



UNIVERSIDADE ESTADUAL DE CAMPINAS  
SISTEMA DE BIBLIOTECAS DA UNICAMP  
REPOSITÓRIO DA PRODUÇÃO CIENTÍFICA E INTELLECTUAL DA UNICAMP

**Versão do arquivo anexado / Version of attached file:**

Versão do Editor / Published Version

**Mais informações no site da editora / Further information on publisher's website:**

<https://www.nature.com/articles/s41598-019-44156-9>

**DOI: 10.1038/s41598-019-44156-9**

**Direitos autorais / Publisher's copyright statement:**

©2019 by Nature Publishing Group. All rights reserved.

DIRETORIA DE TRATAMENTO DA INFORMAÇÃO

Cidade Universitária Zeferino Vaz Barão Geraldo

CEP 13083-970 – Campinas SP

Fone: (19) 3521-6493

<http://www.repositorio.unicamp.br>

# SCIENTIFIC REPORTS



OPEN

## Exploring Collagen Parameters in Pure Special Types of Invasive Breast Cancer

Rodrigo de Andrade Natal<sup>1</sup>, Geisilene R. Paiva<sup>2</sup>, Vitor B. Pelegati<sup>3</sup>, Ludwing Marengo<sup>3</sup>, César A. Alvarenga<sup>4</sup>, Renato F. Vargas<sup>2</sup>, Sophie F. Derchain<sup>5</sup>, Luis O. Sarian<sup>5</sup>, Camille Franchet<sup>6</sup>, Carlos L. Cesar<sup>3,7</sup>, Fernando C. Schmitt<sup>8,9</sup>, Britta Weigelt<sup>10</sup> & José Vassallo<sup>1</sup>

One of the promising tools to evaluate collagen in the extracellular matrix is the second-harmonic generation microscopy (SHG). This approach may shed light on the biological behavior of cancers and their taxonomy, but has not yet been applied to characterize collagen fibers in cases diagnosed as invasive breast carcinoma (BC) of histological special types (IBC-ST). Tissue sections from 99 patients with IBC-ST and 21 of invasive breast carcinoma of no special type (IBC-NST) were submitted to evaluation of collagen parameters by SHG. Tissue microarray was performed to evaluate immunohistochemical-based molecular subtype. In intratumoral areas, fSHG and bSHG (forward-SHG and backward-SHG) collagen parameters achieved their lowest values in mucinous, papillary and medullary carcinomas, whereas the highest values were found in classic invasive lobular and tubular carcinomas. Unsupervised hierarchical cluster analysis and minimal spanning tree using intratumoral collagen parameters allowed the identification of three main groups of breast cancer: group A (classic invasive lobular and tubular carcinomas); group B (IBC-NST, metaplastic, invasive apocrine and micropapillary carcinomas); and group C (medullary, mucinous and papillary carcinomas). Our findings provide further characterization of the tumor microenvironment of IBC-ST. This understanding may add information to build more consistent tumor categorization and to refine prognostication.

Invasive breast carcinoma is considered a combination of heterogeneous diseases, encompassing multiple entities with distinct biological and clinical features. The most common subtype of breast carcinoma is invasive ductal carcinoma of no special type (IBC-NST), formerly known as invasive ductal carcinoma, which accounts for around 50% to 80% of all breast cancers<sup>1-3</sup>. The remaining tumor subtypes are collectively referred to as invasive carcinoma with histological special types (IBC-ST), and account for approximately 25% of invasive breast carcinomas<sup>1-4</sup>. Although pathologists have been aware of the diversity of breast cancers and have endeavored to devise approaches to classify the disease into meaningful groups<sup>3,5-9</sup>, this concept has only been brought to the forefront of breast cancer research after the publication of high-throughput microarray-based discovery studies that unraveled the existence of multiple molecular subtypes<sup>10-13</sup>.

<sup>1</sup>Laboratory of Investigative and Molecular Pathology, CIPED – Faculty of Medical Sciences – State University of Campinas, Rua Tessália Vieira de Camargo, 126, Zip code: 13083-970, Campinas, São Paulo, Brazil. <sup>2</sup>Laboratory of Specialized Pathology, LAPE – Faculty of Medical Sciences – State University of Campinas, Rua Tessália Vieira de Camargo, 126, Zip code: 13083-970, Campinas, São Paulo, Brazil. <sup>3</sup>Department of Quantum Electronics – Institute of Physics “Gleb Wataghin” – State University of Campinas, Rua Sérgio Buarque de Holanda, 777, Zip code: 13083-859, Campinas, São Paulo, Brazil. <sup>4</sup>Instituto de Patologia de Campinas (Private Laboratory), Av. Andrade Neves, 1801, Zip Code: 13070-000, Campinas, São Paulo, Brazil. <sup>5</sup>Department of Obstetrics and Gynecology – Faculty of Medical Sciences – State University of Campinas, Rua Tessália Vieira de Camargo, 126, Zip code: 13083-970, Campinas, São Paulo, Brazil. <sup>6</sup>Department of Pathology, University Cancer Institute, Avenue Irene Joliot Curie, 1, Zip code: 31059, Toulouse, France. <sup>7</sup>Department of Physics, Federal University of Ceará (UFC), Campus do Pici - Bloco 922 - Zip code: 60455-760, Fortaleza, Ceará, Brazil. <sup>8</sup>Institute of Molecular Pathology and Immunology of Porto University (IPATIMUP) – Porto University, Rua Dr. Roberto Frias, s/n, Zip code: 4200-465, Porto, Portugal. <sup>9</sup>National Santé Laboratory, Department of Medicine – L-3555, Dudelange, Luxembourg. <sup>10</sup>Department of Pathology – Memorial Sloan Kettering Cancer Center, York Avenue 1275, Zip code: 10065, New York, USA. Correspondence and requests for materials should be addressed to R.d.A.N. (email: [rodrigo.natal.med@gmail.com](mailto:rodrigo.natal.med@gmail.com)) or J.V. (email: [vassallomeister@gmail.com](mailto:vassallomeister@gmail.com))

Histological types	n (%)	Median (IQR <sup>#</sup> ), years	Histological grade			Lymph node status			Molecular subtype				
			I	II	III	Negative	Positive	Not available	HR+HER2–	HR+HER2+	HR–HER2+	HR–HER2–	Not available
Classic invasive lobular	7 (07.1%)	57 (57–59)	3	2	2	2	1	4	5	—	—	1	1
Tubular	13 (13.1%)	50 (45–64)	13	—	—	9	2	2	8	—	—	—	5
Mucinous	23 (23.2%)	65 (47–74)	17	5	1	11	4	8	15	—	—	—	8
Mucinous A	13 (13.1%)	66 (54–76)	12	1	—	6	3	4	8	—	—	—	5
Mucinous B	10 (10.1%)	63 (42–70)	5	4	1	5	1	4	7	—	—	—	3
Papillary	9 (09.1%)	61 (52–79)	1	5	3	6	1	2	8	—	—	—	1
Micropapillary	6 (06.1%)	48 (42–63)	—	2	4	1	4	1	4	1	—	—	1
Medullary	15 (15.2%)	47 (38–54)	—	2	13	9	4	2	1	—	—	10	4
Typical	6 (06.1%)	48 (36–58)	—	—	6	4	2	—	—	—	—	6	—
Atypical	9 (09.1%)	47 (45–51)	—	2	7	5	2	2	1	—	—	4	4
Metaplastic	9 (09.1%)	51 (33–56)	—	4	5	2	—	7	—	—	—	8	1
Matrix-producing	4 (04.0%)	56 (50–59)	—	2	2	1	—	3	—	—	—	4	—
Squamous cell	4 (04.0%)	40 (35–46)	—	1	3	1	—	3	—	—	—	3	1
Spindle cell	1 (01.0%)	8 (NA)	—	1	—	—	—	1	—	—	—	1	—
Apocrine	17 (17.1%)	61 (46–69)	—	3	14	8	6	3	4	3	4	3	3
<b>Total (IBC-ST)</b>	<b>99 (100.0%)</b>	<b>56 (45–66)</b>	<b>43</b>	<b>29</b>	<b>48</b>	<b>54</b>	<b>25</b>	<b>41</b>	<b>45</b>	<b>04</b>	<b>04</b>	<b>22</b>	<b>24</b>
IBC-NST	21	58 (50–61)	9	6	6	6	3	12	10	—	—	4	7

**Table 1.** Clinicopathological features of breast carcinoma cases. <sup>#</sup>IQR = interquartile range.

IBC-ST refers to the architectural growth of the tumors, defining special morphological and cytological patterns, which have been consistently associated with distinctive clinical presentation and/or outcomes<sup>3</sup>. However, although histological grade identifies prognostic subgroups in IBC-NST, among the IBC-ST some subtypes show high grade but still bear favorable prognosis (e.g. medullary, mucinous and invasive apocrine breast carcinomas)<sup>3</sup>. Interestingly, genetic and transcriptomic features of breast cancers has been correlated with histological grade, and microarray-based genomic signatures for histological grades have been devised<sup>14–17</sup>.

The use of new investigation tools in IBC-ST has been limited partly due to their relative low prevalence with consequent lower interobserver reproducibility, skewing their systematic investigation in class discovery<sup>10–13</sup> and class prediction<sup>16,18–21</sup>. This is the case of tumor microenvironment, especially the extracellular matrix (ECM), which has been intensively studied in tumor progression and outcome<sup>22</sup>. Among the various components of ECM, collagen fibers are the most abundant in breast cancer stroma<sup>23,24</sup>. The restructuring of collagen pattern may be either induced by tumor cells, or occur as a result of an integrated modification between neoplastic cell and stroma, suggesting that progression of a tumor relies not only on changes of cell function, but also on ECM alterations.

One of the promising tools to evaluate collagen in the ECM during cancer progression is the second-harmonic generation microscopy (SHG)<sup>25–27</sup>. This approach may shed light on the biological behavior of cancers and their taxonomy<sup>28,29</sup>. It allows for the identification of collagen maturation process by distinguishing collagen diameters. Thus, backward SHG images exhibit punctate distribution attributable to small-diameter, segmental collagen, which permits recognition of fibrillogenesis in immature tissue by directly imaging backward-propagating SHG (bSHG). By contrast, images from the mature fibrils are identical in the forward (fSHG) and backward directions<sup>30</sup>. Although it has not been applied to IBC-ST, several studies have shown that this is an efficient means to describe common breast diseases, differentiate benign from malignant breast tissue and establish prognosis in invasive ductal carcinoma<sup>31–35</sup>. It should be emphasized that studies also indicate that certain aspects of collagen deposition may lead to a more aggressive behavior of the tumor, like greater amount of collagen, presence of perpendicular fibers to the tumor boundary and thicker collagen fibers<sup>34,36–38</sup>. Thus, the goals of the present study are: (1) to investigate peri- and intratumoral collagen parameters in various histological subtypes of pure IBC-ST, using IBC-NST as standard samples, and (2) correlate these parameters with pathological and clinical features.

## Results

**Clinicopathological features.** Patients' age, histological grade, lymph node status and immunohistochemical-based molecular subtype<sup>39</sup> are shown in Table 1. The median age was 56 years [Interquartile range (IQR) = 45–66 years]. Among all 120 cases, most carcinomas were of histological grade III (n = 48; 40.0%), followed by grade I (n = 43; 35.8%) and grade II (n = 29; 24.2%) carcinomas. The majority of the patients did not present lymph node metastases (n = 54; 45.0%); these were found in 25 patients (20.8%). According to the status of immunohistochemical expression of hormone receptors (HR) and HER2/*neu* protein used as surrogate for molecular types, cases were classified as follows: HR+HER2– was the most frequent subtype (n = 55; 45.8%), followed by HR–HER2– (n = 26; 21.7%), HR+HER2+ (n = 4; 3.3%), and HR–HER2+ (n = 4; 3.3%). In 31 cases (25.8%), patients had no further material for immunohistochemical staining, thus immunohistochemical-based molecular subtyping was not possible for these patients.

Collagen parameters	IBC-NST	Classic invasive lobular	Tubular	Mucinous	Papillary	Micropapillary	Medullary	Metaplastic	Apocrine
<b>Peritumoral</b>									
bSHG collagen quantity <sup>†</sup>	13.61 (0.42)	13.84 (0.46)	13.86 (0.34)	13.58 (0.42)	13.39 (0.31)	13.74 (0.52)	13.34 (0.51)	13.20 (0.47)	13.48 (0.35)
fSHG collagen quantity <sup>†</sup>	14.03 (0.48)	14.19 (0.52)	14.05 (0.35)	13.85 (0.43)	13.75 (0.56)	14.13 (0.52)	13.79 (0.65)	13.75 (0.43)	13.87 (0.30)
bSHG collagen uniformity <sup>†</sup>	5.08 (0.50)	5.40 (0.56)	5.36 (0.39)	5.11 (0.51)	4.81 (0.38)	5.24 (0.56)	4.82 (0.58)	4.64 (0.52)	4.89 (0.42)
fSHG collagen uniformity <sup>†</sup>	5.84 (0.70)	6.07 (0.69)	5.82 (0.47)	5.67 (0.58)	5.37 (0.92)	5.94 (0.58)	5.67 (0.81)	5.59 (0.48)	5.63 (0.49)
bSHG collagen organization <sup>†</sup>	0.16 (0.04)	0.19 (0.04)	0.17 (0.05)	0.15 (0.04)	0.16 (0.05)	0.17 (0.06)	0.16 (0.04)	0.15 (0.03)	0.15 (0.03)
fSHG collagen organization <sup>†</sup>	0.44 (0.07)	0.49 (0.08)	0.48 (0.08)	0.44 (0.08)	0.44 (0.08)	0.49 (0.17)	0.45 (0.05)	0.42 (0.05)	0.45 (0.09)
<b>Intratumoral</b>									
bSHG collagen quantity <sup>†</sup>	11.68 (0.73)	12.21 (0.80)	12.31 (0.66)	10.80 (0.76)	10.64 (0.17)	10.64 (0.50)	10.73 (0.41)	11.26 (0.43)	11.30 (0.74)
fSHG collagen quantity <sup>†</sup>	11.56 (1.04)	12.00 (0.74)	12.13 (0.82)	9.83 (1.68)	9.21 (0.95)	11.40 (0.76)	9.83 (1.14)	11.21 (0.62)	11.02 (1.02)
bSHG collagen uniformity <sup>†</sup>	2.86 (0.86)	3.61 (1.02)	3.50 (0.90)	1.93 (1.01)	1.45 (0.25)	3.21 (0.61)	1.67 (0.53)	2.28 (0.53)	2.44 (0.94)
fSHG collagen uniformity <sup>†</sup>	2.86 (1.31)	3.44 (0.74)	3.25 (1.18)	0.99 (2.37)	-0.42 (1.54)	3.04 (1.05)	0.83 (1.73)	2.44 (1.02)	2.35 (1.60)
bSHG collagen organization <sup>†</sup>	0.08 (0.04)	0.14 (0.06)	0.11 (0.04)	0.04 (0.04)	0.01 (0.02)	0.10 (0.04)	0.04 (0.03)	0.05 (0.04)	0.07 (0.04)
fSHG collagen organization <sup>†</sup>	0.37 (0.12)	0.46 (0.08)	0.41 (0.10)	0.18 (0.16)	0.08 (0.12)	0.44 (0.11)	0.21 (0.14)	0.33 (0.13)	0.34 (0.13)

**Table 2.** Collagen parameters in invasive breast carcinoma. <sup>†</sup>Data in: Mean (Standard deviation).

**Collagen features in breast cancer.** The correlation between collagen parameters (quantity, uniformity and organization) in peritumoral areas was poor, contrasting to high correlation coefficients found in intratumoral areas (Supplementary Table 1 and Fig. 1).

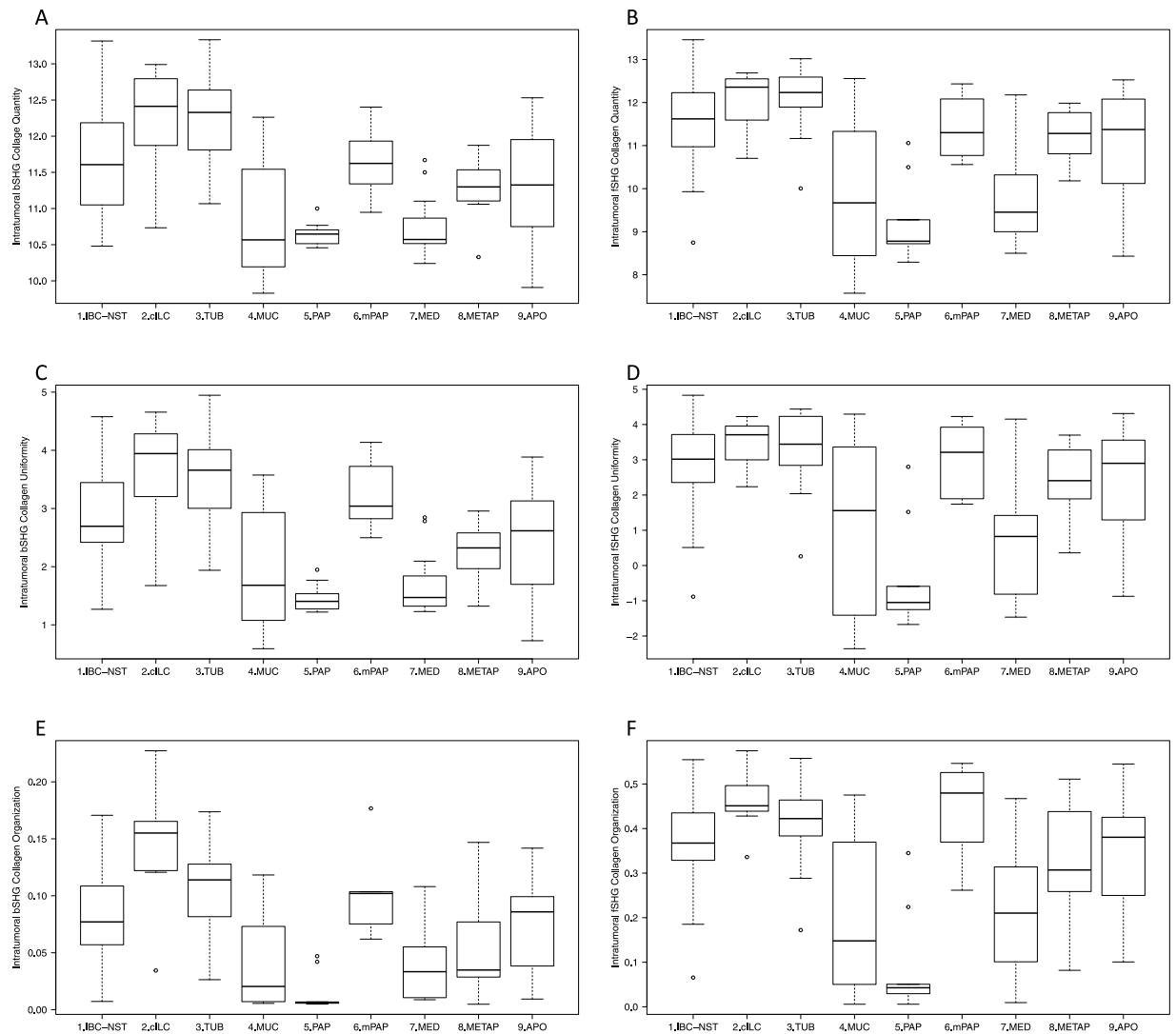
In general, collagen quantity, uniformity and organization were higher in peritumoral regions compared to intratumoral areas (Supplementary Table 2). In classic invasive lobular, tubular, micropapillary and metaplastic carcinomas, there was no statistical difference between fSHG peri- and intratumoral collagen organization. The same was observed for bSHG peri- and intratumoral collagen organization in classic invasive lobular and micropapillary carcinomas.

Peritumoral collagen parameters for both fSHG and bSHG fibers did not differ significantly across histological subtypes (Table 2 and Supplementary Fig. 2). In contrast, in intratumoral areas, fSHG and bSHG collagen quantity, uniformity and organization achieved the lowest values in mucinous, papillary and medullary carcinomas, whereas the highest values for these parameters were found in classic invasive lobular and tubular carcinomas (Table 2, Fig. 1, Supplementary Table 3).

Peritumoral and intratumoral bSHG collagen quantity ( $p = 0.008$  and  $p = 0.003$ , respectively) and uniformity ( $p = 0.009$  and  $p = 0.005$ , respectively) were higher in histological grade I tumors compared to grade III tumors. However, there was no difference between peri- and intratumoral collagen quantity and uniformity, considering histological grade II tumors and grade I or histological grade II and grade III tumors. Peri- and intratumoral collagen organization was similar across histological grades (Supplementary Table 4). Peritumoral collagen quantity, uniformity and organization did not differ in relation to lymph node status and immunohistochemical-based subtype. However, intratumoral collagen bSHG fiber uniformity and organization were higher in HR-HER2- tumors compared to HR-HER2+ tumors (Supplementary Table 5).

The multivariate recursive partitioning method, using intratumoral collagen parameters, showed that fSHG and bSHG quantity, uniformity and organization were lower in mucinous, papillary and medullary tumors ( $p < 0.001$ ). On the other hand, the three parameters were significantly higher in classic invasive lobular and tubular carcinomas ( $p = 0.005$ ) compared to the other histological subtype. IBC-NST, micropapillary, metaplastic and invasive apocrine carcinomas presented an intermediary fSHG and bSHG intratumoral collagen parameters (Supplementary Fig. 3).

Unsupervised hierarchical cluster analysis using intratumoral collagen (Fig. 2 and Supplementary Figs 4 and 5) allowed the identification of three main groups of breast cancer. Group A: high intratumoral collagen quantity, uniformity and organization (classic invasive lobular and tubular carcinomas); group B: intermediary intratumoral collagen quantity, uniformity and organization (invasive ductal, metaplastic, invasive apocrine and micropapillary carcinomas); and group C: low intratumoral collagen quantity, uniformity and organization (medullary, mucinous and papillary carcinomas). Supervised analysis using estrogen receptor (ER) expression or molecular subtype presented the same result. Figure 3 visually summarizes the distinct groups of breast carcinomas according to intratumoral collagen parameters.



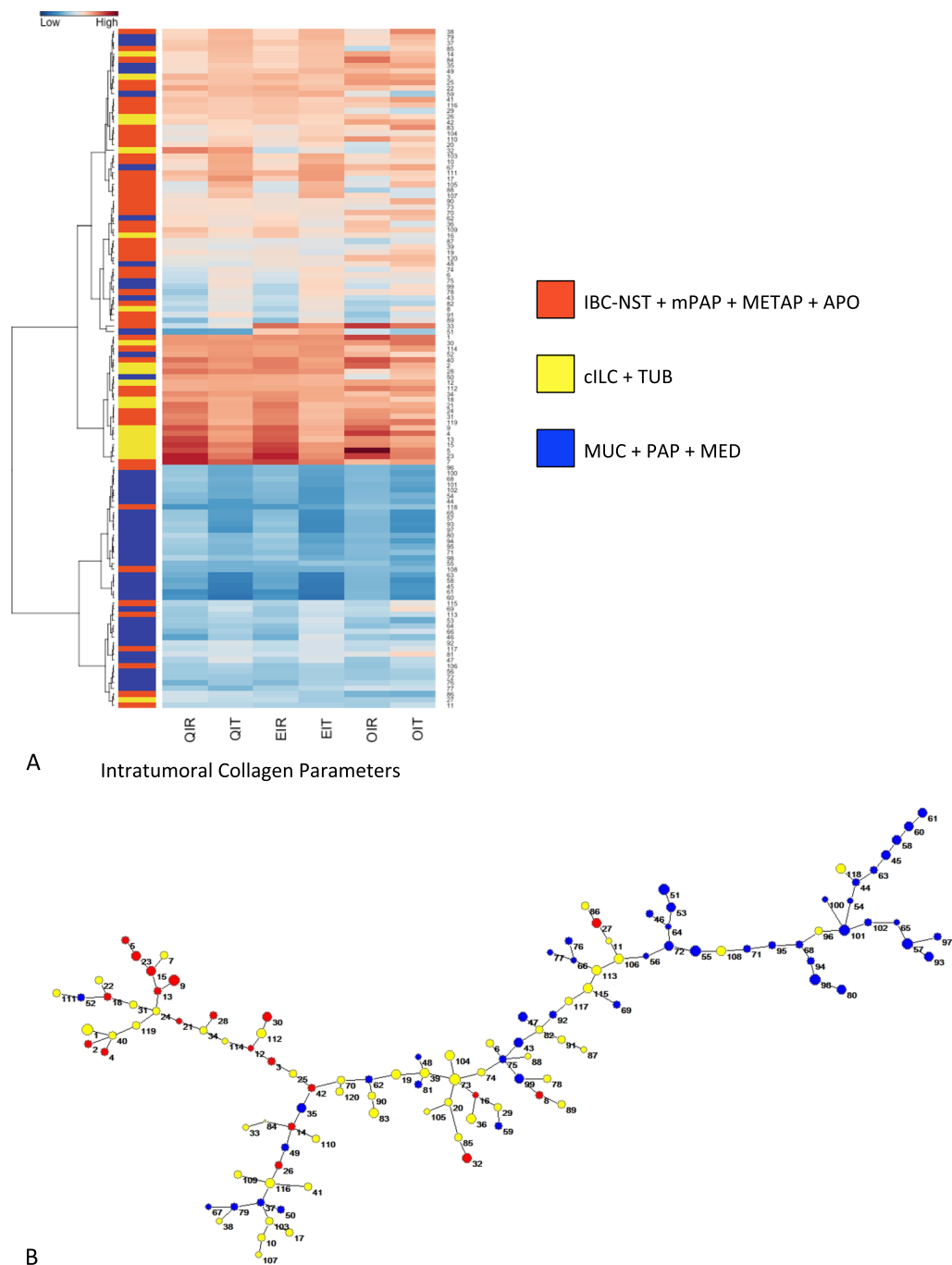
**Figure 1.** Boxplots demonstrating the distribution of intratumoral collagen parameters in histological subtypes: **(A)** bSHG collagen quantity; **(B)** fSHG collagen quantity; **(C)** bSHG collagen uniformity; **(D)** fSHG collagen uniformity; **(E)** bSHG collagen organization; **(F)** fSHG collagen organization. Classic invasive lobular (2. cILC) and tubular (3. TUB) carcinomas presented the highest values for intratumoral collagen parameters. Mucinous (4. MUC), papillary (5. PAP) and medullary (7. MED) presented the lowest values. Invasive ductal carcinoma of no special type (1. IBC-NST), micropapillary (6. mPAP), metaplastic (8. METAP) and invasive apocrine (9. APO) presented intermediary values of intratumoral collagen parameters. The statistically significant relations are shown in Supplementary Table 3. [bSHG, fSHG: backward and forward propagation second harmonic generation, respectively].

**Mucinous, medullary and metaplastic breast carcinoma subclassification.** Collagen parameters did not differ between mucinous A and B breast carcinoma. In addition, peritumoral collagen parameters in typical and atypical medullary were similar. Intratumoral fSHG collagen quantity ( $p = 0.017$ ) and uniformity ( $p = 0.044$ ) were higher in atypical medullary breast cancer; these findings may be extended to bSHG fiber quantity ( $p = 0.054$ ) and uniformity ( $p = 0.054$ ); fSHG and bSHG fiber organization was similar in both tumor subtypes ( $p = 0.198$  and  $p = 0.064$ , respectively).

There were no differences in collagen parameters comparing metaplastic breast carcinoma with squamous cells and with matrix-producing breast carcinomas, except for peritumoral fSHG collagen organization, that was higher in metaplastic breast carcinoma with squamous cells ( $p = 0.017$ ).

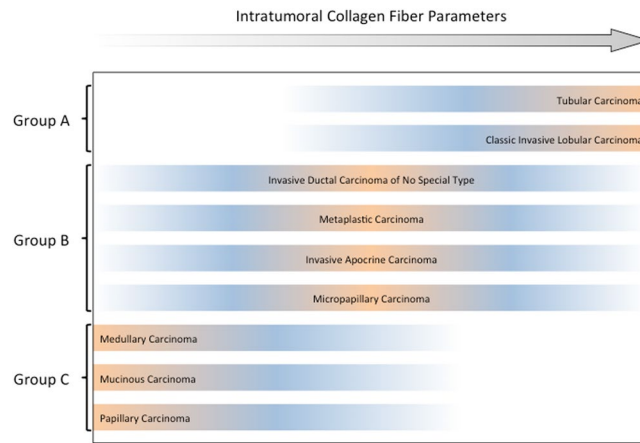
## Discussion

Our results allowed grouping breast carcinomas according to the intratumoral collagen parameters (Fig. 3). Group A, characterized by higher intratumoral collagen quantity, uniformity and organization, was enriched for classic invasive lobular and tubular carcinomas; group B, comprised by intermediary intratumoral collagen quantity, uniformity and organization, was enriched for IBC-NST, metaplastic, invasive apocrine and micropapillary carcinomas; and group C, covered lower intratumoral collagen quantity, uniformity and organization, was enriched for



**Figure 2.** (A) Unsupervised hierarchical clustering of intratumoral collagen parameters using Ward algorithms and Euclidian distance; the lowest values for intratumoral collagen parameters refer to mucinous (MUC), papillary (PAP) and medullary (MED) breast carcinomas (light blue); the highest values refer to classical invasive lobular (cILC) and tubular carcinomas (TUB) (light yellow); the intermediary values refer to invasive ductal carcinoma of no special type (IBC-NST), micropapillary (mPAP), metaplastic (METAP) and invasive apocrine (APO) (light orange), being the latter group very heterogeneous. QIR: bSHG collagen quantity; QIT: fSHG collagen quantity; EIR: bSHG collagen uniformity; EIT: fSHG collagen uniformity; OIR: bSHG collagen organization; and OIT: fSHG collagen organization. (B) Minimal Spanning Tree (MST) of the Euclidian matrix with three defined intratumoral collagen parameters. The MSP gives a spatial representation of what is shown in the heatmap. Cases presenting similar collagen parameters are linked by a line. As result, cases belonging to the same group (blue, yellow and orange) are concentrated in three branches of the tree: blue in the superior part, yellow in the middle and right region and orange in the left branch.

medullary, mucinous and papillary carcinomas. Furthermore, it would be worth highlighting that IBC-NST are not only heterogeneous at the morphologic and molecular level, but also at the clustering analysis performed herein; cases of IBC-NST could be found in any of the possible three groups identified, even if we used a supervised clustering analysis.



**Figure 3.** Distribution of breast cancer histological subtypes per intratumoral collagen parameters. Tubular and invasive lobular breast carcinoma showed the highest intratumoral collagen parameters (quantity, uniformity and organization), while medullary, papillary and mucinous breast carcinoma presented the lowest intratumoral collagen parameters. Invasive ductal, metaplastic, invasive apocrine and micropapillary breast carcinomas presented intermediary intratumoral collagen parameters.

The findings on intratumoral collagen corroborate the disparities between morphological classification and molecular biology. Heterogeneous tumor biology may be found within the same histological subtype, whereas distinct histological categories share common molecular features<sup>39,40</sup>. As such, morphologic subtyping may be challenging not only because of subjectivity, but also as a result of molecular similarities. Peritumoral collagen fibers presented similar parameters among all histological subtypes of breast cancer, suggesting a common endpoint, shared by them, in the mechanism of collagen deposition at tumor periphery. This acknowledgement may support the pivotal interplay of tumor microenvironment in cancer outcome.

Classic invasive lobular and tubular carcinomas, constituting group A in the present study, have been described as showing remarkably similar immunohistochemical and transcriptomic profiles<sup>39</sup>. Notwithstanding, invasive lobular carcinoma can be differentiated from tubular carcinoma based on the expression levels of E-cadherin, which is absent in the former<sup>41,42</sup>. Regarding collagen parameters, both histological subtypes presented similar intratumoral features, and shared the same hierarchical clustering, supporting the acquaintance between them. These findings provide structural evidence for the hypothesis that classic invasive lobular and tubular carcinomas, both frequently members of the low-grade breast carcinoma group, might originate from the same family of low-grade precursors<sup>41</sup>. Lopez-Garcia *et al.* support the existence of a “low-grade breast neoplasia family”; nevertheless, the transcriptomes of these lesions display small, yet important differences, which together with their distinct biological behavior, warrant their separation as discrete entities<sup>43</sup>.

In opposition to the subtypes above, a clear disparity in the clinical, morphological, molecular or immunohistochemical profiles was found within the other two groups of tumors with similar intratumoral collagen parameters, groups B and C<sup>39,40,44</sup>. This is evidence that, although the microenvironment may importantly contribute to progression of cancer in all phases, factors intrinsic to the neoplastic cells may respond for other important features, as the phenotype and genetic characteristics<sup>39,40</sup>.

Subtypes included in our group B have been described as presenting discrepant clinicopathological behavior. Metaplastic and micropapillary carcinomas present mostly poor outcome, in spite of ER status, respectively, negative and positive; apocrine carcinoma, typically ER-negative, present variable outcome<sup>45,46</sup>. In a study of genomic profiling, carcinomas with apocrine differentiation, invasive micropapillary and metaplastic carcinomas showed the highest frequencies of chromosome aberrations (i.e. gene amplifications, gains and losses)<sup>40</sup>. Thus, although this group seems heterogeneous as for clinical outcome and for the variable expression of ER, the collagen parameters behaved similarly, and paralleled the high numbers of chromosome aberrations. It may be assumed that the collagen signature of group B is predictive of higher frequency of finding genetic alterations.

On the other hand, group C, enriched for medullary, mucinous and papillary carcinomas have been related to a more indolent clinical behavior<sup>45</sup>. The three subtypes in group C have been described with similar gross growth pattern in image exams, presenting as a round or oval tumor with well-defined contours<sup>47</sup>. Genetic results demonstrated that papillary carcinomas are a homogenous special histological subtype of breast carcinoma, with good prognosis. They tend to present low rates of lymph node metastasis and low p53 expression, ER-positive, low frequency of gene copy number aberrations and high prevalence of PIK3CA mutations. The genomic profile of its three morphological variants (i.e. encapsulated, solid and invasive papillary) has shown remarkably similar<sup>46</sup>. Likewise, mucinous breast carcinoma display ER-positivity, low level of genetic instability and rare recurrent amplifications<sup>48–51</sup>. These genetic features reasonably approximate both tumors in hierarchic clusters<sup>39</sup>. In opposition, medullary carcinomas have been shown to share the basal-like immunophenotype (i.e. HR negative, HER-2 negative and P-cadherin positive), and high frequency of chromosome aberrations<sup>39,40,52–55</sup>. In spite of this divergence in profiling, medullary carcinomas show, in common with the other two members of group C,

the favorable prognosis and the lower intratumoral collagen quantity, uniformity and organization. As such, this group is homogeneous only in respect to clinical behavior for all three members<sup>45</sup>.

Metaplastic and medullary carcinomas, grouped herein as B and C, respectively, are consistently of triple-negative phenotype, and both were shown to display similar high frequency of gene copy number aberrations and comparable gene expression patterns<sup>39,40,52–55</sup>. In relation to the concordance between the molecular subtyping and other prognostic gene signatures, patients with the basal-like type should be classified as aggressive tumor<sup>9,56</sup>. However, basal-like medullary carcinomas present a good outcome in opposition to metaplastic carcinoma<sup>3</sup>. Then, in contrast to molecular grouping, intrinsic intratumoral collagen parameters clustering proposed in this study assembled these subtypes of breast carcinomas more suitably.

The rarity of the entities analyzed herein represents a drawback. Additional studies of special subtypes of breast carcinoma will be required to validate these findings, and to determine the biological and clinical relevance of clustering these tumors according to the intratumoral collagen parameters. In special, studies on patients' outcome for each individual subtype will require larger cohorts. Also, the long collection period might affect the evaluation of collagen structure, although previous studies pointed out that tissue fixation time did not interfere with the availability of macromolecules<sup>57,58</sup>. Further, the present study evaluated images only with the 40×/1.3 oil immersion magnification, largely used in the literature. It would be interesting to specifically address the comparison of collagen parameters values between different magnifications, in order to evaluate the impact of microscopic resolution for these analyses.

Previous studies have already demonstrated the prognostic value of collagen parameters in breast cancer, showing that high disorganization of collagen fibers, i. e. TACS3, is indicative of poorer prognosis<sup>34,37</sup>. In a study on the evaluation of fibrils deposition in collagen fibers, it was also found that different patterns are related to the molecular subtypes<sup>39</sup>. However, most of them did not emphasize the usage of this technique on IBC-ST, with the exception of the study by Conklin *et al.*, in which collagen parameters were evaluated in invasive lobular carcinoma<sup>34</sup>. In addition, these studies have not compared different collagen signatures in the histological subtypes. Further, the method applied in the present study was quantitative, in contrast to the qualitative or semiquantitative approach of the previous ones. Besides, the results presented herein could provide the basis for future automated analyses with specific mathematical algorithms for classification, as already described<sup>60</sup>. However, this application is still potential, as, to the best of our knowledge, our study represents the first approach to compare collagen parameters in IBC-ST.

In summary, our results do not allow a prognostic stratification, but provide information that could contribute to refine diagnosis. In that sense, an example would be the differential diagnosis between medullary carcinomas and triple-negative IBC-NST, since these entities belong to different groups, as proposed here. Our results provide a step forward to the characterization of the tumor microenvironment of IBC-ST. This understanding may add information to build more consistent tumor categorization, together with molecular findings. Grouping the different histological subtypes according to collagen parameters, as suggested herein, may represent a plausible explanation for the observation that tumors with different genetic profiles share similar clinical behavior. It potentially contributes to the refinement of the microenvironmental phylomorphology, and of the prognostication of breast cancer patients.

## Materials and Methods

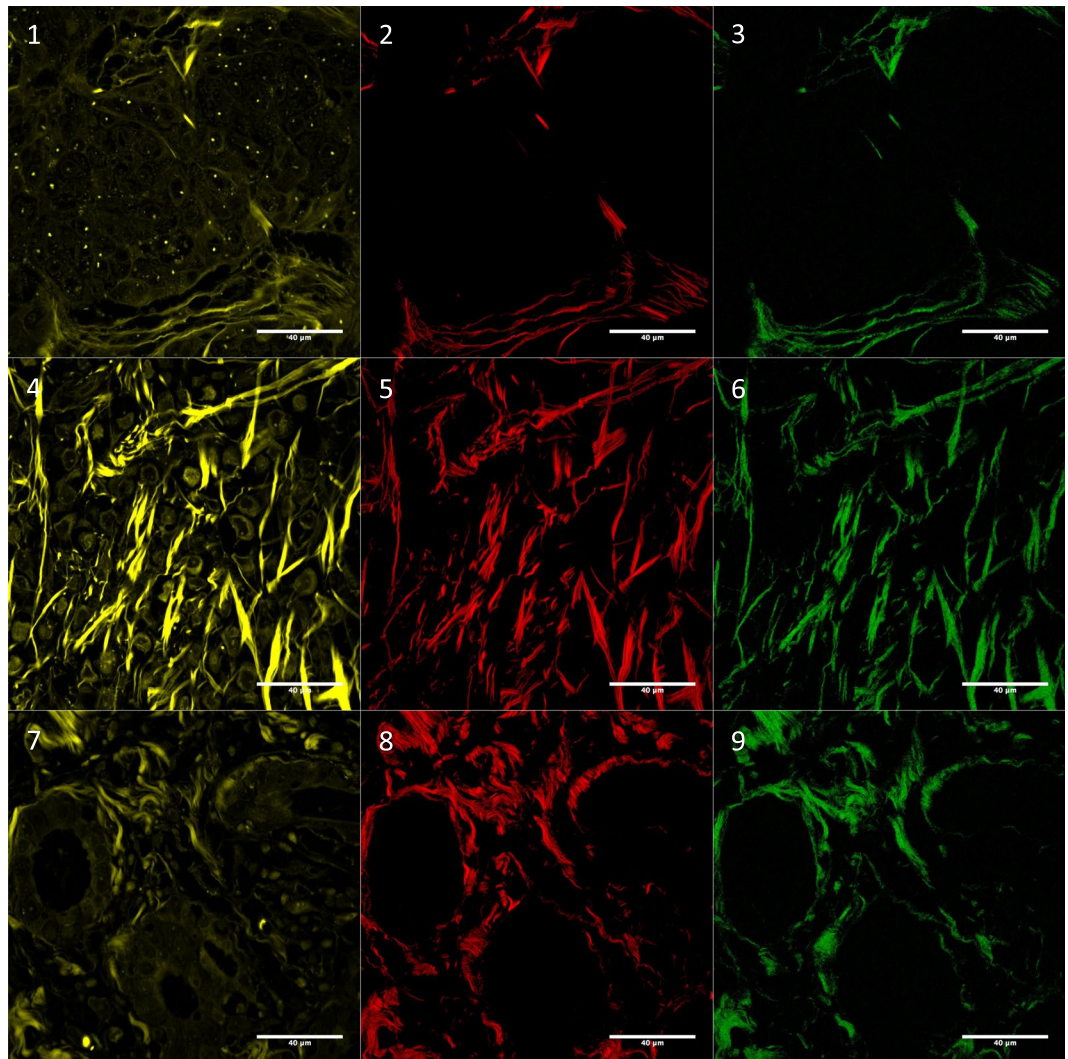
**Case selection.** Formalin-fixed, paraffin-embedded (FFPE) tissue specimens from 120 patients with pure invasive breast carcinoma were consecutively retrieved from four centers of Pathology (two in Brazil, one in Portugal and one in France). Tissues were obtained from primary lumpectomy or mastectomy; no patient had undergone primary chemo- or radiotherapy. Tumors were classified based on the WHO criteria<sup>3</sup>: 21 IBC-NST, 7 classic invasive lobular, 13 tubular, 23 mucinous, 6 invasive micropapillary, 9 invasive papillary, 15 medullary, 9 metaplastic and 17 apocrine carcinomas, from January 2000 to January 2015. Mucinous carcinoma were subdivided into hypocellular mucinous (mucinous A; n = 13) and hypercellular mucinous carcinoma (mucinous B; n = 10) based on the criteria of Capella *et al.*<sup>61</sup>. All cases were reviewed on hematoxylin and eosin-stained (H&E) sections by at least two experienced breast pathologists (GRP, CAA, RVF and FCS). The study is fully compliant with the Declaration of Helsinki (approved by the “Comitê de Ética em Pesquisa da Unicamp”).

In order to confirm diagnosis and evaluate immunohistochemical-based molecular subtype<sup>39,44</sup>, two representative intratumoral areas of 2 mm in diameter from each case were selected to build a tissue microarray (tissue microarray (TMA) builder 20010.02, Histopathology, Pécs, Hungary) for the immunohistochemical analysis. Further, three representative regions from peri- and intratumoral areas were selected on the same H&E sections (4 μm each), and marked for SHG methodology (Supplementary Information 1).

**Immunohistochemical staining and evaluation.** TMA sections were deparaffinized, rehydrated and submitted to antigen retrieval. The primary antibodies used, dilutions and detection methods are briefly described in Supplementary Table 6. Seven cases of medullary carcinoma kindly provided by Camille Franchet were analyzed in whole sections. The evaluation of immunohistochemical results was centrally performed by two experienced breast pathologists (CAA and FCS).

Staining for estrogen-receptor alpha (ER), progesterone receptor (PR) and androgen receptor (AR) were evaluated according to the previously described guidelines<sup>62</sup>. For p63, keratin 5 and P-cadherin, samples were scored as positive when ≥10% cells were immunoreactive<sup>63</sup>. Membrane expression for HER-2 and EGFR was evaluated as previously described<sup>64</sup>. Ki-67 was evaluated by counting 1000 neoplastic cells in areas with highest positivity; cases with <20% positive cells were considered to have low proliferation, and the others, high proliferation<sup>65</sup>. Tumors were classified based on St Gallen International Expert Consensus<sup>65</sup> (Supplementary Information 1).





**Figure 4.** Autofluorescence (01, 04 and 07), fSHG collagen (02, 05 and 08) and bSHG collagen (03, 06 and 09). Images correspond to: invasive ductal carcinoma no special type (01, 02 and 03), classic invasive lobular (04, 05 and 06) and tubular (07, 08 and 09). Autofluorescence was included in this Figure to show tissue architecture. The images were log-transformed. For more examples see Supplementary Fig. 1. [bSHG, fSHG: backward and forward propagation second harmonic generation, respectively].

**Second-harmonic generation (SHG) imaging.** SHG microscopy was performed on a Zeiss LSM 780-NLO inverted confocal system (Carl Zeiss AG, Göttingen, Germany) available at the National Institute of Sciences and Technology on Photonics Applied to Cell Biology (INFABiC, UNICAMP) (Supplementary Information 1, Fig. 4 and Supplementary Fig. 7). Three peri- and three intratumoral areas selected for collagen assessment on H&E stained sections were examined in this condition, and digital images were collected for further evaluation. Peritumoral regions were defined as the fibrous tissue at the tumor borders, in the transition with non-neoplastic tissue. Intratumoral regions corresponded to fibrous bands within groups of neoplastic cells. In cases with multiple tumor sections, the most representative one was chosen.

**Evaluation of collagen fibers.** SHG images were stratified regarding collagen parameters using image pattern analysis methods. Collagen parameter refers to the pixels characterization in the SHG image, being evaluated through three features: quantity, uniformity and organization. Quantitative analysis of collagen parameters was performed in SHG images using ImageJ (<http://imagej.nih.gov/ij/>) and OrientationJ plug-in<sup>66</sup>. For this purpose, 16 representative areas to cover entire image (256 × 256 pixels) were performed in each image. The final value of each parameter represents the average of the values obtained in the three images; the value of each image denotes the average of the 16 representative areas of each image. (Supplementary Information 1). IBC-NST, being the most common subtype of breast cancer, was used as a comparison parameter for the other cases. Therefore, special cases may present collagen parameters equal, higher or lower than IBC-NST. Furthermore, a previous report demonstrated that the area of the cellular component does not represent a confounding variable in the collagen parameters evaluation<sup>67</sup>, so hypo- or hypercellular tumors were evaluated in the same way.

**Statistical analyses.** Collagen quantity and uniformity were log-transformed to base  $e$  for statistical purposes; statistical analyses were performed using R (<https://cran.r-project.org/>). Shapiro-Wilk test was performed to analyze the data distribution of collagen parameters. Comparisons were performed using t-tests or analysis of variance (ANOVA). Pearson's Correlation coefficients were calculated to analyze the correlation between collagen parameters. A multivariate recursive partitioning model was fit to determine groups of collagen between breast cancer subtypes. Hierarchical clustering analysis was performed, using Euclidean distance and Ward's clustering algorithm<sup>49,51,68</sup> and minimal spanning tree (Supplementary Information 1).  $p$  values were corrected using Benjamini and Hochberg adjustment to minimize  $\alpha$  errors, and adjusted  $p$  values less than 0.05 were considered significant.

## References

- Weigelt, B., Geyer, F. C. & Reis-Filho, J. S. Histological types of breast cancer: how special are they? *Mol. Oncol.* **4**, 192–208 (2010).
- Weigelt, B. & Reis-Filho, J. S. Histological and molecular types of breast cancer: is there a unifying taxonomy? *Nat. Rev. Clin. Oncol.* **6**, 718–30 (2009).
- Lakhani, S., Ellis, I., Schnitt, S., Tan, P. & van de Vijver, M. WHO Classification of Tumours of the Breast, Fourth Edition - WHO - OMS (2012).
- Page, D. L. Special types of invasive breast cancer, with clinical implications. *Am. J. Surg. Pathol.* **27**, 832–5 (2003).
- Azzopardi, J. G., Ahmed, A. & Millis, R. R. Problems in breast pathology. *Major Probl. Pathol.* **11**(i–xvi), 1–466 (1979).
- Beckmann, M. *et al.* Biochemical and immunohistochemical detection of the epidermal growth-factor receptor (EGF-R) in breast-tumor specimens. *Int. J. Oncol.* **3**, 389–97 (1993).
- Huvos, A. G., Lucas, J. C. & Foote, F. W. Metaplastic breast carcinoma. *Rare form of mammary cancer.* *N. Y. State J. Med.* **73**, 1078–82 (1973).
- Weigelt, B. *et al.* Mucinous and neuroendocrine breast carcinomas are transcriptionally distinct from invasive ductal carcinomas of no special type. *Mod. Pathol.* **22**, 1401–14 (2009).
- Weigelt, B., Kreike, B. & Reis-Filho, J. S. Metaplastic breast carcinomas are basal-like breast cancers: a genomic profiling analysis. *Breast Cancer Res. Treat.* **117**, 273–80 (2009).
- Hu, Z. *et al.* The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics* **7**, 96 (2006).
- Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747–52 (2000).
- Sørli, T. *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. USA* **98**, 10869–74 (2001).
- Sorlie, T. *et al.* Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc. Natl. Acad. Sci. USA* **100**, 8418–23 (2003).
- Ivshina, A. V. *et al.* Genetic reclassification of histologic grade delineates new clinical subtypes of breast cancer. *Cancer Res.* **66**, 10292–301 (2006).
- Sotiriou, C. & Pusztai, L. Gene-expression signatures in breast cancer. *N. Engl. J. Med.* **360**, 790–800 (2009).
- Sotiriou, C. *et al.* Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J. Natl. Cancer Inst.* **98**, 262–72 (2006).
- Weigelt, B., Baehner, F. L. & Reis-Filho, J. S. The contribution of gene expression profiling to breast cancer classification, prognostication and prediction: a retrospective of the last decade. *J. Pathol.* **220**, 263–80 (2010).
- van't Veer, L. J. *et al.* Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415**, 530–6 (2002).
- van de Vijver, M. J. *et al.* A gene-expression signature as a predictor of survival in breast cancer. *N. Engl. J. Med.* **347**, 1999–2009 (2002).
- Wang, Y. *et al.* Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet (London, England)* **365**, 671–9 (2005).
- Wirapati, P. *et al.* Meta-analysis of gene expression profiles in breast cancer: toward a unified understanding of breast cancer subtyping and prognosis signatures. *Breast Cancer Res.* **10**, R65 (2008).
- Bizzarri, M. & Cucina, A. Tumor and the microenvironment: a chance to reframe the paradigm of carcinogenesis? *Biomed Res. Int.* **2014**, 934038 (2014).
- Keely, P. J., Wu, J. E. & Santoro, S. A. The spatial and temporal expression of the alpha 2 beta 1 integrin and its ligands, collagen I, collagen IV, and laminin, suggest important roles in mouse mammary morphogenesis. *Differentiation.* **59**, 1–13 (1995).
- Lagacé, R., Grimaud, J. A., Schürch, W. & Seemayer, T. A. Myofibroblastic stromal reaction in carcinoma of the breast: variations of collagenous matrix and structural glycoproteins. *Virchows Arch. A. Pathol. Anat. Histopathol.* **408**, 49–59 (1985).
- Burke, K., Tang, P. & Brown, E. Second harmonic generation reveals matrix alterations during breast tumor progression. *J. Biomed. Opt.* **18**, 31106 (2013).
- Campagnola, P. Second harmonic generation imaging microscopy: applications to diseases diagnostics. *Anal. Chem.* **83**, 3224–31 (2011).
- Han, X., Burke, R. M., Zettel, M. L., Tang, P. & Brown, E. B. Second harmonic properties of tumor collagen: determining the structural relationship between reactive stroma and healthy stroma. *Opt. Express* **16**, 1846–59 (2008).
- Mostaço-Guidolin, L. B. *et al.* Collagen morphology and texture analysis: from statistics to classification. *Sci. Rep.* **3**, 2190 (2013).
- Kiss, N. *et al.* Quantitative Analysis on *Ex Vivo* Nonlinear Microscopy Images of Basal Cell Carcinoma Samples in Comparison to Healthy Skin. *Pathol. Oncol. Res.* 1–7, <https://doi.org/10.1007/s12253-018-0445-1> (2018).
- Williams, R. M., Zipfel, W. R. & Webb, W. W. Interpreting second-harmonic generation images of collagen I fibrils. *Biophys. J.* **88**, 1377–86 (2005).
- Tao, Y. K. *et al.* Assessment of breast pathologies using nonlinear microscopy. *Proc. Natl. Acad. Sci. USA* **111**, 15304–9 (2014).
- Wu, Y. *et al.* Identifying three different architectural subtypes of mammary ductal carcinoma *in situ* using multiphoton microscopy. *J. Phys. D. Appl. Phys.* **48**, 405401 (2015).
- Tan, W. J. *et al.* Second harmonic generation microscopy is a novel technique for differential diagnosis of breast fibroepithelial lesions. *J. Clin. Pathol.* **68**, 1033–1035 (2015).
- Conklin, M. W. *et al.* Aligned collagen is a prognostic signature for survival in human breast carcinoma. *Am. J. Pathol.* **178**, 1221–32 (2011).
- Natal, R. A. *et al.* Tumor Biology. 1–12, <https://doi.org/10.1177/1010428318770953> (2018).
- Provenzano, P. P. *et al.* Collagen density promotes mammary tumor initiation and progression. *BMC Med.* **6**, 11 (2008).
- Provenzano, P. P. *et al.* Collagen reorganization at the tumor-stromal interface facilitates local invasion. *BMC Med.* **4**, 38 (2006).
- Natal, R. A. *et al.* Collagen analysis by second-harmonic generation microscopy predicts outcome of luminal breast cancer. *Tumor Biol.* **40**, 101042831877095 (2018).
- Weigelt, B. *et al.* Refinement of breast cancer classification by molecular characterization of histological special types. *J. Pathol.* **216**, 141–150 (2008).
- Horlings, H. M. *et al.* Genomic profiling of histological special types of breast cancer. *Breast Cancer Res. Treat.* **142**, 257–69 (2013).

41. Abdel-Fatah, T. M. A. *et al.* High frequency of coexistence of columnar cell lesions, lobular neoplasia, and low grade ductal carcinoma *in situ* with invasive tubular carcinoma and invasive lobular carcinoma. *Am. J. Surg. Pathol.* **31**, 417–26 (2007).
42. Esposito, N. N., Chivukula, M. & Dabbs, D. J. The ductal phenotypic expression of the E-cadherin/catenin complex in tubulolobular carcinoma of the breast: an immunohistochemical and clinicopathologic study. *Mod. Pathol.* **20**, 130–8 (2007).
43. Lopez-Garcia, M. A. *et al.* Transcriptomic analysis of tubular carcinomas of the breast reveals similarities and differences with molecular subtype-matched ductal and lobular carcinomas. *J. Pathol.* **222**, 64–75 (2010).
44. Alvarenga, C. A. *et al.* Reappraisal of immunohistochemical profiling of special histological types of breast carcinomas: a study of 121 cases of eight different subtypes. *J. Clin. Pathol.* **65**, 1066–71 (2012).
45. Yerushalmi, R., Hayes, M. M. & Gelmon, K. A. Breast carcinoma—rare types: review of the literature. *Ann. Oncol.* **20**, 1763–70 (2009).
46. Vranic, S. *et al.* Apocrine carcinoma of the breast: a comprehensive review. *Histol. Histopathol.* **28**, 1393–409 (2013).
47. Larribe, M., Thomassin-Piana, J. & Jalaguier-Coudray, A. Breast cancers with round lumps: correlations between imaging and anatomopathology. *Diagn. Interv. Imaging* **95**, 37–46 (2014).
48. Piscuoglio, S. *et al.* Integrative genomic and transcriptomic characterization of papillary carcinomas of the breast. *Mol. Oncol.* **8**, 1588–602 (2014).
49. Marchiò, C. *et al.* Mixed micropapillary-ductal carcinomas of the breast: a genomic and immunohistochemical analysis of morphologically distinct components. *J. Pathol.* **218**, 301–15 (2009).
50. Duprez, R. *et al.* Immunophenotypic and genomic characterization of papillary carcinomas of the breast. *J. Pathol.* **226**, 427–41 (2012).
51. Lacroix-Triki, M. *et al.* Mucinous carcinoma of the breast is genomically distinct from invasive ductal carcinomas of no special type. *J. Pathol.* **222**, 282–98 (2010).
52. Reis-Filho, J. S. *et al.* Metaplastic breast carcinomas are basal-like tumours. *Histopathology* **49**, 10–21 (2006).
53. Vincent-Salomon, A. *et al.* Identification of typical medullary breast carcinoma as a genomic sub-group of basal-like carcinomas, a heterogeneous new molecular entity. *Breast Cancer Res.* **9**, R24 (2007).
54. Jacquemier, J. *et al.* Typical medullary breast carcinomas have a basal/myoepithelial phenotype. *J. Pathol.* **207**, 260–8 (2005).
55. Bertucci, F. *et al.* Gene expression profiling shows medullary breast cancer is a subgroup of basal breast cancers. *Cancer Res.* **66**, 4636–44 (2006).
56. Fan, C. *et al.* Concordance among gene-expression-based predictors for breast cancer. *N. Engl. J. Med.* **355**, 560–9 (2006).
57. Craven, R. A. *et al.* Proteomic analysis of formalin-fixed paraffin-embedded renal tissue samples by label-free MS: Assessment of overall technical variability and the impact of block age. *PROTEOMICS - Clin. Appl.* **7**, 273–282 (2013).
58. Kroll, J. *et al.* Isolation of high quality protein samples from punches of formalin fixed and paraffin embedded tissue blocks. *Histol. Histopathol.* **23**, 391–5 (2008).
59. Golaraci, A. *et al.* Changes of collagen ultrastructure in breast cancer tissue determined by second-harmonic generation double Stokes-Mueller polarimetric microscopy. *Biomed. Opt. Express* **7**, 4054 (2016).
60. Campbell, K. R. & Campagnola, P. J. Wavelength-Dependent Second Harmonic Generation Circular Dichroism for Differentiation of Col I and Col III Isoforms in Stromal Models of Ovarian Cancer Based on Intrinsic Chirality Differences. *J. Phys. Chem. B* **121**, 1749–1757 (2017).
61. Capella, C., Eusebi, V., Mann, B. & Azzopardi, J. G. Endocrine differentiation in mucoid carcinoma of the breast. *Histopathology* **4**, 613–30 (1980).
62. Allred, D. C. *et al.* Adjuvant tamoxifen reduces subsequent breast cancer in women with estrogen receptor-positive ductal carcinoma *in situ*: a study based on NSABP protocol B-24. *J. Clin. Oncol.* **30**, 1268–73 (2012).
63. Reis-Filho, J. S. *et al.* Novel and classic myoepithelial/stem cell markers in metaplastic carcinomas of the breast. *Appl. Immunohistochem. Mol. Morphol.* **11**, 1–8 (2003).
64. Milanezi, F., Carvalho, S. & Schmitt, F. C. EGFR/HER2 in breast cancer: a biological approach for molecular diagnosis and therapy. *Expert Rev. Mol. Diagn.* **8**, 417–34 (2008).
65. Coates, A. S. *et al.* Tailoring therapies—improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. *Ann. Oncol.* **26**, 1533–46 (2015).
66. Rezakhanliha, R. *et al.* Experimental investigation of collagen waviness and orientation in the arterial adventitia using confocal laser scanning microscopy. *Biomech. Model. Mechanobiol.* **11**, 461–73 (2012).
67. Ambekar, R., Lau, T.-Y., Walsh, M., Bhargava, R. & Toussaint, K. C. Quantifying collagen structure in breast biopsies using second-harmonic generation imaging. *Biomed. Opt. Express* **3**, 2021–35 (2012).
68. Simpson, P. T. *et al.* Molecular profiling pleomorphic lobular carcinomas of the breast: evidence for a common molecular genetic pathway with classic lobular carcinomas. *J. Pathol.* **215**, 231–44 (2008).

## Acknowledgements

The authors thank the National Institute of Science and Technology on Photonics Applied to Cell Biology (INFABiC) at the State University of Campinas for access to equipment and assistance; INFABiC is co-funded by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (08/57906-3) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (573913/2008-0). C. L. Cesar also acknowledges the grant by FAPESP (11/51959-0) and CNPq (312049/2014-5).

## Author Contributions

R.d.A.N., V.B.P., C.L.C. and J.V. conceived and planned the experiments. R.d.A.N., G.R.P., R.F.V. and F.C.S. carried out the experiments. G.R.P., C.A.A., R.F.V., C.F., F.C.S. and J.V. contributed to sample preparation. R.d.A.N., L.M., S.F.D., L.O.S., C.F., C.L.C., F.C.S., B.W. and J.V. contributed to the interpretation of the results. R.d.A.N. and J.V. took the lead in writing the manuscript. All authors provided critical feedback and helped shape the research, analysis and manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-44156-9>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019