Archives in Oncology Research
ISSN: 3068-5273
ScitechJournals
Open Access

Exploratory Identification of Ethnicity-Associated BRCA1 Variation in Triple-Negative Breast Cancer Using Bioinformatic Analysis

Abstract

Breast cancer is the most commonly diagnosed cancer among women worldwide and remains a global health concern. Triple-negative breast cancer (TNBC) is one of the most aggressive subtypes and lacks effective targeted therapies; therefore, early detection and individualized treatment strategies are critically important. Documented differences in breast cancer incidence, aggressiveness, and mortality across ethnic groups highlight the need to investigate genetic variation among diverse populations. This study examined ethnicity-associated BRCA1 variation by comparing African American and White female patients with ductal TNBC using bioinformatic analysis of sequencing data visualized in the Integrative Genomics Viewer (IGV). Preliminary analysis revealed distinct allele frequency differences at several BRCA1 loci between the two groups, suggesting potential ethnicity-associated genetic patterns. These findings contribute to efforts aimed at improving precision oncology approaches and addressing disparities in breast cancer outcomes. Larger cohort studies are warranted to validate the observed patterns and assess their clinical relevance.

Research Article

Anvita Dogiparthy[1]*, Sreekari Samudrala[1]*, Navdeep Gill[2]

[1]Department of Biology, Halmos College of Arts and Sciences, Nova Southeastern University, Fort Lauderdale, FL, USA

[2]Department of Biology, Halmos College of Arts and Sciences, Nova Southeastern University, Fort Lauderdale, FL, USA

***Correspondence:** *Anvita Dogiparthy and Sreekari Samudrala, Department of Biology, Halmos College of Arts and Sciences, Nova Southeastern University, Fort Lauderdale, FL, USA. . Email:* asdogiparthy@gmail.com sreekarisamudrala8@gmail.com.
Sreekari Samudrala (ORCID):0009-0004-3132-2077.
Anvita Dogiparthy(ORCID): 0009-0003-4220-4677

Received: 14 March, 2026; **Accepted:** 27 March, 2026; **Published:** 06 April, 2026

Introduction

Triple-negative breast cancer (TNBC) accounts for approximately 15–20% of breast cancers and is characterized by the loss of estrogen receptor (ER), progesterone receptor (PR), and HER2 expression. Because TNBC lacks hormonal and HER2 targets, treatment options are limited, and the disease is associated with poor prognosis, high recurrence rates, and reduced therapeutic responsiveness [1].

TNBC disproportionately affects African American women, who are nearly twice as likely to develop this subtype compared to White/Caucasian individuals [2]. This disparity contributes to approximately 40% higher mortality in African American women, despite similar overall breast cancer incidence rates [3]. Understanding the molecular mechanisms underlying these disparities is essential for developing equitable and effective treatment strategies.

Breast Cancer Gene 1 (BRCA1) is a tumor suppressor gene responsible for DNA repair and is implicated in both germline and sporadic TNBC. Approximately 70–80% of breast cancers in individuals with a germline BRCA1 mutation are triple-negative [4]. Even in sporadic TNBC cases, BRCA1 dysfunction, such as loss of heterozygosity or promoter methylation is frequently observed [5]. BRCA1-deficient tumors

demonstrate increased sensitivity to DNA-damaging agents such as platinum compounds and PARP inhibitors, making BRCA1 an important biomarker for therapeutic stratification [6].

Despite these associations, African American women remain underrepresented in genomic studies. Contributing factors include reduced access to genetic counseling, economic barriers, and disparities in insurance coverage. Although the Affordable Care Act expanded coverage for certain preventive services, gaps in the accessibility and affordability of genetic testing persist. These barriers contribute to lower rates of BRCA1/2 testing among African American women, potentially delaying identification of hereditary cancer risk and implementation of personalized therapies. The present study compares African American and White women with ductal TNBC to identify potential ethnicity-associated BRCA1 sequence patterns using publicly available sequencing data and bioinformatic visualization. These patterns could potentially inform targeted interventions and alleviate healthcare disparities [7].

Materials and Methods

Sample Selection

Six TNBC patient samples were analyzed, including three self-identified White patients and three self-identified African American patients. All individuals were diagnosed with invasive ductal carcinoma and were selected from similar age ranges to minimize age-related somatic variation.

Data Acquisition

Genomic sequencing datasets were obtained from the NCBI Sequence Read Archive (SRA). Raw FASTQ reads were aligned to the human reference genome (GRCh38/hg38) using BWA (Burrows–Wheeler Aligner, v0.7.17; https://github.com/lh3/bwa). Variant calling was performed using bcftools (v1.17; https://samtools.github.io/bcftools/). The resulting alignment files (BAM format) were visually inspected using the Integrative Genomics Viewer (IGV, v2.16; https://igv.org/doc/desktop/).

Variant Assessment

Analysis focused on the BRCA1 locus on chromosome 17. Visual inspection of read alignments was performed to identify nucleotide variations. For each position of interest, the proportion of reads containing adenine (A), cytosine (C), guanine (G), or thymine (T) was quantified. Loci showing apparent differences in allele distribution between ethnic groups were flagged for comparison.

Statistical Analysis

Allele counts at each BRCA1 locus were extracted from aligned sequencing reads and converted to allele frequencies for each sample. Due to the small sample size and low read depth at certain positions, the analysis was primarily descriptive. No multiple-testing correction was applied, as the study was designed as a hypothesis-generating pilot analysis.

Results

Across the six TNBC samples, multiple loci within the BRCA1 region showed differences in allele frequencies between African American (AA) and White (Wh) patients. Key observations include:

- Several positions (e.g., chr17:43,093,289; chr17:43,121,623) demonstrated shifts in dominant allele frequency between groups.
- Certain loci exhibited complete allele differences between ethnic groups (e.g., [Table 3], where African American samples showed C while White samples showed A).
- Additional sites showed mixed allele distributions with notable proportional differences.

Representative comparisons are shown in [Tables 1–10]. Overall, the data suggest the presence of localized sequence variability within BRCA1 among the examined samples.

Nucleo-tide	White Count	White %	Strand Bias (White)	AA Count	AA %	Strand Bias (AA)
A	0	0%	0+, 0−	0	0%	0+, 0−
C	10	33%	10+, 0−	1	6%	1+, 0−
G	2	7%	2+, 0−	4	24%	4+, 0−
T	18	60%	18+, 0−	12	71%	8+, 4−
N	0	0%	—	0	0%	—

Table 1: Allele counts, percentages, and strand bias at BRCA1 position chr17:43,093,289 for African American sample AA3 (total reads = 17) and White sample W2 (total reads = 30).

	African American (AA)	White (Wh)
chr17:43,055,582	Total: 4	Total: 6
A	0	0
C	4 (100%, 2+, 2−)	4 (67%, 4+, 0−)
G	0	0
T	0	2 (33%, 2+, 0−)
chr17:43,055,568	Total: 4	Total: 8
A	0	0
C	4 (100%, 2+, 2−)	2 (25%, 2+, 0−)
G	0	0
T	0	6 (75%, 6+, 0−)
chr17:43,055,550	Total: 2	Total: 6
A	0	0
C	2 (100%, 2+, 0−)	4 (67%, 4+, 0−)
G	0	0
T	0	2 (33%, 2+, 0−)
chr17:43,055,540	Total: 2	Total: 6
A	0	0
C	2 (100%, 2+, 0−)	4 (67%, 4+, 0−)
G	0	0
T	0	2 (33%, 2+, 0−)

Table 2: Comparative allele distribution across four BRCA1 loci (chr17:43,055,582; chr17:43,055,568; chr17:43,055,550; chr17:43,055,540) between African American sample AA3 and White sample W2.

	African American (AA)	White (Wh)
chr17:43,051,679	Total: 5	Total: 2
A	0	2 (100%, 2+, 0−)
C	5 (100%, 2+, 3−)	0
G	0	0
T	0	0

Table 3: Allele composition at BRCA1 locus chr17:43,051,679 comparing African American sample AA3 and White sample W2.

	African American (AA)	White (Wh)
chr17:43,090,263	Total: 4	Total: 2
A	0	0
C	2 (50%, 1+, 1−)	2 (100%, 1+, 1−)
G	0	0
T	2 (50%, 1+, 1−)	0
chr17:43,090,262	Total: 4	Total: 2
A	0	0
C	2 (50%, 1+, 1−)	2 (100%, 1+, 1−)
G	0	0
T	2 (50%, 1+, 1−)	0
chr17:43,090,257	Total: 4	Total: 2
A	0	0
C	2 (50%, 1+, 1−)	2 (100%, 1+, 1−)
G	0	0
T	2 (50%, 1+, 1−)	0
chr17:43,090,228	Total: 4	Total: 2
A	0	0
C	2 (50%, 1+, 1−)	2 (100%, 1+, 1−)
G	0	0
T	2 (50%, 1+, 1−)	0

Table 4: Allele distribution across four BRCA1 positions (chr17:43,090,263; chr17:43,090,262; chr17:43,090,257; chr17:43,090,228) for African American sample AA1 and White sample W1.

	African American (AA)	White (Wh)
chr17:43,116,621	Total: 3	Total: 1
A	1 (33%, 1+, 0−)	0
C	0	1 (100%, 1+, 0−)
G	0	0
T	2 (67%, 1+, 1−)	0

Table 5: Allele counts and strand bias at BRCA1 position chr17:43,116,621 comparing African American sample AA2 (total reads = 3) and White sample W1 (total reads = 1).

	African American (AA)	White (Wh)
chr17:43,116,693	Total: 6	Total: 16
A	0	0
C	4 (67%, 3+, 1−)	6 (38%, 6+, 0−)
G	0	0
T	2 (33%, 2+, 0−)	10 (63%, 10+, 0−)

Table 6: Comparison of nucleotide frequencies at BRCA1 locus chr17:43,116,693 between African American sample AA2 (total reads = 6) and White sample W2 (total reads = 16).

Nucleotide	AA2	AA3	Wh1	Wh2
Total reads	22	5	4	8
A	0	0	0	0
C	5 (23%) 3+, 2−	2 (40%) 1+, 1−	3 (75%) 2+, 1−	6 (75%) 6+, 0−
G	0	0	0	0
T	17 (77%) 12+, 5−	3 (60%) 2+, 1−	1 (25%) 1+, 0−	2 (25%) 2+, 0−
N	0	0	0	0

Table 7: Multi-sample allele distribution at BRCA1 position chr17:43,121,623 across African American (AA2, AA3) and White (W1, W2) TNBC samples.

	African American (AA)	White (Wh)
chr17:43,123,005	Total: 3	Total: 6
A	0	0
C	2 (67%, 2+, 0−)	2 (33%, 2+, 0−)
G	0	0
T	1 (33%, 1+, 0−)	4 (67%, 4+, 0−)

Table 8: Allele counts and percentages at BRCA1 locus chr17:43,123,005 comparing African American sample AA3 (total reads = 3) and White sample W2 (total reads = 6).

Nucleotide	AA2	AA3	Wh2
Total reads	3	4	4
A	0	0	0
C	1 (33%) 0+, 1−	3 (75%) 1+, 2−	0
G	0	0	2 (50%) 2+, 0−
T	2 (67%) 0+, 2−	1 (25%) 0+, 1−	2 (50%) 2+, 0−
N	0	0	0

Table 9: Allele composition at BRCA1 position chr17:43,123,136 across African American (AA2, AA3) and White (Wh2) samples.

	African American (AA)	White (Wh)
chr17:43,055,646	Total: 4	Total: 6
A	0	0
C	4 (100%, 2+, 2−)	2 (33%, 2+, 0−)
G	0	0
T	0	4 (67%, 4+, 0−)
chr17:43,055,677	Total: 2	Total: 4
A	0	2 (50%, 2+, 0−)
C	2 (100%, 0+, 2−)	0
G	0	0
T	0	2 (50%, 2+, 0−)

Table 10: Comparative allele distribution at BRCA1 loci chr17:43,055,646 and chr17:43,055,677 between African American sample AA3 and White sample W2.

Discussion

Breast cancer treatment planning depends heavily on tumor subtype, hormone receptor status, and molecular features [8]. TNBC represents a clinically challenging subtype due to its aggressive behavior and lack of targeted therapies. Because BRCA1 plays a central role in TNBC biology, understanding population-specific variation may have implications for precision oncology.

In this exploratory analysis, visual comparison using IGV identified several loci within BRCA1 that differed in allele distribution between African American and White TNBC samples. For example, at chr17:43,121,623 [Table 7], African American samples showed a higher proportion of the T allele (AA2: 77%; AA3: 60%), whereas White samples showed enrichment of the C allele (Wh1: 75%; Wh2: 75%).

Additionally, some loci demonstrated qualitative allele differences rather than simple frequency shifts. In [Table 3], African American sample AA3 exclusively expressed the C allele (100%), while White sample W2 showed only the A allele (100%) at chr17:43,051,679. Similar qualitative differences were observed at positions analyzed in [Table 5,9].

These findings highlight the potential value of multi-omics and population-aware genomic analyses in understanding breast cancer disparities. However, the present study has important limitations, including a very small sample size, lack of formal statistical testing, reliance on visual inspection, absence of functional validation, and potential variability in sequencing depth across loci. Because of these constraints, the observed differences should be considered hypothesis-generating rather than definitive biomarkers.

Limitations

This study has several important limitations. First, the sample size was small (n = 6), which substantially limits statistical power and generalizability. Second, read depth varied across loci, and low coverage at some positions may affect the reliability of allele frequency estimates and could introduce observer bias. Third, the study focused on a limited genomic region within BRCA1 and did not evaluate genome-wide population structure or ancestry-informative markers, which could confound ethnicity-associated interpretations. The findings should therefore be considered exploratory and hypothesis-generating, and the results should be interpreted cautiously and validated in larger cohorts.

Future Directions

This is a preliminary analysis, and future work should include a larger and more diverse sample size with thorough statistical association testing, integration with clinical outcomes, and functional validation of candidate variants. Genome-wide analyses incorporating ancestry-informative markers would also strengthen ethnicity-associated interpretations.

Such efforts will be necessary to determine whether ethnicity-associated BRCA1 variation contributes meaningfully to TNBC disparities.

Conclusion

This exploratory bioinformatic analysis identified several BRCA1 loci exhibiting allele distribution differences between African American and White women with triple-negative breast cancer. While the findings suggest possible ethnicity-associated genetic patterns, the limited sample size and descriptive methodology preclude definitive conclusions. The study underscores the importance of increasing representation of diverse populations in genomic cancer research. Larger, statistically powered investigations integrating multi-omics and clinical data are required to validate potential biomarkers and advance equitable precision oncology for TNBC.

Authors Contribution

Anvita Dogiparthy and Sreekari Samudrala contributed equally to this work and share first authorship.

Acknowledgements

This work was conducted as part of the ExEL Program at Nova Southeastern University (NSU) as an undergraduate class research project.

References

1. Bianchini, Giampaolo, Justin M. Balko, Ingrid A. Mayer, Melinda E. Sanders, and Luca Gianni. “Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease.” Nature reviews Clinical oncology 13, no. 11 (2016): 674-690.
2. Carey, Lisa A., Charles M. Perou, Chad A. Livasy, Lynn G. Dressler, David Cowan, Kathleen Conway, Gamze Karaca et al. “Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study.” Jama 295, no. 21 (2006): 2492-2502.
3. DeSantis, Carol E., Jiemin Ma, Ann Goding Sauer, Lisa A. Newman, and Ahmedin Jemal. “Breast cancer statistics, 2017, racial disparity in mortality by state.” CA: a cancer journal for clinicians 67, no. 6 (2017): 439-448.
4. Atchley, Deann P., Constance T. Albarracin, Adriana Lopez, Vicente Valero, Christopher I. Amos, Ana Maria Gonzalez-Angulo, Gabriel N. Hortobagyi, and Banu K. Arun. “Clinical and pathologic characteristics of patients with BRCA-positive and BRCA-negative breast cancer.” Journal of Clinical Oncology 26, no. 26 (2008): 4282-4288.
5. Turner, Nicholas, Andrew Tutt, and Alan Ashworth. “Hallmarks of’BRCAness’ in sporadic cancers.” Nature reviews cancer 4, no. 10 (2004): 814-819.
6. Lord, Christopher J., and Alan Ashworth. “BRCAness revisited.” Nature Reviews Cancer 16, no. 2 (2016): 110-120.
7. Stern, Alexandra Minna. Telling genes: the story of genetic counseling in America. JHU Press, 2012.
8. Todd, Julia. “Breast Cancer Awareness: How Risk Factors and Screening Methods Impact Detection.” Master’s thesis, Arizona State University, 2023.

Citation: Anvita Dogiparthy, Sreekari Samudrala, Navdeep Gill. “Exploratory Identification of Ethnicity-Associated BRCA1 Variation in Triple-Negative Breast Cancer Using Bioinformatic Analysis.” Arch Oncol Res (2026):101. DOI: 10.59462/3068- 5273.1.1.101