Br J Dermatol 2024; **191**:775–790
https://doi.org/10.1093/bjd/ljae258
Advance access publication date: 25 June 2024

BJD
British Journal of Dermatology
Translational Research

COL10A1 expression distinguishes a subset of cancer-associated fibroblasts present in the stroma of high-risk basal cell carcinoma

Mauro Esposito[1], Laura Yerly[2], Prachi Shukla[1], Victoria Hermes[1], Federica Sella[1], Zsolt Balazs[3,4], Evelyn Lattmann[1], Aizhan Tastanova[1], Patrick Turko[1], Ramon Lang[1], Isabel Kolm[1], Ramon Staeger[1], François Kuonen[2], Michael Krauthammer[3,4], Juerg Hafner[1], Mitchell P Levesque[1] and Gaetana Restivo[1]

[1]Department of Dermatology, University Hospital Zurich, University of Zurich, Zurich, Switzerland
[2]Department of Dermatology and Venereology, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland
[3]Department of Quantitative Biomedicine, University of Zurich, Zurich, Switzerland
[4]Biomedical Informatics, University Hospital of Zurich, Zurich, Switzerland

M.P.L. and G.R. contributed equally to this work.

Correspondence: Gaetana Restivo. Email: gaetana.restivo@usz.ch

Linked Article: Forsthuber and Lichtenberger *Br J Dermatol* 2024; **191**:657–658.

Abstract

Background Basal cell carcinoma (BCC) is the most frequently diagnosed skin cancer and the most common malignancy in humans. Different morphological subtypes of BCC are associated with a low or high risk of recurrence and aggressiveness, but the underlying biology of how the individual subtypes arise remains largely unknown. As the majority of BCCs appear to arise from mutations in the same pathway, we hypothesized that BCC development, growth and invasive potential is also influenced by the tumour microenvironment and, in particular, by cancer-associated fibroblasts (CAFs) and the factors they secrete.

Objectives To characterize the stroma of the different BCC subtypes with a focus on CAF populations.

Methods To investigate the stromal features of the different BCC subtypes, we used laser capture microdissection (LCM) followed by RNA sequencing (RNA-Seq). Fifteen BCC samples from five different 'pure' subtypes (i.e. superficial, nodular, micronodular, sclerosing and basosquamous; $n=3$ each) were selected and included in the analysis. Healthy skin was used as a control ($n=6$). The results were confirmed by immunohistochemistry (IHC). We validated our findings in two independent public single-cell RNA-Seq (scRNA-Seq) datasets and by RNAscope.

Results The stroma of the different BCC subtypes were found to have distinct gene expression signatures. Nodular and micronodular appeared to have the most similar signatures, while superficial and sclerosing the most different. By comparing low- and high-risk BCC subtypes, we found that *COL10A1* is overexpressed in the stroma of sclerosing/infiltrative and basosquamous but not in micronodular high-risk subtypes. Those findings were confirmed by IHC in 93 different BCC and 13 healthy skin samples. Moreover, scRNA-Seq analysis of BCCs from two independent datasets found that the *COL10A1*-expressing population of cells is associated with the stroma adjacent to infiltrative BCC and shows extracellular matrix remodelling features.

Conclusions We identified *COL10A1* as a marker of high-risk BCC, in particular of the sclerosing/infiltrative and basosquamous subtypes. We demonstrated at the single-cell level that *COL10A1* is expressed by a specific CAF population associated with the stroma of infiltrative BCC. This opens up new, tailored treatment options, and suggests *COL10A1* as a new prognostic biomarker for BCC progression.

Lay summary

Basal cell carcinoma ('BCC' for short) is the most common type of cancer in humans. BCC occurs when a certain type of skin cell transforms in the outermost layer of the skin. This is mostly caused by a lot of exposure to sunlight. BCC can appear in different forms, or 'subtypes'. In each subtype, cancer cells grow in a specific way and are visually distinct from the surrounding tissue, known as the 'stroma'. The different subtypes of BCC can have a low or high risk of cancer recurrence and different levels of aggressiveness.

We aimed to find out what the stroma is made up of in the different subtypes of BCC, focusing on the differences between low-risk and high-risk cancers. We measured gene expression in specific areas of tissue and identified one gene called '*COL10A1*' as being overexpressed in the stroma of high-risk BCC, especially in two particular subtypes ('sclerosing' and 'basosquamous' BCC). We confirmed

Accepted: 13 June 2024

this result in BCC biopsies. We also checked other published data and found that *COL10A1* is mostly expressed in a type of cell called a 'fibroblast'. Fibroblasts expressing the *COL10A1* gene are present in the stroma right beside the infiltrative area of BCC.

Our findings could be used to develop more personalized treatment of BCC. The stroma may be a potential anti-cancer target, as well as a new way of testing for BCC progression.

What is already known about this topic?

- Basal cell carcinoma (BCC) morphological subtypes are different in terms of cancer cell growth patterns and the surrounding stroma.
- Infiltrative and nodular BCC morphological subtypes (tumour and stroma) have been characterized and shown to have distinct gene expression profiles.

What does this study add?

- We have characterized the stroma of five different BCC morphological subtypes more extensively and show that low- and high-risk BCC subtypes differ in terms of their stromal gene expression profiles.
- We show that *COL10A1* is expressed in high-risk BCC subtypes and, in particular, by a cancer-associated fibroblast population associated with the stroma adjacent to infiltrative BCC.

What is the translational message?

- We believe our results are novel, actionable and have the potential to change the diagnosis of the most common human cancer as *COL10A1* could be used as a new prognostic biomarker for BCC progression.
- These findings have the potential to lead to a more personalized treatment of BCC, including the stroma, and in particular collagen, as an anticancer target.

Basal cell carcinoma (BCC) is the most common malignancy in humans worldwide.[1,2] Most BCCs are not aggressive; however, if left untreated, invasion and destruction of the surrounding tissue can occur.[3] Rarely, BCCs metastasize and pose a substantial threat to a patient's life.[4] The majority of BCC lesions can be excised by standard surgical procedures or treated topically.[5] In advanced cases, systemic therapies using Hedgehog pathway inhibitors or immune checkpoint inhibitors might be considered.[5] However, neither tumour-targeting nor immune system-engaging agents provide a cure as the majority of patients still show progression of disease after these treatments.[6] New treatments directed at the tumour microenvironment (TME) are in development and are anticipated to improve the management of difficult-to-treat BCCs.[7]

BCCs can appear as different morphological subtypes. The most common subtypes are superficial, nodular, micronodular, basosquamous and sclerosing.[8] In each subtype, cancer cells show a characteristic growth pattern and are visually distinct surrounding tissue, also referred to as the cancer-associated stroma (CAS). The CAS consists of a variety of different cancer-associated fibroblasts (CAFs), as well as proteins and proteoglycans that form the extracellular matrix (ECM).[9] CAFs represent one of the most important players in the TME, influencing disease progression in many different cancer types.[10–13] While most superficial, nodular and micronodular BCC do not show a significant stromal reaction, increased collagen deposition and hypercellularity can be seen in the sclerosing and basosquamous subtypes.[14] Although there have been studies on the gene expression profile of BCC stroma (irrespective of subtype)[15] and comparisons of nodular and infiltrating BCCs,[16] a more extensive study accounting for all the other subtypes is lacking. Here, we show that the stroma of the different BCC morphological subtypes exhibit distinct gene signatures, all of which are dissimilar to healthy skin stroma. Furthermore, we demonstrate that *COL10A1* is expressed in the stroma of most high-risk BCCs and is associated with the invasive niche signature of CAFs.

Materials and methods

Formalin-fixed paraffin-embedded basal cell carcinoma samples

Formalin-fixed paraffin-embedded (FFPE) tissue samples of different BCC subtypes were selected with the help of a board-certified dermatopathologist and retrieved from the archive of the Department of Dermatology, University Hospital Zurich. FFPE blocks of healthy skin samples were also included. Regions of interest were defined for each sample (stained with haematoxylin and eosin) with the help of a dermatopathologist.

Laser capture microdissection

FFPE blocks were sectioned into 10-μm cuts on RNAse-free membrane slides [catalogue no. 50102; Molecular Machines & Industries (MMI), Eching, Germany]. Three cuts were prepared for each sample. The tissue was air-dried overnight, rehydrated in xylol, ethanol 100%, ethanol 95%, and ethanol 70% twice, and stained with cresyl violet [catalogue no. 105235 (Sigma-Aldrich, St. Louis, MO, USA); 1.7 mg mL^{-1} in 100% ethanol], for 1 min each; followed by rehydration

with ethanol 70%, ethanol 95% and ethanol 100% for 5–10 s each; and, finally, rehydrated in ethanol 100% for 1 min. After staining, the slides were left to dry. For microdissection, the MMI laser microdissection system at the ScopeM facility (ETH Zürich, Zürich, Switzerland) was used. Dissected tissue was captured on 0.5 mL Diffuser Isolation Caps (catalogue no. 50202; MMI) and stored at –20 °C.

RNA extraction from formalin-fixed paraffin-embedded tissue

RNA was isolated from FFPE tissue cuts captured by laser capture microdissection (LCM) using a truXTRAC FFPE RNA microTUBE Kit (catalogue no. 520161; Covaris, Woburn, MA, USA) following the kit's instructions and adapting it to the protocol described by Amini *et al.*[17] Briefly, the top layer of the caps containing the captured tissue pieces was removed with a sterile scalpel blade and put into a microTUBE-130 AFA Fiber Pre-Slit Screw-Cap with 110 μL RNA lysis buffer (Covaris). The tubes were sonicated in an E220 Ultrasonicator (Covaris) for 5 min at 20 °C. Subsequently, 10 μL Proteinase K solution (provided with the kit) was added and the samples sonicated again for 10 s. The tubes were then placed in a heat block set to 56 °C for 15 min and heated to 80 °C for 1 h. Samples were treated with DNase I (provided with the kit) for 15 min at room temperature. RNA purification was performed using the RNA Purification Columns supplied in the Covaris kit following the manufacturer's protocol.

RNA sequencing

Sequencing and library preparation were performed at the Functional Genomics Center in Zurich, Switzerland. A SMARTer® Stranded Total RNA-Seq Kit v2 Pico Input Mammalian (Takara Bio, San Jose, CA, USA) was used. Cluster generation and sequencing were performed on a NovaSeq6000 System (Illumina, San Diego, CA, USA) with a run configuration of 100 base pair paired-end reads. A detailed description of library preparation and sequencing can be found in Appendix S1 (see Supporting Information).

RNA sequencing data analysis

Transcript abundances were estimated using Salmon version 1.5.0 in selective alignment mode,[18] with a k-mer length of 31, GC-bias correction, sequence strand bias correction and mapping validation enabled. Raw reads were mapped to a preindexed human genome reference database (index version 5) created from the GRCh38_no_alt_analysis_set from the National Center for Biotechnology Information, downloaded from http://refgenomes.databio.org.

RNA sequencing (RNA-Seq) data were analysed in R (R Foundation for Statistical Computing, Vienna, Austria) using the 'DESeq2' (version 1.36.0) package.[19] In a prefiltering step, only genes with at least 10 counts in at least 3 samples were included. For heatmap visualization and gene clustering, the count data were variance stabilized using regularized-logarithm (rlog) transformation. To determine the significance of differentially expressed genes (DEGs), the Benjamini–Hochberg adjusted *P*-value was considered. Gene set enrichment analysis (GSEA) was performed using the R package 'fgsea' (version 1.22.0) and the Reactome pathway database.[20–22] The list of genes that was given as input was ranked by the result of the Wald test statistic. The minimum size of a gene set to be considered in the analysis was set to 20 genes. To infer the cell-type composition of each sample, the R package 'MCP-counter' was used.[23] The cell types 'neutrophils' and 'NK cells' were grouped together as 'other'.

Discriminative component analysis

Genes from the five bulk RNA DESeq2 differential gene expression results comparing single subtypes with a group of the remaining subtypes were thresholded by an adjusted *P*-value of 0.05. The gene lists were then iteratively thresholded by log2 fold changes (FCs) of > 1 to 8 in steps of 0.5 resulting in 15 gene subsets for each comparison that were then combined based on FC threshold. Lowly expressed genes (lower quantile = 0) were removed from the raw counts before normalization (log10(x+1)), 0-centering and scaling by the SD. The counts were then partitioned based on the gene subsets and iteratively inputted into the discriminative component analysis (DCA) [R package 'dml' (version 1.1.0)][24] to identify the gene set that best separates the five subtypes. The results were plotted and visual separation was used to identify the 134 gene set of log2 FC > 4 as creating the best separation. Heatmap, t-SNE (t-distributed stochastic neighbour embedding) and UMAP (uniform manifold approximation and projection) plots were created to confirm separation by subtype.

Single-cell RNA sequencing analysis

The dataset from Yerly *et al.* was downloaded and analysed as described in the original publication.[16]

The second single-cell RNAseq (scRNAseq) dataset was obtained from Restivo *et al.*[25] Reads were aligned to the human reference genome (GRCh38-3.0.0; 10x Genomics). The 'Seurat' (version 4.0)[26] package for R was used to merge, scale and normalize gene expression data, as well as for clustering, differential gene expression analysis and visualizations. We used SingleR (version 1.4.1)[27] with the Blueprint and ENCODE reference datasets to assign cell types.[28,29] Cell typing was manually curated based on clustering and the expression of known cell-type markers, as previously described.[25] GSEA was performed using the 'fgsea' package for R.[20,21]

Collagen type X immunohistochemistry

Four-micrometre tissue sections were prepared. The tissue was deparaffinized using Bond Dewax Solution (catalogue no. AR9222; Leica Biosystems, Wetzlar, Germany) following the manufacturer's instructions. Antigen retrieval was performed at pH 6 for 30 min and tissue cuts were stained with COLX (X53) primary antibody (catalogue no. 14-9771-82; Invitrogen, Carlsbad, CA, USA) at a dilution of 1 : 300 using a BOND-RXm fully automated stainer (Leica Biosystems). Final colour development was performed using the Bond Polymer Refine Red Detection Kit (catalogue no. DS9390; Leica Biosystems). Coverslips were mounted using Dako mounting medium (catalogue no. CS703; Agilent, Santa Clara, CA, USA) and the stained tissue was imaged using an Aperio ScanScope slide scanner (Leica Biosystems). Staining was quantified in QuPath (version 0.4.1).[30] An area with a margin of 100 μm was defined around each tumour.

To distinguish between tumour and stromal areas, a random trees pixel classifier was trained at a resolution of 0.98 μm per pixel, using Gaussian and Laplacian of Gaussian features at scales of 1.0, 4.0 and 8.0. In areas that were annotated as stroma by this classifier, the percentage of the collagen type X-positive area was calculated using another pixel classifier. Stain vectors for a positive signal and haematoxylin counterstain were estimated for each image before running the analysis. Statistical significance was assessed using a Wilcoxon rank sum test.

RNAscope

The RNAscope assay was performed on 5-μm sections of four FFPE-preserved BCC biopsies from the same cohort used for immunohistochemistry (IHC).[31] The RNAscope Multiplex Fluorescent Detection Reagents_v2 kit (catalogue no. 323110; Advanced Cell Diagnostics, Newark, CA, USA) was used according to the manufacturer's instructions. Standard pretreatment conditions were applied to all samples. An overview of the probes and fluorophores used is provided in Table 1. All fluorophores were used at a dilution of 1 : 1500. Stained slides were imaged on a PhenoImager HT (Akoya Biosciences, Marlborough, MA, USA) using the same exposure time for a given channel. Spectral unmixing was performed in InForm (version 2.6; Akoya Biosciences); images were analysed in QuPath (version 0.4.1).[30] DAPI (4′,6-diamidino-2-phenylindole) was used for nuclei detection and cell segmentation. For each marker, the threshold of mean fluorescence intensity per cell for positive cell detection was determined by visual inspection and the same threshold was applied to all BCC biopsies.

Results

To investigate the stromal features characterizing the different BCC morphological subtypes, we selected FFPE specimens of the superficial, nodular, micronodular, sclerosing/infiltrating and basosquamous subtypes (Table 2). A schematic overview of the project workflow is provided in Figure 1. To analyse specifically stromal cell populations and ECM components directly adjacent to tumour cells or healthy epidermis, we applied LCM followed by RNA-Seq.

Stroma of different basal cell carcinoma subtypes and healthy skin exhibit distinct gene expression profiles

The gene expression profiles of all BCC stroma were compared with that of healthy skin stroma. This comparison revealed a large transcriptional difference between the two populations (Figure 2a). Specifically, 558 genes were significantly upregulated in tumour stroma and 407 were downregulated (Wald test, Benjamini–Hochberg adjusted P-value < 0.05, absolute log2 FC > 1). By clustering the 500 most significant DEGs, specific gene expression signatures were defined for both groups (Figure 2b). Healthy skin stroma showed more homogeneous clusters of over- and underexpressed genes, while expression of this set of genes was more heterogeneous among the different BCC stroma. GSEA revealed a variety of gene sets were significantly enriched (Figure 2c).[20] Interestingly, collagen degradation (Figure 2d) and ECM organization (Figure 2e) were among the positively enriched gene sets in BCC stroma. Inferring the cell-type composition of the samples using MCP-counter revealed that they predominantly contain fibroblasts (Figure S1; see Supporting Information).[23] This prompted us to assess the expression of common CAF markers, many of which were significantly upregulated in the stroma of the BCC samples when compared with healthy skin (Figure 2f). In particular, *FAP*, *TNC*, *ITGA11*, *NG2* (*CSPG4*), *COL11A1*, *SPP1* and *INHBA* showed significantly increased expression in BCC.[32,33] Interestingly, one of the most common markers used to define CAFs, *ACTA2*, was not significantly upregulated in BCC stroma. DCA was performed to explore the heterogeneity of gene expression observed among the different BCC subtypes. This approach – which aimed to determine a gene set that best separated the stroma of the five subtypes – revealed that the nodular and micronodular subtypes showed the highest similarity, while the other three subtypes were more distant from each other and from nodular and micronodular clusters (Figure S2a; see Supporting Information). Clustering the most significantly changed genes produced a specific gene signature for each subtype (Figure S2b). Comparing single subtypes against each other confirmed that nodular and micronodular stroma signatures were the most similar

Table 1 Probes and fluorophores used in the RNAscope assay

Target	Channel	Catalogue no.[a]	Fluorophore
COL10A1	C1	427851	TSA Vivid 570[b]
MMP1	C2	479741-C2	Opal 690[c]
COL11A1	C4	400741-C4	TSA Vivid 520[d]

[a]Advanced Cell Diagnostics, Newark, CA, USA; [b]catalogue no. 7526 (Tocris Bioscience, Bristol, UK); [c]catalogue no. FP1497001KT (Akoya Biosciences, Marlborough, MA, USA); [d]catalogue no. 7523 (Tocris Bioscience).

Table 2 Cohort of formalin-fixed paraffin-embedded blocks of healthy skin or basal cell carcinoma biopsies of different morphological subtypes used for laser capture microdissection

Subtype	Patient ID	Age (years)	Sex	Location
Superficial	S1	59	M	Trunk
Superficial	S2	88	M	Trunk
Superficial	S3	45	F	Limb
Nodular	N1	79	M	Trunk
Nodular	N2	82	F	Trunk
Nodular	N3	66	M	Trunk
Sclerosing	X1	75	F	Face
Sclerosing	X2	76	M	Face
Sclerosing	X3	50	M	Face
Micronodular	M1	63	M	Face
Micronodular	M2	88	F	Scapula
Micronodular	M3	84	F	Trunk
Basosquamous	B1	84	M	Scapula
Basosquamous	B2	73	M	Face
Basosquamous	B3	70	F	Face
Healthy skin	H1	57	M	Trunk
Healthy skin	H2	23	F	Trunk
Healthy skin	H3	78	F	Trunk
Healthy skin	H4	77	M	Trunk
Healthy skin	H5	22	F	Trunk
Healthy skin	H6	47	F	Trunk

F, female; M, male.

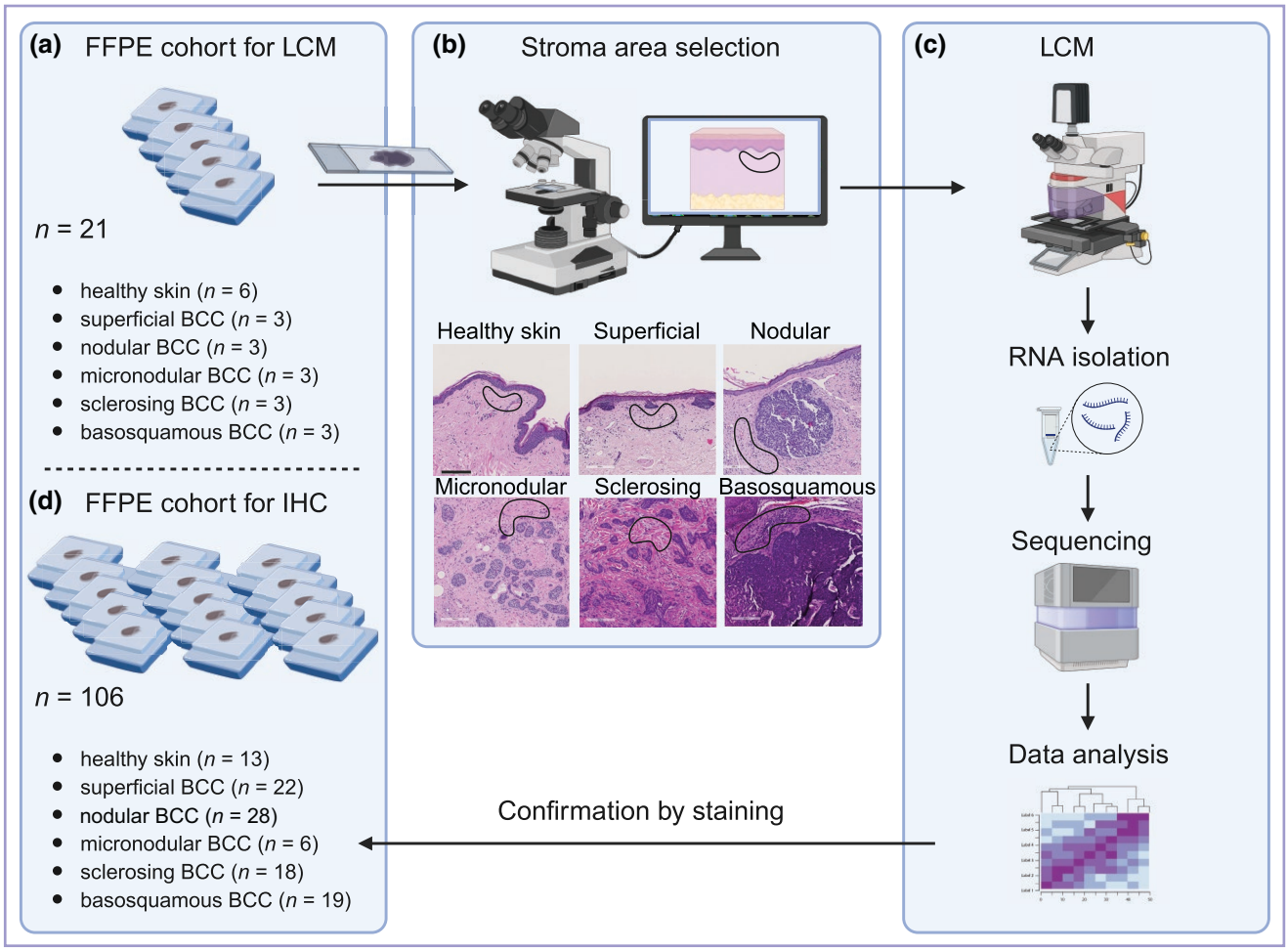

Figure 1 Schematic project workflow. (a) A cohort of formalin fixed paraffin-embedded (FFPE) blocks of different 'pure' basal cell carcinoma (BCC) morphological subtypes (n=15) and healthy skin (n=6) was retrieved from the archive of the dermatology department of the University Hospital Zürich. (b) Areas of interest were selected with the help of a dermatopathologist. (c) The stromal areas of interest of the different BCC subtypes were isolated by laser capture microdissection (LCM) and RNA was extracted. RNA sequencing (RNA-Seq) and data analysis were performed. (d) Data obtained by RNA-Seq were confirmed in a larger cohort by immunohistochemistry (IHC). Figure created with BioRender.com.

(Figure S3a; see Supporting Information), while superficial and sclerosing were the most different as this comparison had the highest number of DEGs (Figure S3b).

COL10A1 is overexpressed in the stroma of high-risk basal cell carcinoma

Morphological subtypes of BCC are classified based on their risk of recurrence. Nodular and superficial BCC are considered to be 'low risk', while micronodular, basosquamous and sclerosing are classified as 'high risk'.[5] To further investigate the stromal gene expression profile of the clinical BCC subtypes, we compared low- vs. high-risk subtypes. This characterization might be important for future treatments as those two risk groups have a different clinical outcome and different treatment options.[5,7] Clustering and heatmap visualization revealed specific signatures discriminating between high- and low-risk subtypes. Considerable variation was present in the two groups (Figure 3a), which led to only four genes being either significantly over- or underexpressed (adjusted P-value < 0.05, absolute log2 FC > 1) (Figure 3b). Among the upregulated genes were *COL10A1*, *KLK4* and *IGHV5-51*, while *CPNE3* was downregulated. Owing to our interest in CAS and, in particular in CAFs, and as *COL10A1* was the only gene to show specific expression in fibroblasts at the RNA and protein levels (Figure S4; see Supporting Information), we decided to investigate this gene in more detail. *COL10A1* encodes collagen type X, which has been shown to be present at high levels in the TME of breast cancer, where it correlates with poor response to chemotherapy;[34] in gastric cancer, where it has been described to be an inducer of epithelial–mesenchymal transition;[35] and in colon cancer, where it has been suggested to be a biomarker for early detection.[36] We examined the expression of *COL10A1* in other cancer types using The Cancer Genome Atlas (TCGA)[37] accessed via Gene Expression Profiling Interactive Analysis (GEPIA).[38] This revealed *COL10A1* to be overexpressed in many cancer types vs. healthy tissue (Figure 3c), and that overall survival in patients with high levels of *COL10A1* was lower than in patients with low *COL10A1* expression (Figure 3d). We then assessed which specific high-risk BCC subtype showed high *COL10A1* expression and found that the sclerosing subtype had the most abundant *COL10A1* transcripts. The basosquamous

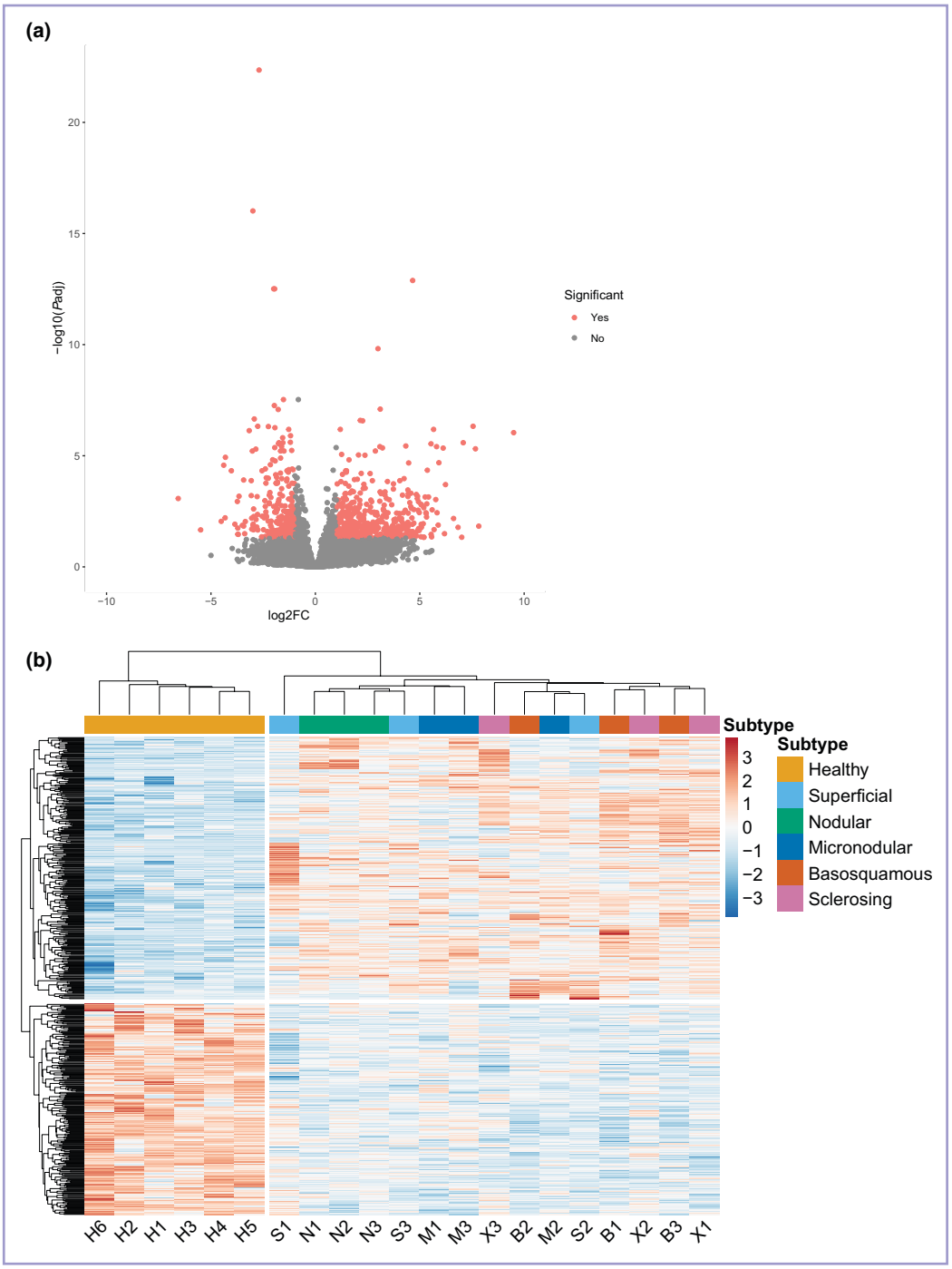

Figure 2 The stroma of different basal cell carcinoma (BCC) morphological subtypes show distinct gene expression profiles. (a) Volcano plot showing genes that are significantly overexpressed in the stroma of BCC compared with healthy skin. Genes with a log2 fold change (FC) > 1 and adjusted *P*-value (*P*adj) < 0.05 were considered to be significant. Significance was determined using a Wald test; *P*-values were adjusted using the Benjamini–Hochberg procedure. (b) Clustering and heatmap visualization of the top 500 most significant genes (as determined by adjusted *P*-value) that are differentially expressed in the stroma of BCC vs. healthy skin. (c) Overview of all significantly enriched pathways (adjusted *P*-value < 0.05), as determined by gene set enrichment analysis (GSEA). Significance was determined by the GSEA algorithm using an empirical phenotype-based permutation test; *P*-values were adjusted for each normalized enrichment score to correct for multiple hypothesis testing. Normalized enrichment score is the enrichment score normalized to the size of a given gene set. Enrichment plots for the pathways of (d) collagen degradation and (e) extracellular matrix organization. The dashed line indicates the maximum deviation from zero of the running sum along the ranked gene list (i.e. the enrichment score). (f) Differential expression analysis for different cancer-associated fibroblast markers. Wald test, Benjamini–Hochberg adjusted *P*-value (**P* < 0.1, ***P* < 0.05, ****P* < 0.01). BCR, B-cell receptor; FCGR, FC gamma receptor; ECM, extracellular matrix; ESR, oestrogen receptor; ER, endoplasmic reticulum; ns, not significant. (*Continued*)

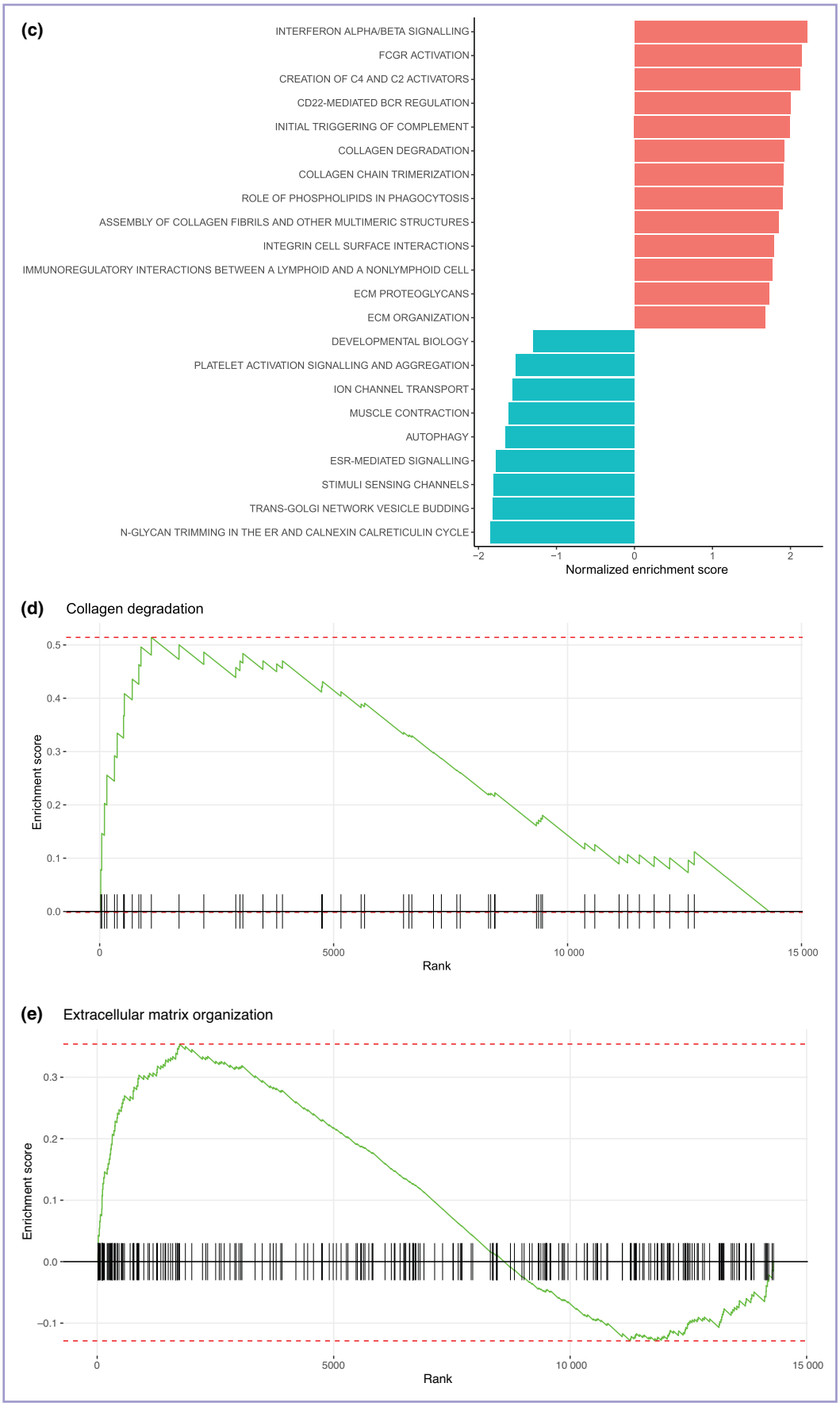

Figure 2 (*Continued*)

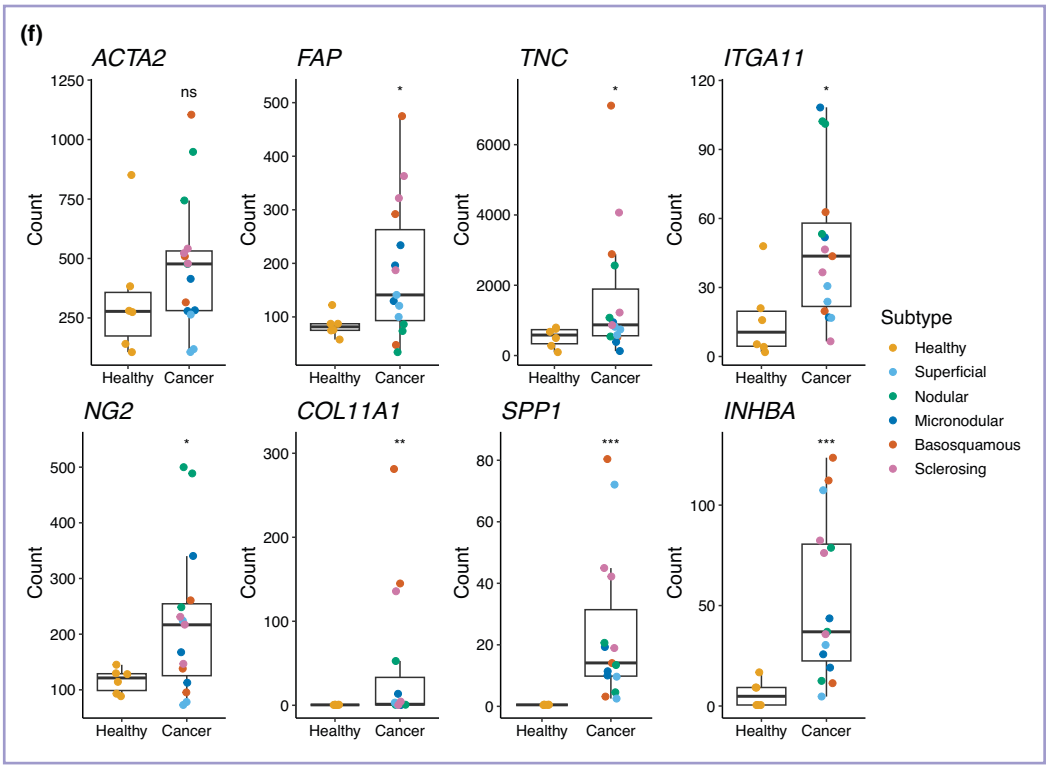

Figure 2 (*Continued*)

subtype also showed expression of *COL10A1*, while the micronodular subtype was negative for *COL10A1* expression (Figure 3e). Collectively, these findings made *COL10A1* a promising candidate for further investigation.

Immunohistochemistry confirms correlation of collagen type X with high-risk basal cell carcinoma

To confirm those findings at the protein level, IHC was performed to visualize the abundance of collagen type X in the stroma of healthy skin samples ($n=13$) and different BCC subtype samples ($n=93$) (Table 3). Comparison of staining in high- vs. low-risk BCC showed a higher protein abundance in the stroma of the high-risk BCC samples (Figure 4a). Comparisons of each subtype with the others and with healthy skin are provided in Figure S5 (see Supporting Information). While the stroma of healthy skin, superficial and micronodular BCC samples were largely negative for collagen type X, variable staining was visible in the stroma of nodular and basosquamous BCC samples, and a consistent staining in the stroma of the majority of sclerosing/infiltrative BCC samples (Figure 4b). Remarkably, nodular BCCs showed strong signals in the regions where cancer cells form smaller nodules that penetrate deeply into the dermis, while large nodules located close to the epidermis were mostly negative. The specificity of the staining is shown in more detail in Figure S6 (see Supporting Information). Taken together, these results corroborate our RNA-Seq results and suggest that collagen type X is associated with a more aggressive BCC phenotype with infiltrative features.

Table 3 Cohort of formalin-fixed paraffin-embedded blocks of healthy skin and basal cell carcinoma subtypes used for immunohistochemistry

	Healthy ($n=13$)	Superficial ($n=22$)	Nodular ($n=28$)	Micronodular ($n=6$)	Basosquamous ($n=19$)	Sclerosing/ infiltrating ($n=18$)	Overall ($n=106$)
Age (years), median (range)	36.0 (18.0–78.0)	77.0 (45.0–94.0)	81.0 (52.0–91.0)	67.0 (52.0–80.0)	79.0 (51.0–92.0)	77.0 (45.0–94.0)	77.0 (18.0–94.0)
Sex							
F	11 (85)	5 (23)	9 (32)	3 (50)	6 (32)	8 (44)	42 (39.6)
M	2 (15)	17 (77)	19 (68)	3 (50)	13 (68)	10 (56)	64 (60.4)
Location							
Head	2 (15)	3 (14)	10 (36)	4 (67)	15 (79)	17 (94)	51 (48.1)
Limbs	0 (0)	6 (27)	7 (25)	0 (0)	2 (11)	0 (0)	15 (14.2)
Trunk	10 (77)	13 (59)	11 (39)	2 (33)	2 (11)	1 (6)	39 (36.8)
Missing	1 (8)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (1.0)

Data are presented as *n* (%) unless otherwise stated. F, female; M, male.

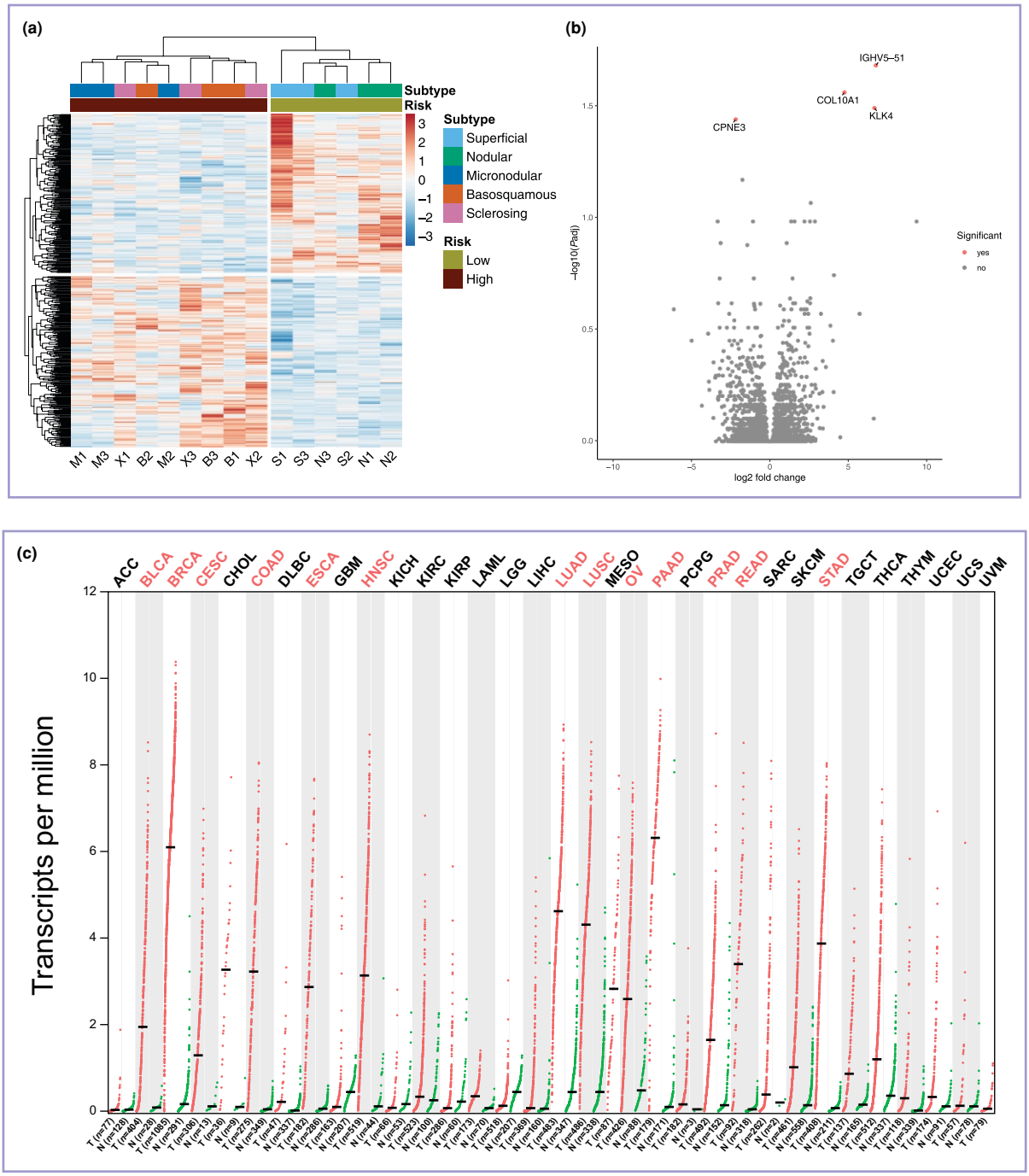

Figure 3 *COL10A1* is overexpressed in the stroma of high-risk basal cell carcinoma (BCC) subtypes. (a) Clustering and heatmap visualization of stromal gene expression of high-risk vs. low-risk BCC subtypes. The top 500 (as determined by adjusted *P*-value) differentially expressed genes (DEGs) are included in the plot. (b) Volcano plot of stromal gene expression of high-risk vs. low-risk BCC subtypes, including the names of significant genes. Genes with a log2 fold change (FC) > 1 and adjusted *P*-value (*P*adj) < 0.05 were considered to be significant. Significance was determined using a Wald test; *P*-values were adjusted using the Benjamini–Hochberg procedure. (c) The expression of *COL10A1* in various tumours is different from that in healthy tissues. Log2 FC: 1; Q-value: 0.01 [data obtained from Gene Expression Profiling Interactive Analysis (GEPIA)]. The cancer types shown in red are those that showed significantly higher expression of *COL10A1* compared with matching healthy tissue (ANOVA). (d) Overall survival of patients, including all cancer types listed in (c) (group cutoff quartile, high: 75%; low: 25%). Data obtained from GEPIA. (e) Normalized transcript counts of *COL10A1* in healthy skin and BCC subtypes. ACC, adrenocortical carcinoma; BLCA, bladder urothelial carcinoma; BRCA, breast invasive carcinoma; CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; COAD, colon adenocarcinoma; DLBC, lymphoid neoplasm diffuse large B-cell lymphoma; ESCA, oesophageal carcinoma; GBM, glioblastoma multiforme; HNSC, head-and-neck squamous cell carcinoma; KIRC, kidney renal clear cell carcinoma; KIRP, kidney renal papillary cell carcinoma; LIHC, liver hepatocellular carcinoma; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; N, normal (healthy) tissue; OV, ovarian serous cystadenocarcinoma; PAAD, pancreatic adenocarcinoma; READ, rectum adenocarcinoma; SKCM, skin cutaneous melanoma; STAD, stomach adenocarcinoma; T, tumour tissue; THYM, thymoma; TPM, transcripts per million; UCEC, uterine corpus endometrial carcinoma; UCS, uterine carcinosarcoma. (*Continued*)

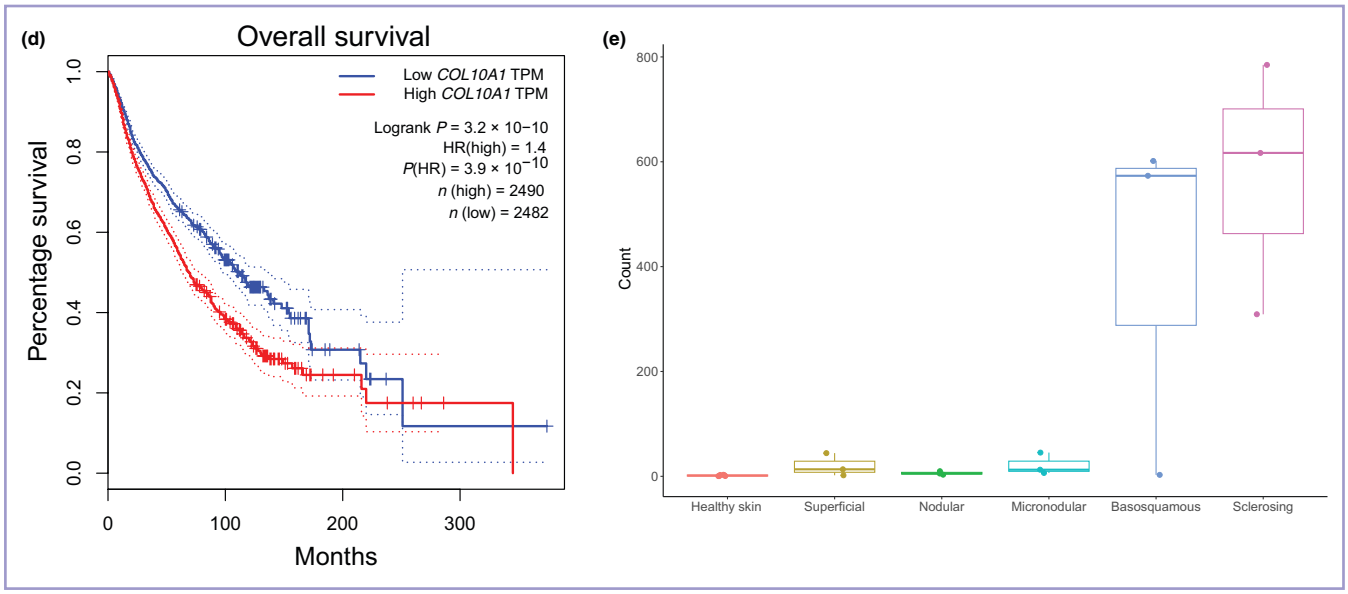

Figure 3 (*Continued*)

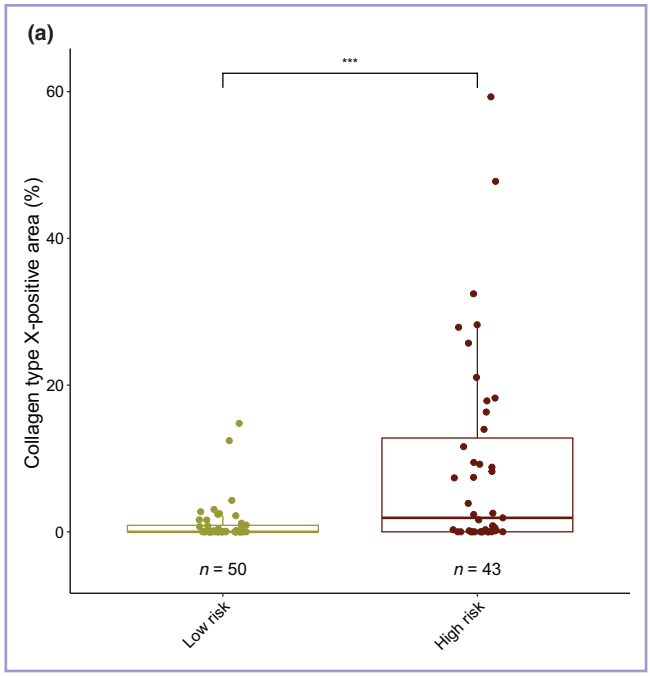

Figure 4 Collagen type X abundance is confirmed at the protein level and correlates with aggressive behaviour in basal cell carcinoma (BCC). (a) Quantification of immunohistochemical staining by QuPath performed on low- and high-risk BCC subtypes to determine the percentage of cancer-associated stroma positive for collagen type X. ***Exact *P*-value = 6.8×10^{-6} (Wilcoxon rank sum test). (b) Micrographs showing haematoxylin and eosin staining and collagen type X staining for healthy skin and the different BCC subtypes. Selected sections are shown at higher magnification in the right panel. (*Continued*)

COL10A1 is expressed by extracellular matrix-remodelling cancer-associated fibroblasts associated with an infiltrative stroma niche

To better characterize the cells expressing *COL10A1* and to gain insights into their function in the TME, we analysed two publicly available human BCC scRNA-Seq datasets.[16,25] In the first dataset,[16] *COL10A1* expression was shown to be specific to fibroblasts (Figure 5a, b). Unsupervised clustering generated three clusters of fibroblasts, with clusters 0 and 1 showing CAF signature enrichment. Remarkably, *COL10A1* was mainly expressed in cluster 1 (Figure 5c–e), a cluster spatially associated with the infiltrative morphology of BCC, as shown by the relative enrichment in the stromal infiltrative signature, as defined by Yerly *et al.* (Figure 5f).[16] Cluster 0 was more associated with a nodular signature and cluster 2 with neither, probably representing the 'normal' fibroblast population (Figure 5g). Moreover, cluster 1 was associated with ECM remodelling (Figure 5h), illustrated by the expression of *MMP11*, *COMP*, *POSTN*, *COL1A1*, *COL11A1*, *COL10A1*, *COL1A2* and *COL5A1* (Figure 5i) and by pathway analysis (Figure 5j). In the second dataset, which consisted of three BCC biopsies (with mixed subtypes), we confirmed that *COL10A1* was mostly expressed by CAFs (Figure S7a, b; see Supporting Information). BCC5613 – described by a dermatopathologist to be 40% sclerosing – contained more CAFs expressing *COL10A1* than the other two samples (Figure S5c). Differential gene expression analysis showed that the CAF cluster expressing *COL10A1* had a higher expression of other collagen types when compared with the CAF cluster negative for *COL10A1*. Also in this dataset, *MMP11*, *COL11A1* and *POSTN* – which have been

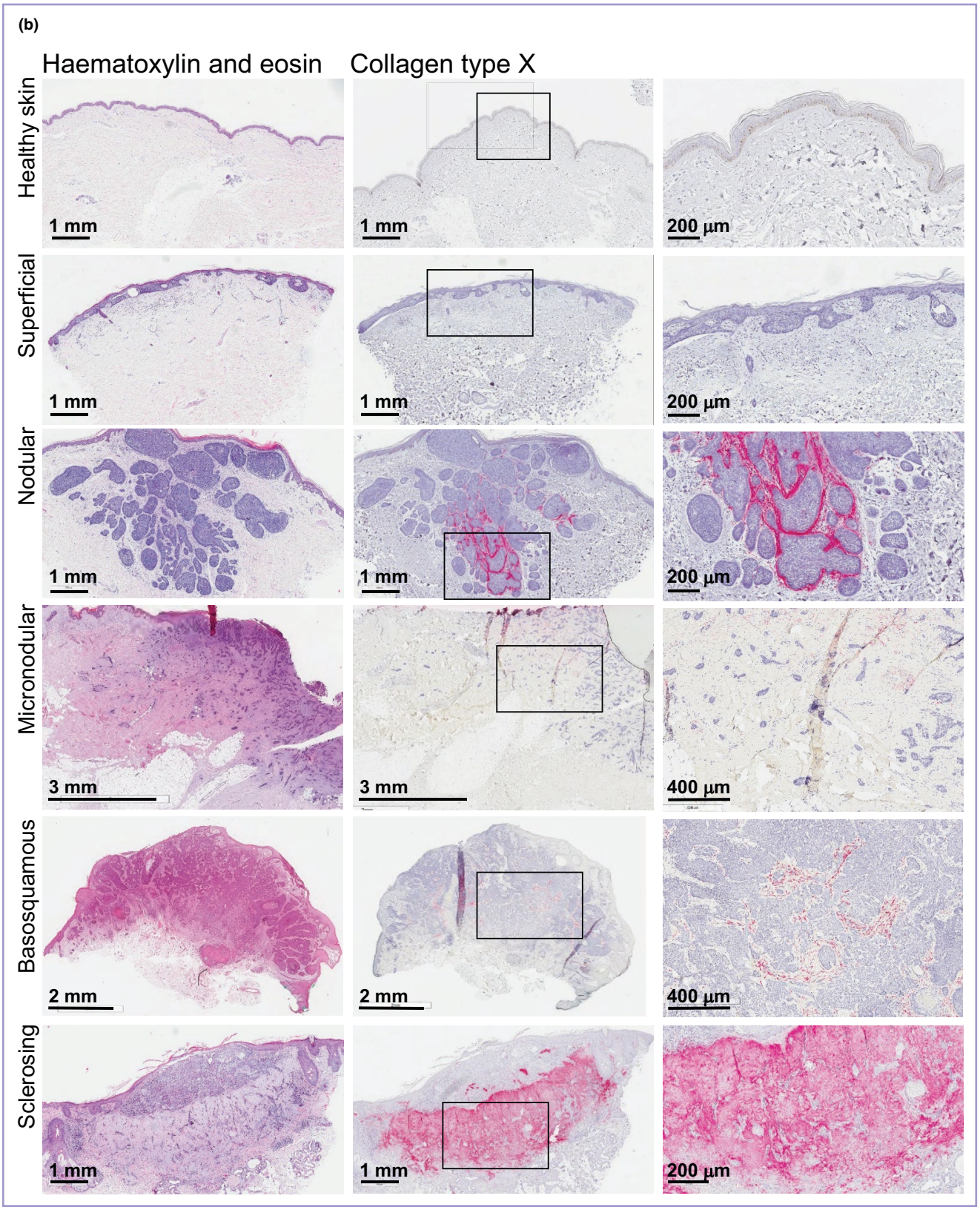

Figure 4 (*Continued*)

described to be overexpressed in the stroma of many cancer types and involved in tumour progression[39–41] – were more represented in the *COL10A1*-positive cluster (Figure S7d). Pathway analysis revealed the presence of several 'ECM remodelling' entries (Figure S7e). Taken together, these data support the potential role of *COL10A1* as a marker of

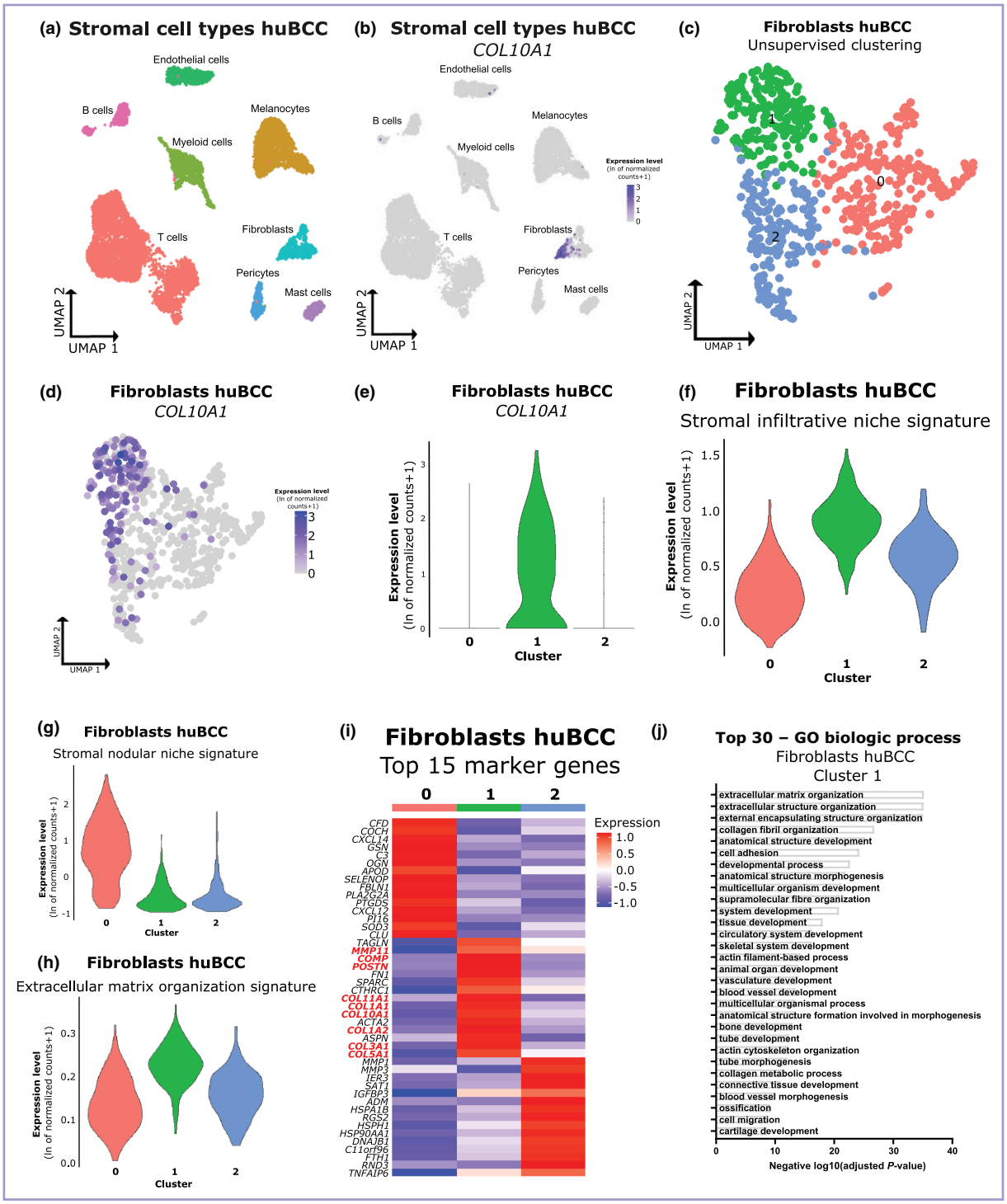

Figure 5 *COL10A1* defines extracellular matrix-remodelling fibroblasts of the invasive niche of basal cell carcinoma (BCC). (a) Uniform manifold approximation and projection (UMAP) plot of stromal cells from Yerly *et al.*[16] BCC single-cell RNA sequencing (scRNA-Seq) dataset coloured according to eight distinct cell types that were annotated using canonical cell type markers. (b) Expression level of *COL10A1* in stromal cells, represented as a colour scale overlaid on the UMAP plot. (c) Unsupervised clustering of the fibroblast subpopulation of Yerly *et al.*'s scRNA-Seq dataset represented as a UMAP plot.[16] (d, e) Expression level of *COL10A1* in the fibroblast subpopulation of the scRNA-Seq dataset, represented as a colour scale overlaid on the (d) UMAP plot and (e) as a violin plot per cell cluster. (f) Average expression level of stromal invasive niche signature (Yerly *et al.*)[16] in the fibroblast subpopulation of the scRNA-Seq dataset, represented as violin plot per cell cluster. (g) Average expression level of stromal nodular signature (Yerly *et al.*)[16] in the fibroblast subpopulation of the scRNA-Seq dataset, represented as a violin plot per cell cluster. (h) Average expression level of the Gene Ontology (GO) biologic process (BP) extracellular matrix organization signature (GO: 0030198) in the fibroblast subpopulation of the scRNA-Seq dataset, represented as a violin plot per cell cluster. (i) Heatmap of the average expression level of the 15 top cluster markers of each cluster in the fibroblast subpopulation of the scRNA-Seq dataset. Genes highlighted in red are members of the GO BP extracellular matrix organization signature (GO: 0030198). (j) Bar graph showing the top 30 GO biologic processes enriched in cluster markers of cluster 1 obtained with gProfiler.[51] huBCC, human BCC.

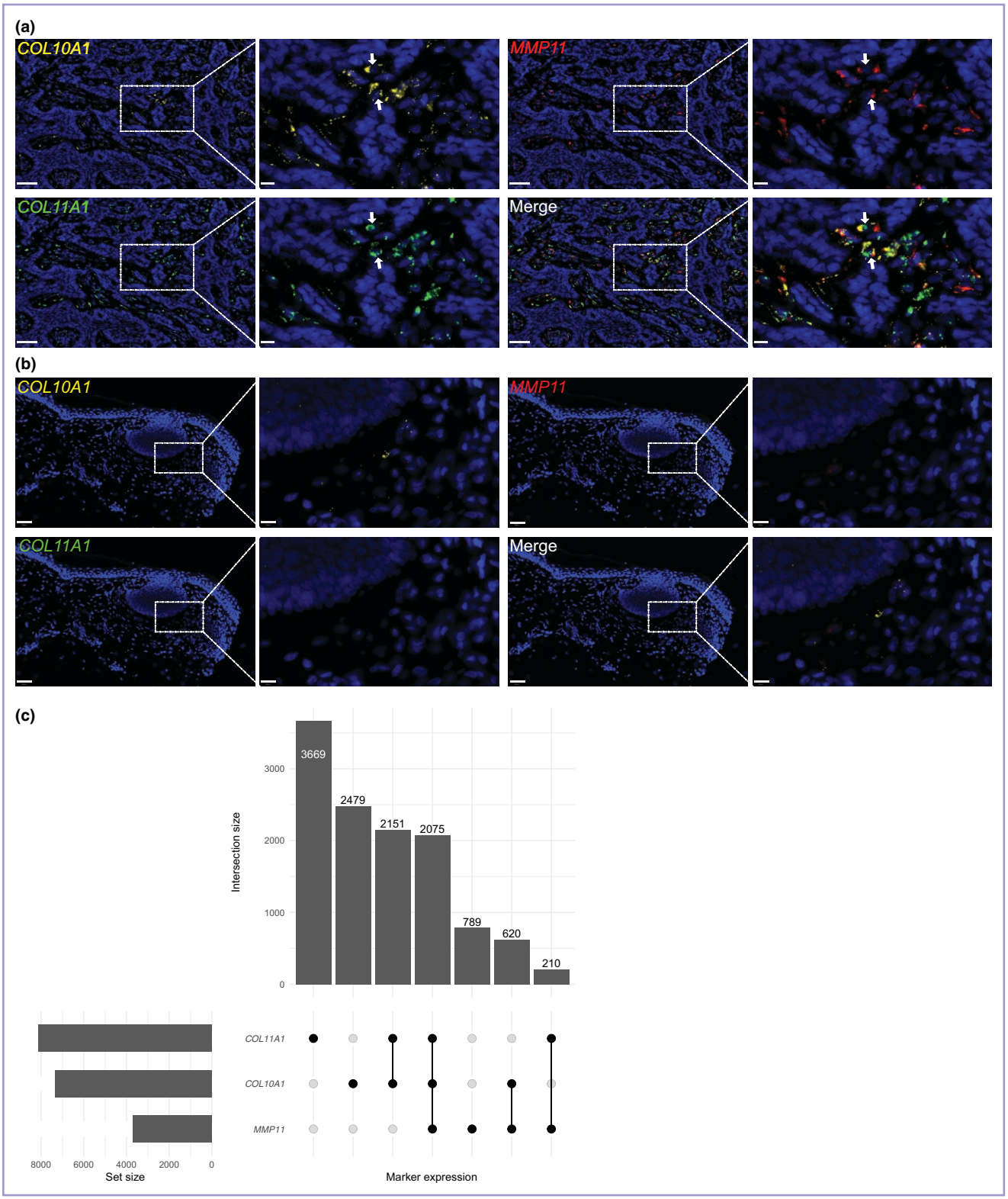

Figure 6 RNAscope confirms the presence of stromal cells in basal cell carcinoma (BCC) tissue expressing *COL10A1* mRNA and other markers associated with extracellular matrix remodelling. (a) Representative fluorescence micrograph showing an RNAscope staining of a sclerosing BCC. Probes targeting *COL10A1*, *COL11A1* and *MMP11* mRNA transcripts were used. Arrows indicate examples for cells that are positive for all three transcripts. Scale bars = 50 μm (10 μm for the higher-magnification images). (b) Representative fluorescence micrograph showing an RNAscope staining of a superficial BCC. Probes targeting *COL10A1*, *COL11A1* and *MMP11* mRNA transcripts were used. Scale bars = 50 μm (10 μm for the higher-magnification images). (c) UpSet plot showing the number of cells that are positive for the three markers and their combinations. Data from four different BCC biopsies are pooled (see Figure S9 for details).

ECM remodelling in CAFs surrounding infiltrating BCCs. We calculated the correlation between the expression of *COL10A1* and the top marker genes of fibroblast cluster 1 (from the dataset of Yerly *et al.*)[16] in other cancer types using the TCGA dataset. A close correlation between *COL10A1* expression and *TAGLN*, *MMP11*, *POSTN* and *COMP* was observed (Pearson correlation, $r > 0.39$, $P = 0$), while the correlation between *COL10A1* with the top marker genes of the remaining fibroblast clusters (such as *CFD* and *MMP1*) was very weak or absent ($r < 0.12$) (Figure S8; see Supporting Information). These data strongly support the association of *COL10A1* in a global ECM remodelling process among many cancer types.

To visualize the fibroblasts of cluster 1 *in situ*, we performed RNAscope staining using probes for *COLA10A1*, *MMP11* and *COL11A1* (*MMP11* and *COL11A1* are ECM remodelling marker genes showing the highest Pearson correlation coefficient with *COL10A1* in the TCGA dataset).[31] The staining showed a strong positive signal for the three probes in stromal areas of the sclerosing BCC (Figure 6a), while only very weak (*COL10A1, MMP11*) or completely absent (*COL11A1*) staining was observed in superficial BCC (Figure 6b). We performed the same RNAscope staining on a nodular BCC with infiltrating nodules and on a basosquamous BCC sample (Figure S9a; see Supporting Information). Quantification of the staining showed a high number of cells positive for all three markers (Figure 6c). Moreover, analysis of the single biopsies showed that the majority of cells were negative, demonstrating the specific expression of those markers in a subset of cells (Figure S9b). Intriguingly, the superficial BCC sample did not show a single cell positive for all three markers (Figure S9b).

Discussion

BCC is the most common cancer in humans and its incidence has been rising over the past few decades.[42] While most BCC tumours can be excised by standard surgical procedures or treated topically, in rare cases the cancer can advance locally or metastasize, requiring a systemic treatment.[5] As most BCCs harbour mutations in the Hedgehog pathway,[43] the most common systemic treatment is Hedgehog pathway inhibitors. Drugs targeting the TME of BCC have recently begun with the use of immunotherapy; however, there has been limited success as resistance is still high.[7] Several distinct morphological subtypes of BCC show differences in tumour growth pattern and the stroma surrounding the tumour. Interestingly, 85% of BCCs harbour mutations in the Hedgehog pathway and – although there are some hints at a distinct mutational profile of the individual subtypes[44] – the underlying biology of how the different subtypes develop remains largely unknown. One possible explanation is that the TME has an influence on tumour growth and infiltrative potential in the different BCC subtypes. To address this question, we collected a cohort of FFPE biopsies from different BCC subtypes. To analyse specific stromal regions of interest, we employed LCM. LCM provides high spatial resolution and allows for the assessment of the histopathology for each slide, an essential feature as the morphology of a tumour can change between consecutive sections. With this technique, we could accurately capture the regions of interest around the tumours and obtain sufficient amounts of RNA for RNA-Seq. Using this approach, we showed that there are many DEGs between the CAS of BCC and the stroma of healthy skin. We identified a distinct gene expression signature for each subtype, which might provide the basis for interesting factors to be investigated and tailored treatments for the different BCC subtypes.

When we compared the stromal gene expression profile of high- vs. low-risk BCC subtypes, four genes – including *COL10A1* – were significantly differentially expressed between the two groups. Interestingly, only two high-risk BCC subtypes showed *COL10A1* overexpression. This might suggest that the BCC subtypes classified as high risk have different biologic mechanisms of recurrence and spreading. The role of *COL10A1* in BCC has not yet been described. However, it is known that the different collagens can influence growth, invasion and treatment resistance in a multitude of cancer types.[45] Moreover, accumulating evidence suggests that they can be a prognostic factor in many cancers.[46] *COL10A1* has been implicated in cancer progression and resistance to treatment in a variety of cancer types.[34–36,47–50] The differential abundance of collagen type X in healthy skin and BCC subtypes was confirmed at the protein level, with the highest expression observed in the stroma of sclerosing/infiltrative and basosquamous BCC. Some nodular BCC also showed a high abundance of collagen type X, suggesting that some of those tumours might acquire invasive properties and transit to a high-risk subtype, especially ones that invade deep into the dermis. Furthermore, scRNA-Seq data of two independent studies showed that *COL10A1* expression is limited to a subpopulation of CAFs associated with ECM remodelling and in the proximity of BCC with invasive features. Moreover, we were able to detect this CAF population *in situ* using an RNAscope. Using probes for *COL10A1, COL11A1* and *MMP1* mRNA transcripts, we confirmed the co-expression of these markers by a stromal cell population present in BCCs with infiltrative properties.

Our study did not define the function of *COL10A1* in BCC, but we have provided important information for the clinical management of aggressive tumours. In addition to their significance as prognostic markers for BCC progression, our findings may allow for a more tailored approach to treating BCC. The use of drugs that target CAFs and factors secreted by them may be of benefit in combination with other routinely used treatments (i.e. Hedgehog pathway inhibitors). Drugs that target collagen are already in clinical trials for several cancer types.[46] Our study may contribute to the development of collagen-targeting therapies to improve the treatment of aggressive BCC.

Acknowledgements

We acknowledge Dr Erin Beebe and Professor Enni Markkanen for introducing us to the laser capture microdissection technique. We thank Dr Phil Cheng for support with the RNA sequencing data analysis; Professor Reinhard Dummer for the valuable clinical inputs; Isabella Dommann and the dermatopathology laboratory of the University Hospital Zurich for their support in the collection of the formalin-fixed paraffin-embedded specimens; and the Functional Genomics Center Zürich and the ScopeM facility of ETH, Zürich. The results shown here are partly based

upon data generated by the Cancer Genome Atlas Research Network (https://www.cancer.gov/tcga).

Funding sources

This work was supported by the SKINTEGRITY.CH collaborative research programme and Monique Dornonville de la Cour – Stiftung.

Conflicts of interest

The authors declare no conflicts of interest.

Data availability

The RNA sequencing data generated in this study have been deposited in the Gene Expression Omnibus (GEO) database under accession code GSE269601. We accessed the dataset reported by Yerly *et al.*[16] deposited in the GEO database under accession code GSE210648. We accessed the dataset reported by Restivo *et al.*[25] deposited to the European Genome-Phenome Archive (EGA) database under accession code EGAS00001005891. We used data from Gene Expression Profiling Interactive Analysis (GEPIA; Tang *et al.*).[24]

Ethics statement

All samples used were surplus material from routine surgeries. Basal cell carcinoma biopsies were provided by the Dermatology Department of the University Hospital Zürich, Switzerland, with the assistance of the SINTEGRITY.CH biobank (EK_647). All experiments conformed to the principles set out in the World Medical Association's Declaration of Helsinki and the Department of Health and Human Services Belmont Report. The use of material for research purposes was approved by the local ethics commissions (Cantonal_Ethic_Commission_Zürich: project no. 2017-00688).

Patient consent

Informed consent was obtained from all patients.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website.

References

1 Verkouteren JAC, Ramdas KHR, Wakkee M, Nijsten T. Epidemiology of basal cell carcinoma: scholarly review. *Br J Dermatol* 2017; **177**:359–72.
2 Madan V, Lear JT, Szeimies R-M. Non-melanoma skin cancer. *Lancet* 2010; **375**:673–85.
3 Goldenberg G, Karagiannis T, Palmer JB *et al.* Incidence and prevalence of basal cell carcinoma (BCC) and locally advanced BCC (LABCC) in a large commercially insured population in the United States: a retrospective cohort study. *J Am Acad Dermatol* 2016; **75**:957–66.
4 Lomas A, Leonardi-Bee J, Bath-Hextall F. A systematic review of worldwide incidence of nonmelanoma skin cancer. *Br J Dermatol* 2012; **166**:1069–80.
5 Peris K, Fargnoli MC, Garbe C *et al.* Diagnosis and treatment of basal cell carcinoma: European consensus–based interdisciplinary guidelines. *Eur J Cancer* 2019; **118**:10–34.
6 Stratigos AJ, Sekulic A, Peris K *et al.* Cemiplimab in locally advanced basal cell carcinoma after hedgehog inhibitor therapy: an open-label, multi-centre, single-arm, phase 2 trial. *Lancet Oncol* 2021; **22**:848–57.
7 Ramelyte E, Restivo G, Mannino M *et al.* Advances in the drug management of basal cell carcinoma. *Expert Opin Pharmacother* 2022; **23**:573–82.
8 Massi D, Cree IA, Elder DE *et al. WHO Classification of Skin Tumours*, 4th edn. World Health Organization, 2018.
9 Balkwill FR, Capasso M, Hagemann T. The tumor microenvironment at a glance. *J Cell Sci* 2012; **125**:5591–6.
10 Chen Y, McAndrews KM, Kalluri R. Clinical and therapeutic relevance of cancer-associated fibroblasts. *Nat Rev Clin Oncol* 2021; **18**:792–804.
11 Friedman G, Levi-Galibov O, David E *et al.* Cancer-associated fibroblast compositions change with breast cancer progression linking the ratio of S100A4$^+$ and PDPN$^+$ CAFs to clinical outcome. *Nat Cancer* 2020; **1**:692–708.
12 Neuwirt H, Bouchal J, Kharaishvili G *et al.* Cancer-associated fibroblasts promote prostate tumor growth and progression through upregulation of cholesterol and steroid biosynthesis. *Cell Commun Signal* 2020; **18**:11.
13 Min K-W, Kim D-H, Noh Y-K *et al.* Cancer-associated fibroblasts are associated with poor prognosis in solid type of lung adenocarcinoma in a machine learning analysis. *Sci Rep* 2021; **11**:16779.
14 Carr RA, Taibjee SM, Sanders DSA. Basaloid skin tumours: basal cell carcinoma. *Curr Diagn Pathol* 2007; **13**:252–72.
15 Sneddon JB, Zhen HH, Montgomery K *et al.* Bone morphogenetic protein antagonist gremlin 1 is widely expressed by cancer-associated stromal cells and can promote tumor cell proliferation. *Proc Natl Acad Sci U S A* 2006; **103**:14842–7.
16 Yerly L, Pich-Bavastro C, Di Domizio J *et al.* Integrated multi-omics reveals cellular and molecular interactions governing the invasive niche of basal cell carcinoma. *Nat Commun* 2022; **13**:4897.
17 Amini P, Ettlin J, Opitz L *et al.* An optimised protocol for isolation of RNA from small sections of laser-capture microdissected FFPE tissue amenable for next-generation sequencing. *BMC Mol Biol* 2017; **18**:22.
18 Patro R, Duggal G, Love MI *et al.* Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* 2017; **14**:417–19.
19 Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014; **15**:1–21.
20 Subramanian A, Tamayo P, Mootha VK *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005; **102**:15545–50.
21 Korotkevich G, Sukhov V, Budin N *et al.* Fast gene set enrichment analysis. *bioRxiv* 2021; https://doi.org/10.1101/060012.
22 Gillespie M, Jassal B, Stephan R *et al.* The reactome pathway knowledgebase 2022. *Nucleic Acids Res* 2022; **50**:D687–92.
23 Becht E, Giraldo NA, Lacroix L *et al.* Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol* 2016; **17**:218.
24 Tang Y, Gao T, Xiao N. 'dml': Distance Metric Learning in 'R'. *J Open Source Softw* 2018; **3**:1036.
25 Restivo G, Tastanova A, Balázs Z *et al.* Live slow-frozen human tumor tissues viable for 2D, 3D, ex vivo cultures and single-cell RNAseq. *Commun Biol* 2022; **5**:1144.
26 Hao Y, Hao S, Andersen-Nissen E *et al.* Integrated analysis of multimodal single-cell data. *Cell* 2021; **184**:3573–87.

27 Aran D, Looney AP, Liu L *et al.* Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat Immunol* 2019; **20**:163–72.
28 Martens JHA, Stunnenberg HG. BLUEPRINT: mapping human blood cell epigenomes. *Haematologica* 2013; **98**:1487–9.
29 ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012; **489**:57–74.
30 Bankhead P, Loughrey MB, Fernández JA *et al.* QuPath: open source software for digital pathology image analysis. *Sci Rep* 2017; **7**:16878.
31 Wang F, Flanagan J, Su N *et al.* RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. *J Mol Diagn* 2012; **14**:22–9.
32 Nurmik M, Ullmann P, Rodriguez F *et al.* In search of definitions: cancer-associated fibroblasts and their markers. *Int J Cancer* 2020; **146**:895–905.
33 Cangkrama M, Wietecha M, Mathis N *et al.* A paracrine activin A–mDia2 axis promotes squamous carcinogenesis via fibroblast reprogramming. *EMBO Mol Med* 2020; **12**:11466.
34 Wang Y, Lu S, Xiong J *et al.* ColXα1 is a stromal component that colocalizes with elastin in the breast tumor extracellular matrix. *Hip Int* 2019; **5**:40.
35 Li T, Huang H, Shi G *et al.* TGF-β1–SOX9 axis-inducible COL10A1 promotes invasion and metastasis in gastric cancer via epithelial-to-mesenchymal transition. *Cell Death Dis* 2018; **9**:849.
36 Solé X, Crous-Bou M, Cordero D *et al.* Discovery and validation of new potential biomarkers for early detection of colon cancer. *PLOS ONE* 2014; **9**:e106748.
37 The Cancer Genome Atlas Program (TCGA). Available at: https://www.cancer.gov/ccg/research/genome-sequencing/tcga (last accessed 26 September 2023).
38 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**:W98–102.
39 González-González L, Alonso J. Periostin: a matricellular protein with multiple functions in cancer development and progression. *Front Oncol* 2018; **8**:225.
40 Ma B, Ran R, Liao H-Y, Zhang H-H. The paradoxical role of matrix metalloproteinase-11 in cancer. *Biomed Pharmacother* 2021; **141**:111899.
41 Salimian N, Peymani M, Ghaedi K *et al.* Collagen 1A1 (COL1A1) and collagen 11A1 (COL11A1) as diagnostic biomarkers in breast, colorectal and gastric cancers. *Gene* 2024; **892**:147867.
42 Dika E, Scarfì F, Ferracin M *et al.* Basal cell carcinoma: a comprehensive review. *Int J Mol Sci* 2020; **21**:5572.
43 Pellegrini C, Maturo MG, Di Nardo L *et al.* Understanding the molecular genetics of basal cell carcinoma. *Int J Mol Sci* 2017; **18**:2485
44 Bonilla X, Parmentier L, King B *et al.* Genomic analysis identifies new drivers and progression pathways in skin basal cell carcinoma. *Nat Genet* 2016; **48**:398–406.
45 Nissen NI, Karsdal M, Willumsen N. Collagens and cancer associated fibroblasts in the reactive stroma and its relation to cancer biology. *J Exp Clin Cancer Res* 2019; **38**:115.
46 Xu S, Xu H, Wang W *et al.* The role of collagen in cancer: from bench to bedside. *J Transl Med* 2019; **17**:309.
47 Giussani M, Landoni E, Merlino G *et al.* Extracellular matrix proteins as diagnostic markers of breast carcinoma. *J Cell Physiol* 2018; **233**:6280–90.
48 Wang X, Bai Y, Zhang F *et al.* Prognostic value of COL10A1 and its correlation with tumor-infiltrating immune cells in urothelial bladder cancer: a comprehensive study based on bioinformatics and clinical analysis validation. *Front Immunol* 2023; **14**:955949.
49 Zhang M, Jin M, Gao Z *et al.* High COL10A1 expression potentially contributes to poor outcomes in gastric cancer with the help of LEF1 and Wnt2. *J Clin Lab Anal* 2022; **36**:e24612.
50 Chen S, Wei Y, Liu H *et al.* Analysis of collagen type X alpha 1 (COL10A1) expression and prognostic significance in gastric cancer based on bioinformatics. *Bioengineered* 2021; **12**:127–37.
51 Raudvere U, Kolberg L, Kuzmin I *et al.* g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res* 2019; **47**:W191–8.