

Analyst

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

ARTICLE

Infrared imaging of MDA-MB-231 breast cancer cell line phenotypes in 2D and 3D cultures

Cite this: DOI: 10.1039/x0xx00000x

Margarita Smolina and Erik Goormaghtigh^{a,*}

Received oothXxxxxxxXXXX,
Accepted oothXxxxxxxXXXX

DOI: 10.1039/x0xx00000x

⁵www.rsc.org/

¹⁰One current challenge in the field of breast cancer infrared imaging is the identification of carcinoma cell subtypes in the tissue. Neither sequencing nor immunochemistry is currently able to provide a cell by cell thorough classification. The latter is needed to build accurate statistical models capable of recognizing the diversity of breast cancer cell lines which may be present in a tissue section. One possible approach to overcome this problem is to obtain the IR spectral signature of well-characterized tumor cell lines in culture. Cultures in three-dimensional matrices appear to generate an environment that mimics better the *in vivo* environment. There are presently series of breast cancer cell lines which have been thoroughly characterized in two- and three-dimensional (2D and 3D) cultures by full transcriptomics analyses. In this work, we describe the methods used to grow, to process and to characterize a triple-negative breast cancer cell line, MDA-MB-231, in 3D laminin-rich extracellular matrix (lrECM) culture and compare it with traditional monolayer cultures and tissue sections. While unsupervised analyses did not completely separate spectra of cells grown in 2D from 3D lrECM cultures, a supervised statistical analysis resulted in an almost perfect separation. When IR spectral responses of epithelial tumor cells from clinical triple-negative breast carcinoma samples were added to these data, a principal component analysis indicated that they cluster closer to the spectra of 3D culture cells than to the spectra of cells grown on a flat plastic substrata. This result is encouraging in view of correlating well-characterized cell line features with clinical biopsies.

³⁰ Introduction

Breast cancer is a global public health issue. It is the most frequently diagnosed malignancy in women in the Western world and the commonest cause of cancer death among European and American women. When performed at its best, basic histopathological examination of breast cancer remains the gold standard in determining patient outcome. During the last two decades, several new clinical and pathological parameters have been used to evaluate the prognosis of breast cancer patients. Analysis of biopsies is required for pathological examination but immunohistochemistry methods are limited by both the volume of tissue available and the fact that they cannot use more than a few probes at the same time on one tissue section. New technologies such as microarray analysis, which enables scientists to look at specific “molecular (gene) signature” or “fingerprint” of a tumor, has allowed gaining new insights into the biology of breast cancer. It has contributed to a better understanding why some women have different clinical outcomes even though the histopathological examination of their tumors provides similar results. Several microarray studies have been remarkably consistent in reproducing similar molecular classification of breast cancer. The collective results conclude that breast tumors can be grouped according to at least three individual subgroups: luminal, basal and ErbB2/HER2 subtypes. These subgroups have distinct clinical outcomes and may respond differently to various therapeutics.^{1–3} These results suggest that molecular profiling could significantly advance our interpretation of breast cancer biology and improve our management of

individual patients. Yet, on the one hand, despite interesting and positive results provided by these gene expression studies, reproducibility and robustness of expression data remain a concern⁴. Even though concordance among microarray platforms improves dramatically after filtering for gene nucleotide sequence identity⁵, the price to pay is a reduced gene expression information retained in the analysis. Furthermore, sample preparation and storage form also a crucial issue for microarray experiment. In fact, RNA is inherently unstable and rapid changes may occur as a result of insults caused by tissue handling or ischemia. On the other hand, there is a relative lack of success of the new molecular clinical tests in terms of clinical outcome.^{6,7} As the expansion of targeted therapies available to breast cancer patients is developing rapidly, exploring alternative strategies for biomarker discovery and individualized therapy is urgent.

⁷⁵ Among the available opportunities for the *ex vivo* analysis of tissues, Fourier transform infrared (FTIR) spectroscopy appears as one of the most relevant tools. It is due to the global information it provides about the molecular content of samples. FTIR spectroscopy is based upon the interaction between the IR radiation and the covalent bonds of molecules in presence. IR spectroscopy exploits the fact that molecules have specific frequencies at which they rotate or vibrate corresponding to discrete energy levels (vibrational modes). Within the mid-infrared range (4000–400 cm⁻¹ or 2.5–25 μm), all organic functions lead to specific IR absorption bands. Each compound has a characteristic set of absorption bands in its infrared spectrum. It is now considered that the FTIR spectrum provides

1 as much information as DNA microarrays as far as diagnostic
2 purposes are concerned. Importantly, all molecular types
3 contribute to the IR spectrum and this contribution depends on
4 the exact molecular structure. For instance, the head group
5 length and unsaturation of membrane lipids contribute all to the
6 IR spectral signature.^{8,9} Similarly, lipid/protein ratio, DNA
7 condensation state and many other parameters can be obtained
8 from the spectra.^{10–12} Besides, the IR spectra account not only
9 for the chemical nature of cell molecules but also for their
10 conformations and are, in particular, very sensitive to protein
11 secondary structure.^{13–17} All together, the various contributions
12 to the FTIR spectrum form a signature of the biochemical
13 composition of the cell that is unique. Interestingly, FTIR
14 spectroscopy is very fast and does not require any labeling or
15 chemical preparation.

16 When coupled with a microscope device, this technique
17 provides spatially resolved information on the various
18 structures present in tissue sections. It makes a significant
19 contribution in histopathology investigations and has been
20 recently recognized as an emerging tool for histopathological
21 studies.^{18–21} Specifically, in breast cancer research, IR
22 spectroscopy has proven its value.^{22–25}

23 The current challenge in the field of breast cancer is the
24 identification of carcinoma cell subtypes in the tissue. The
25 issue is particularly challenging as tumors are often quite
26 heterogeneous. Many clones having different properties may
27 coexist within a small tumor area as demonstrated by single
28 cell sequencing.²⁶ In turn, it remains very challenging to
29 identify the “molecular” or genomic classification of individual
30 cells by FTIR imaging. Even relying on immunohistochemistry
31 labelling just after recording IR images is limited by the
32 possibility to obtain quantifications for many different epitopes
33 on the same tissue section. Highly multiplexed approach could
34 resolve this problem in the future.²⁷ In the absence of an
35 extended spectral database with reliable assignments to
36 carcinoma subtypes, training a model to assign IR spectra to
37 these subtypes in a tissue environment is not possible.

38 One possible approach to overcome this problem is to obtain
39 the IR spectral signature of well-characterized cell lines in
40 culture. There are presently series of breast cancer cell lines
41 which have been thoroughly characterized in 2D and 3D
42 cultures. The availability and relevance of these cell lines (for
43 instance 51 breast cancer cell lines mirroring 145 primary
44 breast tumors²⁸) have been carefully described for breast
45 cancer.^{29,30} Transcriptomic data are available in both 2D and
46 3D cultures. See for instance the *Transcription profiling of 25
47 breast cancer cell lines grown in 2D and 3D tissue culture
48 conditions* published on the EMBL-EBI web site at
49 [http://www.ebi.ac.uk/arrayexpress/browse.html?keywords=E-](http://www.ebi.ac.uk/arrayexpress/browse.html?keywords=E-TABM-244)
50 [TABM-244](http://www.ebi.ac.uk/arrayexpress/browse.html?keywords=E-TABM-244).

51 In a living tissue, cells interact with other cells and with
52 extracellular matrix through biochemical and mechanical cues.
53 These interactions play a crucial major role in maintaining
54 normal homeostasis and functions of the tissue³¹ and, in that
55 way, they regulate cell proliferation, migration and apoptosis.
56 When cultured *ex vivo* on 2D plastic surface, cells lose the
57 possibility to communicate with their microenvironment. Cell
58 cultures in 3D IrECM re-establish interactions with
59 extracellular matrix lost in conventional 2D cultures.³² Thus,
60 they appear to generate a physiologically relevant environment
that is much closer to the environment found in real tissues.³³
For 3D cultures, a tight correlation has been found between the

morphology of the colonies,³⁴ gene expression profile³⁵ and
prognostic.³⁶

⁶⁵ In this work, we describe the methods used to characterize a triple-
negative (ER-, PR- and HER2-) breast cancer cell line, MDA-MB-
231, grown in 3D IrECM culture and compare it with 2D culture
cells and triple-negative carcinoma tissue.

Materials and Methods

70 2D cell culture conditions

Human mammary tumor epithelial cell line MDA-MB-231 was
propagated on standard cell culture plastic in RPMI 1640 medium
(Lonza, Switzerland) supplemented with 10% fetal bovine serum
(FBS) (Life Technologies, USA), 2 mM L-glutamine (Lonza,
⁷⁵ Switzerland), 50 U/mL penicillin and 50 µg/ml streptomycin (Life
Technologies, USA) at 37°C in 5% CO₂. Healthy, early-passage (*p*
= 3 to 5) and less than 75% confluent cells were used, on the one
hand, in 3D culture assays and, on the other hand, were coated with
Matrigel as protective agent before undergoing formalin fixation and
⁸⁰ paraffin embedding (FFPE).

3D IrECM culture system

3D cultures were prepared following the protocol described by Lee
*et al.*³³ 3D embedded culture (*vs* 3D on-top culture) turned out to be
more suitable for further handling, providing an increased number of
cell colonies in each sample. Briefly, single MDA-MB-231 cells
⁸⁵ were suspended in 150 µl growth-factor reduced Matrigel basement
membrane (BM) matrix (9.6 mg/ml; BD Biosciences, USA) at a
density of 1x10⁵ cells/ml and plated into eight-well plastic-
chambered glass microscope slide (0.7 cm²/well; Corning, USA)
⁹⁰ precoated beforehand with 50 µl Matrigel. 3D IrECM cell cultures
were maintained in H14 medium with 5% Matrigel and 1% FBS at
37°C in 5% CO₂ for 7–10 days with medium change every 2–3 days.

By Matrigel, we refer to the solubilized tissue basement membrane
extract derived from the Engelbreth-Holm-Swarm mouse sarcoma.
⁹⁵ It is mainly composed of laminin; other constituents include
collagen IV, heparan sulfate proteoglycans and entactin/nidogen.³⁷
Matrigel is commonly used as a physiologically relevant model of a
laminin-rich BM.

FFPE treatment

¹⁰⁰ Matrigel matrices with 3D cell cultures and 2D culture cells
they derived from were fixed in 10% neutral buffered formalin
solution for 22 h, followed by progressive dehydration and
embedding in paraffin as usual in clinical routine (200 min
protocol). Two 4 µm adjacent sections were cut with a rotary
¹⁰⁵ microtome: one of both was mounted on a BaF₂ slide (Korth
Kristalle GmbH, Germany), deparaffinized and intended for IR
imaging, the other one was stained with hematoxylin and eosin
(H&E) for visual examination. The choice of such a sample
treatment has been guided by the purpose of simplifying
¹¹⁰ comparison with tissue samples.

FTIR data acquisition

The FTIR data were collected using a Hyperion 3000 FTIR
imaging system (Bruker Optics, Ettlingen, Germany), equipped
with a liquid nitrogen cooled 64 x 64 Mercury Cadmium
¹¹⁵ Telluride (MCT) Focal Plane Array (FPA) detector and a 15x
objective (NA = 0.4). The data were collected in transmission
mode from sample regions of 180 x 180 µm². Every element of
the FPA acts as an independent and discrete detector from
which a full spectrum is obtained. The corresponding pixel
¹²⁰ covers an area of 2.8 x 2.8 µm². One FTIR image (unit image)

1 results in 4096 spectra, each one being the average of 256
 2 scans recorded in a spectral range from 3800 to 900 cm^{-1} (ca. 5
 3 minutes). To cover larger sample areas, several FTIR images
 4 were juxtaposed in order to obtain one FTIR map. The
 5 background spectrum, acquired as the average of 512 scans in
 6 the absence of sample or BaF_2 slide. The spectral resolution
 7 was set to 8 cm^{-1} .

8 Six unit images and one 2 x 2 FTIR map were acquired for
 9 each 3D and 2D cell culture sample respectively. Experiments
 10 were performed independently in triplicate: from cell thawing
 11 to FTIR data acquisition.

12 Data analyses

13 Pre-processing of IR spectra

14 All spectra were preprocessed as follows. Water vapor
 15 contribution was subtracted as described previously^{38,39} with
 16 1956-1932 cm^{-1} as reference peak and CO_2 peak was flattened
 17 between 2450 and 2250 cm^{-1} . The spectra were baseline-
 18 corrected. Straight lines were interpolated between the spectra
 19 points at 3620, 2995, 2800, 2395, 2247, 1765, 1724, 1480,
 20 1355, 1144 and 950 cm^{-1} and subtracted from each spectrum.
 21 Spectra were normalized for equal area between 1725 and 1481
 22 cm^{-1} (Amide I and II peaks). The signal-to-noise ratio (S/N)
 23 was then checked on every spectrum. It was required to be
 24 higher than 500:1 with noise defined as the standard deviation
 25 in the 2200-2100 cm^{-1} region of the spectrum and signal
 26 defined as the maximum of the curve between 1750 and 1480
 27 cm^{-1} after subtracting a baseline passing through these two
 28 points. Finally, some rare spectra with normalized absorbance
 29 lower than -5 (negative lobe) and a maximum above 120
 30 (saturation) were discarded. To avoid abrupt refractive index
 31 changes we always selected areas of sample with contiguous
 32 cells. Visual inspection of spectra as well as systematic
 33 screening for negative lobes on the left-hand side of the Amide
 34 I band did not reveal significant dispersive artifacts.⁴⁰⁻⁴⁴ As
 35 Mie scattering corrections rely on simplified models and as
 36 dispersive artifacts were minor in the present work, we
 37 preferred not applying such a correction. Derivation calculation
 38 did not provide more efficiency as also reported elsewhere⁴⁵
 39 and was not applied to preprocessed IR spectra.

40 Statistical analyses of IR spectra

41 In order to observe the intrinsic proximities and distances
 42 within the data set and to group IR spectra according to their
 43 similarity, unsupervised Principal Component Analyses (PCA)
 44 were performed. These multivariate methods enable variable
 45 reduction by building linear combinations of wavenumbers
 46 varying together, called Principal Components (PC). Usually 2
 47 to 6 PCs are sufficient to explain the major proportion of the
 48 original variance of the dataset, reducing the description of
 49 every spectrum to 2 to 6 numbers representing its projection
 50 (scores) on the PCs. Based on these scores, each spectrum can
 51 be represented as a point on a 2D or a 3D PC space.

52 Supervised Partial Least Square Discriminant Analysis (PLS-
 53 DA) was also performed on the data set to extract latent
 54 variables that enable the construction of a factor capable of
 55 predicting a class. It requires *a priori* knowledge about the
 56 classes of spectra and allows both data reduction and
 57 discriminative investigation. This approach consists in the
 58 application of PLS regression resulting in fewer uncorrelated
 59 variables. Linear combination of variables explaining the
 60 membership assignment will then serve as discriminant rules,
 minimizing intragroup and maximizing intergroup separations.

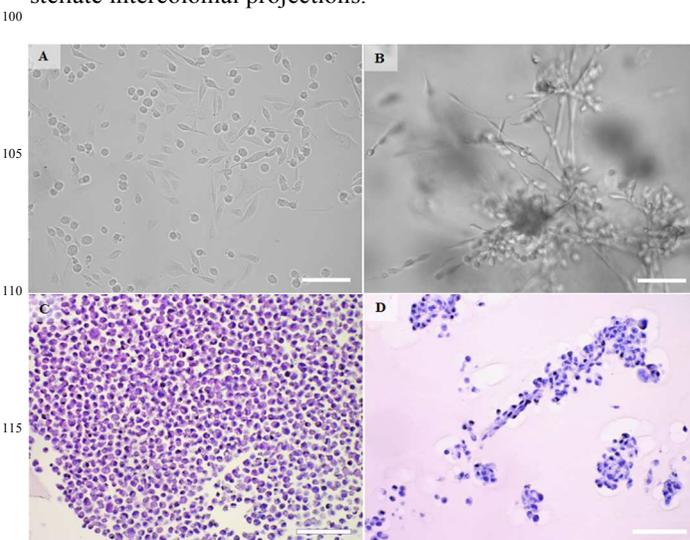
Correction of the IR spectra for water vapor and atmospheric
 CO_2 contribution, baseline subtraction, normalization,
 application of quality filters, PCA and PLS-DA analyses were
 65 carried out by Kinetics, a custom-made program running under
 Matlab (Mathworks, Inc.).

66 Results

67 MDA-MB-239 triple-negative cell line has been grown in 2D
 68 and 3D culture conditions as described by Lee *et al.*³³ We
 69 decided to follow the most common protocol used in the clinic
 70 to prepare cells and tissues. As described in *Materials and*
Methods, cell pellets (2D) and cell colonies (3D) were
 formalin-fixed and paraffin-embedded. Adjacent 4- μm -thick
 sections were prepared for IR imaging and H&E staining. Even
 75 though the FFPE procedure brings some modifications to the
 IR spectra of cells, these modifications are very similar for all
 cell lines (limited loss of lipids and some protein
 conformational changes),⁴⁶ these conditions have also been
 shown to maintain the subtle variations that may exist among
 80 breast cancer cell lines.⁴⁶ Nevertheless, it is worth pointing out
 the importance of lipid metabolism in cancer progression.^{47,48}

81 Comparison of the IR spectra of MDA-MB-231 cell line grown 82 in 2D and 3D cultures

83 When growing on a 2D flat surface of a culture box, MDA-
 84 MB-231 cells form a uniform monolayer after 4-5 days. To
 85 manipulate healthy cells and avoid effects related to
 confluence,⁴⁹ cells were always collected at less than 75%
 confluence. Cell shape before detachment were either round or
 elongated, depending on cell cycle stage,⁵⁰ but all cells became
 90 round-shaped after trypsinization and remained so after FFPE
 treatment (Fig. 1). Even though FFPE processed cancer cells
 grown in 2D accurately reflect the differences that exist among
 different cell lines before FFPE processing,⁴⁶ the 2D phenotype
 is quite distinct from the one found in tissues. MDA-MB-231
 95 cells grown in 3D lrECM-embedded matrix form colonies of
 various sizes and shapes and are phenotypically much closer to
 epithelial cells found in tissue context. Particularly, MDA-MB-
 231 cell line is characterized by its invasive phenotype with
 stellate intercolonial projections.



120 **Figure 1.** Bright-field microscopy images of MDA-MB-231 breast
 cancer cells grown in conventional 2D culture (A and C) and in 3D
 lrECM-embedded culture (B and D). The images A and B have been
 acquired before formalin fixation; the images C and D have been

obtained after FFPE treatment and H&E staining of 4- μm -thick sections. White bars: 100 μm .

Interestingly, Figure 1D also shows that Matrigel matrix in 3D culture, stained by eosin in light pink, is absent from the immediate vicinity of the cell colonies. It is mainly due to post-FFPE processing shrinkage of Matrigel. A comparison between bright-field images of H&E stained, unstained sections and their corresponding IR images is provided in Figure 2. While individual cells can be easily identified on the H&E stained sections and, though less clearly, on the unstained sections, the IR images are characterized by a much poorer resolution as expected from the diffraction-limited resolution at 1654 cm^{-1} ($\lambda \approx 6\text{ }\mu\text{m}$). Considering the numerical aperture ($\text{NA} = 0.4$), diffractions limit resolution to ca. $9\text{ }\mu\text{m}$ at best.⁵¹ As the point-spread-function usually presents side lobes, true resolution is usually even further reduced in standard IR imaging.^{51,52} Raw absorbance spectra are reported in Figure S1 (Supplementary Material).

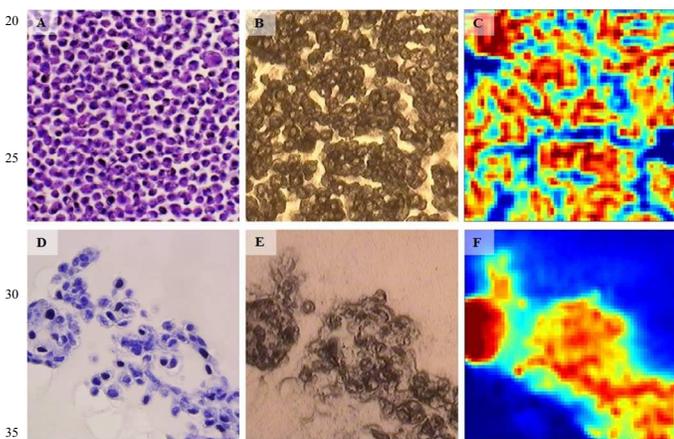


Figure 2. Bright-field microscopy (A-B, D-E) and IR (C, F) images of adjacent 4- μm -thick sections of MDA-MB-231 breast cancer cells grown in 2D (A-C) and 3D (D-F) cultures. For 2D culture, the image A reports the H&A stained section. The adjacent, unstained section image is presented in B. The IR raw image (absorbance at 1654 cm^{-1} , no spectral processing) of the same region is reported in C. For 3D culture, image D reports the H&A stained section. The adjacent, unstained section image is presented in E, and the IR raw image (absorbance at 1654 cm^{-1} , no spectral processing) of the same region is reported in F. One image is $180 \times 180\text{ }\mu\text{m}^2$.

Spectra with sufficient S/N (see *Materials and Methods*) were then extracted manually from the IR images. Three independent cultures provided 600 and 1752 spectra for the 2D and 3D samples respectively. PCA was applied in order to provide an unsupervised view of the data. A score plot is reported in Figure 3.

PC1 is characterized by a strong sigmoidal feature in the Amide I range of the spectrum, with a minimum at 1624 cm^{-1} and a maximum at 1670 cm^{-1} . Considering this shape and the sign of its contribution in Figure 3 (positive for 3D culture cells and negative for 2D culture cells), it indicates that proteins of cells grown in 3D culture have more α -helical and less β -sheet structures than cells grown in 2D culture. This feature is also dominant on a straight difference spectrum obtained by subtracting the mean spectrum calculated for 2D culture cells from the mean spectrum associated with 3D culture cells (not shown). The biological interpretation of this observation remains to be elucidated.

Figure 3 indicates some degree of separation between the 600 MDA-MB-231 breast cancer cells grown in 2D and 3D cultures. The origin of the difference can be appreciated from the shape of the PCs (Figure 4).

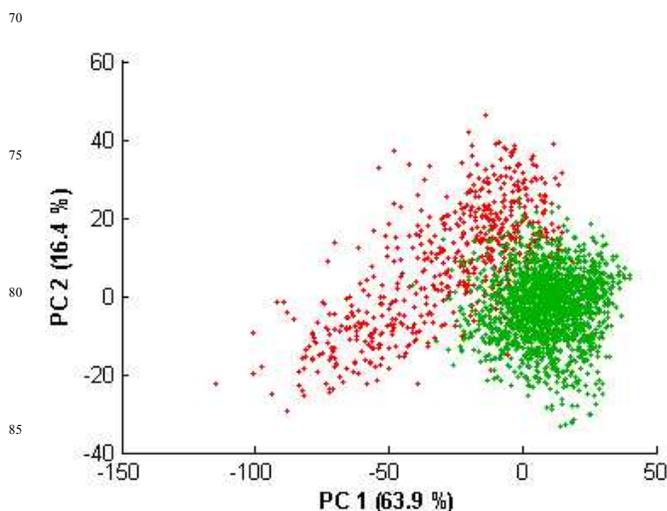


Figure 3. PCA score plot of 600 spectra of MDA-MB-231 breast cancer cells grown in 2D culture (red) and 1752 spectra of cells derived from them and grown in 3D culture (green). Spectra are projected in the PC1 – PC2 space. The fraction of the total variance explained by each PC is indicated on the axes. The analysis has been carried out on the $1800\text{--}1000\text{ cm}^{-1}$ spectral range.

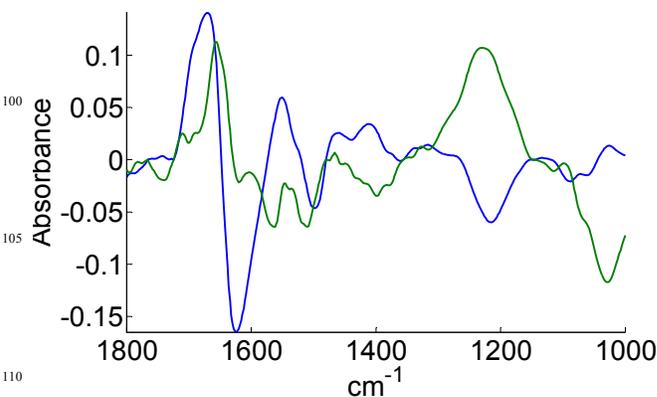


Figure 4. Representation of PC1 (blue) and PC2 (green) obtained after principal component analysis applied on 600 spectra of MDA-MB-231 breast cancer cells grown in 2D culture and 1752 spectra of cells derived from them and grown in 3D culture ($1800\text{--}1000\text{ cm}^{-1}$ range).

Since an unsupervised statistical analysis indicated a trend to separation according to culture type (either 2D or 3D), we attempted to use a supervised method to improve the discrimination. PLS-DA was used to classify all the spectra used in PCA represented in Figure 3. In order to evaluate the contribution of the different spectral regions, a bootstrapping method was set up on 100 cm^{-1} wide spectral ranges. 60% of the spectra were used to build a model and the remaining 40% of the spectra were used to test the quality of this model. The procedure was repeated 40 times by randomly selecting a new training set and a new test set. The procedure was repeated for each spectral interval. Results are summarized in Figure 5.

When run on the entire spectral range, the PLS-DA yielded even better correct assignment scores (not shown), indicating that non redundant information is present in the various sub-regions analyzed in Figure 5.

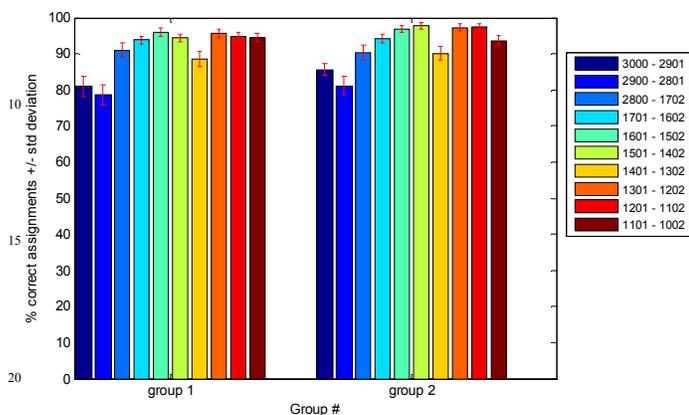


Figure 5. Success rate (in %) of the prediction models as a function of the spectral range selected (see color legend). Interval PLS-DA classification was computed on the spectra described in Figure 3. The percentages of correct assignment by the prediction models are reported on Y-axis; the predicted values are expressed as % of the true values. 60% of the spectral database was used for the training and the remaining 40% as test set. The procedure was repeated 40 times, yielding to a standard deviation for each % value. The entire spectral range between 3000-2800 and 1800-1000 cm^{-1} was investigated by steps of 100 cm^{-1} as indicated in the color legend. Group 1: spectra associated with 2D culture cells; group 2: spectra associated with 3D culture cells.

35 Comparison with epithelial cells in breast cancer tissues

As information able to distinguish cells grown in 2D and 3D is present in their IR spectra, we attempted to compare these two categories of spectra with spectra of carcinoma cells present in a tissue. Sections from three clinical samples were imaged. For the sake of the consistency, three triple-negative tumor cases were selected. A PCA score plot shows that spectra of 2D culture cells form a distinct cluster while spectra of 3D culture and clinical carcinoma cells are largely superimposed on PC#1 (Fig. 6).

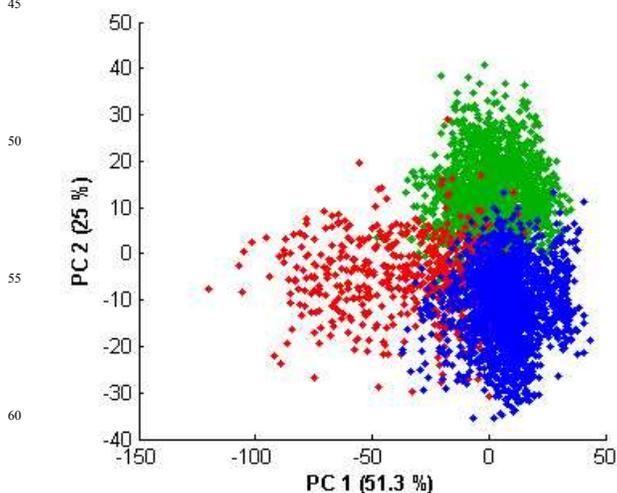


Figure 6. PCA score plot of 600 spectra of cells grown in 2D culture (red), 1752 spectra of cells grown in 3D culture (green) and 1704 spectra of tumor epithelial cells from 3 patients (respectively 504, 648 and 552 spectra) with triple-negative breast carcinoma (blue). The fraction of the total variance explained by PC1 and PC2 is indicated on the axes. The analysis has been carried out on the 1800-1000 cm^{-1} spectral range.

Figure 6 indicates that, even though there is some projection overlap of the spectra of 2D culture cells, on the one hand, with the spectra of 3D culture cells, and, on the other hand, with clinical carcinoma cells, unique characteristics of the two groups are extracted in PC1 which explains about half (51.3%) of the total variance found in the entire dataset.

Furthermore, it appears that after addition to the two existing spectral groups a large number of IR spectra of carcinoma cells recorded in clinical samples, the shape of PC1 remains very similar to the one observed in Figure 4 (Fig. 7).

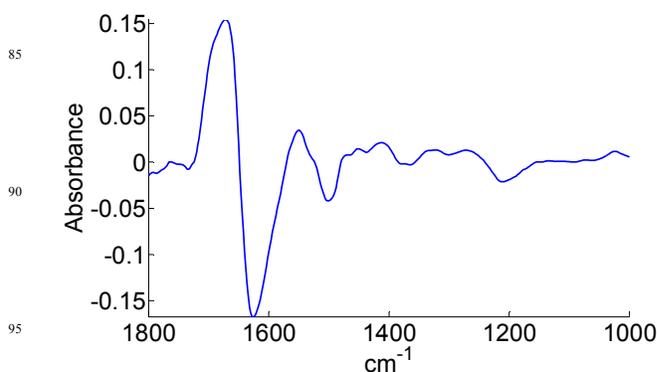


Figure 7. Representation of PC1 obtained after principal component analysis applied on 600 spectra of cells grown in 2D culture, 1752 spectra of cells grown in 3D culture and 1704 spectra of tumor epithelial cells from 3 patients with triple-negative breast carcinoma (1800-1000 cm^{-1} range).

Conclusion

The present paper describes a methodology that allows direct comparison of cell phenotypes grown in 2D monolayer and 3D IrECM cultures. The results show us that FFPE-treated cell cultures preserve their morphological integrity and provide thus satisfactory specimens for IR analysis. Interestingly, bright-field images indicate that at the end of the FFPE processing, the matrix shrinks and leave a space around the cells. After paraffin removal, it results in no contamination of the cell spectra by the extracellular matrix spectra. PCA also demonstrates that Matrigel matrix spectra are completely distinct from the cell spectra (not shown). Phenotypes are, as expected, distinct in 2D and 3D cultures.^{33-36,53,54} It highlights the biochemical changes occurring in studied epithelial cancer cell line according to its growth condition. While unsupervised analysis did not completely separate the two phenotypes, a supervised statistical analysis resulted in an almost perfect separation. When spectra from carcinoma cells present in clinical samples are added to the series, a PCA indicates that they cluster closer to the spectra of 3D culture cells than to the spectra of 2D culture cells. This result is encouraging in view of correlating well-characterized cell line features with biological characteristics of clinical samples. It is obvious that,

even in 3D cultures, cell lines will never have phenotypes absolutely identical to the ones found in tissues. Nevertheless, as long as common features can be extracted, they could be identified in both conditions. 3D cultures could be improved to match better and/or understand better the *in vivo* conditions. The influence of fibroblasts⁵⁵⁻⁵⁹ or lymphocytes^{56,60,61} on breast carcinoma cells has been underlined consistently. Co-cultures with such cells could be easily implemented. Moreover, other well-established breast cancer cell lines should be explored in order to represent as much as possible the diversity of human breast carcinoma forms.

Acknowledgments

This research has been supported by grants from the National Fund for Scientific Research (FRFC 2.4527.10, 2.4526.12 and T.0155.13). E.G. is Research Director with the National Fund for Scientific Research (FNRS) (Belgium), M.S. is Research Fellow supported by grant T.0155.13 from the FNRS (Belgium).

Notes

* Laboratory for the Structure and Function of Biological Membranes, Center for Structural Biology and Bioinformatics, Université Libre de Bruxelles (ULB), Bld du Triomphe 2, CP206/02, B-1050 Brussels, Belgium. E-mail: egoor@ulb.ac.be; Fax: +32-2-650-53-82; Tel: +32-2-650-53-86; * corresponding author

Abbreviations

2D/3D, two-/three-dimensional; BaF₂, Barium Fluoride; BM, basement membrane; ER, estrogen receptors; ErbB2/HER2, Human Epidermal Growth Factor Receptor 2; FBS, fetal bovine serum; FFPE, formalin-fixed, paraffin-embedded; FPA, Focal Plane Array; FTIR, Fourier Transform Infrared Spectroscopy; GFR, Growth Factor Reduced; H&E, hematoxylin-eosin; IrECM, laminin-rich extracellular matrix; MCT, Mercury Cadmium Telluride; PC, Principal Component; PCA, Principal Component Analysis; PLS-DA, Partial Least Square Discriminant Analysis; PR, progesterone receptors; S/N, signal-to-noise ratio

References

- T. Sorlie, R. Tibshirani, J. Parker, T. Hastie, J. S. Marron, A. Nobel, S. Deng, H. Johnsen, R. Pesich, S. Geisler, J. Demeter, C. M. Perou, P. E. Lonning, P. O. Brown, A. L. Borresen-Dale and D. Botstein, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 8418–8423.
- C. Sotiriou, S. Y. Neo, L. M. McShane, E. L. Korn, P. M. Long, A. Jazaeri, P. Martiat, S. B. Fox, A. L. Harris and E. T. Liu, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 10393–10398.
- L. M. Shi, G. Campbell, W. D. Jones, F. Campagne, Z. N. Wen, S. J. Walker, Z. Q. Su, T. M. Chu, F. M. Goodsaid, L. Pusztai, J. D. Shaughnessy, A. Oberthuer, R. S. Thomas, R. S. Paules, M. Fielden, B. Barlogie, W. J. Chen, P. Du, M. Fischer, C. Furlanello, B. D. Gallas, X. J. Ge, D. B. Megherbi, W. F. Symmans, M. D. Wang, J. L. Zhang, H. Bitter, B. Brors, P. R. Bushel, M. Bylesjo, M. J. Chen, J. Cheng, J. Chou, T. S. Davison, M. Delorenzi, Y. P. Deng, V. Devanarayan, D. J. Dix, J. Dopazo, K. C. Dorff, F. Elloumi, J. Q. Fan, S. C. Fan, X. H. Fan, H. Fang, N. Gonzaludo, K. R. Hess, H. X. Hong, J. Huan, R. A. Irizarry, R. Judson, D. Juraeva, S. Lababidi, C. G. Lambert, L. Li, Y. E. Li, Z. G. Li, S. M. Lin, G. Z. Liu, E. K. Lobenhofer, J. Luo, W. Luo, M. N. McCall, Y. Nikolsky, G. A. Pennello, R. G. Perkins, R. Philip, V. Popovici, N. D. Price, F. Qian, A. Scherer, T. L. Shi, W. W. Shi, J. Y. Sung, D. Thierry-Mieg, J. Thierry-Mieg, V. Thodima, J. Trygg, L. Vishnuvajjala, S. J. Wang, J. P. Wu, Y. C. Wu, Q. A. Xie, W. A. Yousef, L. A. Zhang, X. G. Zhang, S. Zhong, Y. M. Zhou, S. Zhu, D. Arasappan, W. J. Bao, A. B. Lucas, F. Berthold, R. J. Brennan, A. Bunes, J. G. Catalano, C. Chang, R. Chen, Y. Y. Cheng, J. A. Cui, W. Czika, F. Demichelis, X. T. Deng, D. Dosymbekov, R. Eils, Y. Feng, J. Fostel, S. Fulmer-Smentek, J. C. Fuscoe, L. Gatto, W. G. Ge, D. R. Goldstein, L. Guo, D. N. Halbert, J. Han, S. C. Harris, C. Hatzis, D. Herman, J. P. Huang, R. V. Jensen, R. Jiang, C. D. Johnson, G. Jurman, Y. Kahlert, S. A. Khuder, M. Kohl, J. Y. Li, M. L. Li, Q. Z. Li, S. Li, J. Liu, Y. Liu, Z. C. Liu, L. Meng, M. Madera, F. Martinez-Murillo, I. Medina, J. Meehan, K. Miclaus, R. A. Moffitt, D. Montaner, P. Mukherjee, G. J. Mulligan, P. Neville, T. Nikolskaya, B. T. Ning, G. P. Page, J. Parker, R. M. Parry, X. J. Peng, R. L. Peterson, J. H. Phan, B. Quanz, Y. Ren, S. Riccadonna, A. H. Roter, F. W. Samuelson, M. M. Schumacher, J. D. Shambaugh, Q. A. Shi, R. Shippy, S. Z. Si, A. Smalter, C. Sotiriou, M. Soukup, F. Staedtler, G. Steiner, T. H. Stokes, Q. L. Sun, P. Y. Tan, R. Tang, Z. Tezak, B. Thorn, M. Tsyganova, Y. Turpaz, S. C. Vega, R. Visintainer, J. von Frese, C. Wang, E. Wang, J. W. Wang, W. Wang, F. Westermann, J. C. Willey, M. Woods, S. J. Wu, N. Q. Xiao, J. Xu, L. Xu, L. Yang, X. A. Zeng, M. Zhang, C. Zhao, R. K. Puri, U. Scherf, W. D. Tong and R. D. Wolfinger, *Nat. Biotechnol.*, 2010, **28**, S5–S16.
- P. K. Tan, T. J. Downey, E. L. Spitznagel, P. Xu, D. Fu, D. S. Dimitrov, R. A. Lempicki, B. M. Raaka and M. C. Cam, *Nucleic Acids Res.*, 2003, **31**, 5676–5684.
- J. E. Larkin, B. C. Frank, H. Gavras, R. Sultana and J. Quackenbush, *Nat. Methods*, 2005, **2**, 337–343.
- P. L. Bedard, M. Ignatiadis, B. Haibe-Kains, S. Loi, C. Criscitiello, C. Desmedt, G. Bontempi, M. J. Piccart and C. Sotiriou, *Cancer Res.*, 2009, **69**, 507S–507S.
- P. L. Bedard and C. Sotiriou, *Embo Mol. Med.*, 2010, **2**, 3–5.
- A. Derenne, T. Claessens, C. Conus and E. Goormaghtigh, in *Encyclopedia of Biophysics*, 2013, pp. 1074–1081.
- A. Derenne, O. Vandersleyen and E. Goormaghtigh, *Biochim. Biophys. Acta*, 2013.
- M. Diem, L. Chiriboga, P. Lasch and A. Pacifico, *Biopolymers*, 2002, **67**, 349–353.
- A. Pevsner and M. Diem, *Appl. Spectrosc.*, 2001, **55**, 1502–1505.
- S. Boydston-White, T. Chernenko, A. Regina, M. Miljkovic, C. Matthaues and M. Diem, *Vib. Spectrosc.*, 2005, **38**, 169–177.
- E. Kleiren, J. M. Ruyschaert, E. Goormaghtigh and V. Raussens, *Spectrosc. Int. J.*, 2010, **24**, 61–66.
- H. H. de Jongh, E. Goormaghtigh, J. M. Ruyschaert and H. H. de Jongh, *Biochemistry*, 1997, **36**, 13603–13610.
- E. Goormaghtigh, V. Cabiaux and J. M. Ruyschaert, *Subcell. Biochem.*, 1994, **23**, 405–450.

- 1 16. E. Goormaghtigh, R. Gasper, A. Benard, A. Goldshtein and V. Raussens, *Biochim Biophys. Acta-Proteins Proteomics*, 2009, **1794**, 1332–1343.
- 2
- 3
- 4
- 5 17. E. Goormaghtigh, J. M. Ruyschaert and V. Raussens, *Biophys. J.*, 2006, **90**, 2946–2957.
- 6
- 7
- 8 18. G. Bellisola and C. Sorio, *Am. J. Cancer Res.*, 2012, **2**, 1–21.
- 9
- 10 19. R. Bhargava, *Anal. Bioanal. Chem.*, 2007, **389**, 1155–69.
- 11
- 12 20. P. Lasch, M. Diem, W. Hänsch and D. Naumann, *J. Chemom.*, 2007, **20**, 209–220.
- 13
- 14 10 21. P. Lasch, L. Chiriboga, H. Yee and M. Diem, *Technol. Cancer Res. Treat.*, 2002, **1**, 1–7.
- 15
- 16
- 17 22. G. J. Ooi, J. Fox, K. Siu, R. Lewis, K. R. Bambery, D. McNaughton and B. R. Wood, *Med. Phys.*, 2008, **35**, 2151–61.
- 18
- 19
- 20 23. B. Bird, M. Romeo, N. Laver and M. Diem, *J. Biophotonics*, 2009, **2**, 37–46.
- 21
- 22 15 24. M. J. Walsh, A. Kajdacsy-Balla, S. E. Holton and R. Bhargava, *Vib. Spectrosc.*, 2012, **60**, 23–28.
- 23
- 24
- 25 25. A. Benard, C. Desmedt, M. Smolina, P. Szternfeld, M. Verdonck, G. Rouas, N. Kheddoumi, F. Rothé, D. Larsimont, C. Sotiriou and E. Goormaghtigh, *Analyst*, 2014, **139**, 1044–56.
- 26
- 27 20 26. N. Navin, J. Kendall, J. Troge, P. Andrews, L. Rodgers, J. McIndoo, K. Cook, A. Stepansky, D. Levy, D. Esposito, L. Muthuswamy, A. Krasnitz, W. R. McCombie, J. Hicks and M. Wigler, *Nature*, 2011, **472**, 90–4.
- 28
- 29
- 30 25 27. C. Giesen, H. A. O. Wang, D. Schapiro, N. Zivanovic, A. Jacobs, B. Hattendorf, P. J. Schüffler, D. Grolimund, J. M. Buhmann, S. Brandt, Z. Varga, P. J. Wild, D. Günther and B. Bodenmiller, *Nat. Methods*, 2014, **11**, 417–22.
- 31
- 32
- 33 30 28. R. M. Neve, K. Chin, J. Fridlyand, J. Yeh, F. L. Baehner, T. Fevr, L. Clark, N. Bayani, J.-P. Coppe, F. Tong, T. Speed, P. T. Spellman, S. DeVries, A. Lapuk, N. J. Wang, W.-L. Kuo, J. L. Stilwell, D. Pinkel, D. G. Albertson, F. M. Waldman, F. McCormick, R. B. Dickson, M. D. Johnson, M. Lippman, S. Ethier, A. Gazdar and J. W. Gray, *Cancer Cell*, 2006, **10**, 515–27.
- 34
- 35
- 36 35 29. D. L. Holliday and V. Speirs, *Breast Cancer Res.*, 2011, **13**, 215.
- 37
- 38
- 39 30 30. T. Vargo-Gogola and J. M. Rosen, *Nat. Rev. Cancer*, 2007, **7**, 659–672.
- 40
- 41
- 42 31. M. J. Bissell, D. C. Radisky, A. Rizki, V. M. Weaver and O. W. Petersen, *Differentiation*, 2002, **70**, 537–546.
- 43
- 44
- 45 40 32. M. J. Bissell, P. A. Kenny and D. C. Radisky, *Cold Spring Harb. Symp. Quant. Biol.*, 2005, **70**, 343–56.
- 46
- 47
- 48 33. G. Y. Lee, P. A. Kenny, E. H. Lee and M. J. Bissell, *Nat. Methods*, 2007, **4**, 359–365.
- 49
- 50
- 51 45 34. J. Han, H. Chang, O. Giricz, G. Y. Lee, F. L. Baehner, J. W. Gray, M. J. Bissell, P. A. Kenny and B. Parvin, *Plos Comput. Biol.*, 2010, **6**.
- 52
- 53 35. P. A. Kenny, G. Y. Lee, C. A. Myers, R. M. Neve, J. R. Semeiks, P. T. Spellman, K. Lorenz, E. H. Lee, M. H. Barcellos-Hoff, O. W. Petersen, J. W. Gray and M. J. Bissell, *Mol. Oncol.*, 2007, **1**, 84–96.
- 54
- 55 50 36. K. J. Martin, D. R. Patrick, M. J. Bissell and M. V Fournier, *PLoS One*, 2008, **3**.
- 56
- 57 37. H. K. Kleinman and G. R. Martin, *Semin. Cancer Biol.*, 2005, **15**, 378–86.
- 58
- 59 55 38. E. Goormaghtigh and J. M. Ruyschaert, *Spectrochim. Acta*, 1994, **50A**, 2137–2144.
- 60 39. E. Goormaghtigh, in *Adv. Biomed. Spectrosc. (Biological and Biomedical Infrared Spectroscopy)*, eds. A. Barth and P. I. Haris, IOS Press, 2009, vol. IOS Press, pp. 104–128.
- 60 40. P. Bassan, A. Kohler, H. Martens, J. Lee, E. Jackson, N. Lockyer, P. Dumas, M. Brown, N. Clarke and P. Gardner, *J. Biophotonics*, 2010, **3**, 609–20.
41. P. Bassan, A. Sachdeva, A. Kohler, C. Hughes, A. Henderson, J. Boyle, J. H. Shanks, M. Brown, N. W. Clarke and P. Gardner, *Analyst*, 2012, **137**, 1370–7.
- 65 42. B. Bird, M. Miljković and M. Diem, *J. Biophotonics*, 2010, **3**, 597–608.
43. K. R. Bambery, B. R. Wood and D. McNaughton, *Analyst*, 2012, **137**, 126–32.
- 70 44. A. Dazzi, A. Deniset-Besseau and P. Lasch, *Analyst*, 2013, **138**, 4191–201.
45. A. Gaigneaux, J. M. Ruyschaert and E. Goormaghtigh, *Appl. Spectrosc.*, 2006, **60**, 1022–1028.
46. M. Verdonck, N. Wald, J. Janssis, P. Yan, C. Meyer, A. Legat, D. E. Speiser, C. Desmedt, D. Larsimont, C. Sotiriou and E. Goormaghtigh, *Breast cancer and melanoma cell line identification by FTIR imaging after formalin-fixation and paraffin-embedding.*, 2013, vol. 138.
- 80 47. R. Katz-Brull, D. Seger, D. Rivenson-Segal, E. Rushkin and H. Degani, *Cancer Res.*, 2002, **62**, 1966–1970.
48. H. Yamaguchi and T. Oikawa, *Oncotarget*, 2010, **1**, 320–8.
49. R. Gasper and E. Goormaghtigh, *Analyst*, 2010, **135**, 3048–3051.
50. A. Derenne, A. Mignolet and E. Goormaghtigh, *Analyst*, 2013, **138**, 3998–4005.
- 85 51. P. Lasch and D. Naumann, *Biochim. Biophys. Acta*, 2006, **1758**, 814–29.
52. E. C. Mattson, M. J. Nasse, M. Rak, K. M. Gough and C. J. Hirschmugl, *Anal. Chem.*, 2012, **84**, 6173–80.

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
53. J. M. Nam, Y. Onodera, M. J. Bissell and C. C. Park, *Cancer Res.*, 2010, **70**, 5238–5248.
54. N. Garamszegi, S. P. Garamszegi, L. A. Shehadeh and S. P. Scully, *Mol. Cancer Res.*, 2009, **7**, 319–29.
- ⁵ 55. S. E. Holton, A. Bergamaschi, B. S. Katzenellenbogen and R. Bhargava, *PLoS One*, 2014, **9**, e96878.
56. Z. I. Khamis, Z. J. Sahab and Q.-X. A. Sang, *Int. J. Breast Cancer*, 2012, **2012**, 574025.
- ¹⁰ 57. S. E. Holton, M. J. Walsh, A. Kajdacsy-Balla and R. Bhargava, *Biophys. J.*, 2011, **101**, 1513–21.
58. M. H. Barcellos-Hoff and D. Medina, *Breast Cancer Res.*, 2005, **7**, 33–6.
59. N. A. Bhowmick, E. G. Neilson and H. L. Moses, *Nature*, 2004, **432**, 332–7.
- ¹⁵ 60. W. H. Fridman, J. Galon, F. Pagès, E. Tartour, C. Sautès-Fridman and G. Kroemer, *Cancer Res.*, 2011, **71**, 5601–5.
61. D. Liao, Y. Luo, D. Markowitz, R. Xiang and R. A. Reisfeld, *PLoS One*, 2009, **4**, e7965.

²⁰

Journal Name

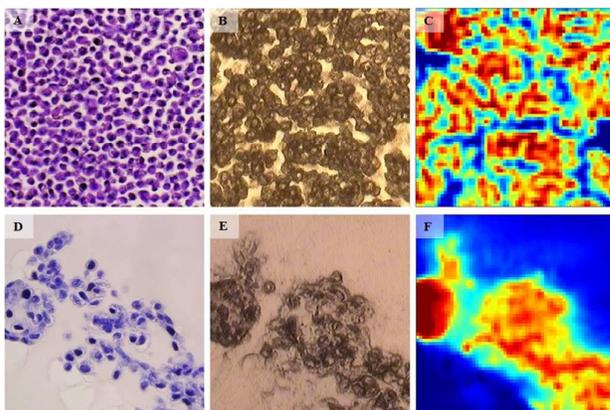
RSC Publishing

ARTICLE

Analyst Accepted Manuscript

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Breast cancer cell line in 2D (top) and 3D (bottom) culture: H&H, unstained bright field and IR image



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60