum indicated clear patterns across the genes (Figure 2). C>T transitions were commonest in ZNF217, with 43 transitions, but

C>A (22), and G>A (18), with A>T and T>G absent. The substitutions seen are more dominant in ZNF750, where A>G (38) and T>C (36) are the majority, but a number of other classes (for example, G>C, T>A) were also not detected. By the way, ZNF703 exhibited far fewer substitutions overall, with no dominant classes; the most frequent substitutions are C>T (6) and T>C (4).

Protein domain distribution and functional enrichment

Mapping mutations at protein sites indicated both domain and non-domain loci where the variants were widespread (Figure 3, Table S2). Mutations in ZNF217 ranged over the entire protein length, including the zinc finger domains, and consisted predominantly of missense and synonymous variants with a single frameshift mutation. ZNF703 mutations were few and far between, rarely grouped together, and were missense or synonymous, with one in-frame deletion in the C-terminal domain. ZNF750 mutations appeared to be distributed throughout the entire protein, with smaller portions localized within and near the zinc finger domain. In addition, consistent with this, the majority of mutations across all genes occurred outside annotated domains, with only a subset overlapping domain regions (Figure 4A, Table S5). ZNF217 had rare domain-related mutations, but none were observed in ZNF703 and ZNF750. Mutation composition analysis confirmed a dominant presence of missense and synonymous variants (Figure 4B), with ZNF217 enriched for missense, ZNF750 for synonymous variants, and ZNF703 demonstrating a more balanced distribution with minor frameshift contribution.

Recurrent mutations and comparison with TCGA

The multiple samples obtained through recurrent mutation analysis exhibited several variants seen in different samples (Figure 5, Table S3). Notably, the highest recurrence recorded was for ZNF750, especially for the synonymous variants known as Ala575= and Pro566=. Val739>Ile in ZNF217 remained the most common non-synonymous recurrent variant. Other prevalent missense variations contained Gln392>Arg, Val478>Ile, and Met235>Val of ZNF750, whereas recurrence in ZNF703 was minimal, with only Pro455= being observed more than once. TCGA comparison revealed concordance and divergence (Table 6, Table S4). ZNF217 was still one of the most common genes altered across the data set, and ZNF703 mutations were also relatively rare. In contrast, although the mutation burden for ZNF750 was high and found in all patients in

this cohort, no mutations were identified for ZNF750 in the TCGA dataset, reflecting a cohort-specific pattern.

Differential Expression of ZNF217, ZNF703, and ZNF750

ZNF217, ZNF703, and ZNF750 were examined for differential expression using normalized RNA-seq data from 23 pairs of tumor and matched normal breast tissue. All three genes were significantly enriched in tumor tissues compared with normal tissues (Table 7). Mean tumor expression was higher than normal in the cases of ZNF217 (10.88 vs 10.23), ZNF703 (11.31 vs 10.39), and ZNF750 (7.46 vs 7.02), and statistically significant positive mean paired differences were observed for each and all of the genes. Tumor expression matched normal in most samples, at 74%, 64%, and 83% for ZNF217, ZNF703, and ZNF750, respectively (Table 7). Paired statistical testing confirmed significant differential expression for ZNF217 ($p = 0.00068$), ZNF703 ($p = 0.00475$), and ZNF750 ($p = 0.00366$), with confidence intervals excluding zero (Table 8). Such differences were highly significant after Benjamini–Hochberg correction (Table 8), suggesting that all three genes are strongly upregulated in tumor samples. Effect size estimates ranged from moderate to large and would affirm the extent to which these differences in expression differ (Table 8). Of the three genes, ZNF703 showed the greatest differential expression, indicating a relatively stronger transcriptional shift in tumor tissues.

Distribution and Variability of Gene Expression Profiles

The gene expression distributions, as visualized, consistently shifted towards higher expression levels in tumor tissues across all 3 genes (Figure 6). Violin plots (Figure 6A) showed wider expression distributions in tumor samples than in matched normal tissues, with the greatest differences in ZNF703 expression. Tumor expression was significantly higher than matched normal tissue for those with ZNF217 ($P = 0.00068$), ZNF703 ($P = 0.00475$), and ZNF750 ($P = 0.00366$). Paired expression plots (Figure 6B) indicated that the majority of patients showed higher tumor expression than matched normal tissue for ZNF217, ZNF703, and ZNF750. ZNF703 exhibited the greatest inter-patient variability: five patients showed lower tumor expression than in matched normal tissue, with one being significantly reduced. On further comparison, reduced tumor expression was reported for four patients for ZNF750 and a minority for ZNF217, with most cases showing elevated tumor expression. Histograms and density plots

(Figure 6C) further supported the observation that tumor expression increased across ranges compared with matched normal tissues, especially for ZNF703, which was more spread out. Pair-matched tumor–normal expression differences analysis (Figure 6D) yielded similar findings: most expression differences were positive across the whole cohort, with ZNF703 showing the largest range of tumor-associated expression changes, while ZNF750 exhibited comparatively smaller but more consistent differences. The Z-score heatmap (Figure 7A) showed coordinated relative overexpression of ZNF217, ZNF703, and ZNF750 mRNA in tumor tissues compared with matched normal tissues. Expression levels were significantly heterogeneous among patients, although the tumor samples showed higher relative expression of all three genes than in matched normal tissues, indicating tumor-related transcriptional dysregulation.

Principal Component Analysis of Gene Expression Profiles

Global expression patterns across samples were assessed with Principal Component Analysis (PCA). The first two principal components accounted for 75.41% and 13.41% of the variance, respectively (Figure 8). Tumor and normal samples were partially separated with PCA, with tumor samples tending to cluster distinctly along the principal component axes. Despite the separation, there was some overlap between normal and tumor samples, reflecting heterogeneity in the expression profile within the cohort. It shows clear distinctions between tumor and normal tissue based on the expression of ZNF217, ZNF703, and ZNF750, but these do not fully separate tumor from normal tissue.

Clinical Associations of ZNF Gene Expression

Correlation analyses, nonparametric group comparisons, and multivariable linear regression were performed to evaluate the associations of tumour expressions of ZNF217, ZNF703, and ZNF750 with specific clinicopathological characteristics. Overall, gene expression was weakly or not significantly associated with age, tumor grade, clinical stage, and receptor status, whereas a positive association with body mass index (BMI) was observed for ZNF703. Scatter plots showing fitted regression lines (Figure 9A) indicated little effect of patient age on the associations of ZNF217, ZNF703, and ZNF750 tumor expression. ZNF217 ($\rho = -0.28$, $P = 0.195$) and ZNF750 ($\rho = -0.20$, $P = 0.359$) indicated weak negative correlations with age, while ZNF703 showed a very weak positive correlation ($\rho = 0.07$, $P = 0.750$). None of the correlations

were statistically significant, although scatter plots of BMI versus tumor expression (Figure 9B) showed a weak positive association across all three genes. ZNF217 had a negligible positive correlation with BMI ($\rho = 0.032$, $P = 0.886$), while ZNF750 showed a weak positive correlation ($\rho = 0.146$, $P = 0.507$). The association between BMI and ZNF703 was strongest ($\rho = 0.391$, $P = 0.065$), indicating a modest, progressive trend towards positive results that approached statistical significance. Across tumor grades, the comparison of gene expression (Figure 9C) showed substantial overlap between Grade 2 and Grade 3 tumors for all three genes. Preliminary Kruskal–Wallis tests exhibited no significant differences in the expression of ZNF217 ($H = 0.0042$, $P = 0.949$), ZNF703 ($H = 0.2667$, $P = 0.606$), or ZNF750 ($H = 1.0667$, $P = 0.302$). Pairwise comparison between Grade 2 and Grade 3 tumors was also nonsignificant, suggesting there was no relationship between tumor grade and differentiation in these genes.

Multivariable Analysis and Integrated Clinical Associations

Two linear regression analyses revealed few independent clinical determinants of gene expression. BMI-associated with an increase of ZNF703 expression ($\beta = 0.194$, $p = 0.025$) was found, and ER-positive status was significantly associated with an increase of ZNF750 expression ($\beta = 1.050$, $p = 0.005$; Table 9, Table S6). No positive predictors have been found for ZNF217. A heatmap of clinical associations summarized these relationships (Figure 11), showing that BMI was positively correlated with ZNF703, and ER status was positively correlated with ZNF750. ZNF217 was poorly correlated with all clinical variables, particularly with PR status ($\rho = -0.36$). These results reveal that, with only a few clinical variables closely correlated with gene expression, the overall effect of clinical variables on ZNF gene expression is weak in this cohort.

Discussion

The study supplies population-based evidence regarding the mutational and transcriptional profiles of ZNF217, ZNF703, and ZNF750 in Kenyan women with breast cancer. We reported that all three genes had somatic changes and substantial tumor-associated upregulation, but differed in mutation burden, predicted functional impact, and clinical association. The importance of these findings is that African groups of breast cancer patients are still

underrepresented in genomic studies, notwithstanding increasing evidence that genomic variation specific to a particular population may affect tumor biology and precision oncology strategies.

The identified mutational differences were apparent at the gene level. ZNF217 showed a high mutation burden, including recurrent missense variants and a high-impact frameshift mutation. This corroborates prior investigations, which portray ZNF217 as an oncogenic modulator in transcriptional dysregulation, tumor progression, endocrine resistance, and poor clinical outcomes [21, 22]. In previous reports on African ancestry groups, it was also suggested that ZNF217-related alterations may contribute to aggressive breast cancer biology [23]. Similarly, the lower mutation burden of ZNF703 was found in only a few patients. This is supported by previous studies demonstrating that ZNF703 mainly contributes to breast cancer through amplification, overexpression, and transcriptional activation rather than frequent coding-sequence mutation [15, 24]. ZNF703 is part of the 8p12 amplicon and has been recognized as an oncogenic driver in luminal B breast cancer related to proliferation and stemness, endocrine resistance, and poor prognosis [16].

ZNF750 exhibited a contrary trend. Even though there were numerous somatic variants in this group, most were synonymous or predicted to be of little functional relevance. This contrasts with earlier literature, which characterizes ZNF750 as a tumor suppressor with roles primarily in epithelial differentiation, repression of invasion-related genes, and inhibition of tumor progression [17]. Previous research has reported that suppression of ZNF750 expression or abnormal function is associated with aggressive disease and poorer outcomes [18]. However, mutation frequency was increased, with no evidence of predominantly disruptive coding mutations in this cohort, further implicating ZNF750 dysregulation in the subgroup and potentially.

The protein-level mapping indicated that mutations were widely distributed across the three genes with limited clustering within the annotated functional domains. This differs from classical hotspot-driven cancer genes but reflects the broader biology of zinc finger proteins, a subset of which can disrupt multiple pathways by modulating transcriptional regulation, DNA binding, protein stability, or regulatory interactions [25, 26]. Comparison with TCGA data suggested a potentially clinically relevant local distinction. Although high mutation rates were observed for

ZNF217 and ZNF750 in this Kenyan cohort, TCGA breast cancer data revealed low mutation frequencies for these two genes. Such differences may indicate variation at the population-level, genomic-level, or variant-screening level, small sample size, or even the inclusion of low-impact variants in the current analysis.

The expression results suggest that transcriptional dysregulation in these genes is also prevalent in the breast cancer population. ZNF217, ZNF703, and ZNF750 were all massively upregulated in tumor tissues compared with matched normal tissues. The upregulation of ZNF217 can be linked to other evidence and is consistent with previous studies that have demonstrated associations among tumor development, metastasis, endocrine resistance, oncogenic signaling pathways, and, in turn, ERBB3-related signaling activation with the expression of factors like ZNF217 [14, 25]. Similarly, the substantial and heterogeneous upregulation of ZNF703 is consistent with its established function as an oncogenic transcription factor, particularly its known role in luminal and endocrine-resistant breast cancer [27, 28]. The upregulation of ZNF750 was thus more unexpected. Previously, ZNF750 was classically referred to as a tumor suppressor that is down-regulated, specifically in breast and other epithelial cancers [29]. A decrease in the expression of ZNF750 is associated with loss of epithelial differentiation and enhanced migration, invasion, and metastasis [17]. Meanwhile, this study detected much greater ZNF750 activity in tumors compared to matched normal tissues.

The independent association of ZNF750 expression with ER-positive status further indicates the possibility of ZNF750 modulating hormone-responsive transcriptional programming in certain tumors. Clinical association analyses showed that most traditional clinicopathological variables, such as age, tumor grade, stage, HER2 status, and PR status, were weakly or not significantly associated with ZNF expression. There was, however, a pair of links between them which stood out. First, ZNF703 expression was positively related to BMI in the multivariable analysis. Conventional studies of ZNF703 have mainly studied amplifications, proliferation, and endocrine resistance [27, 30], where obesity and high BMI have an effect on the biology of breast cancer by inducing inflammation, insulin resistance, adipokine signaling, and estrogen biosynthesis [31]. The effect observed here may therefore imply that metabolic status and ZNF703-mediated transcriptional activity are linked, but that requires confirmation. Second, the expression of ZNF750 was increased in ER-positive tumors. This observation has been

surprisingly neglected in previously reported ZNF750 publications, which have mainly focused on epithelial differentiation and tumor suppressor effects [18, 32]. The present finding of ZNF750 may be associated with hormone receptor biology in breast cancer. Since ER-positive tumors are transcriptionally driven by estrogen receptor signaling, ZNF750 might participate in subtype-specific regulation programs.

There are also limitations associated with this study. A limited sample size, particularly in subgroup analyses by molecular subtype, receptor group, tumor grade, and stage, was associated with poor statistical results. Specifically, the study focused on three candidate genes and thus can be described as inadequate given the overall extent of genomic complexity in breast cancer. The functional consequences of variants were deduced computationally, without experimental validation. Moreover, the lack of an independent validation cohort implies that results provide discovery-level evidence. These results need to be validated in larger, multicentre African cohorts in future studies and to support molecular subtype-stratified analyses.

Future studies should elucidate whether the repeat variants in ZNF217 affect transcriptional activity, whether ZNF703 expression is driven by obesity-dependent signaling pathways, and whether ZNF750 plays context-dependent roles in ER-positive breast cancer. This combination of genomics, transcriptomics, epigenomics, proteomics, and clinical outcomes would provide a more comprehensive understanding of these genes in breast cancer biology.

Conclusion

This study confirmed that ZNF217, ZNF703, and ZNF750 are both mutated and transcriptionally disrupted in Kenyan women with breast cancer. The results affirm the previously established oncogenic significance of ZNF217 and ZNF703, but indicate a more nuanced, context-specific role for ZNF750. These findings add to the body of African breast cancer genomics and support population-based molecular studies to advance equitable precision oncology.

List of abbreviations

BMI Body Mass Index

BQSR Base Quality Score Recalibration

ER Estrogen Receptor

FDR False Discovery Rate

GATK Genome Analysis Toolkit

GEO Gene Expression Omnibus

HER2 Human Epidermal Growth Factor Receptor 2

IGV Integrative Genomics Viewer

LMICs Low- and Middle-Income Countries

NACOSTI National Commission for Science, Technology and Innovation

NCBI National Center for Biotechnology Information

PCA Principal Component Analysis

PoN Panel of Normals

PR Progesterone Receptor

SNP Single Nucleotide Polymorphism

SRA Sequence Read Archive

STAR Spliced Transcripts Alignment to a Reference

TCGA The Cancer Genome Atlas

UniProt Universal Protein Resource

VEP Variant Effect Predictor

WES Whole-Exome Sequencing

ZNF Zinc Finger Protein

Declarations

Ethics approval and consent to participate

This study was conducted in accordance with the Declaration of Helsinki. Ethical approval was obtained from the Mount Kenya University Institutional Ethics Review Committee (IERC Ref: 2018/REC-80 v7) and the National Commission for Science, Technology and Innovation (NACOSTI License No. NACOSTI/P/22/21374). Written informed consent was obtained from all participants prior to enrollment in the study.

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: A Cancer Journal for Clinicians*. 2024;74:229–63. <https://doi.org/10.3322/caac.21834>.
2. Fu M, Peng Z, Wu M, Lv D, Lyu S, Li Y. Assessing the African burden of breast cancer: A demographic analysis using Global Cancer Observatory 2022. *European Journal of Surgical Oncology*. 2025;51:109627. <https://doi.org/10.1016/j.ejso.2025.109627>.
3. Chinedu SN, Ogbodo PN, Effiong ME. Disparities in diagnosis, treatment, and outcomes of early-onset breast cancer: A global perspective. *Critical Reviews in Oncology/Hematology*. 2026;221:105251. <https://doi.org/10.1016/j.critrevonc.2026.105251>.
4. Kim J, Harper A, McCormack V, Sung H, Houssami N, Morgan E, et al. Global patterns and trends in breast cancer incidence and mortality across 185 countries. *Nat Med*. 2025;31:1154–62. <https://doi.org/10.1038/s41591-025-03502-3>.
5. Kenya NCI. Cancer Awareness | National Cancer Institute Kenya. <https://www.ncikenya.go.ke/awareness>. Accessed 6 July 2026.
6. Cancer Today. <https://gco.iarc.who.int/today/>. Accessed 6 July 2026.
7. Mugo N, Contino G. Genomic landscapes of breast cancer in African populations: a systematic review. *npj Breast Cancer*. 2025;11:91. <https://doi.org/10.1038/s41523-025-00809-9>.
8. Ben Rabah C, Sattar A, Ibrahim A, Serag A. A Multimodal Deep Learning Model for the Classification of Breast Cancer Subtypes. *Diagnostics*. 2025;15:995. <https://doi.org/10.3390/diagnostics15080995>.
9. Kim R, Yu J, Lim J, Oh BB-L, Nam SJ, Kim SW, et al. Whole-genome landscapes of 1,364 breast cancers. *Nature*. 2026;649:1282–91. <https://doi.org/10.1038/s41586-025-09812-3>.

10. Sobocińska J, Molenda S, Machnik M, Oleksiewicz U. KRAB-ZFP Transcriptional Regulators Acting as Oncogenes and Tumor Suppressors: An Overview. *International Journal of Molecular Sciences*. 2021;22:2212. <https://doi.org/10.3390/ijms22042212>.
11. Zhao Z, Huang K, Liao Z, Chen B, Chen J, Mei Z. Zinc finger proteins (ZFPs) in health and disease. *Mol Biomed*. 2026;7:19. <https://doi.org/10.1186/s43556-026-00413-8>.
12. Zhou Z, Wu L, Luo J-L, Li Z, Zhong S. Rewiring tumor immunity via zinc finger proteins: a new frontier in cancer immunotherapy. *J Transl Med*. 2025;24:72. <https://doi.org/10.1186/s12967-025-07549-1>.
13. Fahmé P, Ramadan F, Le DT, Thi K-ON, Ghayad SE, Hussein N, et al. The Intricate Interplay between the ZNF217 Oncogene and Epigenetic Processes Shapes Tumor Progression. *Cancers*. 2022;14:6043. <https://doi.org/10.3390/cancers14246043>.
14. Collins C, Rommens JM, Kowbel D, Godfrey T, Tanner M, Hwang S, et al. Positional cloning of ZNF217 and NABC1: Genes amplified at 20q13.2 and overexpressed in breast carcinoma. *Proceedings of the National Academy of Sciences*. 1998;95:8703–8. <https://doi.org/10.1073/pnas.95.15.8703>.
15. Sircoulomb F, Nicolas N, Ferrari A, Finetti P, Bekhouche I, Rousselet E, et al. ZNF703 gene amplification at 8p12 specifies luminal B breast cancer. *EMBO Mol Med*. 2011;3:153–66. <https://doi.org/10.1002/emmm.201100121>.
16. Bazarov AV, Yaswen P. Who is in the driver's seat in 8p12 amplifications? ZNF703 in luminal B breast tumors. *Breast Cancer Res*. 2011;13:308. <https://doi.org/10.1186/bcr2873>.
17. Boxer LD, Barajas B, Tao S, Zhang J, Khavari PA. ZNF750 interacts with KLF4 and RCOR1, KDM1A, and CTBP1/2 chromatin regulators to repress epidermal progenitor genes and induce differentiation genes. *Genes Dev*. 2014;28:2013–26. <https://doi.org/10.1101/gad.246579.114>.
18. Zhou T, Fan J, Wang M, Cheng C, Yuan F, Li J, et al. ZNF750 suppresses esophageal squamous cell carcinoma by maintaining epithelial barrier-driven mucosal immune homeostasis. *Biol Direct*. 2026;21:71. <https://doi.org/10.1186/s13062-026-00787-3>.
19. An G, Feng L, Hou L, Li X, Bai J, He L, et al. A bioinformatics analysis of zinc finger protein family reveals potential oncogenic biomarkers in breast cancer. *Gene*. 2022;828:146471. <https://doi.org/10.1016/j.gene.2022.146471>.
20. Tang W, Zhang F, Byun JS, Dorsey TH, Yfantis HG, Ajao A, et al. Population-specific Mutation Patterns in Breast Tumors from African American, European American, and Kenyan Patients. *Cancer Research Communications*. 2023;3:2244–55. <https://doi.org/10.1158/2767-9764.CRC-23-0165>.
21. Krig SR, Miller JK, Fietze S, Beckett LA, Neve RM, Farnham PJ, et al. ZNF217, a candidate breast cancer oncogene amplified at 20q13, regulates expression of the ErbB3 receptor

tyrosine kinase in breast cancer cells. *Oncogene*. 2010;29:5500–10. <https://doi.org/10.1038/onc.2010.289>.

22. Fahmé P, Ramadan F, Le DT, Thi K-ON, Ghayad SE, Hussein N, et al. The Intricate Interplay between the ZNF217 Oncogene and Epigenetic Processes Shapes Tumor Progression. *Cancers (Basel)*. 2022;14:6043. <https://doi.org/10.3390/cancers14246043>.

23. Ansari-Pour N, Zheng Y, Yoshimatsu TF, Sanni A, Ajani M, Reynier J-B, et al. Whole-genome analysis of Nigerian patients with breast cancer reveals ethnic-driven somatic evolution and distinct genomic subtypes. *Nat Commun*. 2021;12:6946. <https://doi.org/10.1038/s41467-021-27079-w>.

24. Zhang X, Mu X, Huang O, Wang Z, Chen J, Chen D, et al. ZNF703 promotes triple-negative breast cancer cells through cell-cycle signaling and associated with poor prognosis. *BMC Cancer*. 2022;22:226. <https://doi.org/10.1186/s12885-022-09286-w>.

25. Gosztyla ML, Zhan L, Olson S, Wei X, Naritomi J, Nguyen G, et al. Integrated multi-omics analysis of zinc-finger proteins uncovers roles in RNA regulation. *Molecular Cell*. 2024;84:3826-3842.e8. <https://doi.org/10.1016/j.molcel.2024.08.010>.

26. An G, Feng L, Hou L, Li X, Bai J, He L, et al. A bioinformatics analysis of zinc finger protein family reveals potential oncogenic biomarkers in breast cancer. *Gene*. 2022;828:146471. <https://doi.org/10.1016/j.gene.2022.146471>.

27. Klæstad E, Sawicka JE, Engstrøm MJ, Ytterhus B, Valla M, Bofin AM. ZNF703 gene copy number and protein expression in breast cancer; associations with proliferation, prognosis and luminal subtypes. *Breast Cancer Res Treat*. 2021;186:65–77. <https://doi.org/10.1007/s10549-020-06035-0>.

28. Zhang X, Mu X, Huang O, Wang Z, Chen J, Chen D, et al. ZNF703 promotes triple-negative breast cancer cells through cell-cycle signaling and associated with poor prognosis. *BMC Cancer*. 2022;22:226. <https://doi.org/10.1186/s12885-022-09286-w>.

29. Cassandri M, Butera A, Amelio I, Lena AM, Montanaro M, Mauriello A, et al. ZNF750 represses breast cancer invasion via epigenetic control of prometastatic genes. *Oncogene*. 2020;39:4331–43. <https://doi.org/10.1038/s41388-020-1277-5>.

30. Wang G, Wang J, Li C, Mu X, Mu Q, Zhang X, et al. ZNF703 promotes Triple-Negative breast cancer cell progression and in combination with STK11 predicts disease recurrence (ZS - TNBC Model). *Gene*. 2025;942:149258. <https://doi.org/10.1016/j.gene.2025.149258>.

31. Poggi C. Breast Cancer Risk Factors. In: Poggi C, editor. *Breast Imaging Techniques for Radiographers*. Cham: Springer Nature Switzerland; 2024. p. 9–11. https://doi.org/10.1007/978-3-031-63314-0_2.

32. Butera A, Cassandri M, Rugolo F, Agostini M, Melino G. The ZNF750–RAC1 axis as potential prognostic factor for breast cancer. *Cell Death Discov*. 2020;6:135. <https://doi.org/10.1038/s41420-020-00371-2>.

FIGURES

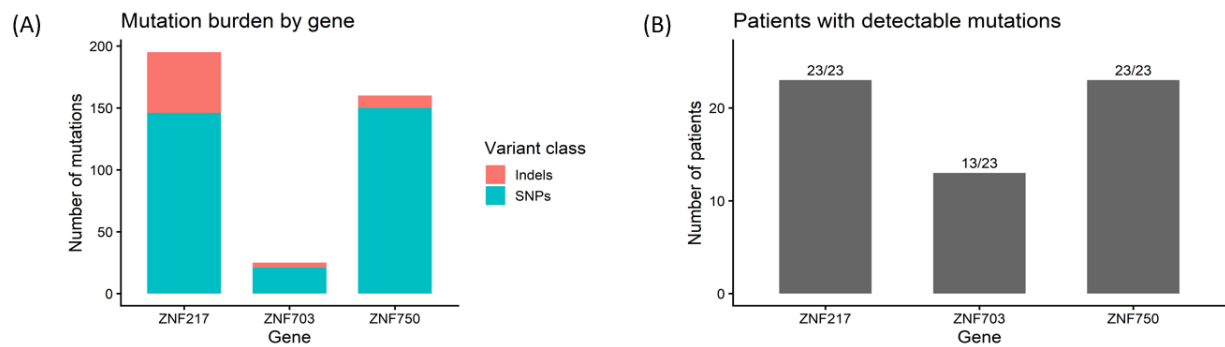


Figure 1: Mutation profile of ZNF217, ZNF703, and ZNF750 genes in 23 patients.

(A) Mutation burden by gene showing the number of single nucleotide polymorphisms (SNPs) and indels, (B) Coding consequence composition illustrating the distribution of synonymous, missense, and frameshift variants, (C) Number of patients with detectable mutations per gene.

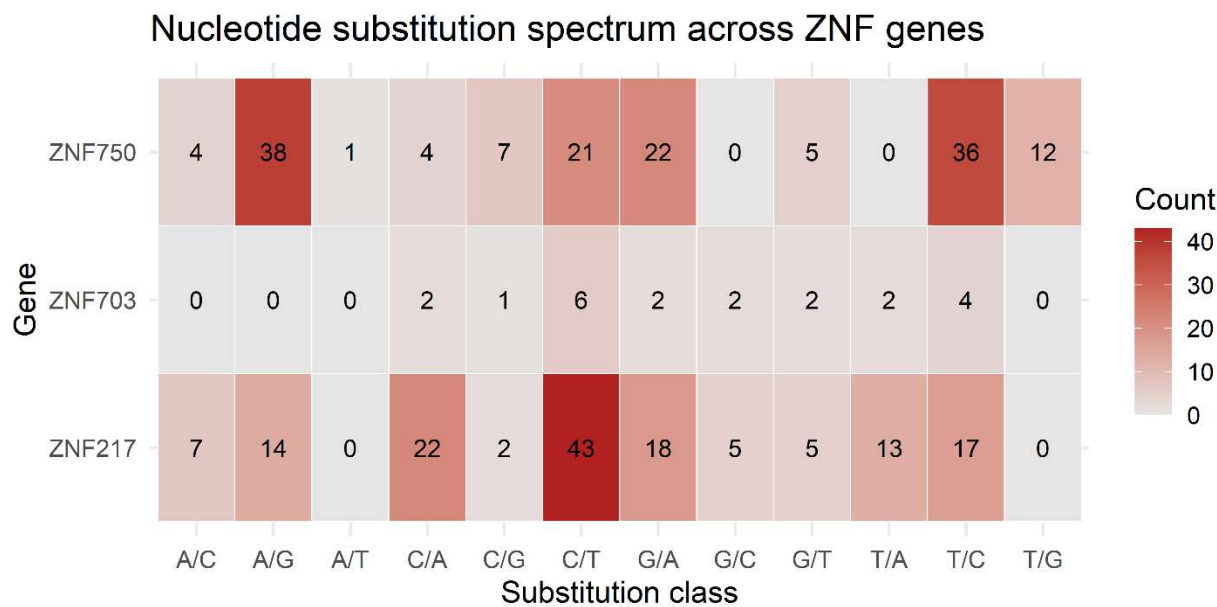


Figure 2: Nucleotide substitution spectrum across ZNF217, ZNF703, and ZNF750 genes in 23 patients. Heatmap showing the distribution and frequency of nucleotide substitution classes

(A/C, A/G, A/T, C/A, C/G, C/T, G/A, G/C, G/T, T/A, T/C, T/G) across the three genes, with color intensity representing mutation counts.

Distribution of coding mutations across protein domains

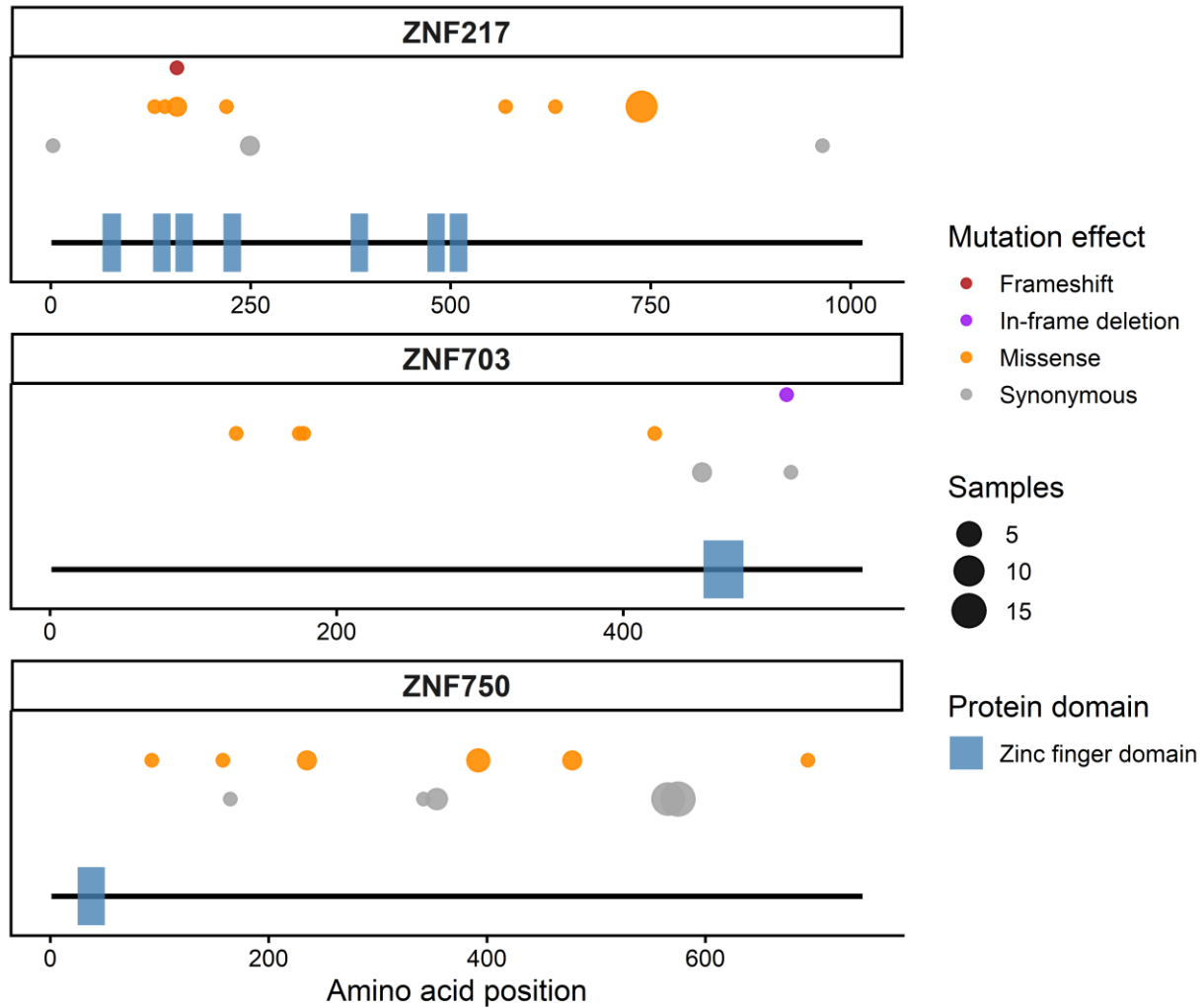


Figure 3: Distribution of coding mutations across protein domains in ZNF217, ZNF703, and ZNF750 genes. Lollipop plots showing the positions of coding mutations along the protein sequences of each gene, with annotated functional domains indicated.

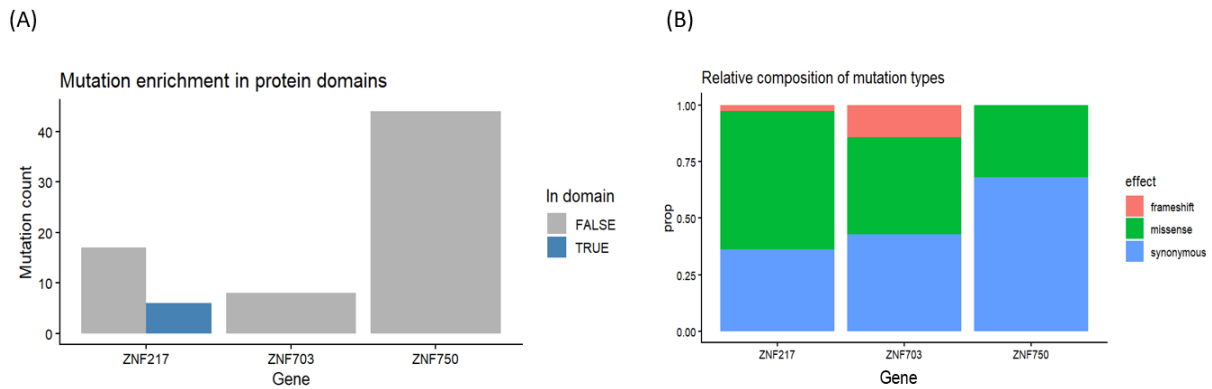


Figure 4: Functional characterization of mutations across ZNF217, ZNF703, and ZNF750 genes.

(A) Mutation enrichment in protein domains showing the number of mutations occurring within and outside annotated functional domains for each gene, (B) Relative composition of mutation types illustrating the proportion of synonymous, missense, and frameshift variants across the three genes.

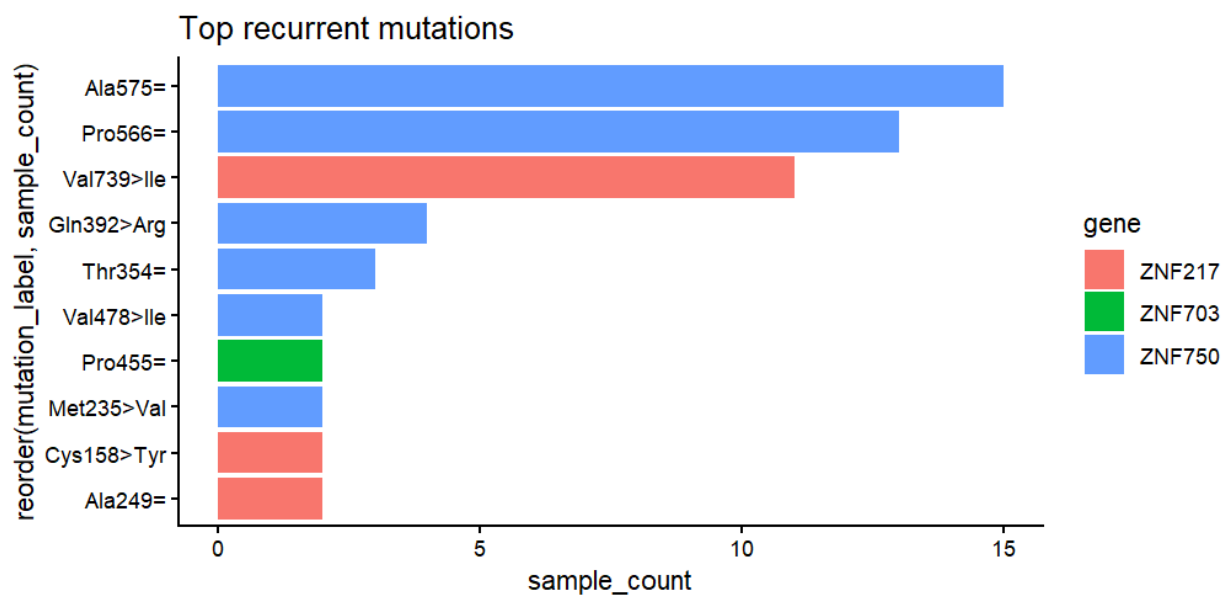


Figure 5: Recurrent mutations across ZNF217, ZNF703, and ZNF750 genes in 23 patients.

Bar plot showing the most frequently observed amino acid substitutions across the three genes,

with bar length representing the number of samples harboring each mutation and colors indicating gene origin.

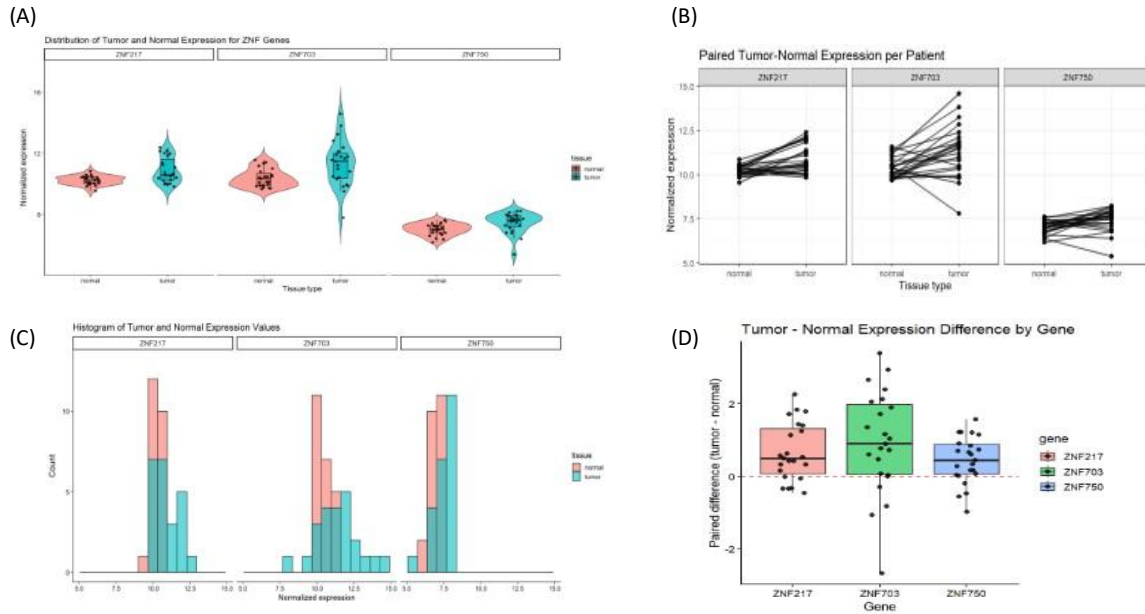
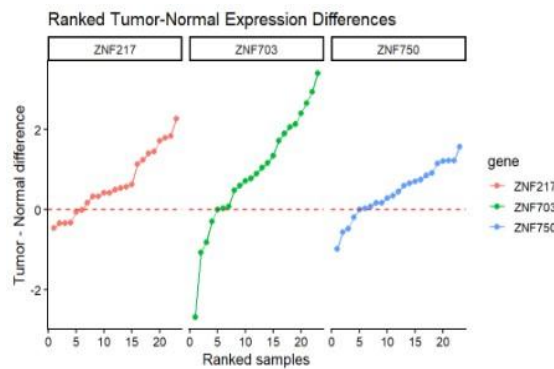


Figure 6. Differential expression patterns of ZNF217, ZNF703, and ZNF750 in paired breast tumor and normal tissues. (A) Violin plots showing the distribution of normalized gene expression levels in tumor and matched normal samples for each gene. (B) Paired tumor-normal expression profiles for individual patients. (C) Histograms depicting the distribution of normalized expression values in tumor and normal tissues. (D) Boxplots of paired expression differences (tumor minus normal) for each gene.

(A)



(B)

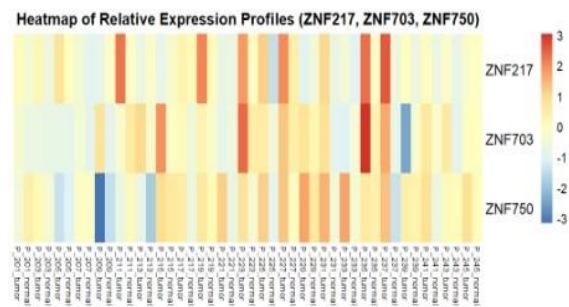


Figure 7. Patient-level expression differences and relative expression patterns of ZNF217, ZNF703, and ZNF750. (A) Ranked tumor–normal expression differences for each gene across individual patients. (B) Heatmap of z-score–normalized expression values for ZNF217, ZNF703, and ZNF750 across all samples, with color scale indicating relative expression levels.

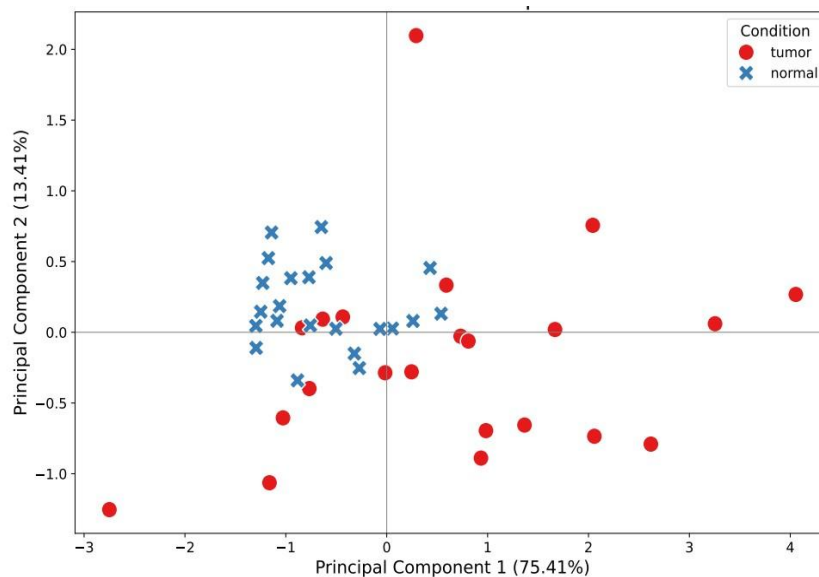


Figure 8. Principal component analysis of gene expression profiles. Scatter plot of the first two principal components based on normalized expression values of ZNF217, ZNF703, and ZNF750, with samples colored by condition (tumor and normal). Axes indicate the proportion of variance explained by each component.

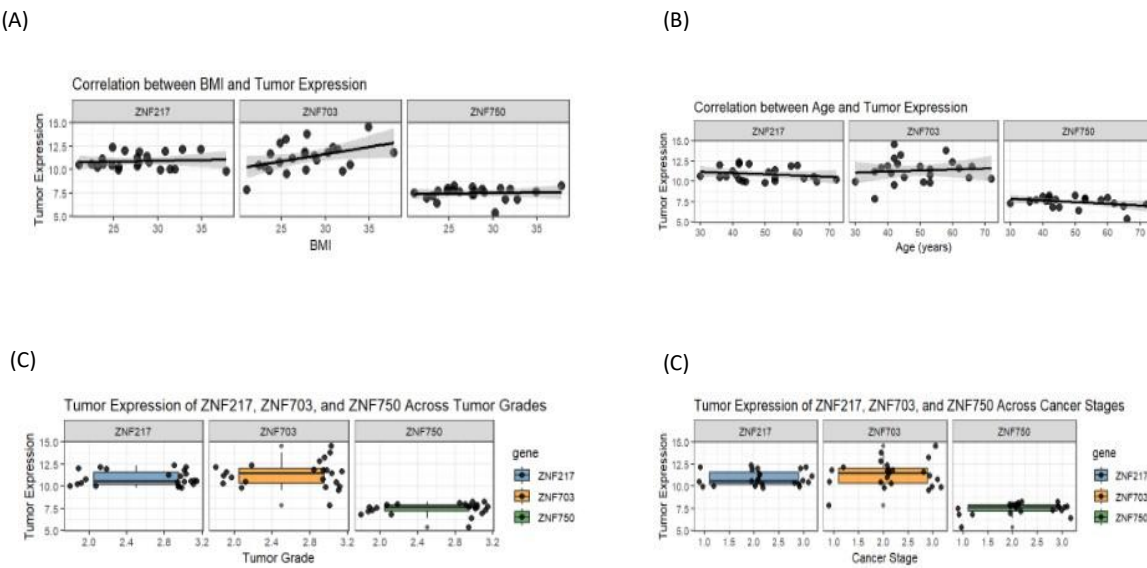
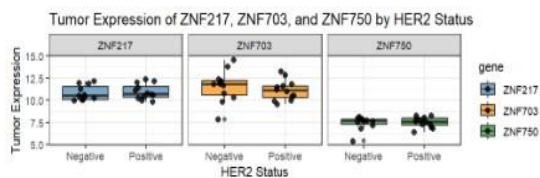
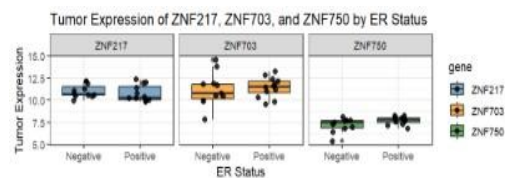


Figure 9. Associations between gene expression and clinical variables. (A) Scatter plots showing the relationship between body mass index (BMI) and tumor expression of ZNF217, ZNF703, and ZNF750. (B) Scatter plots showing the relationship between age and tumor expression for each gene. (C) Boxplots of tumor expression across tumor grades. (D) Boxplots of tumor expression across cancer stages.

(A)



(B)



(C)

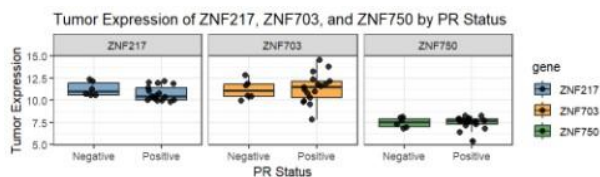


Figure 10. Gene expression across receptor status categories. (A) Boxplots of tumor expression of ZNF217, ZNF703, and ZNF750 stratified by HER2 status. (B) Boxplots of tumor expression stratified by estrogen receptor (ER) status. (C) Boxplots of tumor expression stratified by progesterone receptor (PR) status.

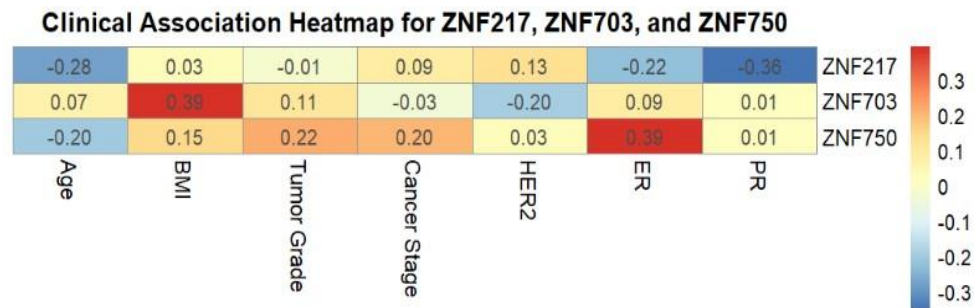


Figure 11. Clinical association heatmap of ZNF217, ZNF703, and ZNF750 expression. Heatmap showing correlation coefficients between tumor expression of ZNF217, ZNF703, and ZNF750 and clinical variables, including age, body mass index (BMI), tumor grade, cancer stage, HER2 status, estrogen receptor (ER) status, and progesterone receptor (PR) status. Color scale represents the strength and direction of correlations.

TABLES

Table 1: Summary Statistics of Continuous Variables

Variable	Mean	Median	Minimum	Maximum	Standard Deviation
Age (years)	48.7	45	30	72	10.8
BMI (kg/m ²)	27.8	27.7	21.1	37.8	4.4

Table 2: Summary Distribution of Clinical and Pathological Characteristics

Variable	Category	Frequency (n)	Percentage (%)
Age Group (years)	< 40	5	21.7
	40–49	7	30.4
	50–59	6	26.1
	≥ 60	5	21.7
BMI Category	Normal (18.5–24.9)	8	34.8
	Overweight (25–29.9)	9	39.1
	Obese (≥30)	6	26.1
Tumor Grade	Grade 2	8	34.8
	Grade 3	15	65.2
Cancer Stage	Stage I	5	21.7
	Stage II	11	47.8
	Stage III	7	30.4
HER2 Status	Positive	12	52.2
	Negative	11	47.8
PR Status	Positive	16	69.6
	Negative	7	30.4
ER Status	Positive	12	52.2
	Negative	11	47.8

Table 3: Combined Receptor Subtypes

Subtype	Frequency (n)	Percentage (%)
Triple Positive (ER+, PR+, HER2+)	8	34.8
Triple Negative (ER–, PR–, HER2–)	5	21.7

Other Combinations	10	43.5
--------------------	----	------

Table 4: Gene Level Mutation Summary

Gene	Total Mutations	Total SNPS	Total Indels	Mutated Patients	Total Patients	Synonymous	Missense	Frameshift	Recurrent Coding Variants	Dominant Substitution
ZNF217	195	146	49	23	23	13	22	1	3	C/T
ZNF703	25	21	4	13	23	3	3	1	1	C/T
ZNF750	160	150	10	23	23	36	17	0	6	A/G

Table 5: Recurrent Coding Variants

Gene	Mutation Label	Effect	Ammino Acid Position	Sample Count	Domain Overlap	Overlapping Domain
ZNF217	Val739>Ile	Missense	739	11	FALSE	
ZNF217	Cys158>Tyr	Missense	158	2	TRUE	C2H2-type 3
ZNF217	Ala249=	Synonymous	249	2	FALSE	
ZNF703	Pro455=	Synonymous	455	2	FALSE	
ZNF750	Ala575=	Synonymous	575	15	FALSE	
ZNF750	Pro566=	Synonymous	566	13	FALSE	
ZNF750	Gln392>Arg	Missense	392	4	FALSE	
ZNF750	Thr354=	Synonymous	354	3	FALSE	
ZNF750	Met235>Val	Missense	235	2	FALSE	
ZNF750	Val478>Ile	Missense	478	2	FALSE	

Table 6: Combined Mutation Comparison

Gene	Mutated Patients	Total Patients	Mutation Frequency	Dataset	Mutation Type	Count
ZNF217	15	23	0.652174	This study	Missense	17
ZNF217	15	23	0.652174	This study	Synonymous	4
ZNF750	16	23	0.695652	This study	Missense	11
ZNF750	16	23	0.695652	This study	Synonymous	31
ZNF217	7	994	0.007042	TCGA	Missense	6
ZNF217	7	994	0.007042	TCGA	Nonsense	1
ZNF703	1	994	0.001006	TCGA	Missense	2

Table 7: ZNF Expression Descriptive Statistics

Gene	n	Mean Tumor	SD Tumor	Mean Normal	SD Normal	Mean Difference	SD Diff	Median Difference	Tumor Higher (%)
ZNF217	23	10.8847	0.83289	10.2301	0.2969	0.6546	0.7948	0.4888	73.9
ZNF703	23	11.3147	1.5103	10.3862	0.5878	0.9285	1.4177	0.8957	64.3
ZNF750	23	7.4563	0.6622	7.0198	0.3909	0.4365	0.6440	0.4391	82.6

Table 8: ZNF Expression Descriptive Statistics

Gene	p-value	95% CI (Lower)	95% CI (Upper)	Mean Difference	SD Difference	Cohen's d	Benjamin–Hochberg	
							Raw p-value	Adjusted p-value
ZNF217	0.0006	0.3109	0.9984	0.6547	0.7949	0.8236	0.0009	0.0029
ZNF703	0.0047	0.3154	1.5416	0.9285	1.4178	0.6549	0.0043	0.0043
ZNF750	0.0036	0.1581	0.7151	0.4366	0.6441	0.6778	0.00345	0.0043

Table 9: Multiple Regression Summary

Gene	Variable	Estimate (β)	Std. Error	t-Statistic	p-value
ZNF217	Intercept	11.747	2.496	4.706	0.0003
	Age	-0.009	0.019	-0.478	0.640
	BMI	0.027	0.050	0.539	0.598
	Tumor Grade	-0.403	0.505	-0.798	0.437
	Cancer Stage	0.174	0.321	0.543	0.595
	HER2 Positive	0.072	0.481	0.149	0.883
	ER Positive	-0.357	0.551	-0.647	0.528
	PR Positive	-0.442	0.562	-0.786	0.444
ZNF703	Intercept	1.394	3.854	0.362	0.723
	Age	0.025	0.030	0.827	0.421
	BMI	0.194	0.078	2.497	0.025
	Tumor Grade	0.969	0.780	1.243	0.233
	Cancer Stage	0.702	0.495	1.417	0.177
	HER2 Positive	-1.084	0.742	-1.461	0.165
	ER Positive	0.943	0.851	1.107	0.286
	PR Positive	-0.960	0.867	-1.107	0.286
ZNF750	Intercept	5.979	1.441	4.148	0.0009
	Age	-0.018	0.011	-1.565	0.139
	BMI	0.017	0.029	0.577	0.573
	Tumor Grade	0.466	0.292	1.599	0.131
	Cancer Stage	0.345	0.185	1.864	0.082
	HER2 Positive	-0.566	0.277	-2.038	0.060
	ER Positive	1.050	0.318	3.298	0.005

Gene	Variable	Estimate (β)	Std. Error	t-Statistic	p-value
	PR Positive	-0.506	0.324	-1.560	0.140

SUPPLEMENTARY MATERIALS

Table S1: ZNF mutation summary

Table S2: ZNF mutation positions

Table S3: Amino acid change summary

Table S4: TCGA mutation dataset

Table S5: ZNF protein domains

Table S6: Combined ZNF expression and pathological and clinical data

Supplementary Files

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

- [TableS1ZNFmutationssummary.xlsx](#)
- [TableS2ZNFmutationpositions.xlsx](#)
- [TableS3Aminoacidchangessummary.xlsx](#)
- [TableS5Znfproteindomains.xlsx](#)
- [TableS4TCGAmutationdataset.xlsx](#)
- [TableS6CombinedZNFExpressionandPathologicalClinicalData.xlsx](#)