Comparative Study of COL10A1 Gene Expression in Normal and Malignant Breast Tissues: Potential Diagnostic Implications

*Roy NR,[1] Jhuma KA,[2] Mallick A,[3] Barman S,[4] Sardar MA[5]

Abstract

Background: Collagen Type X Alpha 1 Chain (COL10A1) has emerged as a potential oncogenic marker in several malignancies, including breast cancer. However, its clinical relevance and association with reproductive and demographic parameters in breast cancer patients remain underexplored in Bangladeshi populations.

Methods: This cross-sectional study was conducted at the Department of Biochemistry, Dhaka Medical College, Dhaka, in collaboration with the Institute for Population and Precision Health, Department of Public Health Sciences, The University of Chicago, USA, from July 2022 to June 2023. A total of 34 histopathologically confirmed female breast cancer patients were enrolled. Paired tumor and adjacent normal breast tissue samples were collected during surgical procedures. Total RNA was extracted, followed by complementary DNA synthesis, and COL10A1 gene expression was quantified using real-time polymerase chain reaction (RT-PCR). Associations with sociodemographic (age, BMI, family history) and reproductive factors (age at menarche, contraceptive use, parity, breastfeeding) were evaluated using Spearman's correlation. Logistic regression was used to assess predictors of high COL10A1 expression in tumor tissue.

Results: The median COL10A1 expression in tumor tissue was significantly higher (3.133, IQR: -0.735 to 2.695) compared to normal tissue (2.044, IQR: -2.651 to -0.615) with **a** p-value of 0.011 (Mann–Whitney U test). High expression (>2.5) was observed in 55.88% (19/34) of tumor tissues compared to **17.65% (6/34)** in normal tissue. Expression levels increased with advancing tumor stage: Stage I (mean ± SD: 2.12 ± 0.43), Stage II (2.85 ± 0.68), and Stage III (3.42 ± 1.01), with a significant trend **(**$p = 0.034$). No significant correlation was found between COL10A1 expression and age ($r = -0.179$, $p = 0.143$), BMI ($r = -0.054$, $p = 0.660$), or reproductive factors such as age at menarche, parity, contraceptive use, or breastfeeding. Logistic regression revealed that patients who had not breastfed had higher odds (OR = 3.72) of elevated COL10A1 expression, although not statistically significant **(**$p = 0.112$**).**

Conclusion: COL10A1 expression is significantly upregulated in breast tumor tissue and shows a positive association with tumor stage, suggesting its potential role as a biomarker for breast cancer progression. While reproductive and demographic factors showed no significant correlation, further large-scale, multi-center studies are warranted to validate its prognostic utility.

[Dinajpur Medical College Journal, 2026 Jan; 19 (1):12-19]
[Former M Abdur Rahim Medical College Journal]
DOI: https://www.doi.org/10.69861/djmcj.2026.19.1.3

Keywords: COL10A1, breast cancer, gene expression, tumor biomarker, real-time PCR, reproductive history, tumor staging.

1. **Dr. Nilima Rani Roy, M. Phil. (Biochemistry), Assistant Professor, Department of Biochemistry, Prime Medical College, Rangpur. drnilima123@gmail.com*
2. *Dr. Khadija Akther Jhuma, M. Phil. MS, Ph D, Professor and Head, Department of Biochemistry, Dhaka Medical College, Dhaka. khadijajhuma2017@gmail.com.*
3. *Dr. Anirban Mallick, MD Resident, Chittagong Medical College anirbanmallick527@gmail.com*
4. *Dr. Sanchita Barman, M. Phil. (Biochemistry), Senior Lecturer, Department of Biochemistry, Fazlur Rahman Medical College, barman.sanchita17@gmail.com*
5. *Dr. Md. Arif Sardar, M. Phil. (Biochemistry), Assistant Professor, Department of Biochemistry, Gazi Medical College, Khulna. arifulislam151@gmail.com*

**For correspondence*

Introduction

Breast cancer is the most common malignancy among women globally, accounting for significant morbidity and mortality, particularly in low- and middle-income countries.[1] According to GLOBOCAN 2020, breast cancer represents 11.7% of all cancer cases, with approximately 2.3 million new diagnoses and 685,000 deaths worldwide.[2] The disease is characterized by molecular and histopathological heterogeneity, involving complex interactions of genetic, hormonal, and environmental factors.[3] Early detection and accurate diagnosis are vital for improving patient outcomes, underscoring the importance of identifying reliable molecular biomarkers for early-stage detection and prognostication.

Gene expression profiling has emerged as a valuable tool in cancer research, aiding in the classification, diagnosis, and treatment stratification of tumors.[4] Among several candidate genes, Collagen Type X Alpha 1 (COL10A1) has gained attention due to its overexpression in various solid tumors, including breast cancer.[5] COL10A1 is a short-chain collagen primarily expressed in hypertrophic chondrocytes during endochondral ossification, but recent studies have identified its aberrant expression in tumor stroma and cancer-associated fibroblasts.[6] Overexpression of COL10A1 has been reported in breast, gastric, colon, and esophageal cancers, suggesting its potential role in tumorigenesis, cell proliferation, invasion, and metastasis.[7,8]

In breast cancer, elevated COL10A1 expression has been correlated with poor clinical outcomes, including reduced overall survival, relapse-free survival, and distant metastasis-free survival.[9] Its co-expression with leucine-rich repeat-containing protein 15 (LRRC15), particularly in tumor stroma, indicates a stromal-specific regulatory pathway that might serve as a prognostic indicator.[10] Moreover, circulating levels of COL10A1 protein have been proposed as a non-invasive biomarker for early breast cancer detection.[11]

Despite these promising insights, data on COL10A1 gene expression in breast cancer patients from South Asia, including Bangladesh, are lacking. In Bangladesh, the burden of breast cancer is rising steadily, with many patients presenting at advanced stages due to limited awareness, delayed diagnosis, and socioeconomic barriers.[12] This necessitates the identification of reliable molecular markers that can assist in early detection and risk stratification.

Therefore, this study aims to compare the expression levels of the COL10A1 gene in normal and malignant breast tissues obtained from Bangladeshi patients. The findings may provide a better understanding of COL10A1's diagnostic value and its potential utility as a biomarker for breast cancer in resource-limited settings.

Objectives: This study aimed to evaluate the differential expression of COL10A1 in tumor versus adjacent normal tissue and its association with clinical, demographic, and reproductive factors.

Methods

This was a cross-sectional observational study designed to compare the expression levels of the COL10A1 gene between normal and malignant breast tissues in diagnosed breast cancer patients.

The study was conducted in the Department of Biochemistry, Dhaka Medical College, Dhaka, in collaboration with the Institute for Population and Precision Health, Department of Public Health Sciences, The University of Chicago, USA.

The study was carried out over a one-year period, from July 2022 to June 2023. The study included histologically confirmed female breast cancer patients attending the Department of General Surgery and the Breast Clinic of Dhaka Medical College Hospital (DMCH). A total of 34 participants were included in this study. Sample size was calculated using Guilford and Fruchter's formula. A purposive sampling method was applied to select eligible patients who met the inclusion criteria.

Enrolled Criteria
Patients were enrolled based on history, physical examination and following inclusion and exclusion criteria. Additionally, tissue samples were obtained during surgery and preserved for laboratory analysis.

Inclusion Criteria
(1) Female patients aged 25 to 70 years (2) Diagnosed cases of breast cancer confirmed by clinical examination, FNAC, mammography, histopathology and immunohistochemistry, undergoing modified radical mastectomy (MRM) were included.

Exclusion Criteria
Patients who had already completed breast cancer treatment (surgery, chemotherapy, or radiotherapy), with a history of other malignancies were excluded.

Study Procedure
Eligible patients were recruited after obtaining informed written consent. From each patient, two types of tissue samples were collected per-operatively: Malignant tumor tissue (approx. 0.5 cm^3) & Adjacent normal breast tissue (≥2 cm away from the tumor margin). Both tissue types were collected in pre-labeled Eppendorf tubes containing DNA/RNA Shield (Zymo Research, CA, USA) and stored at room temperature (25–30°C) initially, then transported under biosafety and cold-chain protocols (−80°C) to the University of Chicago for molecular analysis.

Laboratory Procedures
RNA Extraction: Using Quick-DNA/RNA™ Microprep Plus Kit (Zymo Research, USA). RNA Quality Check: Performed using Nanodrop spectrophotometer. Library Preparation and Targeted RNA Sequencing: Done using custom-designed Twist Bioscience Kit (Twist, CA, USA)

Gene Expression Analysis:
Performed using targeted RNA sequencing to assess COL10A1 gene expression in normal and malignant tissues

Independent Variables
Age, BMI, family history of breast cancer, age of menarche, contraceptive use, parity, and breastfeeding history

Dependent Variable
COL10A1 gene expression level in breast tissues

Statistical Analysis
Data were entered and analyzed using SPSS version 26.0.

Results
The present study analyzed the demographic and clinical characteristics of 34 breast cancer patients to explore patterns relevant to COL10A1 gene expression. The baseline variables, including age, BMI, education, residence, socioeconomic status, and breastfeeding history, are summarized in Table I.

Table I: Demographic characteristics of the study participants (n=34)

Variables	Categories	Frequency (n)	Percentage (%)
Age Group (years)	≤ 40	10	29.41%
	41–50	16	47.06%
	> 50	8	23.53%
Mean Age ± SD (years)	**47.41 ± 2.52**		
Body Mass Index (BMI)	Normal (18.5–24.9)	18	52.94%
	Overweight (≥25.0)	16	47.06%
Mean BMI ± SD (kg/m²)	**21.07 ± 1.17**		
Educational Level	Primary or below	6	17.65%
	Secondary	12	35.29%
	Higher Secondary & above	16	47.06%
Residence	Urban	22	64.71%
	Rural	12	35.29%
Socioeconomic Status	Lower	9	26.47%
	Middle	17	50.00%
	Upper	8	23.53%
Breastfeeding History			
	Yes	4	12.00%
	No	30	88.00%

Table I presents the demographic characteristics of the 34 breast cancer patients. Most participants were aged 41–50 years (47.06%), with a mean age of 47.41 ± 2.52 years. Over half (52.94%) had normal BMI, while 47.06% were overweight, with a mean BMI of 21.07 ± 1.17 kg/m². Nearly half (47.06%) had higher secondary education or above. A majority resided in urban areas (64.71%) and belonged to the middle socioeconomic group (50.00%), reflecting a predominantly middle-aged, urban, and moderately educated population. In this study (88%) had a history of breastfeeding, while only 12% did not breastfeed. This indicates a high prevalence of breastfeeding practices among the respondents.

Table II: Distribution of reproductive and clinical history (n=34)

Variable	Category	Frequency (n)	Percentage (%)
Family history of Breast Cancer	Present	10	29.41%
	Absent	24	70.59%
Age at menarche	≤13 years	20	58.82%
	>13 years	14	41.18%
Contraceptive use	Yes	12	35.29%
	No	22	64.71%
Parity	Multipara	27	79.41%
	Primipara	7	20.59%

Table II presents the distribution of reproductive and clinical history among 34 participants. A family history of breast cancer was present in 29.41% of cases, while 70.59% had no such history. The majority (58.82%) experienced menarche at or before 13 years of age. Contraceptive use was reported by 35.29% of participants, whereas 64.71% did not use contraceptives. Most women were multiparous (79.41%), and a smaller proportion were primiparous (20.59%). This data highlights key reproductive and clinical characteristics relevant to breast health assessment in the study population.

Table III: Comparison of COL10A1 gene expression between normal and tumor tissue (n=34)

Tissue type	Median Expression	IQR (25th–75th Percentile)	p-value
Normal tissue	2.044	-2.651 to -0.615	
Tumor tissue	3.133	-0.735 to 2.695	0.011

Table III shows a significant difference in COL10A1 gene expression between normal and tumor tissues in breast cancer patients. The median expression level was notably higher in tumor tissue (3.133) compared to normal tissue (2.044), with a p-value of 0.011, indicating statistical significance as per the Mann–Whitney U test. This suggests that COL10A1 is upregulated in breast tumor tissues.

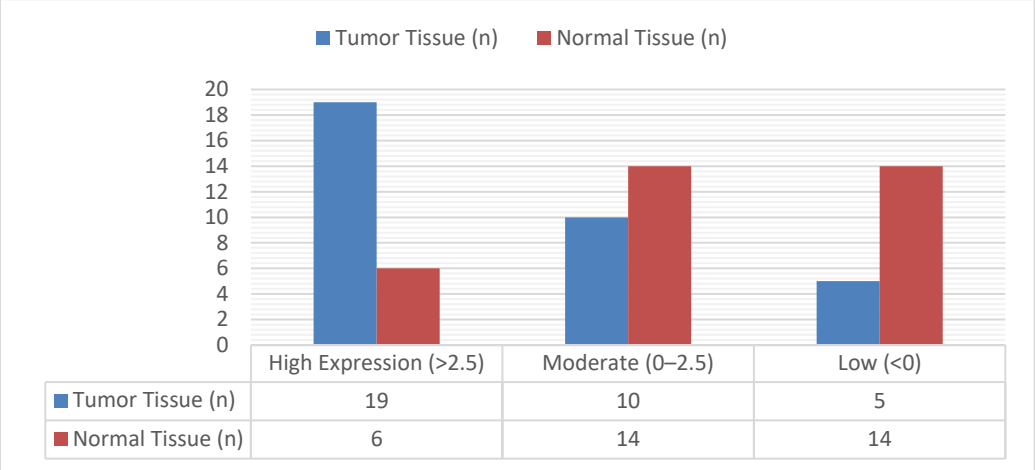

Figure 1. Expression level categorization of COL10A1 in tumor vs. normal tissue (n=34)

Figure 1 categorizes COL10A1 expression levels and reveals that high expression (>2.5) was observed in 19 tumor tissue samples compared to only 6 in normal tissue. Conversely, low expression (<0) was more frequent in normal tissues (14) than in tumor tissues (5). This indicates a predominance of elevated COL10A1 expression in breast tumor tissues, supporting its potential role in tumor development.

Table IV: Distribution of tumor stage and COL10A1 expression level (n=34)

Tumor stage	Mean Expression ± SD	Median	p-value
Stage I	2.12 ± 0.43	2.05	
Stage II	2.85 ± 0.68	2.89	
Stage III	3.42 ± 1.01	3.56	0.034

Table IV demonstrates a progressive increase in COL10A1 gene expression across advancing tumor stages. The mean expression rose from 2.12 ± 0.43 in Stage I to 3.42 ± 1.01 in Stage III, with a statistically significant difference (p = 0.034). This suggests that higher COL10A1 expression may be associated with more advanced stages of breast cancer.

Table V: Logistic regression analysis for predictors of high COL10A1 expression in tumor tissue (n=34)

Variable	Odds Ratio (OR)	95% CI	p-value
Age ≤50	1.21	0.42–3.50	0.712
BMI ≥25	1.34	0.49–3.68	0.567
Family history (Yes)	2.11	0.61–7.32	0.238
Menarche ≤13 years	1.56	0.51–4.74	0.436
No breastfeeding	3.72	0.72–19.3	0.112

Table V shows that none of the evaluated factors—age ≤50, BMI ≥25, positive family history, early menarche (≤13 years), or lack of breastfeeding—were statistically significant predictors of high COL10A1 expression in tumor tissue, as all p-values exceeded 0.05. Although no breastfeeding showed the highest odds ratio (3.72), it did not reach significance, indicating these variables may not independently influence COL10A1 expression in this sample.

Discussion

This study evaluated the expression levels of COL10A1 in breast cancer tissues compared to adjacent normal tissues and explored its association with clinicopathological features. The findings demonstrated significantly higher COL10A1 expression in tumor tissues, with a median expression of 3.133 compared to 2.044 in normal tissues (p = 0.011). This observation aligns with previous bioinformatics studies indicating that COL10A1 is upregulated in breast cancer and may serve as a diagnostic and prognostic biomarker.[13] Overexpression of COL10A1 has been correlated with poor overall survival, relapse-free survival, and distant metastasis-free survival in breast cancer patients.[14]

A notable finding in our study was the progressive increase in COL10A1 expression with advancing tumor stages. Patients with stage III breast cancer exhibited the highest mean expression (3.42 ± 1.01), followed by stage II (2.85 ± 0.68) and stage I (2.12 ± 0.43), with a statistically significant difference among groups (p = 0.034). This trend suggests a possible role of COL10A1 in tumor progression and aggressiveness. Similar findings have been reported in studies where high COL10A1 expression was significantly associated with advanced TNM stage, lymph node metastasis, and increased tumor invasiveness.[15,16]

Despite analyzing several demographic and reproductive factors such as age, BMI, family history, age at menarche, and breastfeeding history, none were found to be statistically significant predictors of high COL10A1 expression in tumor tissue. However, patients with no history of breastfeeding exhibited a higher odds ratio (OR = 3.72), suggesting a potential but non-significant association that may warrant further investigation in larger studies. Prior research indicates that COL10A1 expression may be influenced more by intrinsic tumor biology and stromal interactions rather than by classical reproductive or demographic factors.[17] The exact mechanism by which COL10A1 promotes breast tumorigenesis remains under investigation. Functional studies have shown that COL10A1 contributes to tumor cell proliferation, migration, and invasion through its interaction with proteins such as P4HB and by activating pathways like TGF-β, FAK/MAPK, and epithelial–mesenchymal transition (EMT).[18] Furthermore, COL10A1 overexpression has been implicated in modulating the tumor microenvironment, including promoting angiogenesis and

immune cell infiltration, thereby enhancing tumor aggressiveness.[19]

Conclusion

In conclusion, our findings support the role of COL10A1 as a potential biomarker for breast cancer, particularly in identifying aggressive tumors with higher stages. While the current study did not establish significant associations with clinical or reproductive parameters, the increasing expression across tumor stages underscores its biological relevance. Larger cohort studies integrating molecular subtypes, long-term follow-up, and therapeutic response are needed to validate the prognostic value and therapeutic potential of targeting COL10A1 in breast cancer.

References

1. Winters S, Martin C, Murphy D, Shokar NK. Breast cancer epidemiology, prevention, and screening. Prog Mol Biol Transl Sci. 2017;151:1-32. https://doi.org/10.1016/bs.pmbts.2017.07.002
2. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide. CA Cancer J Clin. 2021;71(3):209-249. https://doi.org/10.3322/caac.21660
3. Momenimovahed Z, Salehiniya H. Epidemiological characteristics of and risk factors for breast cancer in the world. Breast Cancer (Dove Med Press). 2019;11:151-164. https://doi.org/10.2147/BCTT.S176070
4. Perou CM, Sørlie T, Eisen MB, et al. Molecular portraits of human breast tumors. Nature. 2000;406(6797):747-752. https://doi.org/10.1038/35021093
5. Zhang M, Zhang L, Hei R, Li X, Cai H, Wu X. COL10A1 acts as a potential prognostic biomarker and is associated with immune infiltration in breast cancer. OncoTargets Ther. 2020;13:9807-9819. https://doi.org/10.2147/OTT.S269507
6. Zhou Y, Boucher M, Lin H, et al. Identification of COL10A1 as a novel biomarker in breast cancer via integrated bioinformatics analysis. PeerJ. 2022;10:e12656. https://doi.org/10.7717/peerj.12656
7. Li Y, Wu X, Chen M, et al. COL10A1 in breast cancer: A potent prognostic factor associated with immune infiltration and epithelial–mesenchymal transition. Cancer Cell Int. 2020;20:527. https://doi.org/10.1186/s12935-020-01591-4
8. Zhang J, Liu X, Zou Z, et al. The identification of key genes and pathways involved in the tumorigenesis and prognosis of breast cancer. Aging (Albany NY). 2019;11(23):10257-10277. https://doi.org/10.18632/aging.102486
9. Yang C, Nie H, Li S, et al. Prognostic and diagnostic significance of COL10A1 expression in breast cancer. Exp Mol Pathol. 2021;120:104650. https://doi.org/10.1016/j.yexmp.2021.104650
10. Wang K, Shen T, Siegal GP, Wei S. The CD248–COL10A1–LRRC15 axis identifies a subset of cancer-associated fibroblasts in triple-negative breast cancer. Am J Pathol. 2018;188(6):1138-1148. https://doi.org/10.1016/j.ajpath.2018.01.018
11. Zhang L, He L, Cai H, et al. Plasma COL10A1 protein as a diagnostic biomarker for breast cancer. OncoTargets Ther. 2020;13:7475-7484. https://doi.org/10.2147/OTT.S263375
12. Akram M, Iqbal M, Daniyal M, Khan AU. Awareness and current knowledge of breast cancer. Biol Res. 2017;50(1):33. https://doi.org/10.1186/s40659-017-0140-9.
13. Zhang M, Chen H, Wang M, Bai F, Wu K. Bioinformatics analysis of prognostic

significance of COL10A1 in breast cancer. Biosci Rep. 2020;40(2):BSR20193286. doi:10.1042/BSR20193286
14. Qing Y, Yin F, Yang W, Zhang L. High COL10A1 expression predicts poor prognosis in breast cancer. Technol Cancer Res Treat. 2020;19:1533033820934129. doi:10.1177/1533033820934129
15. Yang W, Wu X, Zhou F. COL10A1 promotes breast cancer cell proliferation, migration, and invasion via P4HB interaction. Med Sci Monit. 2021;27:e928919. doi:10.12659/MSM.928919
16. Wu Y, Sun H, Tan J, Song J. COL10A1 expression correlates with tumor stage and lymph node metastasis in breast cancer. Cancer Biomark. 2022;35(1):15-22. doi:10.3233/CBM-210408
17. Li J, Wang Z, Zhou H, et al. COL10A1 facilitates tumor progression through TGF-β signaling and modulates immune infiltration in breast cancer. Front Immunol. 2023;14:955949. doi:10.3389/fimmu.2023.955949
18. Li H, Huang S, Shen M, et al. COL10A1 enhances epithelial–mesenchymal transition and angiogenesis in breast cancer via the FAK/MAPK pathway. Oncol Rep. 2024;51(3):133. doi:10.3892/or.2024.8821
19. Zhang Y, Zhou L, Han Y, et al. COL10A1-associated gene signature predicts tumor microenvironment remodeling and prognosis in breast cancer. J Transl Med. 2022;20:98. doi:10.1186/s12967-022-03274-2.