Unveiling Collagen’s Role in Breast Cancer: Insights into Expression Patterns, Functions and Clinical Implications

Xia Li¹, Yue Jin¹, Jian Xue²

¹Department of Molecular Diagnosis, Northern Jiangsu People’s Hospital, Yangzhou, People’s Republic of China; ²Department of Emergency Medicine, Yizheng People’s Hospital, Yangzhou, People’s Republic of China

Correspondence: Yue Jin, Department of Molecular Diagnosis, Northern Jiangsu People’s Hospital, No. 98, Nantong West Road, Yangzhou, Jiangsu Province, People’s Republic of China, Email yyyue_king@163.com

Abstract: Collagen, the predominant protein constituent of the mammalian extracellular matrix (ECM), comprises a diverse family of 28 members (I–XXVIII). Beyond its structural significance, collagen is implicated in various diseases or cancers, notably breast cancer, where it influences crucial cellular processes including proliferation, metastasis, apoptosis, and drug resistance, intricately shaping cancer progression and prognosis. In breast cancer, distinct collagens exhibit differential expression profiles, with some showing heightened or diminished levels in cancerous tissues or cells compared to normal counterparts, suggesting specific and pivotal biological functions. In this review, we meticulously analyze the expression of individual collagen members in breast cancer, utilizing Transcripts Per Million (TPM) data sourced from the GEPIA2 database. Through this analysis, we identify collagens that deviate from normal expression patterns in breast cancer, providing a comprehensive overview of their expression dynamics, functional roles, and underlying mechanisms. Our findings shed light on recent advancements in understanding the intricate interplay between these aberrantly expressed collagens and breast cancer. This exploration aims to offer valuable insights for the identification of potential biomarkers and therapeutic targets, thereby advancing the prospects of more effective interventions in breast cancer treatment.

Keywords: collagen, breast cancer, extracellular matrix, prognostic marker

Introduction

Breast cancer, characterized by its heterogeneity, has become the leading cause of cancer-related deaths among women worldwide, presenting significant challenges to clinical management due to its high morbidity and mortality rates.¹ Despite considerable advancements in breast cancer research, the identification of specific and effective targets remains elusive, rendering the treatment of breast cancer an ongoing challenge. Consequently, there is an urgent need to identify effective diagnostic, prognostic and therapeutic targets.

Recent advances in the understanding of breast cancer progression implicate key processes such as epithelial–mesenchymal transition (EMT), the presence of cancer stem cells (CSC), and dysregulation of ECM functions. Among these, ECM is crucial for cell proliferation, differentiation, and tissue homeostasis maintenance. Serving as a pivotal component of the tumor microenvironment, it orchestrates complex tissue interactions influencing tumor growth and metastasis. Acting as a substrate for cell adhesion and a reservoir for growth factors, ECM plays indispensable roles in cancer development and metastasis. It is extensively studied in the context of breast cancer due to its disrupted regulation.²

Within ECM, proteins, glycoproteins, proteoglycans, and polysaccharides collectively contribute, with collagen standing out as the most abundant component.³ Collagen constitutes approximately 30% of mammalian total protein, predominantly found in bone and skin. Apart from its structural role in determining tissue mechanical properties and shape, collagen acts as a physical barrier to cell migration and regulates proliferation in both normal and cancer cells. Alterations in its expression can lead to various fibrotic diseases. In breast cancer, increased mammographic density,
associated with elevated collagen concentration, is considered to be a significant risk factor for breast cancer development. The expression, organization, and post-translational modifications of collagen types have been reported to be profoundly altered, contributing to a desmoplastic reaction and a stiffened stroma, which in turn influence tumor cell behavior, angiogenesis, and metastatic potential. Meanwhile, collagen restructuring manifests pathological hallmarks of cancer progression, influencing fiber organization, matrix deposition, stiffness, proteolytic products, immune recruitment, and cell phenotype. At present, strategies aimed at inhibiting collagen synthesis or deposition have shown promise in suppressing breast cancer progression.

Typically, the collagen superfamily identified in humans comprises at least 28 types (I–XXVIII), exhibiting distinct structural classifications, including fibrous collagen, basal membrane collagen, microfibrous collagen, anchored collagen, hexagonal reticular collagen, non-fibrous collagen, and transmembrane collagen (Table 1). Collagen XXIX is identified

<table>
<thead>
<tr>
<th>Collagen Type</th>
<th>α Chains</th>
<th>Gene Name</th>
<th>Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I collagen</td>
<td>α1 (I), α2 (I)</td>
<td>COL1A1, COL1A2</td>
<td>Fibril-forming collagen</td>
</tr>
<tr>
<td>Type II collagen</td>
<td>α1 (II)</td>
<td>COL2A1</td>
<td>Fibril-forming collagen</td>
</tr>
<tr>
<td>Type III collagen</td>
<td>α1 (III)</td>
<td>COL3A1</td>
<td>Fibril-forming collagen</td>
</tr>
<tr>
<td>Type IV collagen</td>
<td>α1 (IV), α2 (IV), α3 (IV), α4 (IV), α5 (IV), α6 (IV)</td>
<td>COL4A1, COL4A2, COL4A3, COL4A4, COL4A5, COL4A6</td>
<td>Network-forming collagen</td>
</tr>
<tr>
<td>Type V collagen</td>
<td>α1 (V), α2 (V), α3 (V), α4 (V)</td>
<td>COL5A1, COL5A2, COL5A3</td>
<td>Fibril-forming collagen</td>
</tr>
<tr>
<td>Type VI collagen</td>
<td>α1 (VI), α2 (VI), α3 (VI), α4 (VI), α5 (VI), α6 (VI)</td>
<td>COL6A1, COL6A2, COL6A3, COL6A4, COL6A5, COL6A6</td>
<td>Beaded filament-forming collagen</td>
</tr>
<tr>
<td>Type VII collagen</td>
<td>α1 (VII)</td>
<td>COL7A1</td>
<td>Anchoring fibrils</td>
</tr>
<tr>
<td>Type VIII collagen</td>
<td>α1 (VIII), α2 (VIII)</td>
<td>COL8A1, COL8A2</td>
<td>Network-forming collagen</td>
</tr>
<tr>
<td>Type IX collagen</td>
<td>α1 (IX), α2 (IX), α3 (IX)</td>
<td>COL9A1, COL9A2, COL9A3</td>
<td>Fibril-associated collagen</td>
</tr>
<tr>
<td>Type X collagen</td>
<td>α1 (X)</td>
<td>COL10A1</td>
<td>Network-forming collagen</td>
</tr>
<tr>
<td>Type XI collagen</td>
<td>α1 (XI), α2 (XI), α3 (XI)</td>
<td>COL11A1, COL11A2, COL11A3 (COL2A1)</td>
<td>Fibril-forming collagen</td>
</tr>
<tr>
<td>Type XII collagen</td>
<td>α1 (XII)</td>
<td>COL12A1</td>
<td>Fibril-associated collagen</td>
</tr>
<tr>
<td>Type XIII collagen</td>
<td>α1 (XIII)</td>
<td>COL13A1</td>
<td>Membrane collagen</td>
</tr>
<tr>
<td>Type XIV collagen</td>
<td>α1 (XIV)</td>
<td>COL14A1</td>
<td>Fibril-associated collagen</td>
</tr>
<tr>
<td>Type XV collagen</td>
<td>α1 (XV)</td>
<td>COL15A1</td>
<td>Multiplexins</td>
</tr>
<tr>
<td>Type XVI collagen</td>
<td>α1 (XVI)</td>
<td>COL16A1</td>
<td>Fibril-associated collagen</td>
</tr>
<tr>
<td>Type XVII collagen</td>
<td>α1 (XVII)</td>
<td>COL17A1</td>
<td>Membrane collagen</td>
</tr>
<tr>
<td>Type XVIII collagen</td>
<td>α1 (XVIII)</td>
<td>COL18A1</td>
<td>Multiplexins</td>
</tr>
<tr>
<td>Type XIX collagen</td>
<td>α1 (XIX)</td>
<td>COL19A1</td>
<td>Fibril-associated collagen</td>
</tr>
<tr>
<td>Type XX collagen</td>
<td>α1 (XX)</td>
<td>COL20A1</td>
<td>Fibril-associated collagen</td>
</tr>
<tr>
<td>Type XXI collagen</td>
<td>α1 (XXI)</td>
<td>COL21A1</td>
<td>Fibril-associated collagen</td>
</tr>
<tr>
<td>Type XXII collagen</td>
<td>α1 (XXII)</td>
<td>COL22A1</td>
<td>Fibril-associated collagen</td>
</tr>
<tr>
<td>Type XXIII collagen</td>
<td>α1 (XXIII)</td>
<td>COL23A1</td>
<td>Membrane collagen</td>
</tr>
</tbody>
</table>

(Continued)
Table 1 (Continued).

<table>
<thead>
<tr>
<th>Collagen Type</th>
<th>α Chains</th>
<th>Gene Name</th>
<th>Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type XXIV collagen</td>
<td>α1 (XXIV)</td>
<td>COL24A1</td>
<td>Fibril-forming collagen</td>
</tr>
<tr>
<td>Type XXV collagen</td>
<td>α1 (XXV)</td>
<td>COL25A1</td>
<td>Membrane collagen</td>
</tr>
<tr>
<td>Type XXVI collagen</td>
<td>α1 (XXVI)</td>
<td>COL26A1</td>
<td>Beaded filament-forming collagen</td>
</tr>
<tr>
<td>Type XXVII collagen</td>
<td>α1 (XXVII)</td>
<td>COL27A1</td>
<td>Fibril-forming collagen</td>
</tr>
<tr>
<td>Type XXVIII collagen</td>
<td>α1 (XXVIII)</td>
<td>COL28A1</td>
<td>Beaded filament-forming collagen</td>
</tr>
</tbody>
</table>

Notes: *The α4 (V) is solely synthesized by Schwann cells. **The α4 (VI) does not exist in humans. ***The α5 (VI) is designated as α1(XXIX).

as a novel epidermal collagen, but its encoding gene COL29A1 is shown to be identical to the COL6A5 gene, and the α1 (XXIX) chain corresponds to the α5 (VI) chain. 7,8 Each collagen can form homotrimers or heterotrimers composed of three α chains. Recent research has identified several collagen family members as novel biological targets for breast cancer. In this comprehensive review, we analyze the expression patterns of all collagens in breast cancer, emphasizing the nuanced roles and mechanisms of distinct members, particularly the most abundant type I collagen, highly expressed collagen type X alpha 1 (COL10A1) and collagen type XI alpha 1 (COL11A1), and other minor yet significant collagens. We highlight recent insights into the interactions of these collagens with breast cancer, underscoring their potential as promising biomarkers and therapeutic targets in breast cancer management.

Collagen Expression Profile in Breast Cancer

The alteration of the collagen expression profile is a quintessential hallmark of breast cancer, significantly influencing tumor progression and metastasis. In this malignancy, collagen undergoes profound alterations not only in abundance but also in its molecular integrity, thereby reshaping the tumor microenvironment and modulating the phenotypic attributes of neoplastic cells. 9,10

A critical hallmark of breast cancer pathophysiology is the augmented density of collagen, particularly type I collagen, which is the most abundant collagen in the stroma. 11 Elevated concentration of this collagen type is routinely identified proximal to neoplastic sites. Such an escalation in collagen density is intimately linked with increased tissue rigidity, which in turn precipitates malignant transformation and abets the migration and invasiveness of tumor cells. Beyond the mere increase in density, the compositional spectrum of collagen also undergoes significant alteration, such as type IV collagen. Suppression of collagen type IV alpha 2 (COL4A2) significantly suppresses the migration and proliferation of triple-negative breast cancer (TNBC) cells. 12 Variations in less abundant collagen types, such as types XV and XVIII, are implicated in the angiogenic and lymphangiogenic processes within the tumor milieu. 13,14 Moreover, differences in orientation and organization of collagen also alter its effects on tumor invasion. For example, linearized (fibrillar) collagen I induces cellular phenotypes consistent with an invasive behavior, while high-density non-fibrillar collagen I induces tumor-suppressive attributes. 15

Leveraging TPM data derived from the GEPIA2 database, we stratified collagens into two distinct categories based on their expression profiles in breast cancer tissues: those exhibiting comparative overexpression and those showing underexpression (as depicted in Figure 1). Predominantly, a greater spectrum of collagen types exhibits heightened expression in breast cancer tissues, with COL10A1, COL11A1, collagen type I alpha 1 (COL1A1), collagen type XXII alpha 1 (COL22A1) and collagen type III alpha 1 (COL3A1) emerging as the five most significantly overexpressed. In the subsequent review, we opted to analyze several other collagens apart from COL22A1, given its limited study in breast cancer. Among these low expressed collagens, collagen type XXV alpha 1 chain (COL25A1) and collagen type XVII alpha 1 (COL17A1) show prominent differences, followed by collagen type VI alpha 6 (COL6A6), collagen type IV alpha 3 (COL4A3), and collagen type IV alpha 6 (COL4A6). These expressed collagens often correlate with a better prognosis in breast cancer. Thus, in this review, we also provide a summary of the top 5 collagens exhibiting poor
expression in breast cancer. Moreover, owing to the paucity of research on relatively under-expressed collagens in breast cancer, we conducted a combined analysis of their family members.

**Specific Collagen Families in Breast Cancer**

**Type I Collagen**

Type I collagen, the most abundant type, exists in extra parenchymal tissue (meninges, choroid plexus stroma), blood vessels, and parenchyma of the subependymal layer in the brain. It is also found in skin, bone, tendon, ligament, blood vessel wall, and teeth and widely distributed in the interstitial tissue of connective tissue. Type I collagen is a heterotrimer composed of three α-chains encoded by COL1A1 and the alpha 2 chain (COL1A2) genes. Functionally, type I collagen operates through collagen fibers, thereby exerting profound effects on physiological and pathological processes at the cellular, tissue, or organismal level. In breast cancer, analysis of data from the GEPIA2 database reveals an upregulation in both COL1A1 and COL1A2 expression levels. Meanwhile, it is indicated to augment the aggressive characteristics of breast cancer cells.

**Collagen Type I Alpha 1 Chain (COL1A1)**

COL1A1, a major constituent of type I collagen and one of the most abundant proteins in the human body, is widely distributed in the interstitium of parenchymal organs and connective tissue. Produced by interstitial fibroblasts, COL1A1 plays a crucial role in epithelial–mesenchymal transition, intricately linked to malignant tumorigenesis. In
various cancers, including breast cancer, COL1A1 has been identified as a cancer-promoting factor, influencing cell proliferation, metastasis, apoptosis, cisplatin resistance, and overall cancer progression and prognosis. Currently, it has been seen as a novel predictive biomarker in metastatic lung cancer,21 lung adenocarcinoma,22 gastric cancer,23 mesothelioma,24 hepatocellular cancer,25 and breast cancer.

In breast cancer cell lines, studies reveal that COL1A1 is predominantly expressed in cytoplasm, with higher expression in less invasive cell lines (MCF7 and T47D) compared to more invasive triple-negative cell lines (MDA-MB-231 and BT549).26 It is also detected in tumor samples and found significantly higher in tumor samples compared with the corresponding normal samples.27 The expression of COL1A1 is found to be associated with estrogen receptor or progesterone receptor (ER/PR) expression and metastasis status. In ER+ breast cancer patients, higher COL1A1 expression correlates with aggressive cellular behavior and poorer prognosis. Knockdown of COL1A1 in breast cancer cells can limit the proliferative and invasive ability of breast cancer cells.26 Furthermore, higher COL1A1 expression also predicts increased cisplatin-based chemotherapy response rates.26 It is predicted to be radiation-associated differentially expressed genes, with another two collagens COL3A1 and COL1A2.28 In fact, the intricate relationship between COL1A1 and radiation has been elucidated in several studies. It serves as a pivotal radio-resistance factor, exerting radio-resistance effects through intricate signaling cascades such as miR-29a or the Caspase-3/PI3K/AKT pathways.29,30 Subsequent to radiation exposure, discernible alterations in COL1A1 expression have been confirmed.31 Current investigations underscore the upregulation of COL1A1 as a significant risk factor for the onset of secondary nonbreast diseases in breast cancer patients following radiation therapy.28,32 These insights offer a potential therapeutic avenue for discerning radioresistance in breast cancer.

Functionally, COL1A1 can serve as a direct target of microRNAs (miRNAs) and be regulated by circular RNAs (circRNAs). For instance, it is targeted by miR-1184 and positively modulated by circ_0000523, promoting cell proliferation, cell cycle progression, migration and invasion of cancer cells.33 It is also identified as a direct downstream target gene of miR-196b-5p, and its overexpression partly abrogates miR-196b-5p-mediated inhibition of proliferation and migration in MDA-MB-231 and MDA-MB-468 breast cancer cells.34 COL1A1 still serves as a direct target of miR-328-3p and could be up-regulated by overexpression of hsa_circRNA_002178 in breast cancer.35 Furthermore, RhoBTB3 can regulate breast cancer progression by controlling collagen deposition, specifically, knockdown of RhoBTB3 regulates breast cancer cell proliferation and invasion, accompanied by the reduction of COL1A1.36 MRTF-A physically interacts with the promoter of COL1A1 to facilitate the acetylation of chromatin and the recruitment of RNA polymerase II in breast cancer.37 In addition, COL1A1 is associated with immune and tumor microenvironment. It is strongly associated with tumor immune infiltration of CD4+ T and CD8+ T cells, neutrophils, macrophages and dendritic cells.38 Down-regulation of COL1A1 inhibits exosome secretion, possibly via inhibiting COL I and upregulating CAV-1, thereby inhibiting tumor-associated fibroblast activation and matrix remodeling in the tumor microenvironment.39 Beyond that, COL1A1 is also implicated in breast cancer metastasis through interactions with RhoC-ROCK and transforming growth factor-β (TGF-β) signaling pathways, RIP1-RIP3-MLKL, MMP pathways, and WNT/planar cell polarity (PCP) signaling pathway.40,41 Pathogenic variants in COL1A1 are also found.42,43

Collagen Type I Alpha 2 Chain (COL1A2)

In contrast, COL1A2 is extensively used in experimental models to study collagen I biosynthesis, osteoporosis and bone diseases.44,45 Its abnormal expression has been reported in various diseases, including heart failure,46 melanoma,47 gastric cancer,48 breast cancer, hepatocellular carcinoma, colorectal carcinoma,49,50 pancreatic cancer,51 and glioblastoma.52 Defective COL1A2 can alter the structural integrity of the ECM and lead to cardiomyopathy in adulthood53 and is also positively associated with immune infiltration and tumor immune escape.49 Overexpression of COL1A2 affects proliferation, migration, and invasion of cancer cell lines.50

Although less studied in breast cancer compared to COL1A1, COL1A2 is identified as a hub gene associated with the survival outcomes of human epidermal growth factor receptor 2 (HER2)-positive patients by constructing the protein–protein interaction (PPI) network.34 Upregulated in breast cancer tissues, COL1A2 is correlated with reduced overall and recurrence-free survival,55 and its expression is elevated by radiation.28 Research on its mechanism in breast cancer is
ongoing, with previous studies indicating potent induction of COL1A2 transcription by TGF-β, involving Smad family proteins or Smad-binding elements.56

**Collagen Type X Alpha 1 Chain (COL10A1)**

COL10A1, a member of the type X collagen family, stands out as a secreted short-chain collagen, serving as a major component of the interstitial matrix. It shares notable homology with collagen VIII and exhibits significant upregulation in various tumor types, displaying restricted expression in normal tissues.37 Functioning as a gene associated with disease progression, COL10A1 plays a pivotal role in regulating the proliferation, migration, and invasion of tumor cells.58,59 Studies have shown that COL10A1 suppression remarkably inhibits cell proliferation, migration, and invasion in breast cancer cells.

At present, considerable attention has been given to COL10A1 as a potential diagnostic and prognostic predictor for breast cancer. Its concentration in the plasma can discriminate breast cancer patients from those with benign disease.60 Across different breast cancer subtypes, COL10A1 is significantly overexpressed, showing poor overall survival (OS), relapse-free survival (RFS), distant metastasis-free survival (DMFS) and disease-free survival (DFS).61 Furthermore, the tumor microenvironment-specific expression of collagen X, together with its localization in the vasculature, also facilitate its use as a novel target for the diagnosis and treatment of diverse solid tumor types.57

Although overexpression of COL10A1 in breast cancer is well established, its specific regulatory mechanisms in tumor progression are still undereexplored. Currently, ongoing research has revealed its impact on breast cancer cells by modulating immune-infiltrating cells.62 It is positively associated with immune cell infiltration, including B cell, CD8+ T cell, CD4+ T cell, macrophage, neutrophil, and dendritic cell.63 COL10A1’s direct interaction with Prolyl 4-hydroxylase beta polypeptide (P4HB) has been identified, leading to the upregulation of P4HB expression, thereby promoting malignant progression in breast cancer.64 Additionally, COL10A1 is enriched in TGF-β signaling pathway, with 15-leucine-rich repeat-containing membrane protein (LRRC15) identified as a correlated gene. Strong correlation between the expression of LINC01614 and COL10A1 (R² = 0.6929) is also observed.65 However, more in-depth mechanistic studies are warranted.

**Collagen Type XI Alpha 1 Chain (COL11A1)**

COL11A1, one of the three alpha chains of type XI collagen, is a minor fibrillar collagen primarily expressed in the cartilage.66 It serves as the most abundant component of the interstitial extracellular matrix, playing a crucial role in ECM and tissue integrity. Elevated expression of COL11A1 is observed in various human tumors like ovarian cancer, breast cancer, pancreatic cancer, lung adenocarcinoma, colorectal cancer, prostate adenocarcinoma and so on. In some hematological malignancies, COL11A1 overexpression is associated with better prognosis. However, in solid tumors, its high level is often linked to aggressive tumor phenotype and poor prognosis. In breast cancer, high level of COL11A1 is linked to metastasis, tumor aggressiveness, chemoresistance, malignant relapses and a poor survival. Machine learning algorithms also designate COL11A1 as a potential therapeutic target in breast cancer, with its high expression strongly correlated with poor prognosis and invasive recurrence in breast ductal carcinoma in situ.67–69 Furthermore, it is previously described as a diagnostic marker to differentiate between invasive and non-invasive breast cancer.70

Mechanistically, COL11A1 is targeted by several microRNA and long non-coding RNA (lncRNA), like miR-139-5p and miR-4458 (Table 2). Overexpression of miR-139-5p inhibits the proliferation and promotes the apoptosis of breast cancer cells by suppressing COL11A1 expression.71 miR-4458 mimetic transfection and si-COL11A1 inactivate the DDR2/SRC signaling pathway in estrogen receptor-positive breast cancer.72 COL11A1 is also associated with microRNA let-7b, which can be promoted by caudal-type homeobox 2 (CDX2) and exert an inhibitory effect on the proliferation, migration, and metastasis of breast cancer cells by inhibiting COL11A1 expression.73 Moreover, COL11A1 is selected as the N6-methyladenosine (m6A) target gene and the downstream gene of Wilms tumor gene 1 (WT1).74,75 Its regulation is involved in miR-29,76 TGF-β1 signaling,77 AKT/c/EBPβ/PDK1 axis78 and other collagens, such as collagen type XI alpha 2 (COL11A2), collagen type XI alpha 3 (COL11A3), collagen type V alpha 1 (COL5A1) and collagen type V alpha 2 (COL5A2).79
Additionally, COL11A1 is observed to be upregulated in recurrent tumors, promoting tumor recurrence post chemotherapy. It can mediate resistance of cancer cells to some drugs, which is particularly informative for clinical work. COL11A1 can mediate resistance of cancer cells to cisplatin through multiple mechanisms. For example, it upregulates inhibitors of Apoptosis Proteins (IAP) expression to evade cisplatin-induced apoptosis, especially XIAP, BIRC2 and BIRC3. It also drives cisplatin resistance by relying on TWIST1 expression and fatty acid metabolism.

By regulating TWIST1 expression, it still confers paclitaxel resistance, a commonly used chemotherapy drug. An in-depth understanding of the mechanism of COL11A1 in drug-resistance holds significant clinical implications for breast cancer treatment.

Collagen Type III Alpha 1 Chain (COL3A1)
COL3A1, encoding the pro-alpha 1 chains of type III collagen, stands as a homotrimeric fibril-forming collagen within connective tissues. COL3A1 is aberrantly overexpressed in various cancers, including bladder cancer, gastric cancer, and breast cancer, and consistently associated with worse prognosis, advanced tumor stage, local recurrence, invasion, tumor-infiltrating immune cells (TIICs) recruitment, ECM-receptor interaction, regulation of actin cytoskeleton, and adhesion pathways. Previous research has linked COL3A1 to the malignant potential of breast cancer, showing elevated expression in TNBC tissues and cells. It has also been identified as a key gene associated with brain metastases in breast cancer.

In addition to being significantly correlated with overall survival of breast cancer patients, COL1A1 is associated with breast cancer distant metastasis and death after surgery and systemic treatment. Silencing of COL3A1 can inhibit the proliferation, invasion, migration, and immune escape of breast cancer cells.

Nevertheless, to date, the specific regulatory mechanisms of COL3A1 remain elusive. It has been reported to exert its effects through the mitogen-activated protein kinase (MAPK) signaling pathway. COL3A1 also shows significant correlation with prolyl-4-hydroxylase α subunit 2 (P4HA2) during breast cancer development and progression. Moreover, it is identified as a potential target gene of methyltransferase-like 3 (METTL3), displaying a negative correlation. Downregulation of METTL3 contributes to the mobility of triple-negative breast cancer cells through m6A methylation-mediated COL3A1 up-regulation. Knocking down COL3A1 leads to reduced protein expression of PD-L1, suggesting a reverse correlation between PD-L1 upregulation and the inhibitory effects mediated by COL3A1 knockdown.

### Table 2: Expression, Targeted Molecules and Functions of COL11A1 in Breast Cancer

<table>
<thead>
<tr>
<th>Targeted Molecule</th>
<th>Role of COL11A1</th>
<th>Mechanism</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Immune infiltration</td>
<td>Is abnormally upregulated in breast cancer and affects the prognosis of breast cancer patients.</td>
<td>Participates in the regulation of tumor immune infiltration.</td>
<td>[67,80]</td>
</tr>
<tr>
<td>MiR-139-5p</td>
<td>Is highly expressed in breast cancer and promotes the cell proliferation.</td>
<td>MiR-139-5p overexpression inhibits the proliferation and promotes the apoptosis of breast cancer cells by inhibiting the expression of COL11A1.</td>
<td>[71]</td>
</tr>
<tr>
<td>microRNA Let-7</td>
<td>Its over-expression affects cell proliferation, invasion, migration, and metastasis.</td>
<td>CDX2 promotes let-7b expression, which may exert the inhibitory effect via repressing the expression of COL11A1.</td>
<td>[73]</td>
</tr>
<tr>
<td>miR-4458, DDR2/SRC signaling pathway</td>
<td>A directly interacting protein of miR-4458.</td>
<td>Transfection with miR-4458 mimic and COL11A1 knockdown inhibit MCF7 cell migration and invasion, inactivating the DDR2/SRC signaling pathway</td>
<td>[72]</td>
</tr>
<tr>
<td>Oncostatin M (OSM)</td>
<td>As the ECM protein.</td>
<td>Binds to OSM to activate Signal Transducer and Activator of Transcription (STAT) signaling.</td>
<td>[81]</td>
</tr>
</tbody>
</table>
Collagen Type XVII Alpha 1 Chain (COL17A1)

COL17A1, a type-II single-transmembrane protein predominantly expressed in epidermal basal keratinocytes, serves as a structural component of the dermoepidermal junction and plays important roles in the anchoring of hemidesmosomes at the dermal-epidermal junction. Although extensively studied in skin diseases, especially junctional epidermolysis bullosa, COL17A1 has been identified as a regulator of differentiation, cell migration, skin inflammation, and cancer development. It promotes the formation of multilayered, transformed epithelia and is a crucial regulator for the clonal expansion of transformed cells within multilayered epithelia, thus being potential target for early diagnosis and preventive treatment for precancerous lesions. COL17A1 can orchestrate the stem cell-centric aging program of the epithelial mini-organ. It is also essential for the collective migration of keratinocytes and for reepithelialization by enhancing keratinocyte stem cell motility. Generally, its mutation is prevalent in junctional epidermolysis bullosa and amelogenesis imperfecta. Beyond that, its dysregulation appears to occur in numerous cancers, like squamous cell carcinoma, melanoma, pancreatic carcinoma, thyroid cancer, colorectal cancer, lung cancer, nasopharyngeal cancer, cervical cancer, salivary gland cancer and breast cancer. Expression of COL17A1 in tumor tissue is closely related to patient survival, but the positive or negative correlation depends on the cancer type.

In breast cancer, research on COL17A1 is limited but significant. Unlike some collagens, it is reported to be under-expressed in ductal breast cancers, correlating with higher TNM staging, increased invasion, and postmenopausal status. COL17A1 overexpression demonstrates an anti-proliferative effect on breast cancer cells through mTOR deactivation. Its effects on the AKT/mTOR signaling pathway involve deactivation of AKT, mTOR, and downstream effectors 4EBP1. Meanwhile, COL17A1 is considered a favorable prognostic marker and a novel p53 transcriptional target. Its promoter is found hypermethylated in breast cancer, highlighting its potential as a target for therapeutic intervention. Despite its importance, research on COL17A1 in breast cancer remains limited, emphasizing the need for further exploration.

Type VI Collagen Family

Analyzing the TPM data indicates that expression of COL6A6 in the normal group is more than twice the expression in tumor group, showing a significant difference between these two groups. COL6A6, an extracellular matrix protein belonging to the collagen type VI (COL6) family, is evolutionarily conserved and expressed in various human tissues, including lung, kidney, liver, spleen, thymus, heart and skeletal muscle. Recent studies outside breast cancer have revealed its inhibitory effects on the growth and metastasis of pituitary adenoma, non-small cell lung cancer. Its expression is inversely associated with pathological stage, tumor stage, and lymph node metastasis. Furthermore, COL6A6 causes retinitis pigmentosa in patients with autosomal dominant transmission. Research have suggested that COL6A6 inhibits the progression of cancer cells through the JAK signaling pathway and PI3K-Akt pathway and is positively associated with the infiltration of B cells, T cells, neutrophils and dendritic cells.

In fact, relatively little research has been conducted on COL6A6 in breast cancer, although differential expression analysis through the GEPIA2 database indicates significant differences between breast cancer tissue and normal tissue. It is screened as the differentially expressed gene between breast cancer and non-tumor tissues in the unique mRNA fingerprint of breast cancer in Lebanese women. Studies focusing on axillary lymph node metastasis in triple-negative breast cancer have identified COL6A6 and other genes regulating cell microenvironment interaction as down-regulated genes. Additionally, its high expression level is also found significantly correlated with an early pathological stage of breast cancer.

Studies on other members of the COL6 family in breast cancer are also limited, despite their essential roles in tumor initiation and progression. Collagen type VI alpha 1 (COL6A1) has been implicated in Fzd7-Wnt5b signaling, and mediates Fzd7-Wnt5b-induced mesenchymal-like stemness in breast cancer cell lines and tissues. It is correlated with breast cancer brain metastasis in HER2 expression. COL6A1 and other types of collagens, collagen type VI alpha 2 (COL6A2) and collagen type VI alpha 3 (COL6A3) are identified as key genes associated with the overall survival of breast cancer patients. COL6A2, a fat-related collagen, is involved in Liver kinase B1 (LKB1) regulation of the tumor microenvironment through fibril matrix remodeling and suppression of adipogenesis in MDA-MB-231 COL6A3, widely present in most connective tissues, is considered a predictive marker of poor prognosis. It is included in a subset...
of 6 EMT genes to predict triple-negative breast cancer metastasis\textsuperscript{119} and also associated with breast cancer brain metastasis in HER2 expression. Collagen type VI alpha 5 (COL6A5), with restricted mRNA expression to a few tissues (including lung, testis, and colon), is strongly correlated with worse overall survival in human breast cancer patients.\textsuperscript{120} In cancer-associated stroma, transcriptomic upregulation of COL6A5 is validated.\textsuperscript{120,121} While the roles of COL6A6 and other COL6 family members in breast cancer are still emerging, their varied functions and associations with cancer progression warrant further investigation.

**Type IV Collagen Family**

The type IV collagen family, a major component of the basement membrane, is ubiquitously present and encoded by six distinct genes from α1 (IV) to α6 (IV), which are located on three different chromosomes. These genes exhibit different expression patterns, with collagen type IV alpha 1 (COL4A1) and COL4A2 being ubiquitous, while collagen type IV alpha 5 (COL4A5) and COL4A6 are specific to the basement membrane of mammary duct and lobule, epidermis, prostate gland, smooth muscle cells, and epithelium of the alimentary tract.\textsuperscript{122} The other isoforms also exhibit restricted distribution. Type IV collagen plays important roles in cell adhesion, migration, differentiation, growth and cancer progression.

In invasive ductal carcinoma, highly expressed COL4A1 influences proliferation and colony formation, with its knockdown resulting in reduced cell viability and cell cycle arrest.\textsuperscript{123,124} Its low expression in the tumor cells of breast cancer patients is found to significantly reduce the overall survival and relapse-free survival rates of neoadjuvant chemotherapy patients.\textsuperscript{125} COL4A2 is overexpressed in TNBC cells, and suppression of this gene can lead to significant reduction in cell proliferation and migration level.\textsuperscript{12} COL4A2 can be targeted by miR-29b and its degradation promotes invasion in MCF7 cells.\textsuperscript{126} COL4A3 is identified as a key angiogenesis-related gene in breast cancer, with surgery inducing its upregulation in patient.\textsuperscript{127,128} Deregulation of collagen type IV alpha 4 (COL4A4) is also validated by real-time quantitative PCR in breast tumor-derived endothelial cells.\textsuperscript{129} COL4A5, preferentially expressed in luminal-type breast cancer and regulated by estrogen receptor-α, contributes to impaired cell growth and tumor development capability upon ablation.\textsuperscript{130} In premetastatic lungs and breast cancer-conditioned lung fibroblasts, primary breast tumors can stimulate increased COL4A5 expression.\textsuperscript{131} For COL4A6, its high expression is negatively correlated with survival and a risk factor for RFS of TNBC patients.\textsuperscript{132,133} Down-regulated COL4A6 is significantly associated with insensitive responses to paclitaxel-based therapy.\textsuperscript{134} Additionally, it is a target of TNBC-long non-coding RNAs in breast cancer.\textsuperscript{132}

In summary, the type IV collagen family plays crucial roles in breast cancer progression. Their diverse expression patterns and functions across subtypes highlight the complexity of their involvement in cancer development. The detailed nature and complexity of each collagen’s role in breast cancer warrant ongoing research for a more comprehensive understanding of their specific contributions to disease progression and potential therapeutic interventions.

**Discussion**

Over the past few decades, the molecular characterization of twenty-eight collagen types has provided invaluable insights into their diverse functions. Comprised of three homo- or hetero-trimeric polypeptide chains (α chains), collagens contribute significantly to the molecular architecture, shape, and mechanical properties of tissues, with fibrillar collagens assuming a structural role in tissue organization. In contrast, membrane collagens (such as XIII, XVII, XXV, and XXVIII) exhibit distinct functions, predominantly expressed in neurons, neuronal structures, or neuronal tissues, implicating them in various diseases and tumors.\textsuperscript{5} Their functions are continuously explored through different methods. Recent studies have underscored the pivotal role of collagens in regulating tumor cell growth, differentiation, migration, and their involvement in disorders resulting from mutations. Remarkably, collagens also exhibit a multifaceted involvement in pregnancy and assume a dual role in postpartum-associated breast cancer (PABC). They act as a physical barrier against tumor cells invading the ECM. In the context of postpartum-associated breast cancer, postpartum collagen demonstrates a protective effect, adopting an anti-proliferative TACS1 architecture, which has been postulated to contribute to the protective effect of pregnancy.\textsuperscript{15,135} However, their increased deposition during involution also promotes breast cancer development through the stiffening of the ECM producing tension within the microenvironment.\textsuperscript{136} Additionally, postpartum collagen with anti-proliferative properties can undergo conversion into
an involution-like pro-tumorigenic collagen under the influence of a collagen-dependent oncogene, protease pappalysin-1 (PAPP-A).\textsuperscript{137} The sporadic presence of PAPP-A is sufficient to convert post-partum collagen into a high TACS3
involution-like collagen architecture, which is specifically correlated with worst prognosis in breast cancer patients
and is a stronger predictor for metastasis than tumor stage.\textsuperscript{135,138}

Despite considerable strides in understanding collagen functions, particularly in breast cancer, numerous questions
persist, particularly regarding their roles and underlying mechanisms in tumor progression and drug resistance. This
review categorizes the 28 collagens into two fractions—those relatively highly expressed and those poorly expressed
in breast cancer tissues. Notably, COL11A1 emerges as a noteworthy collagen highly expressed in breast cancer, demon-
strating significant associations with prognosis and the aggressive behavior of invasive breast cancer. Its differential
expression across normal, preinvasive, and invasive tumors suggests distinct roles at various stages.\textsuperscript{139} Intriguingly,
COL11A1 may serve as a marker for radiotherapy resistance, with high stromal expression correlating with increased
recurrence rates despite radiotherapy. This collagen also mediates resistance to cisplatin and paclitaxel, underscoring its
potential clinical significance once the underlying mechanisms are elucidated. At present, therapeutic strategies targeting
COL11A1’s upstream or downstream molecules have been proposed; however, direct targeting remains an unexplored
avenue, emphasizing the need for tailored therapies. In addition, as mentioned earlier, COL11A1 overexpression is
associated with poor prognosis in solid tumors, but with better prognosis in some hematologic malignancies. The
differences in mechanisms between them warrant further investigation.

Another gene significantly upregulated in breast cancer is COL10A1, which plays a crucial role in prognosis and
treatment. Its potential diagnostic utility is underscored by its ability to distinguish between breast cancer patients and
those with benign diseases, highlighting its role as a potential diagnostic predictor. Despite its substantial expression
difference in breast cancer, research on COL10A1 lags behind that of some other collagens, such as COL1A1. Further
investigations are warranted to deepen our understanding of this collagen in breast cancer. Additionally, the expression of
COL22A1 is significantly different between tumor samples and the paired normal tissue samples, but its expression in
both groups is relatively low compared to some other collagens. Meanwhile, numerous collagens which exhibited
relatively low expression in breast cancer also remain understudied. For instance, COL25A1, identified as a collagen-
like Alzheimer’s amyloid plaque component precursor, is exclusively found in breast tumors but not in normal tissues.\textsuperscript{140}

Comprehensive studies regarding its function and mechanisms in breast cancer are lacking. Similarly, collagens like
COL6A6 and COL4A3, with low expression in breast tumors, are differentially expressed genes between breast cancer
and non-tumor tissues, underscoring the need for further exploration.

The clinical significance and application prospects stand as pivotal facets in collagen research. Our findings
accentuate the pivotal role of collagens in diagnosis, prognosis, and treatment. Notably, COL10A1 and COL11A1
have emerged as potential diagnostic and therapeutic markers for breast cancer. COL6A3 has been linked to the
mediation of chemoresistance and cisplatin resistance.\textsuperscript{141,142} Beyond these, collagens, including COL1A1, COL3A1,
COL10A1, and COL11A1, have been implicated in prognosis, with their elevated expression correlating with poorer
outcomes. Conversely, for COL4A3 and COL17A1, their higher expression levels correlate positively with a favorable
prognosis in breast cancer.\textsuperscript{107,128} Integrating the expression patterns of these collagens into the prognostic assessment of
breast cancer patients can furnish clinicians with more precise prognostic information, facilitating the formulation of
tailored treatment strategies.

These observations in this review underscore the considerable predictive value of collagens in monitoring cancer
processes and recurrence. Future research avenues should employ innovative approaches to unravel the functions and
mechanisms of collagens in breast cancer. Additionally, considering the upstream regulators, downstream effectors,
receptors, and tumor microenvironment of these collagens is crucial, as these factors may play pivotal roles in mitigating
the pro-tumor effects of collagens. Currently, advancements in imaging techniques, such as multiphoton microscopy
and second harmonic generation imaging, have facilitated the detailed visualization and quantification of collagen
alterations in breast cancer, thus augmenting our comprehension of the tumor microenvironment. A comprehensive
understanding of collagen dynamics in breast cancer holds significant promise for enhancing the diagnosis, prognosis,
and treatment modalities of the cancer.
Author Contributions
All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding
This work was supported by the Yue Jin’s Postdoctoral Excellence Program of Jiangsu Province, China (2022ZB895).

Disclosure
The authors report no conflicts of interest in this work.

References


46. Huang K, Zhang X, Duan J, et al. STAT4 and COL1A2 are potential diagnostic biomarkers and therapeutic targets for heart failure comorbid with depression. Brain Res Bull. 2022;184:68–75. doi:10.1016/j.brainresbull.2022.03.014


