Identification of potential prognostic biomarkers for node-negative breast tumours by proteomic analysis: A multicentric 2004 national PHRC study

FRANÇOISE DESCOTES[1], PASCAL JÉZÉQUEL[2], FRÉDÉRIQUE SPYRATOS[3], LOÏC CAMPION[4], CATHERINE GRENOT[1], FLORENCE LEREBOURS[5], MARIO CAMPONE[6], CATHERINE GUÉRIN-CHARBONNEL[4], DIDIER LANOË[2], MARJORIE ADAMS[1], JEAN ANDRÉ[1], ANTOINE CARLIOZ[7], PIERRE-MARIE MARTIN[7], AGNÈS CHASSEVENT[8], MARIE-LISE JOURDAN[9], CATHERINE GUETTE[8], ISABELLE ZANELLA-CLEON[10] and GABRIEL RICOLLEAU[2]

[1]Department of Biochemistry Molecular Biology, Hospices Civils de Lyon (HCL), CHU Lyon Sud, F-69495 Pierre Bénite Cedex; [2]Department of Oncology Biology, René Gauducheau Cancer Center, F-44805 Nantes - Saint Herblain Cedex; [3]Department of Oncogenetics, Institut Curie - Hôpital René Huguenin, F-92210 St-Cloud; [4]Biostatistics Unit, René Gauducheau Cancer Center, F-44805 Nantes - Saint Herblain Cedex; [5]Department of Medical Oncology, Institut Curie - Hôpital René Huguenin, F-92210 St-Cloud; [6]Department of Medical Oncology, René Gauducheau Cancer Center, F-44805 Nantes - Saint Herblain Cedex; [7]Transfer Laboratory of Biological Oncology, AP-HM, Faculty of Medicine North, F-13916 Marseille Cedex 20; [8]Paul Papin Cancer Center, F-49933 Angers Cedex 9; [9]Laboratory of Oncology, CORAD, CHU Bretonneau, F-37044 Tours Cedex; [10]Center for Microanalysis of Proteins, UMS3444 Lyon-Gerland Sud BioSciences, Institute of Biology and Chemistry of Proteins, F-69007 Lyon, France

Received December 22, 2011; Accepted February 20, 2012

DOI: 10.3892/ijo.2012.1456

Abstract. We used a 2D-electrophoresis (2-DE) proteomic approach to identify novel biomarkers in node-negative breast cancers. This retrospective study focused on a population of patients with ductal pN0M0 tumours. A subset of patients who developed metastases and in whose tumours were found high levels of uPA and PAI-1 (metastatic relapse, MR: n=20) were compared to another subset in whom no metastatic relapse occurred and whose tumours were found to have low levels of uPA and PAI-1 (no relapse, NR: n=21). We used a 2-DE coupled with MS approach to screen cytosol fractions using two pH-gradient scales, a broad scale (3.0-11.0) and a narrower scale focussing in on a protein rich region (5.0-8.0). This study was conducted on 41 cytosol specimens analyzed in duplicate on two platforms. The differential analysis of more than 2,000 spots in 2-DE gels, obtained on the two platforms, allowed the identification of 13 proteins which were confirmed by western blotting. Two proteins, GPDA and FABP4 were down-regulated in the MR subset whereas all the others were up-regulated. An *in silico* analysis revealed that *GMPS* (GUAA), *GAPDH* (G3P), *CFL1* (COF1) and *FTL* (FRIL), the most informative genes, displayed a proliferation profile (high expression in basal-like, HER2$^+$ and luminal B molecular subtypes). Inversely, similar to FABP4, GPD1 [GPDA] displayed a high expression in luminal A subtype, a profile characteristic of tumour suppressor genes. Despite the small size of our cohort, the 2-DE analysis gave interesting results which were confirmed by the *in silico* analysis showing that some of the corresponding genes had a strong prognostic impact in breast cancer, mostly because of their link with proliferation: *GMPS, GAPDH, FTL* and *GPD1*. A validation phase on a larger cohort is now needed before these biomarkers could be considered for use in clinical practice.

Correspondence to: Dr Françoise Descotes, Department of Biochemistry Molecular Biology, Pavillon 3D, Hospices Civils de Lyon (HCL), CHU Lyon Sud, F-69495 Pierre Bénite Cedex, France
E-mail: francoise.descotes@chu-lyon.fr

Key words: breast cancer, proteomic, 2D electrophoresis, node negative, prognosis, biomarkers

Introduction

Patients diagnosed with sporadic breast cancer have been shown to have markedly different clinical outcome and underlining the need for further molecular characterization of this disease to allow improved patient monitoring. Approximately 70% of patients with node-negative breast cancer will be metastasis-free 10 years after initial treatment without adjuvant chemotherapy, while 30% will develop metastatic relapse (MR). Traditional prognostic markers (age at diagnosis, tumour size, hormonal

receptor status, tumour grade) are not sufficient for precise risk-group discrimination in breast cancer (1). Current guidelines, such as those of St Gallen or NIH, recommend adjuvant systemic chemotherapy for more than 90% of node-negative breast cancer patients, which results in over-treatment and increases morbidity and treatment costs (2-5). Converging data have shown the importance of the two components of the plasminogen activator system, uPA and PAI-1, in predicting the risk of metastasis of node-negative breast cancer (6,7). However, despite their power, uPA and PAI-1 are not 100% reliable at predicting metastatic risk. Therefore research in node-negative breast cancer now focuses on identifying more powerful prognostic markers to help direct treatment decisions by subtyping of patients into risk subgroups who may or may not benefit from adjuvant therapy. The era of large-scale science linked both to recent technological advances and to the availability of complete genetic information, has boosted the search for new biomarkers. Today, techniques involving high-throughput screening of molecular compartments such as the genome, the transcriptome and the proteome can be accessed by many research groups. Proteome analysis is currently accomplished by a combination of two-dimensional gel electrophoresis (2-DE) to separate and visualize proteins and mass spectrometry (MS) or tandem MS (MS/MS) to identify the selected protein spots. Important technical advances regarding 2-DE and protein MS have increased the sensitivity, reproducibility and throughput of proteome analysis while creating an integrated technology. While much effort has been devoted to screening the breast cancer tumour transcriptome to identify prognostic markers, the breast cancer tumour proteome has been less investigated (8-12). To our knowledge, aside one of our works based on SELDI-TOF-MS profiling, no other study using high throughput proteomic approaches has focused on node-negative breast cancer tumours to identify prognostic biomarkers (13).

In the present study, we used a 2-DE coupled with MS approach to screen the cytosol fraction of node-negative breast cancers. We focused on a population of patients with ductal pN0M0 tumours showing clinically homogeneous behaviour but which were biologically heterogeneous potentially leading to their opposing outcomes. We compared patients with extreme levels of uPA and PAI-1 associated with opposite outcomes to identify robustly differentially expressed proteins: patients with tumours containing high levels of uPA and PAI-1 and who developed metastases (MR) were compared to those with tumours containing low levels of uPA and PAI-1 and who did not suffer metastatic relapse (NR). These very restrictive selection criteria associated with at least 10-years follow-up required the recruitment of patients from 4 medical centres.

This study was performed on two proteomic platforms with experience using 2-DE situated in Nantes and Lyon and was conducted blind using two complementary methodological approaches to optimize results.

Materials and methods

Patient selection and tumour characteristics. This study retrospectively included 41 selected women with sporadic node-negative primary breast tumours, diagnosed and treated in 4 medical centres (Institut Curie - René Huguenin Hospital, St Cloud; Centre René Gauducheau, St Herblain; Centre Hospitalier Lyon sud; AP-HM, Marseille), between 1981 and 1998. Median age at diagnosis was 54 years (range 36-76). Informed consent was obtained from patients to use their surgical specimens and clinicopathological data for research purposes, as required by the French Committee for the Protection of Human Subjects. All patients had an infiltrating ductal carcinoma and showed no evidence of distant metastasis at the time of diagnosis. None had received chemotherapy, endocrine therapy or radiation therapy prior to surgery. Treatment decisions were based solely on consensus recommendations at the time of diagnosis. Patients were followed-up for MR (nodal or distant metastasis; local relapses were excluded). For the study we selected from a population of patients with invasive ductal node-negative tumours without adjuvant chemotherapy, two groups as different as possible both from an evolutionary and a biological perspective to identify differentially expressed proteins. Twenty patients (MR subset) with metastatic relapse and whose tumours had high levels of uPA and PAI-1 (median time to metastasis 2.6 years; range 0.48-8.39) were compared with 21 patients (NR subset) without metastatic relapse and whose tumours had low levels of uPA and PAI-1. For this second group, the median follow-up was 15.25 years (range 11.26-21.99). The clinicopathologic characteristics of the patients are indicated in Table I.

Cytosol preparation. The cytosolic fractions used in this study were initially prepared for hormone receptor assays at the time of surgery in the four centres of sample collection using the same protocol by homogenizing fragments in the following buffer: 10 mM Tris-HCl, pH 7.4, 1.5 mM EDTA, 10 mM sodium molybdate, 0.5 mM dithiotreitol, 10% glycerol. Homogenates were centrifuged at 105,000 g for 1 h and supernatants were aliquoted and thereafter stored in liquid nitrogen until use. For this study, protein concentrations were remeasured retrospectively in one laboratory using the Bradford method (14).

Selection of the two studied subsets. Each of the four centres measured uPA and PAI-1 in their samples by the enzyme-linked immunoassay (uPA: Imubind no. 894; PAI-1, Imubind no. 821, American Diagnostica) using the method developed within their respective laboratory (15,16). In each laboratory, cytosols were classified into 4 quartile groups based on levels of uPA and PAI-1. From cytosols with low uPA and PAI-1 (1st quartile) and high uPA and PAI-1 (4th quartile), two subsets of samples were retained: respectively 21 NR and 20 MR as defined in 'Patient selection and tumour characteristics'.

Two-dimensional gel electrophoresis

Team Nantes. For the first dimension, 24-cm Immobiline DryStrips [ReadyStrips IPG Strips 24 cm, pH 3.0-11.0 non-linear (NL) (analytical gels and preparative gels), GE Healthcare, Aulnay sous Bois, France] were passively rehydrated overnight with 450 μl Destreak buffer (GE Healthcare) and isoelectric focusing (IEF) was performed using a PROTEAN IEF Cell electrophoresis system (Bio-Rad, Hercules, CA) at 19°C. Samples containing a mix of IEF markers (Bio-Rad) and either 30 μg of cytosolic proteins for analytical gels or 200 μg for preparative gels were diluted in a buffer composed of 7 M urea, 2 M thiourea, 4% w/v CHAPS, 0.4% Triton X-100, 0.5% IPG buffer pH 3.0-11.0 NL (GE Healthcare), 0.28% DTT,

Table I. Clinicopathologic characteristics of the 41 studied patients.

Characteristics	No metastatic relapse subset	Metastatic relapse subset	p-value
n	21	20	
Median age at diagnosis (years)	54 (37-71)	55 (36-76)	0.548
Tumour size (mean, mm)	15 (10-67)	25 (15-40)	0.002
Tumour size (mm)			
pT <20	14	6	0.029
pT ≥20	7	14	
SBR grade[a]			
I	4	3	0.220
II	14	9	
III	3	8	
Estrogen receptor receptor status			
Positive	15	14	1.000
Negative	6	6	
Progesterone receptor status			
Positive	10	10	1.000
Negative	11	10	
Adjuvant therapy			
Radiotherapy	18	11	0.043
Hormonotherapy	1	2	0.606
NPI[b] (mean score)	3.3 (2.2-4.5)	3.5 (2.4-4.8)	0.017
NPI grade			
1	14	8	0.121
2	7	12	
3	0	0	
AOL[c] 10-year-relapse risk	25 (15-50)	37 (25-47)	0.010

[a]SBR, Scarff Bloom Richardson; [b]NPI, Nottingham prognostic index; [c]AOL, Adjuvant! Online.

and traces of Orange G dye to obtain a final volume of 80 µl before being cuploaded on strips. The linear ramping mode of the IEF voltage for analytical and preparative gels was applied according to the following parameters: 250 V for 1 h and 10,000 V for 1 h with rapid ramping until a total of 160,000 Vh was reached. Following IEF, the strips were collected in a plastic bag and frozen at -80˚C until use.

For the second dimension, the strips were first equilibrated at room temperature for 15 min in 10 ml of 50 mM Tris-HCl pH 8.8 buffer containing 6 M urea, 2% SDS, 30% glycerol and 65 mM DTT (reduction step) and then transferred to 10 ml of 50 mM Tris-HCl pH 8.8 buffer containing 6 M urea, 2% SDS, 30% glycerol and 135 mM iodoacetamide for 20 min (alkylation step). The strips were washed with electrophoresis buffer (25 mM Tris-HCl pH 8.3, 192 mM glycine, 0.1% SDS), fixed with 0.5% agarose on top of a 24x20-cm polyacrylamide gradient gel (9-18%) and run on an Ettan™ DALTsix electrophoresis system (GE Healthcare) according to the manufacturer's recommendations. The electrophoresis was performed at 40 V for 1 h then at 300 V for 16 h (50 mA maximum) at 10˚C, and was stopped when the LMW marker orange G dye (Sigma-Aldrich, Saint-Quentin Fallavier, France) used as a reference had disappeared. Each sample was treated in at least two independent 2-DE replicates.

Team Lyon. For the first dimension, samples containing a mixture of ampholytes and either 30 µg of cytosolic proteins for analytical gels or 200 µg for preparative gels were mixed in a rehydration buffer composed of 7 M urea, 2.3 M thiourea, 4% w/v CHAPS, 0.24% Triton X-100, 1X Biolyte 3/10 ampholyte (Bio-Rad), 20 mM DTE, and traces of bromophenol blue to obtain a final volume of 350 µl. IPG strips [17 cm pH 5.0-8.0 (Bio-Rad)] were passively rehydrated. IEF was performed using the PROTEAN IEF Cell electrophoresis system (Bio-Rad). The linear ramping mode of the IEF voltage for analytical and preparative gels was applied according to the following parameters: 50 V for 1 h, 200 V for 1 h, 1000 V for 1 h and 10,000 V during 2 h until a total of 40,000 Vh was obtained, before a final step of 500 V for 18 h. After IEF, the strips were collected and frozen at -80˚C until use.

For the second dimension, the strips were first equilibrated at room temperature for 10 min in 6 ml of 375 mM Tris-HCl pH 8.8 buffer containing 6 M urea, 2% SDS, 20% glycerol and 133 mM DTE (reduction step) before being transferred to 6 ml of 375 mM Tris-HCl pH 8.8 buffer containing 6 M urea, 2% SDS, 20% glycerol and 135 mM iodoacetic acid for 10 min (alkylation step). The strips were then washed with electrophoresis buffer (25 mM Tris-HCl pH 8.3, 192 mM glycine, 5 mM sodium thiosulfate, 0.2% SDS), fixed with 0.5% agarose on top of a 19x20 cm polyacrylamide gradient gel (9-18%) and run on a Protean II Xi Multi Cell electrophoresis system (Bio-Rad) according to the manufacturer's recommendations. The electrophoresis was performed at 50 V for 1 h and at 250 V for 16 h (35 mA/gel) at 8˚C. For each sample, 2-DE was repeated at least twice in independent replicates.

Detection of proteins by silver staining. In both laboratories, analytical gels were silver-stained (17). For preparative gels, the silver-staining protocol was performed as described by O'Connell and Stults (18). When necessary, the preparative gels were stored in 1% acetic acid at 4˚C for no more than 30 days, until the excision of protein spots.

Image analysis

Team Nantes. The stained gels were digitized using a densitometer (ImageScanner, Bio-Rad) and analyzed with the Melanie 3

computer software (GeneBio, Geneva, Switzerland). Spot detection, quantification and matching were performed with Melanie 3 tools. After choosing manually 420 common spots as landmarks, the alignment and matching was performed. Spot intensities given as spot volumes were normalized to the total spot volume. A master gel was prepared by the combination of two sub-master gels, MR-pool and NR-pool. On one side, 1.5 μg of each of the 20 MR samples were mixed to obtain the MR-pool and on the other the 21 NR samples for the NR-pool. The master gel was then constructed by numerical combination of the two sub-master gels.

Team Lyon. The stained gels were digitized using a densitometer (ImageScanner, GE Healthcare) and analyzed with the Progenesis SameSpots version v3.2 (Nonlinear Dynamics, Newcastle upon Tyne, UK). Before spot detection, all the images were aligned (100% matching) to a reference image. From 2,015 aligned spots in all images, the statistical analysis performed with the Progenesis SameSpots software with three criteria: fold >1.5, p-Anova <0.05 and power >0.8, permitted the selection of spots differentially expressed between the 2 subsets.

Statistical analysis

Discriminating spots identification. The first aim was to identify robust spots displaying statistically significant differences in intensity according to patient outcome. For each patient, the 2,404 Nantes and the 2,015 Lyon relative spot intensity values were extracted twice and the two replicates with the highest correlation coefficient were kept (r>0.80). Both individual values and the mean of the two values were then used for statistical analysis. In order to strengthen the accuracy of the spots selection in this small population, it was necessary to use methods with repeated permutations. The local false discovery rate (FDR) and q-value of each spot were calculated by means of SAM software distributed by Stanford University (19). The FDR and permutation p-values of each spot were also calculated by means of BRB-Arrays 3.8.1 developed by Dr Richard Simon and BRB-ArrayTools Development Team.

List of discriminating spots. Spots in the Nantes and Lyon datasets showing significance (based on FDRs $<10^{-3}$ and/or permutation p- or q-value $<10^{-3}$) with both statistical tools in both series and in the averaged series were selected for further analyses. Nantes and Lyon selected spots were pooled. Then, a top-14 spots was determined based on best FDR scores.

Biomarker identification and database search

Team Nantes. Spots were recovered using the Ettan Spot picker (GE Healthcare) and individually treated with the Proteoextract All-in-one Trypsin Digestion kit (Calbiochem, Merck Chemicals Ltd., Nottingham, UK) according to the manufacturer's instructions. Peptide digests were concentrated on C18-Zip Tips (Millipore, Billerica, MA, USA) and eluted with 1.5 μl α-cyano-4-hydroxycinnamic acid matrix solution (in 50% acetonitrile, 0.05% trifluoroacetic acid) onto a MALDI-TOF target plate (Opti-TOF 384 well insert, AB Sciex, France). Peptide spectra acquisition was realized on the 4800 MALDI TOF/TOF analyzer (AB Sciex). After screening the sample position in MS-positive reflector mode using 1500 laser shots, the fragmentation of automatically-selected precursors was performed with 1 kV collision energy using air as the collision gas (pressure of $2x10^{-6}$ Torr). MS spectra were acquired between m/z 800 and 4000. Up to 12 of the most intense ion signals having an S/N >12 were selected as precursors for MS/MS acquisition. Peaklist generation and protein identification were performed by the ProteinPilot™ Software V 2.0 (AB Sciex) using the Paragon algorithm. Each MS/MS spectrum was searched across all species and then against the Homo Sapiens (Swissprot-Uniprot) selected database. The searches were run with the fixed modification of iodoacetamide-labelled cysteine parameters. Other parameters such as tryptic cleavage specificity, precursor ion mass accuracy and fragment ion mass accuracy are MALDI 4800 built-in functions of ProteinPilot software. The preliminary protein identifications obtained automatically from the software were inspected manually for conformation prior to acceptance.

Team Lyon

In-gel digestion. Silver-stained protein spots were excised from the preparative gels and destained according to the method of Gharahdaghi *et al* (20). Proteolytic digestion was done in-gel using 5-10 μl of a trypsin solution (10 ng/μl in 50 mM NH_4HCO_3; sequence-grade trypsin, Promega, France) for 45 min at 50˚C. The resulting tryptic peptides from the supernatant fraction were recovered. A second extraction step was performed using 10-15 μl of $H_2O/CH_3CN/HCOOH$ (60/36/4; v/v/v) for 30 min at 30˚C. Finally, all extracts were pooled and dried in a vacuum concentrator and then resuspended in 0.3% trifluoroacetic acid (12 μl) for LC-MS/MS analyses.

NanoLC/nanospray/tandem mass spectrometry (LC-ESI/MS/MS). Experiments were performed on a Q-STAR XL instrument (AB Sciex). The information dependent acquisition (IDA) mode allowed peptide ions within a m/z 400-1600 survey scan mass range to be analyzed for subsequent fragmentation. MS/MS spectra were acquired in the m/z 65-2000 range for the +2 to +4 charged ions. The collision energy was automatically set by the software (Analyst 1.1, AB Sciex) and was related to the mass and the charge of the precursor ion. Tryptic peptides were separated using an Ultimate-nanoLC (Dionex, France) with a C18 PepMap micro-precolumn (5 μm; 100 Å; 300 μm x 5 mm; Dionex) and a C18 PepMap nanocolumn (3 μm; 100 Å; 75 μm x 150 mm; Dionex). The chromatographic separation was developed using a linear 30 min gradient from 0 to 60% B, where solvent A was 0.1% HCOOH in H_2O/CH_3CN (95/5) and solvent B was 0.1% HCOOH in H_2O/CH_3CN (20/80) at 300 nl/min flow rate. Protein identification was achieved by the Paragon Algorithm from the ProteinPilot software (AB Sciex) against the SwissProt database limited to the human species.

Western blotting. Cytosolic proteins (20 μg) were electrophoresed on a 12.5% polyacrylamide gel at 400 V for 2 h according to the standard protocol (21). The proteins were then transferred to nitrocellulose for 7 min at 20 V using an iBlot Dry Blotting System (Invitrogen, Carlsbad, CA) according to the manufacturer's protocol. After an overnight incubation at 4˚C in a blocking buffer (25 mM Tris-HCl, pH 7.4, 150 mM NaCl, 0.05% Tween-20) containing 5% non-fat dry milk, the membranes were incubated for 2 h at room temperature in the above buffer containing 1% non-fat dry milk and one of the primary antibodies at an appropriate dilution (Table II). After washing for 30 min with

Table II. Western blotting conditions.

Protein	Reference	Dilution	Host
ARPC3	PAB6621, Abnova, Taipei, Taiwan	1:80	Goat
COF1	PAB5302, Abnova	1:400	Rabbit
IF5A1	NB600-1077, Novus Biologicals, Litlleton, CO, USA	1:1000	Rabbit
FABP4	H00002167-A01, Abnova	1:1000	Mouse
FRIL	FERL14A, Alpha Diagnostic, San Antonio, TX, USA	1:125	Rabbit
FUBP1	H00008880-A01, Abnova	1:500	Mouse
G3P	H00002597-A01, Abnova	1:1500	Mouse
GUAA	H00008833-A01, Abnova	1:250	Mouse
GPDA	H00002819-A01, Abnova	1:1000	Mouse
NNMT	H00004837-A01, Abnova	1:250	Mouse
PPIB	SP5113P, Acris Antibodies GmbH, Herford, Germany	1:333	Rabbit
SNX6	AP1464A, Abgent, San Diego CA, USA	1:50	Rabbit
TAGL	sc-53932, Santa Cruz Biotechnology, Santa Cruz, CA, USA	1:200	Mouse

blocking buffer without milk, the membranes were incubated first for 90 min with anti-mouse IgG (1:4000, no. E0354, Dako, Glostrup, Denmark), anti-rabbit IgG (1:4000, no. E0432, Dako) or anti-goat IgG (1:4000, no. E0466, Dako) biotinylated secondary antibodies, and for 1h with streptavidin biotinylated horseradish peroxydase (HRP) conjugate (GE Healthcare). The chemiluminescent signals were visualized using the SuperSignal West Pico chemiluminescent substrate (Thermo Fisher Scientific, Lafayette, CO) according to the manufacturer's protocol and the light emission was captured with the LAS 4000 camera (Fujifilm Medical Ltd., Tokyo, Japan) and signal intensity was analysed by the Multi Gauge software v3.0 (Fujifilm). The membrane was reprobed with anti-β-actin (1:50000, no. A5541, Sigma-Aldrich) for loading control.

In silico analysis. Prognostic informativity of these protein biomarkers was explored at the RNA level by means of bc-GenExMiner v2.0 web tool (22). Briefly, this user-friendly, web-based application based on breast cancer DNA microarray results permits the evaluation of gene prognostic informativity by means of targeted and exhaustive analysis. The prognostic impact of genes is evaluated by means of a Cox proportional hazards model. A targeted analysis performs gene-based survival analyses for each cohort separately and all cohorts pooled together (Table III); cohorts may be split by N and estrogen receptor (ER) status and analyses may be based on 5 different event criteria. In an exhaustive analysis, the prognostic informativity is evaluated for each of the possible pools corresponding to every combination of population (N and ER status) and event criteria. Furthermore, bc-GenExMiner provides gene expression maps according to molecular subtypes.

Results

The same 41 breast cancer cytosols (21 NR and 20 MR) were analyzed by 2-DE analysis using complementary methodological approaches, using a broad pH range (3.0-11.0) in the first dimension and a narrower pH range (5.0-8.0) focusing in on a protein-rich region. All cytosols were screened twice in two independent tests and the gels were independently analyzed with Melanie or Progenesis SameSpots Software. The two sets of 2-DE gels obtained using the broad range (NL pH 3.0-11.0, Nantes) and the narrower range (pH 5.0-8.0, Lyon) presented 2,404 spots and 2,015 spots respectively. Representative 2-DE-gels obtained in the 2 site conditions are displayed in Fig. 1.

Statistical analysis and protein identification. The spots of two series of 2-DE gels were compared with robust permutation techniques (based on FDRs $<10^{-3}$ and/or 10,000 permutation p- or q-value $<10^{-3}$) permitting the selection of 65 and 26 spots, for Lyon and Nantes, respectively (Fig. 1). Among the 91 spots, 14 spots were finally retained based on FDR best scores: 9 spots from pH 5.0-8.0 and 5.0 from pH 3.0-11.0 (Table IV) and their identity revealed using mass spectrometry and database search. Among these proteins, transgelin (TAGL) was found by both teams. Other proteins were: nicotinamide N-methyltransferase (NNMT), glutamine amidotransferase (GUAA), ferritin light chain (FRIL), actin related protein 2/3 complex, subunit 3 (ARPC3), glyceraldehyde-3-phosphate dehydrogenase (G3P), peptidyl-prolyl *cis-trans* isomerase B (PPIB), far upstream element-binding protein 1 (FUBP1), glycerol-3-phosphate dehydrogenase (GPDA), eukaryotic translation initiation factor 5A-1 (IF5A1), fatty acid-binding protein, adipocyte type (FAPB4), sorting nexin-6 (SNX6) and cofilin 1 (COF1) (Table V).

Western blotting. The differential expression of these proteins between the NR and MR subgroups tested in 2-DE, was validated by western blot analysis (Fig 2 and Table VI). As shown by 2-DE analysis most of the identified proteins (n=11) were overexpressed in the MR group. By contrast, FABP4 and GPDA were overexpressed in the NR group.

In silico analysis

Targeted analysis. Expression results at the RNA level with MR as the event were concordant with 2-DE ones concerning expression in node-negative breast cancer patients for *GMPS*

Table III. Genomic data included in bc-GenExMiner v2.0.

Authors (refs.)	N patients	Study code	Platform name	Platform code	DNA chip	N unique genes (2009)	bc-GenExMiner version
Van de vijver *et al* (10)	295	Rosetta2002	Agilent		25k oligo custom	12221	1.0
Sotiriou *et al* (40)	99	PNAS1732912100	NCI		8k cDNA custom	4039	1.0
Ma *et al* (41)	59	GSE1378	Arcturus	GPL1223	22k oligo custom	15246	1.0
Minn *et al* (42)	82	GSE2603	Affymetrix	GPL96	HG-U133A	12778	1.0
Pawitan *et al* (43)	159	GSE1456	Affymetrix	GPL96-GPL97	HG-U133A + B	17956	1.0
Wang *et al* (12)	286	GSE2034	Affymetrix	GPL96	HG-U133A	12778	1.0
Weigelt *et al* (44)	89	GSE2741	Agilent	GPL885-GPL887-GPL1390	011521 Human 1A G4110A-012097 Human 1A v2 G4110B-Human 1A oligo UNC custom	13524	1.0
Bild *et al* (45)	158	GSE3143	Affymetrix	GPL91	HG-U95A v2	8877	1.0
Chin *et al* (46)	112	E_TABM_158	Affymetrix	A-AFFY-76	HG-U133A v2	12778	1.0
Ivshina *et al* (47)	249	GSE4922	Affymetrix	GPL96-GPL97	HG-U133A + B	17956	1.0
Desmedt *et al* (48)	198	GSE7390	Affymetrix	GPL96	HG-U133A	12778	1.0
Loi *et al* (49)	401	GSE6532	Affymetrix	GPL96-GPL97-GPL570	HG U133A + B + P2	20763	1.0
Minn *et al* (50)	58	GSE5327	Affymetrix	GPL96	HG-U133A	12778	1.0
Naderi *et al* (51)	135	HG-U133A	Agilent	A-AGIL-14	Human 1A oligo G4110A	14730	1.0
Anders*et al* (52)	75	GSE7849	Affymetrix	GPL91	HG-U95A v2	8877	1.0
Chanrion *et al* (53)	155	GSE9893	MlRG	GPL5049	Human 21k v12.0	13182	1.0
Loi *et al* (54)	77	GSE9195	Affymetrix	GPL570	HG-U133P2	20254	1.0
Schmidt *et al* (55)	200	GSE11121	Affymetrix	GPL96	HG-U133A	12778	1.1
Calabro *et al* (56)	139	GSE10510	DKFZ	GPL6486	35k oligo	16842	1.0
Jézéquel *et al* (57)	252	GSE11264	UMGC-IRCNA	GPL4819	9k cDNA custom	5776	1.0
Zhang *et al* (58)	136	GSE12093	Affymetrix	GPL96	HG-U133A	12778	1.1

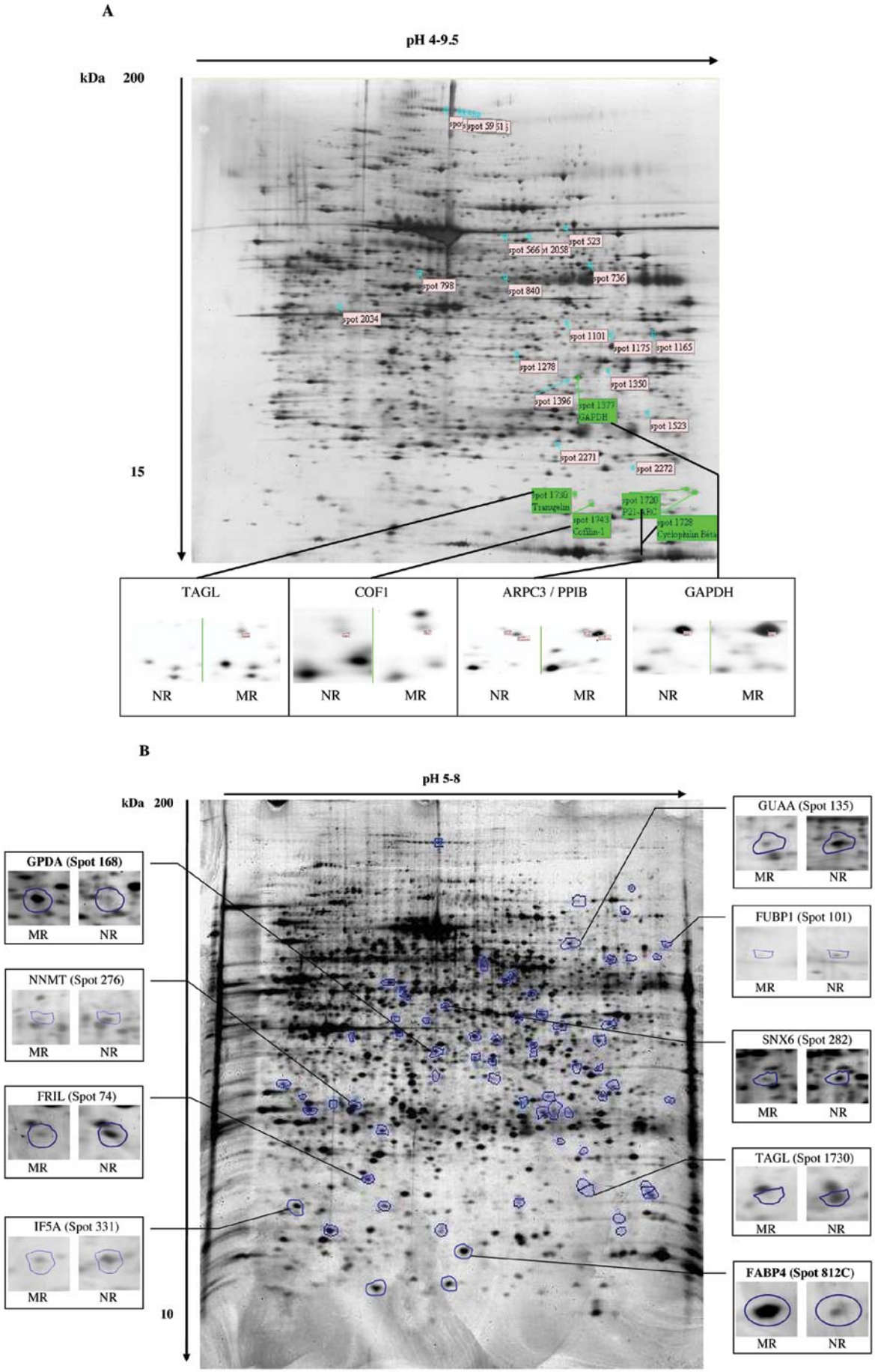
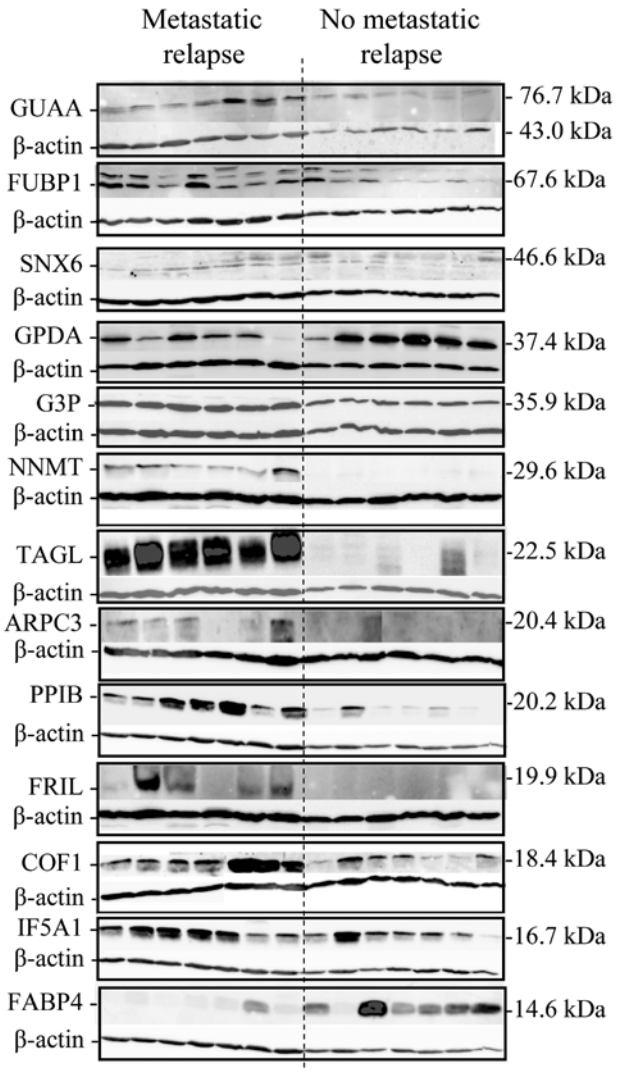


Figure 1. Typical 2-DE gels. Representative 2-DE pattern [non-linear pH 3.0-11.0 (A) or pH 5.0-8.0 (B) in the first dimension and 9-18% SDS-PAGE in the second] of cytosol protein content from a node-negative breast tumour. Proteins are silver-stained. Discriminating protein spots submitted to MS identification are displayed with a numbered tag. MR, metastatic relapse; NR, no metastatic relapse as defined in Materials and methods.

Table IV. Identification of top-13 selected proteins: statistical classification based on FDR best scores and protein characteristics.

	Statistical analysis								
	BRB-array		SAM						
Class	FDR[a]	Perm-p[b]	Local FDR	q-value	Protein	Spot no.	pI	MW	Team
1	10^{-7}	10^{-7}	$5x10^{-5}$	0	TAGL	20	6.55	22,414	Lyon
						1730	7.43	21,552	Nantes
2	10^{-7}	10^{-7}	$2x10^{-5}$	0	NNMT	276	5.56	29,650	Lyon
3	$8x10^{-5}$	10^{-7}	0	0	GUAA	135	6.42	77,001	Lyon
4	$4x10^{-6}$	10^{-7}	$1.3x10^{-2}$	0	FRIL	74	5.55	19,888	Lyon
5	$5x10^{-6}$	10^{-7}	0	0	ARPC3	1720	9.23	22,291	Nantes
6	10^{-5}	10^{-7}	$1.1x10^{-2}$	0	G3P	1377	7.48	36,000	Nantes
7	$3x10^{-5}$	10^{-7}	$6.7x10^{-2}$	0	PPIB	1728	9.50	21,650	Nantes
8	$8x10^{-5}$	$2x10^{-4}$	0	0	GPDA	168	5.88	37,439	Lyon
9	$3x10^{-5}$	10^{-7}	0	0	IF5A1	331	5.10	17,401	Lyon
10	$2x10^{-5}$	10^{-4}	0	0	FUBP1	101	7.16	67,560	Lyon
11	$2x10^{-4}$	10^{-4}	$5.4x10^{-2}$	0	SNX6	282	5.81	46,649	Lyon
12	$2x10^{-5}$	10^{-7}	$1.9x10^{-2}$	0	FABP4	812-C	6.80	14,588	Lyon
13	10^{-7}	10^{-7}	$5x10^{-5}$	0	COF1	1743	7.69	20,298	Nantes

[a]FDR, false discovery rate; [b]Perm-p, 10,000 permutation p-value; pI, experimental value of isoelectric point; MW, experimental value of molecular weight.



(GUAA), *GAPDH* (G3P), *CFL1* (COF1), *GPD1* (GPDA). A tendency was found for *SNX6, NNMT* and *FTL* (FRIL).

Exhaustive analysis. Prognostic informativity screening in different breast cancer cohorts displayed decreasing performance from *GMPS* to *GAPDH, CFL1, GPD1, FTL, TAGLN, NNMT* and *SNX6*. No prognostic informativity was found for *ARPC3, FABP4, PPIB, FUBP1* and *EIF5A*. Nottingham Prognostic Index (NPI)- and Adjuvant! Online (AOL)-adjusted Cox proportional hazards models showed that six markers (*TAGLN, NNMT, GMPS, FTL, FUBP1* and *CFL1*) bore independent prognostic informativity (23,24); the best ones being: *CFL1* (COF1), *FTL* (FRIL) and *NNMT*. Four of the five most prognostic genes [*GMPS* (GUAA), *GAPDH* (G3P), *CFL1* (COF1) and *FTL* (FRIL)] displayed a proliferation profile (high expression in basal-like, $HER2^+$ and luminal B molecular subtypes). These results are in concordance with numerous works showing that prognosis in breast cancer is largely dependent on proliferation status (25,26). Inversely, similar to *FABP4, GPD1* (GPDA) displayed a high expression in luminal A subtype and a low expression in the three other subtypes. This profile is characteristic of tumour suppressor genes. This possibility was previously evoked for FABP4 for the majority of bladder cancer (27). Gene expression correlation analyses showed a moderate correlation between *TAGLN* and *NNMT* [r(Pearson) = 0.59; p<0.001], and *GDP1* (GPDA) and *FABP4* [r(Pearson) = 0.57; p<0.001]. Except for 4 selected proteins (FUBP1, SNX6, PPIB and EIF5A), which

Figure 2. Thirteen biomarkers analyzed by western blotting in node-negative breast cancer patients with metastatic relapse (MR) or with no metastatic relapse (NR).

Table V. Results obtained from mass-spectrometry identification of selected spots.

Spot no.	Accession no.	Mascot score[a]	Prot score[b]	% cov.	m/z	Peptide sequence	Protein name and symbol	Molecular function
20	Q01995		4.35	14.4	964.49	AAEDYGVIK	Transgelin, TAGL	Muscle organ development; actin binding protein
					1209.55	KYDEELEER		
					1294.61	EFTESQLQEGK		
1730		99			1204.72	TLMALGSLAVTK		
		99			1976.08	DMAAVQRTLMALGSLAVTK		
276	P40261		7.96	20.1	1084.63	GPEKEEKLR	Nicotinamide N-methyl transferase, NNMT	Methyltransferase
					2174.01	MESGFTSKDTYLSHFNPR		
					1311.83	VKGPEKEEKLR		
					912.61	QAVKQVLK		
					815.47	EAVEAAVK		
					875.53	FSSLPLGR		
135	P49915		12.01	41.0	1866.87	IMYDLTSKPPGTTEWE	Glutamine amido transferase, GUAA	Purine biosynthesis
					2774.43	VINAAHSFYNGTTTLPISDEDRTPR		
					3050.61	VRELFVQSEIFPLETPAFAIKEQGFR		
					1258.65	SGNIVAGIANESK		
					1470.89	MVTEIKKIPGISR		
					1533.83	TLNMTTSPEEKRK		
					1312.66	QADFEAHNILR		
74	P02792		3.70	21.1	1606.81	LGGPEAGLGEYLFER	Ferritin light chain, FRIL	Ferric iron binding, cellular membrane organization
					833.46	ALFQDIK		
1720	O15145	99			1210.75	LIGNMALLPIR	Actin-related protein 2/3 complex, subunit 3, ARPC3	Actin binding, structural constituent of cytoskeleton
1377	P04406	99			1763.83	LISWYDNEFGYSNR	Glyceraldehyde-3-phosphate dehydrogenase, G3P	Glycolysis, protein binding
		99			1319.72	LEKPAKYDDIKK		
1728	P23284	63			1842.95	SIYGERFPDENFKLK	Peptidyl-prolyl cis-trans isomerase B, PPIB	Peptide binding
		99			812.46	VYFDLR		

Table V. Continued.

Spot no.	Accession no.	Mascot score[a]	Prot score[b]	% cov.	m/z	Peptide sequence	Protein name and symbol	Molecular function
101	Q96AE4		2.5	4.8	1335.73	IGGNEGIDVPIPR	Far upstream element-binding protein 1, FUBP1	RNA binding, protein binding, regulation of transcription
168	**P21695**		5.83	15.5	1014.63	LISEVIGER	**Glycerol-3-phosphate dehydrogenase, GPDA**	**NAD or NADH binding, protein homodimerization activity**
					984.53	GVDEGPNGLK		
					890.53	KVAEAFAR		
					986.62	ANATGISLIK		
					870.51	LQGPETAR		
331	P63241		2.51	20.8	1297.74	VHLVGIDIFTGK	Eukaryotic translation initiation factor 5A-1, IF5A1	Induction of apoptosis, elongation factor
					1340.68	EDLRLPEGDLGK		
812C	**P15090**		7.38	34.8	1062.64	EVGVGFATRK	**Fatty acid-binding protein, adipocyte type, FAPB4**	**Fatty acid binding, transport process**
					2363.16	LVSSENFDDYMKEVGVGFATR		
					1034.58	WDGKSTTIK		
					1446.67	LVSSENFDDYMK		
					1190.70	WDGKSTTIKR		
282	Q9UNH7		3.52	5.9	1750.94	SSLPNFKQNEFSVVR	Sorting nexin-6, SNX6	Protein transport
					1093.59	VSELFDKTR		
1743	P23528	99			1337.62	YALYDATYETK	Cofilin 1, COF1	Actin binding, anti-apoptosis
		99			1380.75	KAVLFCLSEDKK		

In bold characters, down-regulated protein spots in the metastatic relapse subset (MR). [a]Probability based Mowse score [ions score = 10*log (P), where P is the probability that the observed match is a random event. Protein scores >55 are considered significant]. [b]Percent confidence, expressed in ProtScore units with ProtScore >2 percent confidence >99%. The m/z correspond to $(M+H)^{1+}$ for the spots 1730, 1720, 1377, 1728, 1743 and to the calculated experimental M for the other spots (20, 276, 135, 74, 101, 168, 331, 812C and 282).

Table VI. Mann-Whitney p-values of western blot analysis.

Protein name	p-value*
TAGL	<0.001
NNMT	<0.001
GUAA	0.025
FRIL	0.015
ARPC3	0.003
G3P	0.025
PPIB	<0.001
GPDA	0.028
IF5A1	0.039
FUBP1	0.009
SNX6	0.008
FABP4	<0.001
COF1	0.021

[a]Mann-Whitney test.

displayed discordant protein/RNA expression patterns and/or low prognostic performance, all other candidates showed promising results.

Discussion

This retrospective study focused on a population of patients with ductal pN0M0 tumours: a subset of patients whose tumours have high levels of uPA and PAI-1 and who had developed metastases were compared to another subset of patients with tumours presenting low levels of uPA and PAI-1 and who did not relapse. We used a 2-DE coupled with MS approach to screen cytosol fractions using two pH scales, a broad scale (3.0-11.0) and a narrower scale focusing in on a protein rich region (5.0-8.0). The differential analysis of 2-DE gels allowed the identification of 13 proteins which were confirmed by western blotting. Two proteins, GPDA and FABP4 were down-regulated in the MR subset while all the others were up-regulated.

The quest for novel cancer biomarkers currently involves different '-omic' compartment screenings to increase success rate. Integrated approaches may also strengthen the selection and/or validation of molecular candidates. In this study, we chose proteomic analysis to identify prognostic markers of node-negative breast cancer; analytical validation was performed by means of western blotting. *In silico* study then confirmed that 9 out of the 13 selected proteins with concordant RNA expression patterns had a significant prognostic informativity at the RNA level in independent cohorts. However, the discovery of potential new markers is far easier than their transfer into clinical practice, the latter needing their prior validation in large independent cohorts awarded only those markers displaying greater or equal strength than already validated biomarkers and clinical indexes. Markers correlated with proliferation (GUAA, G3P, COF1 and FRIL), which is the most common molecular pathway linked to breast cancer prognosis, seem to be good candidates. Recently, a clear relation has been found in mammary cell migration and breast cancer metastasis between AURKA, which is considered as the prototypic marker of proliferation, and COF1. In a parallel work, the same patients were screened by means of SELDI-TOF-MS (28) and FRIL was validated as a prognostic marker. A significant correlation between FRIL values quantified by SELDI-TOF-MS and 2-DE was obtained [r(Spearman) = 0.63, $p<10^{-4}$]. NPI- and AOL-adjusted analyses demonstrated that the three markers and their corresponding genes [COF1 (*CFL1*), NNMT (*NNMT*) and FRIL (*FTL*)] have independent prognostic informativity. For these reasons, we hypothesize that COF1 and FRIL might be the most interesting protein markers within our list.

From a basic research point of view, some of the selected markers demonstrated molecular or pathway interactions. Beyond proliferation, two other markers [FRIL (*FTL*) and FABP4] point out macrophages as possible effectors of breast cancer aggressiveness, as we recently underlined for FRIL (*FTL*) (28), FABP4 has been primarily regarded as an adipocyte- and macrophage-specific protein, however, recent studies do suggest its broader expression in endothelial cells of capillaries and small veins, the main angiogenic compartment of the vasculature. In macrophages, FABP4 participates towards the regulation of inflammatory activity and cholesterol trafficking. Its physiopathological role in cancer could be independent or synergic with macrophages. FABP4 expression is down-regulated in breast cancer cell lines in comparison with normal breast cell lines (29). In bladder transitional cell carcinoma, FABP4 protein and RNA have been found significantly decreased with increasing tumour stage and histological grade, and it has demonstrated a good prognostic value (27,30).

Three of our markers (ARPC3, COF1 and TAGL) belong to the actin microfilament system, which is considered the engine behind cellular migration (31). COF1, which is activated by AURKA, interacts with TAGL and ARPC3 (Arp2/3 complex) (32,33). A clear relation has been demonstrated in breast cancer between COF1 and tumour invasion and metastasis via aberrant active cell migration (33).

One bibliographical study was conducted to explain *in silico* the moderate correlation observed between *NNMT* and *TAGLN* genes. No link between these markers and breast cancer was underlined in the literature but three genomic studies showed significant expression of both genes in different physiological or pathological conditions. Both were overexpressed in metastatic serous papillary ovarian tumours versus primary ovarian serous carcinomas (34), in mononuclear cells of human umbilical cord blood versus its mesenchymal stem cell subpopulation (35), and in adipocytes from obese versus non-obese Pima Indians (36).

Before concluding, two points need to be raised. First, our results demonstrate that *TAGLN* (TAGL) is not a suppressor gene as has been previously pointed out in different cancers. We found the same pattern of expression on both 2-DE platforms, which was confirmed by western blotting, and *in silico* study showed that this gene was overexpressed in pejorative disease evolution. Second, *GAPDH* (G3P) has long been considered as a housekeeping gene and used as a reference for real-time quantitative PCR. Our results show that its protein has a prognostic value in breast cancer, and is overexpressed at the protein and RNA levels in cases of metastatic relapse. Some authors have already underlined this fact and concluded that *GAPDH* gene should not be used as a control RNA in human breast cancer (37-39).

In conclusion, the selection of these proteins by means of 2-DE, which permitted the differentiation of more than 2,000 protein spots per cytosolic fraction, must be considered as a preliminary exploratory study. This first phase has however led to the identification of 13 potential breast cancer biomarkers. Despite the fact that this training cohort is too small to definitively conclude on the prognostic role of these proteins, *in silico* analyses have shown that some of their corresponding genes have a strong prognostic impact in breast cancer, mostly because of their link with proliferation: *GMPS* (GUAA), *GAPDH* (G3P), *FTL* (FRIL) and *GPD1* (GDPA). A validation phase on a larger cohort is now needed before these biomarkers can be considered for use in clinical practice.

Acknowledgements

This study was supported by the French Ministry of Health (PHRC National 2004). English was reviewed by an independent company.

References

1. Fitzgibbons PL, Page DL, Weaver D, *et al*: Prognostic factors in breast cancer. College of American Pathologists Consensus Statement 1999. Arch Pathol Lab Med 124: 966-978, 2000.
2. Adjuvant therapy for breast cancer. In: NIH consensus statement, pp1-35, 2000.
3. Goldhirsch A, Glick JH, Gelber RD, Coates AS and Senn HJ: Meeting highlights: International Consensus Panel on the Treatment of Primary Breast Cancer. Seventh International Conference on Adjuvant Therapy of Primary Breast Cancer. J Clin Oncol 19: 3817-3827, 2001.
4. Goldhirsch A, Glick JH, Gelber RD and Senn HJ: Meeting highlights: International Consensus Panel on the Treatment of Primary Breast Cancer. J Natl Cancer Inst 90: 1601-1608, 1998.
5. Goldhirsch A, Wood WC, Gelber RD, Coates AS, Thurlimann B and Senn HJ: Meeting highlights: updated international expert consensus on the primary therapy of early breast cancer. J Clin Oncol 21: 3357-3365, 2003.
6. Harbeck N, Kates RE and Schmitt M: Clinical relevance of invasion factors urokinase-type plasminogen activator and plasminogen activator inhibitor type 1 for individualized therapy decisions in primary breast cancer is greatest when used in combination. J Clin Oncol 20: 1000-1007, 2002.
7. Look MP, van Putten WL, Duffy MJ, *et al*: Pooled analysis of prognostic impact of urokinase-type plasminogen activator and its inhibitor PAI-1 in 8377 breast cancer patients. J Natl Cancer Inst 94: 116-128, 2002.
8. Bertucci F, Houlgatte R, Benziane A, *et al*: Gene expression profiling of primary breast carcinomas using arrays of candidate genes. Hum Mol Genet 9: 2981-2991, 2000.
9. Bertucci F, Nasser V, Granjeaud S, *et al*: Gene expression profiles of poor-prognosis primary breast cancer correlate with survival. Hum Mol Genet 11: 863-872, 2002.
10. Van de Vijver MJ, He YD, van't Veer LJ, *et al*: A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 347: 1999-2009, 2002.
11. Van't Veer LJ, Dai H, van de Vijver MJ, *et al*: Gene expression profiling predicts clinical outcome of breast cancer. Nature 415: 530-536, 2002.
12. Wang Y, Klijn JG, Zhang Y, *et al*: Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet 365: 671-679, 2005.
13. Ricolleau G, Charbonnel C, Lode L,*et al*: Surface-enhanced laser desorption/ionization time of flight mass spectrometry protein profiling identifies ubiquitin and ferritin light chain as prognostic biomarkers in node-negative breast cancer tumors. Proteomics 6: 1963-1975, 2006.
14. Bradford MM: A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72: 248-254, 1976.
15. Bouchet C, Hacene K, Martin PM, *et al*: Dissemination risk index based on plasminogen activator system components in primary breast cancer. J Clin Oncol 17: 3048-3057, 1999.
16. Descotes F, Riche B, Saez S, *et al*: Plasminogen activator inhibitor type 1 is the most significant of the usual tissue prognostic factors in node-negative breast ductal adenocarcinoma independent of urokinase-type plasminogen activator. Clin Breast Cancer 8: 168-177, 2008.
17. Rabilloud T: Silver staining of 2-D electrophoresis gels. Methods Mol Biol 112: 297-305, 1999.
18. O'Connell KL and Stults JT: Identification of mouse liver proteins on two-dimensional electrophoresis gels by matrix-assisted laser desorption/ionization mass spectrometry of in situ enzymatic digests. Electrophoresis 18: 349-359, 1997.
19. Tusher VG, Tibshirani R and Chu G: Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci USA 98: 5116-5121, 2001.
20. Gharahdaghi F, Weinberg CR, Meagher DA, Imai BS and Mische SM: Mass spectrometric identification of proteins from silver-stained polyacrylamide gel: a method for the removal of silver ions to enhance sensitivity. Electrophoresis 20: 601-605, 1999.
21. Laemmli UK: Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685, 1970.
22. Jezequel P, Campone M, Gouraud W, *et al*: bc-GenExMiner: an easy-to-use online platform for gene prognostic analyses in breast cancer. Breast Cancer Res Treat 131: 765-775, 2012.
23. Galea MH, Blamey RW, Elston CE and Ellis IO: The Nottingham Prognostic Index in primary breast cancer. Breast Cancer Res Treat 22: 207-219, 1992.
24. Ravdin PM, Siminoff LA, Davis GJ, *et al*: Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. J Clin Oncol 19: 980-991, 2001.
25. Loussouarn D, Campion L, Leclair F, *et al*: Validation of UBE2C protein as a prognostic marker in node-positive breast cancer. Br J Cancer 101: 166-173, 2009.
26. Wirapati P, Sotiriou C, Kunkel S, *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.
27. Boiteux G, Lascombe I, Roche E, *et al*: A-FABP, a candidate progression marker of human transitional cell carcinoma of the bladder, is differentially regulated by PPAR in urothelial cancer cells. Int J Cancer 124: 1820-1828, 2009.
28. Jezequel P, Campion L, Spyratos F, *et al*: Validation of tumour-associated macrophage ferritin light chain as a prognostic biomarker in node-negative breast cancer tumours: a multicentric 2004 national PHRC study. Int J Cancer (Epub ahead of print) doi: 10.1002/ijc.26397, 2011.
29. Hammamieh R, Chakraborty N, Barmada M, Das R and Jett M: Expression patterns of fatty acid binding proteins in breast cancer cells. J Exp Ther Oncol 5: 133-143, 2005.
30. Celis JE, Ostergaard M, Basse B, *et al*: Loss of adipocyte-type fatty acid binding protein and other protein biomarkers is associated with progression of human bladder transitional cell carcinomas. Cancer Res 56: 4782-4790, 1996.
31. Lambrechts A, van Troys M and Ampe C: The actin cytoskeleton in normal and pathological cell motility. Int J Biochem Cell Biol 36: 1890-1909, 2004.
32. DesMarais V, Macaluso F, Condeelis J and Bailly M: Synergistic interaction between the Arp2/3 complex and cofilin drives stimulated lamellipod extension. J Cell Sci 117: 3499-3510, 2004.
33. Wang LH, Xiang J, Yan M, *et al*: The mitotic kinase Aurora-A induces mammary cell migration and breast cancer metastasis by activating the Cofilin-F-actin pathway. Cancer Res 70: 9118-9128, 2005.
34. Bignotti E, Tassi RA, Calza S, *et al*: Gene expression profile of ovarian serous papillary carcinomas: identification of metastasis-associated genes. Am J Obstet Gynecol 196: 245 e241-211, 2007.
35. Jeong JA, Hong SH, Gang EJ, *et al*: Differential gene expression profiling of human umbilical cord blood-derived mesenchymal stem cells by DNA microarray. Stem Cells 23: 584-593, 2005.
36. Lee YH, Nair S, Rousseau E, *et al*: Microarray profiling of isolated abdominal subcutaneous adipocytes from obese vs non-obese Pima Indians: increased expression of inflammation-related genes. Diabetologia 48: 1776-1783, 2005.
37. Lu H, Zhang Y, Roberts DD, Osborne CK and Templeton NS: Enhanced gene expression in breast cancer cells in vitro and tumors in vivo. Mol Ther 6: 783-792, 2002.

38. Revillion F, Pawlowski V, Hornez L and Peyrat JP: Glyceraldehyde-3-phosphate dehydrogenase gene expression in human breast cancer. Eur J Cancer 36: 1038-1042, 2000.
39. Valenti MT BF, Dalle Carbonare L, Azzarello G, Zenari S, Zanatta M, Balducci E, Vinante O and Lo Cascio V: The effect of bisphosphonates on gene expression: GAPDH as a housekeeping or a new target gene? BMC Cancer 6: 49, 2006.
40. Sotiriou C, Neo SY, McShane LM, *et al*: Breast cancer classification and prognosis based on gene expression profiles from a population-based study. Proc Natl Acad Sci USA 100: 10393-10398, 2003.
41. Ma XJ, Wang Z, Ryan PD, *et al*: A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. Cancer Cell 5: 607-616, 2004.
42. Minn AJ, Gupta GP, Siegel PM, *et al*: Genes that mediate breast cancer metastasis to lung. Nature 436: 518-524, 2005.
43. Pawitan Y, Bjöhle J, Amler L, *et al*: Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts. Breast Cancer Res 7: R953-R964, 2005.
44. Weigelt B, Hu Z, He X, *et al*: Molecular portraits and 70-gene prognosis signature are preserved throughout the metastatic process of breast cancer. Cancer Res 65: 9155-9158, 2005.
45. Bild AH, Yao G, Chang JT, *et al*: Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature 439: 353-357, 2006.
46. Chin K, DeVries S, Fridlyand J, *et al*: Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. Cancer Cell 10: 529-541, 2006.
47. Ivshina AV, George J, Senko O, *et al*: Genetic reclassification of histologic grade delineates new clinical subtypes of breast cancer. Cancer Res 66: 10292-10301, 2006.
48. Desmedt C, Piette F, Loi S, *et al*: Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series. Clin Cancer Res 13: 3207-3214, 2007.
49. Loi S, Haibe-Kains B, Desmedt C, *et al*: Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade. J Clin Oncol 25: 1239-1246, 2007.
50. Minn AJ, Gupta GP, Padua D, *et al*: Lung metastasis genes couple breast tumor size and metastatic spread. Proc Natl Acad Sci USA 104: 6740-6745, 2007.
51. Naderi A, Teschendorff AE, Barbosa-Morais NL, *et al*: A gene-expression signature to predict survival in breast cancer across independent data sets. Oncogene 26: 1507-1516, 2007.
52. Anders CK, Acharya CR, Hsu DS, *et al*: Age-specific differences in oncogenic pathway deregulation seen in human breast tumors. PLoS One 3: e1373, 2008.
53. Chanrion M, Negre V, Fontaine H, *et al*: A gene expression signature that can predict the recurrence of tamoxifen-treated primary breast cancer. Clin Cancer Res 14: 1744-1752, 2008.
54. Loi S, Haibe-Kains B, Desmedt C, *et al*: Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen. BMC Genomic 9: 239, 2008.
55. Schmidt M, Böhm D, von Törne C, *et al*: The humoral immune system has a key prognostic impact in node-negative breast cancer. Cancer Res 68: 5405-5413, 2008.
56. Calabrò A, Beissbarth T, Kuner R, *et al*: Effects of infiltrating lymphocytes and estrogen receptor on gene expression and prognosis in breast cancer. Breast Cancer Res Treat 116: 69-77, 2009.
57. Jézéquel P, Campone M, Roché H, *et al*: A 38-gene expression signature to predict metastasis risk in node-positive breast cancer after systemic adjuvant chemotherapy: a genomic substudy of PACS01 clinical trial. Breast Cancer Res Treat 116: 509-520, 2009.
58. Zhang Y, Sieuwerts AM, McGreevy M, *et al*: The 76-gene signature defines high-risk patients that benefit from adjuvant tamoxifen therapy. Breast Cancer Res Treat 116: 303-309, 2009.