https://doi.org/10.15407/exp-oncology.2023.01.028

Shreshtha Malvia [1], Chintamani Chintamani [2], Ramesh Sarin [3], Uma S Dubey [4], Sunita Saxena [5], Sarangadhara Appala Raju Bagadi [1,*]

1 Tumor Biology Division, ICMR-National Institute of Pathology, New Delhi, 110029, India
[2] Department of Surgery, Safdarjung Hospital, New Delhi, 110029, India
[3] Department of Surgery, Indraprastha Apollo Hospital, New Delhi, 110076, India
[4] Department of Biological Sciences, Birla Institute of Technology and Sciences, Pilani, Rajasthan, 333031,
[5] Consultant, Department of Health Research, New Delhi, 110001 & Ex-Director National Institute of Pathology-ICMR Safdarjang Hospital Campus
* Correspondence: E-mail: bsaraju@gmail.com, sappalarajubagadi.nip@gov.in

ABERRANT EXPRESSION OF *COL14A1*, *CELRS3*, and *CTHRC1* IN BREAST CANCER CELLS

Background. Collagens are the major component of the ECM and are involved in regulation of tumor microenvironment in breast and several other cancers. **Aim:** To analyze the transcript level expression of *COL1A1, COL5A1, COL10A1, COL11A1, COL12A1, COL14A1, CTHRC1,* and *CELRS3* genes and the clinical relevance of their differential expression in breast cancer (BC). **Materials and Methods.** The transcript level expression of the genes was analyzed using the quantitative real-time PCR (qPCR) in tumor tissue of 60 BC patients. **Results.** Overexpression of *COL1A1, COL5A1, COL10A1, COL11A1, COL12A1, CTHRC*, and *CELRS3* and down-regulated expression of *COL14A1* were observed. *COL14A1* down-regulation was associated with aggressive, basal, and Her-2/neu BC subtypes ($p = 0.031$). Overexpression of *CELSR3* was found to be associated with the older age of the patients (> 55 years, $p = 0.049$). Further analysis with the TCGA BC data set has shown a concordance in the differential expression of the above genes. Furthermore, overexpression of *CTHRC1* was associated with poor overall survival (OS), particularly with poor prognosis ($p = 0.00042$) for the luminal BC subtype. On the other hand, *CELSR3* overexpression was associated with mucinous tumors and also with poor prognosis in post-menopausal women. *In silico* target prediction identified several BC-associated miRNAs and members of miR-154, -515, and -10 families to have a likely regulatory role in the above ECM genes. **Conclusion.** The present study shows that the expression of *COL14A1* and *CTHRC1* may serve as potential biological markers for the detection of basal BC and prognostic marker in luminal BC.

Keywords: breast cancer, extracellular matrix, overexpression, down-regulation, collagens.

Citation: Malvia S, Chintamani C, Sarin R, Dubey US, Saxena S, Raju Bagadi SA. Aberrant Expression of COL14A1, CELRS3 and CTHRC1 in Breast Cancer. *Exp Oncol.* 2023; 45(1): 28-43. https://doi.org/10.15407/exp-oncology.2023.01.028

Breast cancer (BC) is the most common cancer among women with 2.1 million new cancer cases being diagnosed in 2018 worldwide [1]. Although there is the continuous implementation of targeted therapies and treatment strategies for BC, patients with metastatic BC have around 27% 5-year overall survival (OS) rate as compared to non-metastatic BC, which has around 85% 5-year OS rate.

The extracellular matrix (ECM) plays a crucial role in BC development and progression [2]. ECM components such as proteoglycan, structural proteins, and glycoprotein undergo extensive reorganization, thus disrupting the fine balance of ECM signals and leading to cancer progression. The most abundant proteins in ECM are collagens that promote flexibility of the connective tissues and their increased deposition and crosslinking leads to stiffness in cancer tissues, which further promote metastasis [2]. The collagens are known to be aberrantly expressed in BC and are associated with poor prognosis [2—4]. High levels of collagen crosslinking are associated with the development of metastatic lesions. ECM stiffness promotes malignancy and enhances tumor aggression in BC patients [5]. Aggressive BC subtypes, i.e., basal and HER2/neu possess stiffer ECM than that of luminal subtypes A and B [5].

Earlier we demonstrated differential expression of collagen genes such as *COL1A1, COL5A1, COL10A1, COL11A1, COL12A1, COL14A1, CTHRC1*, and *CELSR3* in BC using microarray [4]. In the present study, we validated the differential expression of collagen genes using quantitative real-time PCR (qPCR) to determine the differential expression and its clinical relevance in Indian BC patients to determine the utility of their expression as biomarkers for BC.

Materials and Methods

Patients and tissue samples. In the present study, 60 patients with histologically confirmed BC admitted to Safdarjung Hospital or Indraprastha Apollo Hospital, New Delhi, India, during 2008—2012 were enrolled (Table 1). The institutional ethical committees of Safdarjung Hospital and Indraprastha Apollo Hospital approved the present study. Informed consent was taken from all the patients, and standard guidelines and regulations were followed in performing the experiments. The age of patients who were enrolled in the study ranged between 25—80 years, 30 patients of which were premenopausal and 30 patients were postmenopausal women (Table 1). Patients who received neoadjuvant therapy were excluded from the present study to avoid the effect of chemotherapeutic agents on the expression of genes being analyzed. The staging of BC was done according to the American Joint Committee on Cancer (AJCC) guidelines. Immunohistochemistry (IHC) was used to determine the expression of estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2/neu) as described elsewhere [6]. Stratification of the tumor samples was done in subtypes, namely the luminal subtype comprising ER and/or PR-positive and HER2/neu negative or ER-positive with any PR and HER2/neu positive samples, the basal subtype comprising ER, PR, and HER2/neu negative samples, and the HER2/neu overexpressing tumors subtype comprising ER, PR negative and HER2/neu positive samples. All tissue samples were snap-frozen in liquid nitrogen immediately after the modified radical mastectomy or after incision/trucut biopsy for RNA isolation and stored in RNALater solution (Ambion, USA) at –80 °C.

Total RNA extraction. Total RNA was routinely extracted using TRIzol reagent following the manufacturer's protocol (Life Technologies, USA). The quantity and quality of the RNA obtained were determined using Nanodrop (Thermo Fischer Scientific, USA) and Agilent 2100 Bioanalyzer (Agilent Technologies, USA), respectively.

TCGA data analysis on GEPIA and UALCAN. The web-based tools, GEPIA [7] (Gene Expres-

sion Profiling Interactive Analysis) consisting of 1085 breast tumors, and 291 controls and UALCAN [8] consisting of 1097 breast tumors and 114 controls were used to determine differential expression of all the genes from TCGA (The Cancer Genomic Atlas) database. Their association with clinicopathological features was also determined using the above tools.

Table 1. **Clinicopathological characteristics of BC patients under study**

AGE	Number of cases, %
Menopausal status	
Premenopausal	30 (50)
Postmenopausal	30 (50)
Tumor stage	
Stage I	2 (3.3)
Stage II	24 (40)
Stage III	30 (50)
Stage IV	3 (5)
Undetermined	1 (1.6)
IHC subtypes	
Luminal	21 (35)
Basal	13 (21.6)
HER2/neu overexpressing	23 (38.3)
Undetermined	3 (6.6)
Estrogen receptor	
Positive	21 (35)
Negative	36 (60)
Undetermined	3 (5)
Progesterone receptor	
Positive	21 (35)
Negative	36 (60)
Undetermined	3 (5)
HER2/neu overexpressing	
Positive	32 (53.3)
Negative	25 (41.6)
Undetermined	3 (5)
Chemotherapy/ radiotherapy received	
Yes	0
No	60 (100)

Real-time PCR. qPCR was done for *COL1A1, COL5A1, COL10A1, COL11A1, COL12A1, COL14A1, CELSR3, CTHRC1*, and *18S* gene, which was used as an endogenous control for normalization of the qPCR data (Table 2). The differential expression of these genes was determined in 60 BC samples. RNA obtained from 38 distant normal breast tissues was pooled and used as a control; two human mammary total RNAs (obtained from Ambion) were also used as controls for determination of transcript expression in BC (Supplementary Table S1) [1].

The cDNA was synthesized following the manufacturer's protocol, briefly, 1 μg of total RNA from tumors and controls, 250 ng of random hexamers, and 200 U SuperScript III RT (Invitrogen, Thermo Fisher Scientific, USA) in a total reaction volume of 20 μL, at 50 °C for 50 min. The qPCR was performed using 1 μL of cDNA along with 5.0 μL of SYBR green mix (2X) and 0.25 pm of gene-specific primers in a total volume of 10 μL. To ensure lack of any nonspecific amplification due to genomic DNA contamination, a non-RT control was also used during qPCRs. All the samples were run in triplicate in a 96 well plate on a Step One Real-Time PCR machine (ABI, Foster City, USA) with initial denaturation at 95 °C for 2 min, 40 cycles at 95° C for 10 s, and 60 °C for 1 min. The specificity for all the primers was confirmed by the melting curve analysis using StepOne software v.2.3 (ABI, Foster City, USA). The mean Ct obtained for each gene was normalized with endogenous controls to obtain ΔCt, and further fold change (FC) was obtained by the ΔΔCt method on Data assist software v.3.01 (ABI, Foster City, USA).

Statistical analysis. The Mann — Whitney U and one-way Kruskal — Wallis tests were used along with Bonferroni correction for pairwise post-hoc analysis to determine the significance of differentially expressed genes and their asso-

[1] Supplementary materials are posted at http://dx.doi.org/10.13140/RG.2.2.25153.43364

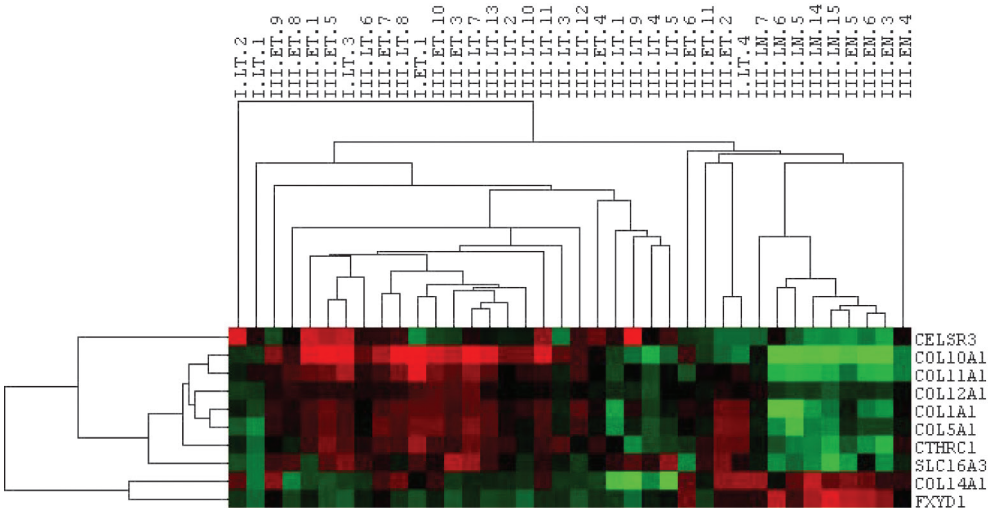

Fig. 1. ECM genes showing differential expression in BC cells. Red represents up-regulated genes and green shows down-regulated genes (by microarray analysis)

ciation with various clinicopathological parameters using SPSS. All the graphs were plotted using Graph Pad Prism software. The Kaplan — Meier survival analysis was used to determine the overall survival of breast cancer patients ($p = 0.005$) for patients from TCGA database. The median of expression in tumors and controls was used for statistical analysis for all the genes.

MicroRNA prediction. Prediction of miRNAs was done using miRWalk [9], and a list of validated miRNAs was obtained from miRCancer database, from where miRNAs that are reported to be either up-regulated or down-regulated in BC were obtained. The up-regulated or down-regulated miRNAs, which targeted 3'-UTR or CDS of the collagens or anion transport genes, were further filtered using binding energy ≤–18 kcal/mol. A total of 341 miRs that were experimentally verified to be either up-regulated or down-regulated were downloaded from the miRCancer database (Supplementary Table S2). The miRNAs which might potentially target the ECM genes that were selected for the present study were predicted using the miRWalk software. For up-regulated genes (*COL1A1, COL5A1, COL10A1, COL11A1, COL12A1*, and *CELSR3*), down-regulated miRNAs were

Table 2. **List of primers for all the genes used for qPCRs**

Gene	Primers
COL1A1_F	GCTTGGTCCACTTGCTTGAA
COL1A1_R	GTGTGGAGAAAGGAGCAGAAAG
COL5A1_F	GTGCCTTCCCGATGGATTAAAG
COL5A1_R	TGGGTTTCCGTGCATGATTG
COL10A1_F	TATGACCCAAGGACTGGAATC
COL10A1_R	CCAAACATGAGTCCCTTTCAC
COL11A1_F	CTGGTCATCCTGGGAAAGAA
COL11A1_R	GACACCATCTGCTCCCTTTA
COL12A1_F	GGAGACCACCGAAGAAGTAAA
COL12A1_R	CTGGAGCTTGAGTCCAAGTAA
COL14A1_F	GAGGAGGTTGTCCTGAAAGA
COL14A1_R	CATCATCAAGTACGGCCAATAG
CTHRC1_F	GAAGGAATTGGTGCTGGATTAG
CTHRC1_R	GCGAGAAACTGAATTCCATCC
CELSR3_F	CAGTCAGTATCCAGGTGATGGT
CELSR3_R	CACTGAGCCCACAATGCTATTC
18S_F	GGATCCATTGGAGGGCAAGT
18S_R	CCCAAGATCCAACTACGAGCTT

used; for down-regulated genes (*COL14A1* and *FXYD1*), up-regulated miRNAs from the list were used, and for those miRNA where up- and down-regulation were shown by different studies, they were included in both groups for target prediction.

Results

A distinct cluster of ECM genes was obtained upon unsupervised hierarchical clustering of the above genes using the microarray-based gene expression profiling of BC earlier in our laboratory (Fig. 1, Supplementary Table S3). In the present study, qPCR analysis of the genes *COL1A1, COL5A1, COL10A1, COL11A1, COL12A1, CTHRC1*, and *CELSR3* was done in 60 BC samples.

Quantitative real-time analysis of ECM gene expression in BC patients. qPCR analysis transcript expression of *COL1A1, COL5A1, COL10A1, COL11A1, COL12A1, COL14A1, CTHRC1*, and *CELSR3* genes showed that most of the collagens analyzed in the present study were overexpressed, except for *COL14A1*, which was found to be down-regulated in the tumors. *COL1A1* was overexpressed in most of the tumors with a mean fold change of 102.2 ± ± 20.35 ($p = 0.0048$) (Table 3). Similarly, *COL5A1* showed overexpression with a mean fold change of 13.21 ± 3.4, *COL10A1* with a mean fold change of 39.7 ± 14.39, *COL11A1* with a mean fold change of 2159 ± 485.6 ($p = 0.0039$), *COL12A1* with a mean fold change of 107.37 ± 22.8 ($p = 0.006$). On the other hand, *COL14A1* was found to be down-regulated in most of the tumors with a mean fold change of –2.2 ± 1.4. *CTHRC1*, which is a collagen triple helix repeat containing molecule, known to regulate the expression of collagen 1 and 3, was found to be overexpressed with a mean fold change of 102.4 ± 15.95 ($p = 0.0039$). *CELSR3*, a Seven-Pass G-Type Receptor, was also found to be overexpressed in the tumors with a mean fold change of 6.5 ± 1.5 (Table 3).

Analysis of differential expression of ECM genes in BC from TCGA database. In order to study the clinical relevance of the expression of the above genes in a larger number of patients, we analyzed their expression using the TCGA dataset with the UALCAN web portal in 1097 BC patients and compared with 114 normal samples. Overexpression of the *COL1A1, COL5A1, COL10A1, COL11A1, COL12A1, CTHRC1*, and *CELSR3* genes and underexpression of the *COL14A1* gene were observed in BC from TCGA dataset (Fig. 2), which is in concordance with the present study.

Table 3. **Expression of ECM genes in BC patients**

Gene	Mean FC ± SEM by qPCR	*p* (qPCR)	FC_Total BC by microarray	Adjusted *p*-value (microarray)
COL1A1	102.2 ± 20.35	0.0048	3.7654	7.15E-05
COL5A1	13.21 ± 3.4	0.278	2.729898	0.002791
COL10A1	39.7 ± 14.39	0.068	23.2815	5.96E-06
COL11A1	2159 ± 485.6	0.0039	5.122635	1.18E-06
COL12A1	107.37 ± 22.8	0.006	1.567429	0.013473
COL14A1	–2.2 ± 1.4	0.378	–2.80683	0.02485
CTHRC1	102.4 ± 15.95	0.0039	2.91235	0.000735
CELSR3	6.5 ± 1.5	0.21	2.813078	0.022869

Note: FC — fold change.

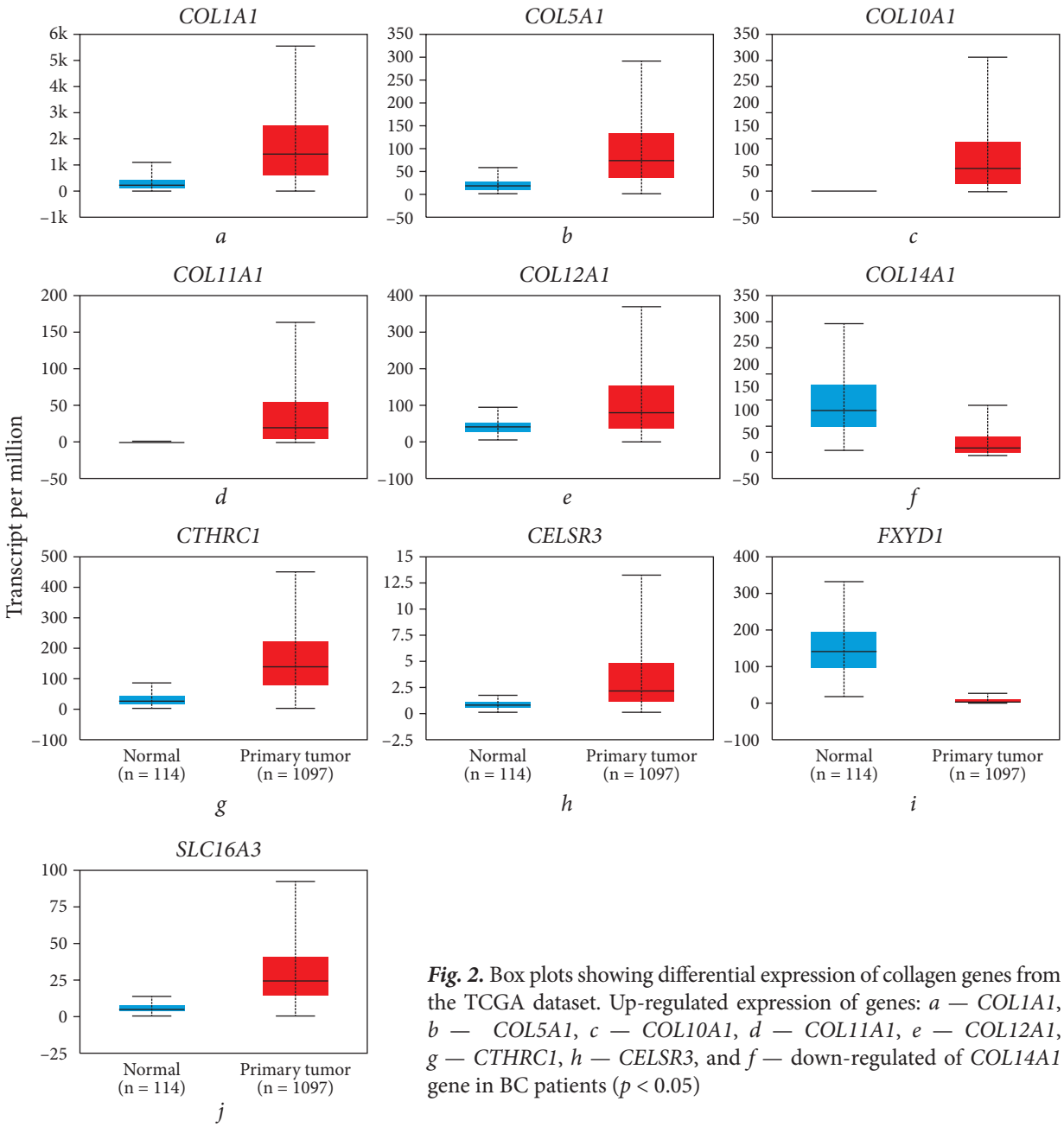

Fig. 2. Box plots showing differential expression of collagen genes from the TCGA dataset. Up-regulated expression of genes: *a* — *COL1A1*, *b* — *COL5A1*, *c* — *COL10A1*, *d* — *COL11A1*, *e* — *COL12A1*, *g* — *CTHRC1*, *h* — *CELSR3*, and *f* — down-regulated of *COL14A1* gene in BC patients ($p < 0.05$)

Clinicopathological association of differential expression of ECM genes in BC. Statistical analysis was done for the differential expression of the above genes and their association with various clinicopathological parameters, to identify if there is any clinical relevance of their differential expression (Supplementary Table S4). Overexpression of *COL12A1* was associated with hormone receptor (ER, PR) positivity ($p = 0.05$) in BC. Analysis of the relationship between the differential expression of these genes and the BC subtypes (luminal, basal, and HER-2) has shown that down-regulated expression of *COL14A1* ($p = 0.031$) was associated

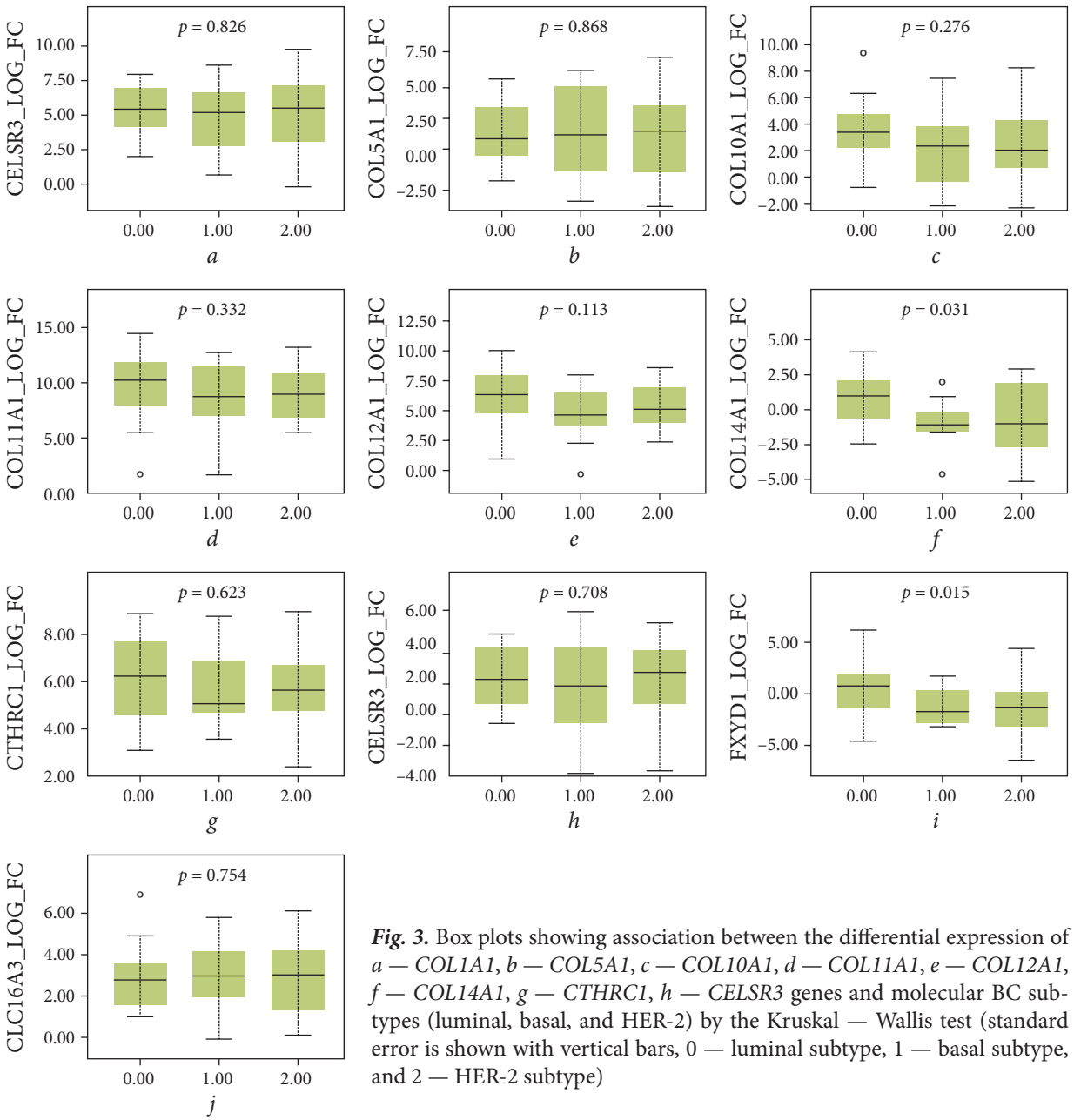

Fig. 3. Box plots showing association between the differential expression of *a* — *COL1A1*, *b* — *COL5A1*, *c* — *COL10A1*, *d* — *COL11A1*, *e* — *COL12A1*, *f* — *COL14A1*, *g* — *CTHRC1*, *h* — *CELSR3* genes and molecular BC subtypes (luminal, basal, and HER-2) by the Kruskal — Wallis test (standard error is shown with vertical bars, 0 — luminal subtype, 1 — basal subtype, and 2 — HER-2 subtype)

with the basal subtype, and HER-2 subtypes were compared to the luminal subtype, according to the Kruskal — Wallis one-way analysis and post-hoc analysis (Fig. 3, Supplementary Table S4). Further, overexpression of *CELSR3* was found to be associated with the older age group (>55 years of age, $p = 0.049$), compared to early onset tumors (Table 4).

Clinicopathological association of differential expression of ECM genes in BC from TCGA database. We also analyzed the association of the differential expression of *COL1A1*, *COL5A1*, *COL10A1*, *COL11A1*, *COL12A1*, *COL14A1*, *CTHRC1*, and *CELSR3* with various clinicopathological features to find clinical relevance of their differential expression using UALCAN.

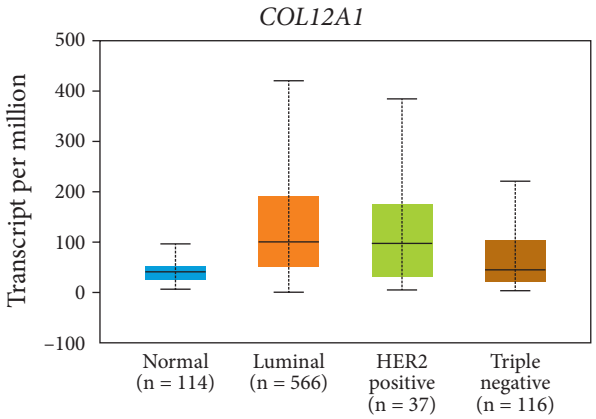

Fig. 4. Box plots showing differential expression of *COL12A1* in different BC subtypes in the TCGA dataset. *COL12A1* was found to be significantly overexpressed in the luminal BC subtype as compared to triple-negative subtype, $p = 1.6e^{-12}$

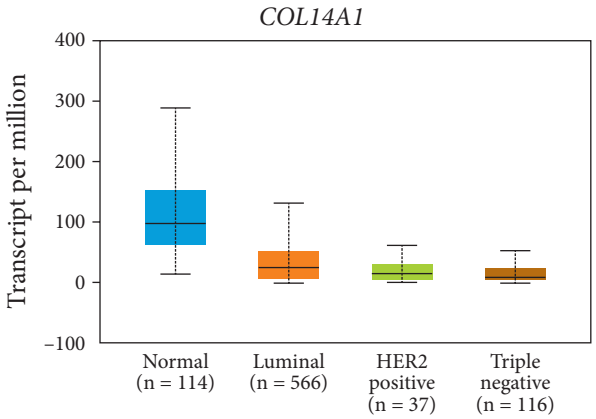

Fig. 5. Box plots showing differential expression of *COL14A1* in all BC subtypes in TCGA dataset. *COL14A1* was found to be significantly underexpressed in triple-negative BC subtype as compared to luminal subtype, $p = 1.5e^{-11}$

Overexpression of *COL12A1* was found to be significantly higher in luminal BC (n = 566) (101 transcripts per million (TPM)) as compared to triple negative BC (n = 116), which showed 44 TPM ($p = 1.6e^{-12}$) (Fig. 4). Underexpression of COL14A1 was found to be associated with triple negative BCs which showed 7.7 TPM compared to 24.2 TPM ($p = 1.5e^{-11}$) in the luminal subtype (Fig. 5). The overexpression of *CELSR3* was found to be higher in an older group (61—80 years), 2.099 TPM, compared to the expression 1.69 TPM in a younger group (21—40 years) ($p = 0.038$) (Fig. 6, *a*). Overexpression of this gene was observed with increasing age of the patients, from the age of 21—40 years (n = 97) with 1.69 TPM, 41—60 years (n = 505) with 2.23 TPM, 61–80 years (n = 431) with 2.1 TPM up to 81–100 years (n = 54) with 2.1 TPM (Fig. 6, *a*). On the other hand, *CELSR3* overexpression was found to be associated with mucinous tumors (invasive ductal carcinoma *vs.* mucinous, $p = 0.001$, invasive lobular carcinoma *vs.* mucinous, $p = 0.000084$, and mucinous vs. metaplastic, $p = 0.0095$) (Fig. 6, *b*).

Survival analysis. We have analyzed the prognostic significance of the differential expression

Table 4. **Fold change (FC) of ECM genes and their association with various clinicopathological parameters**

Gene	Mean FC ± SEM			*p*
	ER+	ER–		
COL12A1 with ER/PR	183.94 ± 57.99	67.82 ± 13.88		0.05
	Luminal	Basal	HER2/neu	
COL14A1 with subtypes	2.12 ± 1.19	–3.85 ± 2.09	–7.14 ± 4.74	0.031
	Early age onset	Late age onset		
CELSR3 with age	2.90 ± 1.53	9.85 ± 2.44		0.049

Note: FC — fold change.

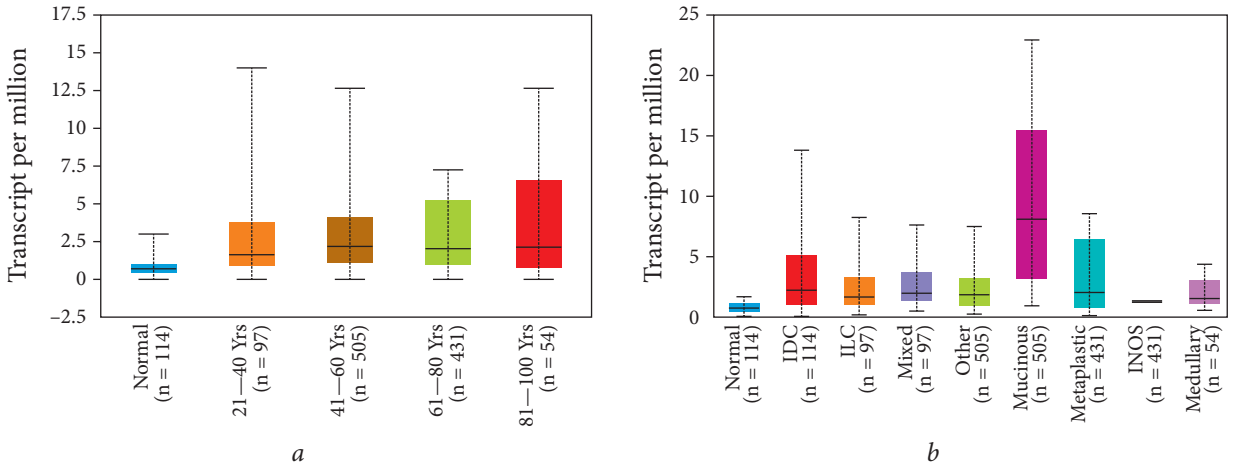

Fig. 6. Box plots showing the expression of *CELSR3* in the TCGA dataset: *a* — in different age groups of BC. *CELSR3* was found to be significantly overexpressed in the age group 61—80 years compared to the expression in the age group 21—40, $p = 0.038$; *b* — in different histological subtypes (IDC — invasive ductal carcinoma; ILC — invasive lobular carcinoma). *CELSR3* overexpression with mucinous tumors suggesting it be a potential biomarker for detection of mucinous tumors

of the above ECM genes selected in the present study in BC patients from TCGA dataset. The TCGA data has shown that BC patients with higher expression of *CTHRC1* had poor OS as compared to patients with lower *CTHRC1* expression ($p = 0.024$) among the genes analyzed in the present study (Fig. 7, *a*). Further, analysis for different subtypes showed that overexpression of *CTHRC1* is significantly associated with poor prognosis ($p = 0.00042$) for luminal subtype tumors (Fig. 7, *b*). Similarly, *CELSR3* overexpression was found to be associated with poor prognosis in post-menopausal women (Fig. 7, *c*).

Potential miRNAs targeting ECM genes. Upon analysis, 98 miRNAs were established to have one or more genes as their targets (Supplementary Table S5). *COL12A1* was predicted to be likely target for a maximum of 66 miRNAs, followed by 46 miRNA for *CELSR3*, 43 miRNA for *COL5A1*, 42 miRNA for *COL11A1*, 40 miRNA for *COL1A1*, 37 miRNA for *COL10A1* and 29 miRNA for *SLC16A3*, 27 miRNA for *COL14A1* and 12 miRNA for *FXYD1* (Supplementary Table S6). We further filtered the miRNAs based on the number of hits each miRNA has for each gene (Supplementary Table S7). It worth noting that hsa-miR-1301-3p, hsa-miR-409-3p, hsa-miR-485-3p, hsa-miR-485-5p, hsa-miR-502-5p, hsa-miR-520d-3p, and hsa-miR-601 were found to have a maximum of 6 genes as targets (Supplementary Table S5) suggesting a high probability of their involvement in the regulation of the predicted target genes.

Then we tried to identify the miRNAs families that are co-expressed and may be involved in the co-regulation of collagen genes in a synchronized manner. MiR-154 family consisting of hsa-mir-485, hsa-mir-409, hsa-mir-410, and hsa-mir-300 was found to have the highest number (6) of genes: *COL1A1, COL10A1, COL5A1, COL11A1, COL12A1*, and *CELSR3* as targets (Fig. 8, *a*). Interestingly, it was observed that the miR-515 family consisting of hsa-miR-520a-3p, hsa-miR-519d-3p, and hsa-miR-520d-3p targeted 6 genes (Fig. 8, *b*). Mir-10 family consisting of hsa-mir-99a, hsa-mir-125a, and hsa-mir-125b, was found to target 6 genes as well (Fig. 8, *c*), suggesting a probable role of the above miR families in the regulation of the collagen and anion transport genes in a coordinated manner.

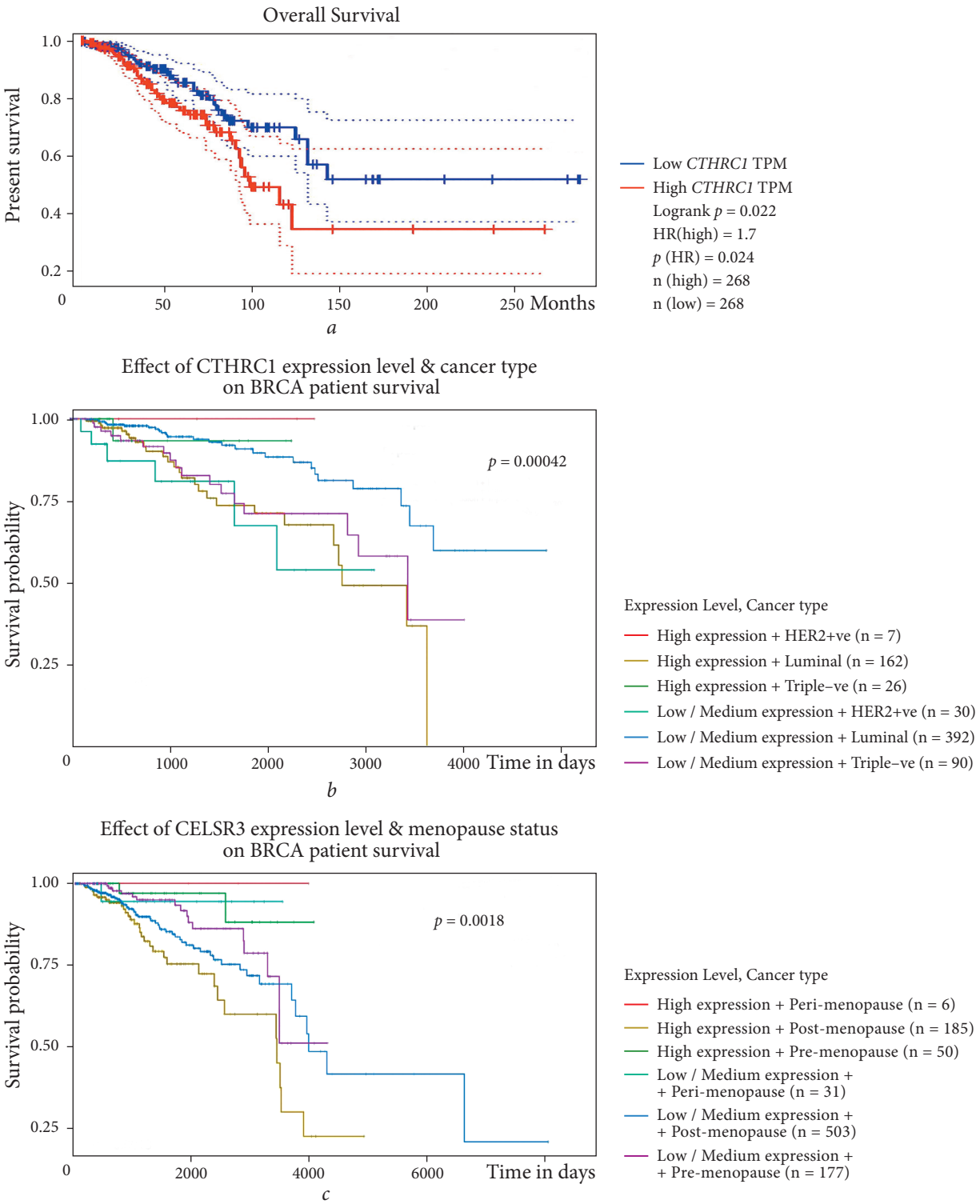

Fig. 7. Survival plots of BC patients: *a* — the worse survival of BC patients with overexpression of *CTHRC1* mRNA in TCGA dataset ($p = 0.024$); *b* — poor survival of luminal BC subtype patients with overexpression of *CTHRC1* mRNA; *c* — poor survival of post-menopausal women with overexpression of *CELSR3* mRNA

Discussion

In the present study, collagen genes *COL1A1*, *COL5A1*, and *COL11A1* were found to be overexpressed at the mRNA level using qPCR; the expression of these genes at mRNA level is reported for the first time in BC patients from India. Overexpression of *COL1A1*, *COL5A1*, and *COL11A1* has been reported to be associated with poor prognosis in several cancers including breast carcinoma [3, 10—14]. The overexpression of *COL1A1*, *COL5A1*, and *COL11A1* observed in the present study is in concordance with the literature: *COL12A1* was overexpressed significantly in BC, and also the overexpression of *COL12A1* protein was observed in BC cell lines, that is, *COL12A1* was overexpressed significantly in BC (by qPCR data), and the overexpression of the protein was also observed using dbDEPC and in BC cell lines [15]. The expression of *COL12A1* was found to be higher in tumors, it was also found to be overexpressed in BC patients by TCGA data. Its overexpression has shown an association with luminal BC, which has not been reported so far, suggesting it to be a potential biomarker for the luminal tumors if tested in the large number of patients. Its overexpression has been reported for other cancers such as gastric [16] and colorectal [17] cancers and has shown an association with poor survival for lung cancer patients [18].

On the other hand, in the present study, *COL14A1* was found to be underexpressed. *In silico* analysis of this gene in the TCGA has also shown its down-regulation in BC, supporting our observation. So far, there has been no report on the expression of this gene in BC at the mRNA level and hence this study reports its down-regulated expression in BC patients and association with basal tumors. Its expression has been reported to be up-regulated in cervical cancer [19] at mRNA level, however, for esophageal [20] soft tissue sarcoma [21] and renal carcinoma [22] promoter methylation of *COL14A1*, down-regulated expression of its mRNA has been reported. Goto et al. [23] have reported its up-regulation in metastatic BC tissues using LC/MS. There may be tissue specific and stage specific expression of this gene in different cancers contributing to overexpression in some cancers and down-regulation in others; in metastatic BC, the expression of this gene needs to be verified at both transcript and protein levels in metastatic tumors to clarify such differences. Analysis using the TCGA database consisting of 1085 BC patients has shown the association

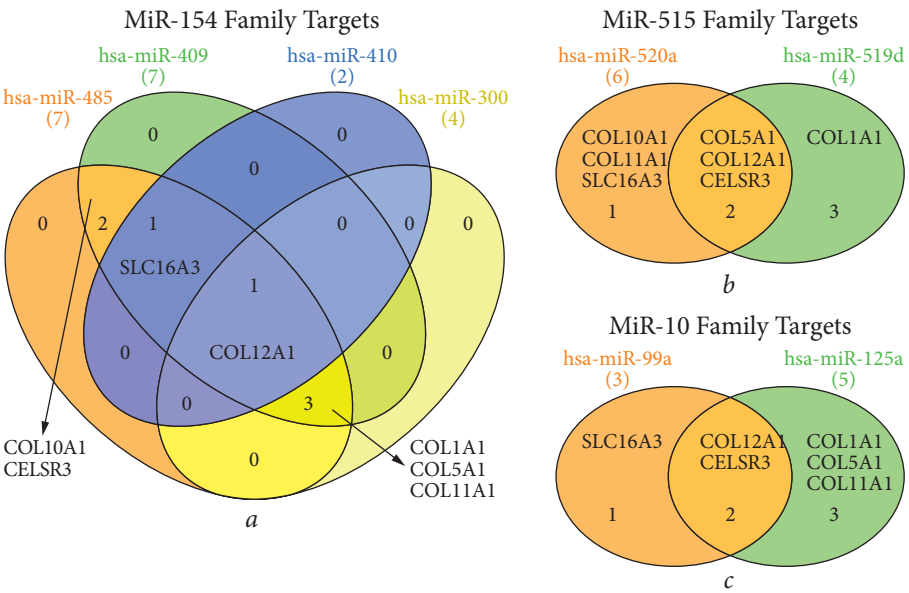

Fig. 8. Venn diagrams: *a* — miR 154 family miRNAs, hsa-miR-300 hsa-miR-409 hsa-miR-410 and hsa-miR-485 and their target genes; *b* — miR family 515 miRNAs, hsa-miR-519, and hsa-miR-520a and their targets genes; *c* — miR family 10, hsa-miR-125a, and hsa-miR-99a and their target genes

with down-regulated expression of *COL14A1* in triple-negative BC, which is aggressive in nature suggesting its potential utility as a biomarker.

CTHRC1 transcripts were found to be overexpressed in BC using both qPCR, and further analysis in the TCGA database confirmed its overexpression associated with the poor OS of BC patients, suggesting it to be a potential biomarker for prognosis. Luminal tumors are often susceptible to chemotherapy, however a subset of them is recurrent, which is often difficult to identify compared to basal tumors. *CTHRC1* mRNA expression in BC patients may serve as a prognostic factor to identify patients with a poor prognosis. *CTHRC1* has been reported to be highly overexpressed in several cancers such as lung, gastric, hepatocellular, colorectal, esophageal, ovary, breast, and thyroid cancers [24, 25]. Collagen triple helix repeat containing I *(CTHRC1)* inhibits the synthesis of collagen I and is involved in repairing damaged blood vessel tissues by limiting collagen matrix deposition and improving cell migration, and hence overexpression of this gene might be helpful in migration and metastasis of cancer cells [25].

Cadherin EGF LAG seven-pass G-type receptor, *CELSR3*, is a member of the cadherin family belonging to the flamingo protein subfamily, and is known to play an important role in the WNT/PCP signaling pathway, which, in turn, controls the polarity of tissues and cell migration [26]. In the present study, *CELSR3* was found to be overexpressed using qPCR, and the analysis in the TCGA data set confirmed its overexpression in BC. To our knowledge, this study reports the overexpression of *CELSR3* in BC for the first time. Further, overexpression of *CELSR3* was found to be significantly associated with older age (≥55 years), suggesting its utility as a potential biomarker in BC occurring in older age. A similar trend was observed in TCGA data set as well with significant down-regulation in the age group of 61—80 years supporting the observations of the present study. The *CELSR3* overexpression was found to be associated with poor prognosis for post-menopausal patients, suggesting it can serve as a prognostic factor for predicting risk for postmenopausal women. Expression of this gene was also found to be associated with mucinous tumors, suggesting it to be a potential marker for detection of mucinous tumors, if verified at the protein level. The functional role of *CELSR3* in breast carcinogenesis is not known, which needs investigating experimentally.

We hypothesized that miRNAs can be one of the factors resulting in differential expression of the ECM genes observed in the present study, hence, we predicted the miRNAs that are likely to taget these genes. MiR-124-3p is one of the miRNAs predicted to target the *COL10A1* gene with 6 hits, suggesting a probability for the involvement of this miRNA in regulating the gene (Supplementary Table S7). *COL10A1* gene was found to be up-regulated in the present study, and MiR-124-3p has been shown to function as a tumor suppressor in BC by targeting CBL [27], and to play a tumor suppressor role in bladder [28], gastric [29], and esophageal [30] cancers. Similarly, miR-485-3p and miR-520d-3p had a maximum of 8 hits against *COL12A1*, suggesting their likely involvement in the deregulated expression of this gene, which needs verifying experimentally. However, inhibition of any of the above ECM genes has not been experimentally shown so far.

MicroRNA belonging to a family are often expressed in a coordinated fashion and regulate the expression of genes associated with similar functions. In the present study, we predicted miRNAs belonging to the miR-154, miR-515, and miR-10 miRNA families to target ECM and anion transport genes analyzed in the present study. The miR-154 family consists of total 19 miRNAs, of which 4 miRNAs were found to target the ECM genes analyzed (Fig. 8, *a*), suggesting a likely role of this miR family in the regulating the ECM and anion transport genes. MiR-154 and other members of this family have shown to facilitate EMT, tumor growth, proliferation, and bone metasta-

sis in BC and several other cancers such as prostate, thyroid, laryngeal squamous cell carcinoma, hepatocellular and gastric cancers [31—38]. MiR-515 family consists of a total of 42 miRNAs, and using miRwalk, 3 miRNAs were predicted to target *COL11A1, COL5A1, COL12A1, CELSR3, COL10A1*, and *COL1A1* genes. MiR-515 is down-regulated in BC and several other cancers [39—46] (Fig. 8, b). MiR-10 family consists of a total of 8 miRNAs, of which 3 miRNAs were found to target *SLC16A3, COL12A1, CELSR3, COL11A1, COL1A1*, and *COL5A1* collagen and anion transport genes analyzed in the present study (Fig. 8, *c*). The miR-10/100 family of miRNAs is evolutionarily ancient and highly conserved, considering both their sequence similarity and genomic location in the Hox gene clusters [47]. Ma et al. [48] identified miR-10b to be highly expressed in metastatic BC cells as compared to non-metastatic cells. MiR-154, miR-515, and miR10 family members could potentially regulate the selected ECM genes, which needs verifying experimentally.

To sum up, the present study has evaluated the differential expression of *COL1A1, COL5A1, COL10A1, COL11A1, COL12A1, CLERS3, CTHRC1*, and *COL14A1* genes involved in ECM at the transcript level in Indian BC patients. In this study, we have found an association of the overexpression of *COL12A1* with the less aggressive ER/PR positive tumors of the luminal subtype, whereas the underexpression of *COL14A1* was found to be associated with the basal BC subtype, suggesting them to be candidate biomarkers for detecting luminal and basal tumors, respectively. This study has demonstrated for the first time the down-regulated expression of *CELSR3* mRNA in BC, and its association with the old age of patients. Further, overexpression of this gene was found to be associated with poor prognosis for post-menopausal patients in TCGA data set, hence this gene may serve as a biomarker for cancer detection in groups of old patients. *CTHRC1* gene transcripts were found to be overexpressed in BC cells, its overexpression was found to be inversely associated with the OS of BC patients analyzed from the TCGA data set. Strikingly, overexpression of *CTHRC1* was associated with poor survival in the luminal BC cases, suggesting it to be a potential biomarker for predicting the outcome in patients with the luminal BC subtype. In silico target prediction has identified a likely involvement of miR-154, -515, -10 families in regulation of the ECM genes analyzed in the present study.

Acknowledgment

This work was funded by Department of Biotechnology (DBT), India. Authors thank DBT for funding the experimental work.

Conflict of Interest

Authors Dr. S. A. Raju Bagadi and Dr. Sunita Saxena have received a grant from Department of Biotechnology (DBT), India. The authors declare that there is no conflict of interests.

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018; **68**: 394-424. doi: 10.3322/caac.21492
2. Eble JA, Niland S. The extracellular matrix in tumor progression and metastasis. Clin Exp Metastasis 2019; **36**: 171-198. doi: 10.1007/s10585-019-09966-1
3. Liu J, Shen JX, Wu HT et al. Collagen 1A1 (COL1A1) promotes metastasis of breast cancer and is a potential therapeutic target. Discov Med 2018; **25**: 211-223.
4. Malvia S, Bagadi SAR, Pradhan D, et al. Study of gene expression profiles of breast cancers in Indian women. Sci Rep 2019; **9**: 10018. doi: 10.1038/s41598-019-46261-1

5. Acerbi I, Cassereau L, Dean I, et al. Human breast cancer invasion and aggression correlates with ECM stiffening and immune cell infiltration. Integr Biol (Camb) 2015; **7**: 1120-1134.doi: 10.1039/c5ib00040h
6. Mishra AK, Agrawal U, Negi S, et al. Expression of androgen receptor in breast cancer & its correlation with other steroid receptors & growth factors. Indian J Med Res 2012; **135**: 843-852.
7. Tang Z, Li C, Kang B, et al. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res 2017; **45**: W98-W102. doi: 10.1093/nar/gkx247
8. Chandrashekar DS, Bashel B, Balasubramanya SAH, et al. UALCAN: A portal for facilitating tumor subgroup gene expression and survival analyses. Neoplasia 2017; **19**: 649-658. doi: 10.1016/j.neo.2017.05.002
9. Sticht C, De La Torre C, Parveen A, et al. miRWalk: An online resource for prediction of microRNA binding sites. PloS One 2018; **13**: e0206239. doi: 10.1371/journal.pone.0206239
11. Zhang Z, Fang C, Wang Y, et al. COL1A1: A potential therapeutic target for colorectal cancer expressing wild-type or mutant KRAS. Int J Oncol 2018; **53**: 1869-1880. doi: 10.3892/ijo.2018.4536
12. Liu S, Liao G, Li G. Regulatory effects of COL1A1 on apoptosis induced by radiation in cervical cancer cells. Cancer Cell Int 2017; **17**: 73. doi: 10.1186/s12935-017-0443-5
13. Li G, Li X, Yang M, et al. Prediction of biomarkers of oral squamous cell carcinoma using microarray technology. Sci Rep 2017; **7**: 42105. doi: 10.1038/srep42105
14. Chen FF, Zhang SR, Peng H et al. Integrative genomics analysis of hub genes and their relationship with prognosis and signaling pathways in esophageal squamous cell carcinoma. Mol Med Rep 2019; **20**: 3649-3660. doi: 10.3892/mmr.2019.10608
15. Deng T, Gong YZ, Wang XK, et al. Use of genome-scale integrated analysis to identify key genes and potential molecular mechanisms in recurrence of lower-grade brain glioma. Med Sci Monit 2019; **25**: 3716-3727. doi: 10.12659/MSM.913602
16. Yen TY, Haste N, Timpe LC, et al. Using a cell line breast cancer progression system to identify biomarker candidates. J Proteomics 2014; **96**: 173-183. doi: 10.1016/j.jprot.2013.11.006
17. Jiang X, Wu M, Xu X, et al. COL12A1, a novel potential prognostic factor and therapeutic target in gastric cancer. Mol Med Rep 2019; **20**: 3103-3112. doi: 10.3892/mmr.2019.10548
18. Liu J, Li H, Shen S, et al. Alternative splicing events implicated in carcinogenesis and prognosis of colorectal cancer. J Cancer 2018; **9**: 1754-1764. doi: 10.7150/jca.24569
19. Chudasama D, Bo V, Hall M, et al. Identification of cancer biomarkers of prognostic value using specific gene regulatory networks (GRN): a novel role of RAD51AP1 for ovarian and lung cancers. Carcinogenesis 2018; **39**: 407-417. doi: 10.1093/carcin/bgx122
20. Sun D, Jin H, Zhang J, et al. Integrated whole genome microarray analysis and immunohistochemical assay identifies COL11A1, GJB2 and CTRL as predictive biomarkers for pancreatic cancer. Cancer Cell Int 2018; **18**: 174. doi: 10.1186/s12935-018-0669-x
21. Li X, Zhou F, Jiang C, et al. Identification of a DNA methylome profile of esophageal squamous cell carcinoma and potential plasma epigenetic biomarkers for early diagnosis. PloS One 2014; **9**: e103162. doi: 10.1371/journal.pone.0103162
22. Renner M, Wolf T, Meyer H, et al. Integrative DNA methylation and gene expression analysis in high-grade soft tissue sarcomas. Genome Biol 2013; **14**: r137. doi: 10.1186/gb-2013-14-12-r137
23. Morris MR, Ricketts C, Gentle D, et al. Identification of candidate tumour suppressor genes frequently methylated in renal cell carcinoma. Oncogene 2010; **29**: 2104-2117. doi: 10.1038/onc.2009.493
24. Goto R, Nakamura Y, Takami T, et al. Quantitative LC-MS/MS Analysis of proteins involved in metastasis of breast cancer. PloS One 2015; **10**: e0130760. doi: 10.1371/journal.pone.0130760
25. Wu Q, Yang Q, Sun H. Role of collagen triple helix repeat containing-1 in tumor and inflammatory diseases. J Cancer Res Ther 2017; **13**: 621-624. doi: 10.4103/jcrt.JCRT_410_17
26. Kim JH, Baek TH, Yim HS, et al. Collagen triple helix repeat containing-1 (CTHRC1) expression in invasive ductal carcinoma of the breast: the impact on prognosis and correlation to clinicopathologic features. Pathol Oncol Res 2013; **19**: 731-737. doi: 10.1007/s12253-013-9636-y
27. Katoh M. WNT/PCP signaling pathway and human cancer (review). Oncol Rep 2005; **14**: 1583-1588.
28. Wang Y, Chen L, Wu Z, et al. miR-124-3p functions as a tumor suppressor in breast cancer by targeting CBL. BMC Cancer 2016; **16**: 826. doi: 10.1186/s12885-016-2862-4
29. Wang JR, Liu B, Zhou L, et al. MicroRNA-124-3p suppresses cell migration and invasion by targeting ITGA3 signaling in bladder cancer. Cancer Biomark 2019; **24**: 159-172. doi: 10.3233/CBM-182000

30. Liu F, Hu H, Zhao J, et al.miR-124-3p acts as a potential marker and suppresses tumor growth in gastric cancer. Biomed Rep 2018; **9**: 147-155. doi: 10.3892/br.2018.1113
31. Zeng B, Zhang X, Zhao J, et al.The role of DNMT1/hsa-miR-124-3p/BCAT1 pathway in regulating growth and invasion of esophageal squamous cell carcinoma. BMC Cancer 2019; **19**: 609. doi: 10.1186/s12885-019-5815-x
32. Qin C, Zhao Y, Gong C, et al. MicroRNA-154/ADAM9 axis inhibits the proliferation, migration and invasion of breast cancer cells. Oncol Lett 2017; **14**: 6969-6975. doi: 10.3892/ol.2017.7021
33. Xu H, Fei D, Zong S, et al. MicroRNA-154 inhibits growth and invasion of breast cancer cells through targeting E2F5. Am J Transl Res 2016; **8**: 2620-2630.
34. Gururajan M, Josson S, Chu GC, et al. miR-154* and miR-379 in the DLK1-DIO3 microRNA mega-cluster regulate epithelial to mesenchymal transition and bone metastasis of prostate cancer. Clin Cancer Res 2014; **20**: 6559-6569. doi: 10.1158/1078-0432.CCR-14-1784
35. Niu JT, Zhang LJ, Huang YW, e t al. MiR-154 inhibits the growth of laryngeal squamous cell carcinoma by targeting GALNT7. Biochem Cell Biol 2018; **96**: 752-760. doi: 10.1139/bcb-2018-0047
36. Wang M, Cai WR, Meng R, et al.miR-485-5p suppresses breast cancer progression and chemosensitivity by targeting surviving. Biochem Biophys Res Commun 2018; **501**: 48-54. doi: 10.1016/j.bbrc.2018.04.129
37. Kang M, Ren MP, Zhao L, et al. miR-485-5p acts as a negative regulator in gastric cancer progression by targeting flotillin-1. Am J Transl Res 2015; **7**: 2212-2222. doi:
38. Zhang CS, Lin Y, Sun FB, et al. miR-409 down-regulates Jak-Stat pathway to inhibit progression of liver cancer. Eur Rev Med Pharmacol Sci 2019; **23**: 146-154. doi: 10.26355/eurrev_201901_16758
39. Yu J, Xie F, Bao X, et al. miR-300 inhibits epithelial to mesenchymal transition and metastasis by targeting Twist in human epithelial cancer. Mol Cancer 2014; **13**: 121. doi: 10.1186/1476-4598-13-121
40. Pardo OE, Castellano L, Munro CE, et al. miR-515-5p controls cancer cell migration through MARK4 regulation. EMBO Rep 2016; **17**: 570-584. doi: 10.15252/embr.201540970
41. Zhang X, Zhou J, Xue D, et al. MiR-515-5p acts as a tumor suppressor via targeting TRIP13 in prostate cancer. Int J Biol Macromol 2019; **129**: 227-232. doi: 10.1016/j.ijbiomac.2019.01.127
42. Yu J, Tan Q, Deng B, et al. The microRNA-520a-3p inhibits proliferation, apoptosis and metastasis by targeting MAP3K2 in non-small cell lung cancer. Am J Cancer Res 2015; **5**: 802-811.
43. Li J, Wei J, Mei Z, et al. Suppressing role of miR-520a-3p in breast cancer through CCND1 and CD44. Am J Transl Res 2017; **9**: 146-154.
44. Su H, Ren F, Jiang H, et al. Upregulation of microRNA-520a-3p inhibits the proliferation, migration and invasion via spindle and kinetochore associated 2 in gastric cancer. Oncol Lett 2019; **18**: 3323-3330. doi: 10.3892/ol.2019.10663
45. Li D, Song H, Wu T, et al. MiR-519d-3p suppresses breast cancer cell growth and motility via targeting LIM domain kinase 1. Mol Cell Biochem 2018; **444**: 169-178. doi: 10.1007/s11010-017-3241-4
46. Li YY, Shao JP, Zhang SP, et al. miR-519d-3p inhibits cell proliferation and invasion of gastric cancer by down-regulating B-cell lymphoma 6. Cytogenet Genome Res 2018; **154**: 12-19. doi: 10.1159/000487372
47. Ren Z, Yang T, Ding J, et al. MiR-520d-3p antitumor activity in human breast cancer via post-transcriptional regulation of spindle and kinetochore associated 2 expression. Am J Transl Res 2018; **10**: 1097-1108.
48. Tehler D, Hoyland-Kroghsbo NM, Lund AH. The miR-10 microRNA precursor family. RNA Biol 2011; **8**: 728-734. doi: 10.4161/rna.8.5.16324
49. Ma L, Teruya-Feldstein J, Weinberg RA. Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. Nature 2007; **449**: 682-688. doi: 10.1038/nature06174

Submitted: December 23, 2021

Шрешта Малвія [1], Чінтамані Чінтамані [2], Рамеш Сарін [3], Ума С Дубей [4], Суніта Саксена [5], Сарангадхара Аппала Раджу Багаді [1,]*

[1] Відділ біології клітин, ICMR Національний інститут патології, Нью-Делі, Індія
[2] Госпіталь Сафдардхунг, Нью-Делі, Індія
[3] Госпіталь Індрапрастха Аполлон, Нью-Делі, Індія
[4] Бірла науково-технологічний інститут, Пілані, Раджастан, Індія
[5] Департамент досліджень з охорони здоров'я, Нью-Делі, Індія

АБЕРАНТНА ЕКСПРЕСІЯ ГЕНІВ *COL14A1*, *CELRS3* ТА *CTHRC1* В КЛІТИНАХ РАКУ ГРУДНОЇ ЗАЛОЗИ

Стан питання. Колагени становлять основний компонент екстрацелюлярного матрикса і залучені до регуляції пухлинного мікрооточення. Вони можуть диференційно експресуватися в клітинах раку грудної залози з різним транскриптомним профілем. **Мета.** Проаналізувати експресію генів *COL1A1, COL5A1, COL10A1, COL11A1, COL12A1, COL14A1, CTHRC1* та *CELRS3* на рівні мРНК та можливе клінічне значення такої диференційної експресії при раку грудної залози. **Матеріали та методи.** Рівень транскрипції визначали в пухлинній тканині 60 хворих на рак грудної залози методом кількісної ПЛР в реальному часі. **Результати.** Продемонстровано надекспресію *COL1A1, COL5A1, COL10A1, COL11A1, COL12A1, CTHRC1* та *CELRS3* і знижену експресію *COL14A1*. Зниження експресії COL14A1 асоціювалось із більш агресивними базальним та Her-2/neu підтипами раку грудної залози ($p = 0{,}031$). Надекспресія *CELSR3* відзначалась у хворих старшого віку (понад 55 років, $p = 0{,}049$). Аналіз даних з бази атласу ракового геному підтвердив зазначений характер диференційної експресії досліджуваних генів. Надекспресія *CTHRC1* асоціюється з гіршою загальною виживаністю та поганим прогнозом люмінального підтипу раку грудної залози. Надекспресія *CELSR3* асоціюється з муцинозними пухлинами та несприятливим прогнозом у постменопаузних жінок. *In silico* було визначено мішені деяких мікроРНК, асоційованих із раком грудної залози. Виявилось, що мікроРНК родин miR-154, -515 та -10 відіграють певну регуляторну роль відносно досліджених колагенових генів. **Висновки.**Експресія генів *COL14A1* та *CTHRC1* може розглядатись як потенційний біологічний маркер для виявлення базального раку грудної залози та прогнозування виживаності у хворих на рак грудної залози люмінального підтипу.

Ключові слова: рак грудної залози, екстрацелюлярний матрикс, надекспресія, знижена експресія, колагени.