Analyst

PAPER

View Article Online
View Journal

Cite this: DOI: 10.1039/c4an01833h

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

Margarita Smolina and Erik Goormaghtigh*

Received 13th October 2014,
Accepted 17th December 2014

DOI: 10.1039/c4an01833h

www.rsc.org/analyst

One 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 that may be present in a tissue section. One possible approach for overcoming 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, at present, series of breast cancer cell lines that 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 because of correlating well-characterized cell line features with clinical biopsies.

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 a specific "molecular (gene) signature" or "fingerprint" of a tumor, have enabled gaining new insights into the biology of breast cancer. They have contributed to a better understanding of 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 classifications of breast cancer. The collective results conclude that breast tumors can be grouped into 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 the management of individual patients. Yet, on the one hand, despite interesting and positive results provided by these gene expression studies, the reproducibility and robustness of expression data remain a concern.[4] Even though concordance among microarray platforms improves dramatically after filtering for gene nucleotide sequence identity,[5] the price to pay is the reduced gene expression information retained in the analysis. Furthermore, sample preparation and storage also form a crucial issue for microarray experiments. 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

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

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c4an01833h

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.

Among the available opportunities for the *ex vivo* analysis of tissues, Fourier transform infrared (FTIR) spectroscopy appears as one of the most relevant tools. This 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 the biological molecules. 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 as much information as DNA microarrays as far as diagnostic purposes are concerned. Importantly, all molecular types contribute to the IR spectrum and this contribution depends on the exact molecular structure. For instance, the head group length and unsaturation of membrane lipids contribute all to the IR spectral signature.[8,9] Similarly, the lipid/protein ratio, DNA condensation state and many other parameters can be obtained from the spectra.[10–12] Besides, the IR spectra account not only for the chemical nature of cell molecules but also for their conformations and are, in particular, very sensitive to the protein secondary structure.[13–17] Taken together, the various contributions to the FTIR spectrum form a signature of the biochemical composition of the cell that is unique. Interestingly, FTIR spectroscopy is very fast and does not require any labeling or chemical preparation.

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

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

One possible approach to overcome this problem is to obtain the IR spectral signature of well-characterized cell lines in culture. There are, at present, series of breast cancer cell lines which have been thoroughly characterized in 2D and 3D cultures. The availability and relevance of these cell lines (for instance, 51 breast cancer cell lines mirroring 145 primary breast tumors[28]) have been carefully described for breast cancer.[29,30] Transcriptomic data are available in both 2D and 3D cultures. See for instance the *Transcription profiling of 25 breast cancer cell lines grown in 2D and 3D tissue culture conditions* published on the EMBL-EBI website at http://www.ebi.ac.uk/arrayexpress/browse.html?keywords=E-TABM-244.

In a living tissue, cells interact with other cells and with the extracellular matrix through biochemical and mechanical cues. These interactions play a crucial role in maintaining normal homeostasis and functions of the tissue[31] and, in that way, they regulate cell proliferation, migration and apoptosis. When cultured *ex vivo* on the 2D plastic surface, cells lose the possibility of communicating with their microenvironment. Cell cultures in 3D lrECM re-establish interactions with the extracellular matrix lost in conventional 2D cultures.[32] Thus, 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 profile,[35] and prognostic.[36]

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 lrECM culture and compare it with 2D culture cells and triple-negative carcinoma tissue.

Materials and methods

2D cell culture conditions

The 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, Switzerland), 50 U mL^{-1} penicillin and 50 μg mL^{-1} streptomycin (Life Technologies, USA) at 37 °C in 5% CO_2. 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 a protective agent before undergoing formalin fixation and paraffin embedding (FFPE).

The 3D lrECM 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 were suspended in 150 μl growth-factor reduced Matrigel basement membrane (BM) matrix (9.6 mg ml^{-1}; BD Biosciences, USA) at a density of 1×10^5 cells ml^{-1} and plated onto an eight-well plastic-chambered glass microscope slide (0.7 cm^2 per well; Corning, USA) precoated beforehand with 50 μl Matrigel. 3D lrECM cell cultures were

maintained in H14 medium with 5% Matrigel and 1% FBS at 37 °C in 5% CO_2 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. 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

Cell colonies embedded in Matrigel or cells collected from 2D cultures 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 microtome: one of them was mounted on a BaF_2 slide (Korth Kristalle GmbH, Germany), deparaffinized and intended for IR imaging, and 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 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 × 64 Mercury Cadmium Telluride (MCT) Focal Plane Array (FPA) detector and a 15× objective (NA = 0.4). The data were collected in transmission mode from sample regions of 180 × 180 μm^2. Every element of the FPA acts as an independent and discrete detector from which a full spectrum is obtained. The corresponding pixel covers an area of 2.8 × 2.8 μm^2. One FTIR image (unit image) results in 4096 spectra, each one being the average of 256 scans recorded in a spectral range from 3800 to 900 cm^{-1} (*ca.* 5 minutes). To cover larger sample areas, several FTIR images were juxtaposed in order to obtain one FTIR map. The background spectrum, acquired as the average of 512 scans in the absence of a sample or a BaF_2 slide. The spectral resolution was set to 8 cm^{-1}.

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

Data analyses

Pre-processing of IR spectra. All spectra were preprocessed as follows. The water vapor contribution was subtracted as described previously[38,39] with 1956–1932 cm^{-1} as the reference peak and the CO_2 peak was flattened between 2450 and 2250 cm^{-1}. The spectra were baseline-corrected. Straight lines were interpolated between the spectra points at 3620, 2995, 2800, 2395, 2247, 1765, 1724, 1480, 1355, 1144 and 950 cm^{-1} and subtracted from each spectrum. Spectra were normalized for equal areas between 1725 and 1481 cm^{-1} (amide I and II peaks). The signal-to-noise ratio (S/N) was then checked on every spectrum. It was required to be higher than 500 : 1 with noise defined as the standard deviation in the 2200–2100 cm^{-1} region of the spectrum and signal defined as the maximum of the curve between 1750 and 1480 cm^{-1} after subtracting a baseline passing through these two points. Finally, some rare spectra with normalized absorbance lower than −5 (negative lobe) and a maximum above 120 (saturation) were discarded. To avoid abrupt refractive index changes we always selected areas of the sample with contiguous cells. Visual inspection of spectra as well as systematic screening for negative lobes on the left-hand side of the amide I band did not reveal significant dispersive artifacts.[40–44] As Mie scattering corrections rely on simplified models and as dispersive artifacts were minor in the present work, we preferred to not apply such a correction. Derivation calculations did not provide more efficiency as also reported elsewhere[45] and was not applied to preprocessed IR spectra.

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

A Supervised Partial Least Square Discriminant Analysis (PLS-DA) was also performed on the data set to extract latent variables that enable the construction of a factor capable of predicting a class. It requires *a priori* knowledge of the classes of spectra and allows both data reduction and discriminative investigation. This approach consists in the application of PLS regression resulting in fewer uncorrelated variables. A linear combination of variables explaining the 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 contributions, baseline subtraction, normalization, application of quality filters, PCA, and PLS-DA analyses were carried out using Kinetics, a custom-made program running on Matlab (Mathworks, Inc.).

Results

The MDA-MB-239 triple-negative cell line has been grown under 2D and 3D culture conditions as described by Lee *et al.*[33] We decided to follow the most common protocol used in the clinic to prepare cells and tissues. As described in the Materials and methods section, 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 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 breast cancer cell lines.[46] Nevertheless, it is worth pointing out the importance of lipid metabolism in cancer progression.[47,48]

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

When growing on a 2D flat surface of a culture box, MDA-MB-231 cells form a uniform monolayer after 4–5 days. To manipulate healthy cells and avoid effects related to confluence,[49] cells were always collected at less than 75% confluence. The cell shape before detachment was either round or elongated, depending on cell cycle stage,[50] but all cells became 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 that found in tissues. MDA-MB-231 cells grown in the 3D lrECM-embedded matrix form colonies of various sizes and shapes and are phenotypically much closer to the epithelial cells found in the tissue context. Particularly, the MDA-MB-231 cell line is characterized by its invasive phenotype with stellate intercolonial projections.

Interestingly, Fig. 1D also shows that the 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 and unstained sections and their corresponding IR images is provided in Fig. 2. While individual cells can be easily identified on the H&E stained sections and, although 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} ($\lambda \approx 6$ μm). Considering the numerical aperture (NA = 0.4), diffractions limit the resolution to *ca.* 9 μm 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 Fig. S1 (ESI†).

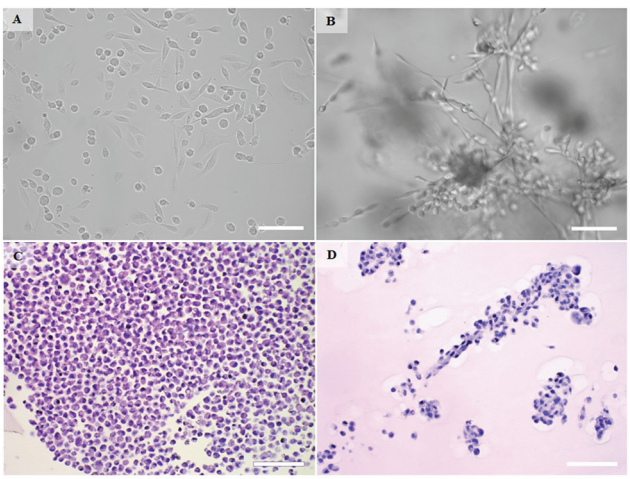

Fig. 1 Bright-field microscopy images of MDA-MB-231 breast cancer cells grown in conventional 2D culture (A and C) and in 3D lrECM-embedded culture (B and D). Images (A) and (B) have been acquired before formalin fixation; the images (C) and (D) have been obtained after FFPE treatment and H&E staining of 4 μm-thick sections. White bars: 100 μm.

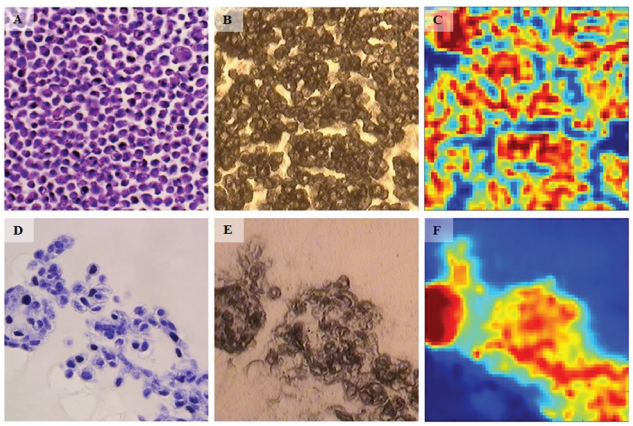

Fig. 2 Bright-field microscopy (A, B, D, E) and IR (C, F) images of adjacent 4 μm-thick sections of MDA-MB-231 breast cancer cells grown in 2D (A–C) and 3D (D–F) cultures. For 2D culture, image A shows the H&A stained section. The adjacent, unstained section is presented in image B. The IR raw image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

Spectra with sufficient S/N (see the Materials and methods section) 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 Fig. 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 Fig. 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 those grown in 2D culture. This feature is also dominant in 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.

Fig. 3 indicates some degree of separation between the 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 (Fig. 4).

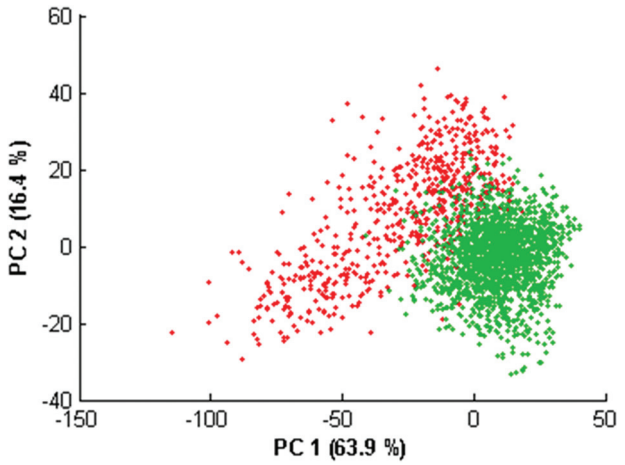

Fig. 3 PCA score plots of 600 spectra of MDA-MB-231 breast cancer cells grown in 2D culture (red) and 1752 spectra of cells derived from them and grown in 3D culture (green). Spectra are projected in the PC1–PC2 space. The fraction of the total variance explained by each PC is indicated on the axes.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

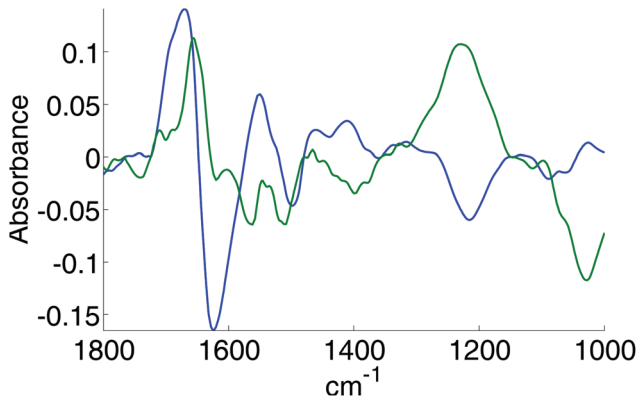

Fig. 4 Representation of PC1 (blue) and PC2 (green) obtained after principal component analyses applied to 600 spectra of MDA-MB-231 breast cancer cells grown in 2D culture and 1752 spectra of cells derived from them and grown in 3D culture (1800–1000 cm^{-1} range).

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 Fig. 3. In order to evaluate the contribution of the different spectral regions, a bootstrapping method was set up in 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 Fig. 5.

When run on the entire spectral range, the PLS-DA yielded even better correct assignment scores (not shown), indicating that nonredundant information is present in the various subregions analyzed in Fig. 5.

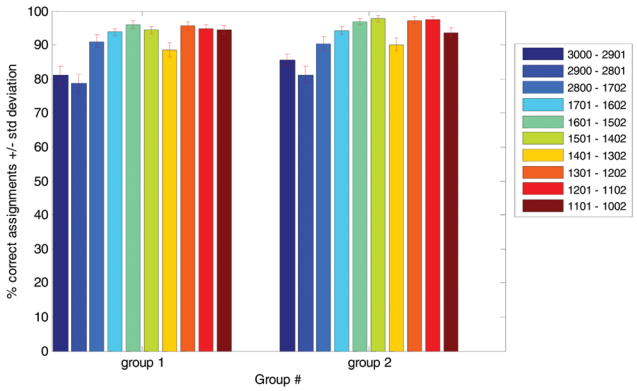

Fig. 5 Success rate (in %) of the prediction models as a function of the spectral range selected (see the color legend). Interval PLS-DA classification was computed on the spectra described in Fig. 3. The percentages of correct assignment by the prediction models are shown on the *Y*-axis; the predicted values are expressed as % of the true values. 60% of the spectral database was used for the training and the remaining 40% as the test set. The procedure was repeated 40 times, yielding 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.

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 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).

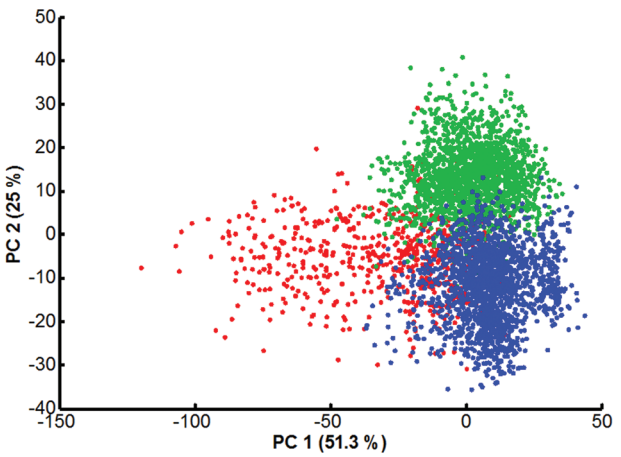

Fig. 6 PCA score plot of 600 spectra of cells grown in 2D culture (red), 1752 spectra of cells grown in 3D culture (green) and 1704 spectra of tumor epithelial cells from 3 patients (504, 648 and 552 spectra, respectively) with triple-negative breast carcinoma (blue). The fraction of the total variance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

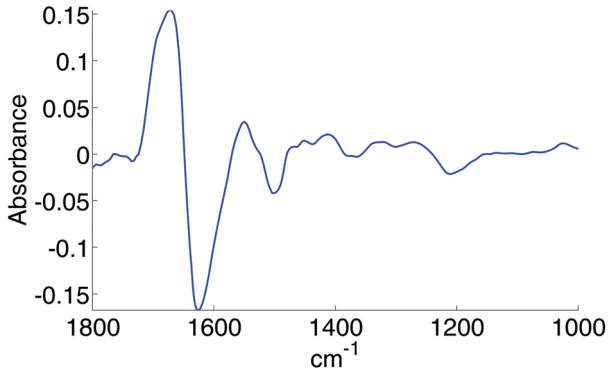

Fig. 7 Representation of PC1 obtained after principal component analyses applied to 600 spectra of cells grown in 2D culture, 1752 spectra of cells grown in 3D culture and 1704 spectra of tumor epithelial cells from 3 patients with triple-negative breast carcinoma (1800–1000 cm^{-1} range).

Fig. 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%) the total variance found in the entire dataset.

Furthermore, it appears that after adding 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 that observed in Fig. 4 (Fig. 7).

Conclusion

The present paper describes a methodology that allows direct comparison of cell phenotypes grown in 2D monolayer and 3D lrECM 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 leaves a space around the cells. After paraffin removal, there is 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 the studied epithelial cancer cell line according to its growth conditions. While an 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 those found in tissues. Nevertheless, as long as common features can be extracted, they could be identified under both conditions. 3D cultures could be improved to match better and/or understand better the *in vivo* conditions. The influence of fibroblasts[55–59] 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.

Abbreviations

2D/3D	Two-/three-dimensional
BaF_2	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
lrECM	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

Acknowledgements

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

Notes and references

1 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.
2 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.
3 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. Buness, 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.

4 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.

5 J. E. Larkin, B. C. Frank, H. Gavras, R. Sultana and J. Quackenbush, *Nat. Methods*, 2005, **2**, 337–343.

6 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.

7 P. L. Bedard and C. Sotiriou, *EMBO Mol. Med.*, 2010, **2**, 3–5.

8 A. Derenne, T. Claessens, C. Conus and E. Goormaghtigh, in *Encyclopedia of Biophysics*, 2013, pp. 1074–1081.

9 A. Derenne, O. Vandersleyen and E. Goormaghtigh, *Biochim. Biophys. Acta*, 2014, **1841**, 1200–1209.

10 M. Diem, L. Chiriboga, P. Lasch and A. Pacifico, *Biopolymers*, 2002, **67**, 349–353.

11 A. Pevsner and M. Diem, *Appl. Spectrosc.*, 2001, **55**, 1502–1505.

12 S. Boydston-White, T. Chernenko, A. Regina, M. Miljkovic, C. Matthaus and M. Diem, *Vib. Spectrosc.*, 2005, **38**, 169–177.

13 E. Kleiren, J. M. Ruysschaert, E. Goormaghtigh and V. Raussens, *Spectrosc. Int. J.*, 2010, **24**, 61–66.

14 H. H. de Jongh, E. Goormaghtigh, J. M. Ruysschaert and H. H. de-Jongh, *Biochemistry*, 1997, **36**, 13603–13610.

15 E. Goormaghtigh, V. Cabiaux and J. M. Ruysschaert, *Subcell. Biochem.*, 1994, **23**, 405–450.

16 E. Goormaghtigh, R. Gasper, A. Benard, A. Goldsztein and V. Raussens, *Biochim. Biophys. Acta, Proteins Proteomics*, 2009, **1794**, 1332–1343.

17 E. Goormaghtigh, J. M. Ruysschaert and V. Raussens, *Biophys. J.*, 2006, **90**, 2946–2957.

18 G. Bellisola and C. Sorio, *Am. J. Cancer Res.*, 2012, **2**, 1–21.

19 R. Bhargava, *Anal. Bioanal. Chem.*, 2007, **389**, 1155–1169.

20 P. Lasch, M. Diem, W. Hänsch and D. Naumann, *J. Chemom.*, 2007, **20**, 209–220.

21 P. Lasch, L. Chiriboga, H. Yee and M. Diem, *Technol. Cancer Res. Treat.*, 2002, **1**, 1–7.

22 G. J. Ooi, J. Fox, K. Siu, R. Lewis, K. R. Bambery, D. McNaughton and B. R. Wood, *Med. Phys.*, 2008, **35**, 2151–2161.

23 B. Bird, M. Romeo, N. Laver and M. Diem, *J. Biophotonics*, 2009, **2**, 37–46.

24 M. J. Walsh, A. Kajdacsy-Balla, S. E. Holton and R. Bhargava, *Vib. Spectrosc.*, 2012, **60**, 23–28.

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**, 1044–1056.

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**, 90–94.

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**, 417–422.

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**, 515–527.

29 D. L. Holliday and V. Speirs, *Breast Cancer Res.*, 2011, **13**, 215.

30 T. Vargo-Gogola and J. M. Rosen, *Nat. Rev. Cancer*, 2007, **7**, 659–672.

31 M. J. Bissell, D. C. Radisky, A. Rizki, V. M. Weaver and O. W. Petersen, *Differentiation*, 2002, **70**, 537–546.

32 M. J. Bissell, P. A. Kenny and D. C. Radisky, *Cold Spring Harbor Symp. Quant. Biol.*, 2005, **70**, 343–356.

33 G. Y. Lee, P. A. Kenny, E. H. Lee and M. J. Bissell, *Nat. Methods*, 2007, **4**, 359–365.
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.
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.
36 K. J. Martin, D. R. Patrick, M. J. Bissell and M. V. Fournier, *PLoS One*, 2008, 3.
37 H. K. Kleinman and G. R. Martin, *Semin. Cancer Biol.*, 2005, **15**, 378–386.
38 E. Goormaghtigh and J. M. Ruysschaert, *Spectrochim. Acta, Part A*, 1994, **50**, 2137–2144.
39 E. Goormaghtigh, in *Adv. Biomed. Spectrosc. (Biological and Biomedical Infrared Spectroscopy)*, ed. A. Barth and P. I. Haris, IOS Press, 2009, pp. 104–128.
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**, 609–620.
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**, 1370–1377.
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**, 126–132.
44 A. Dazzi, A. Deniset-Besseau and P. Lasch, *Analyst*, 2013, **138**, 4191–4201.
45 A. Gaigneaux, J. M. Ruysschaert 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.
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**, 320–328.
49 R. Gasper and E. Goormaghtigh, *Analyst*, 2010, **135**, 3048–3051.
50 A. Derenne, A. Mignolet and E. Goormaghtigh, *Analyst*, 2013, **138**, 3998–4005.
51 P. Lasch and D. Naumann, *Biochim. Biophys. Acta*, 2006, **1758**, 814–829.
52 E. C. Mattson, M. J. Nasse, M. Rak, K. M. Gough and C. J. Hirschmugl, *Anal. Chem.*, 2012, **84**, 6173–6180.
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**, 319–329.
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.
57 S. E. Holton, M. J. Walsh, A. Kajdacsy-Balla and R. Bhargava, *Biophys. J.*, 2011, **101**, 1513–1521.
58 M. H. Barcellos-Hoff and D. Medina, *Breast Cancer Res.*, 2005, **7**, 33–36.
59 N. A. Bhowmick, E. G. Neilson and H. L. Moses, *Nature*, 2004, **432**, 332–337.
60 W. H. Fridman, J. Galon, F. Pagès, E. Tartour, C. Sautès-Fridman and G. Kroemer, *Cancer Res.*, 2011, **71**, 5601–5605.
61 D. Liao, Y. Luo, D. Markowitz, R. Xiang and R. A. Reisfeld, *PLoS One*, 2009, **4**, e7965.

