

# Prediction of Recurrence and Survival for Triple-Negative Breast Cancer (TNBC) by a Protein Signature in Tissue Samples<sup>\*S</sup>

Mario Campone<sup>‡§</sup>, Isabelle Valo<sup>§</sup>, Pascal Jézéquel<sup>‡</sup>, Marie Moreau<sup>¶</sup>, Alice Boissard<sup>§</sup>, Loïc Champion<sup>‡</sup>, Delphine Loussouarn<sup>||</sup>, Véronique Verriele<sup>§</sup>, Olivier Coqueret<sup>§¶</sup>, and  Catherine Guette<sup>§\*\*</sup>

To date, there is no available targeted therapy for patients who are diagnosed with triple-negative breast cancers (TNBC). The aim of this study was to identify a new specific target for specific treatments. Frozen primary tumors were collected from 83 adjuvant therapy-naïve TNBC patients. These samples were used for global proteome profiling by iTRAQ-OFFGEL-LC-MS/MS approach in two series: a training cohort ( $n = 42$ ) and a test set ( $n = 41$ ). Patients who remains free of local or distant metastasis for a minimum of 5 years after surgery were classified in the no-relapse group; the others were in the relapse group. OPLS and Kaplan–Meier analyses were performed to select candidate markers, which were validated by immunohistochemistry. Three proteins were identified in the training set and validated in the test set by Kaplan–Meier method and immunohistochemistry (IHC): TrpRS as a good prognostic markers and DP and TSP1 as bad prognostic markers. We propose the establishment of an IHC test to calculate the score of TrpRS, DP, and TSP1 in TNBC tumors to evaluate the degree of aggressiveness of the tumors. Finally, we propose that DP and TSP1 could provide therapeutic targets for specific treatments. *Molecular & Cellular Proteomics* 14: 10.1074/mcp.M115.048967, 2936–2946, 2015.

Triple-negative breast cancers (TNBC)<sup>1</sup> are defined by a lack of expression of estrogen (ER), progesterone (PR), and

HER2/neu receptors and account for about 15% of all breast cancers. This subtype is associated with poor prognosis (1) in terms of distant free survival (DFS) and overall survival (OS), and to date, there is no clinically available targeted therapy for patients diagnosed with TNBC. Because of the absence of specific treatment guidelines for this group of patients, TNBC are managed with standard adjuvant chemotherapy (2), which, however, seems to be less effective in those cancers. In order to improve survival, it is important to determine new specific-targeted treatment.

A proteomic analysis has several inherent advantages over a genomic approach in that measured mRNA levels do not necessarily correlate to corresponding protein levels. In addition, protein detection is probably also more reflective of the tumor microenvironment. Several proteomic studies have been conducted on TNBC (3–5), but no proteomic study was conducted on large cohorts including the clinical outcome of the patients, except a recent comparative proteome analysis that identified a 11-protein signature for aggressive TNBC in a large cohort of 93 microdissected tumors (6). Although microdissection was necessary to elucidate the contribution of TNBC cells, it did not reflect the tumor with its microenvironment that is increasingly described as fundamental to explain the tumor outcome. Thus, it is now recognized that carcinomas derive from phenomena that occur in tissues, not in individual cancer cells. From this perspective, the microenvironment becomes an integral, essential part of the tumor (7,

tein kinase 1; FASP, filter aided sample preparation; fMLP, N-formyl-methionyl-leucyl-phenylalanine; IDH, isocitrate dehydrogenase; IFN $\gamma$ , Interferon gamma; IPG, Immobilized pH gradient; iTRAQ, isobaric tags for relative and absolute quantification; M, metastasis; M0, no distant spread is found; N0, cancer has not spread to nearby lymph node; N1, cancer has spread to 1 to 3 axillary lymph node(s); N3, cancer has spread to 10 or more axillary lymph nodes; NADP, nicotinamide adenine dinucleotide phosphate; OCT, optimal cutting temperature compound; OPLS, supervised orthogonal partial least square analysis; pI, isoelectric point; S/N, ratio signal/noise; T1, tumor size is 2 cm or less across; T2, tumor size is more than 2 cm but not more than 5 cm across; TrpRS<sup>tryptophan</sup>, tryptophan synthetase; TSP1, thrombospondin-1; UICC, International Union Against Cancer; VE-cadherin, vascular endothelial-cadherin; WARS, tryptophanyl-TRNA synthetase; WHO, World Health Organization.

<sup>‡</sup>René Gauducheau ICO Cancer Center, Inserm U892, CNRS 6299, Bd J. Monod, 44805 Saint Herblain Cedex, France; <sup>§</sup>Paul Papin ICO Cancer Center, Inserm U892, CNRS 6299, 2 rue Moll, 49933 Angers Cedex 9, France; <sup>¶</sup>Angers University, 4 Boulevard de Lavoisier, Angers, 49000, France; <sup>||</sup>INSERM U892, CNRS 6299, IRT-UN, 8 quai Moncousu, 44007 Nantes Cedex, France

Received February 10, 2015, and in revised form, July 21, 2015

Published, MCP Papers in Press July 24, 2015, DOI 10.1074/mcp.M115.048967

Author contributions: M.C. C.G. and designed research; I.V., P.J., M.M., A.B., L.C., D.L., V.V., and C.G. performed research; O.C. and C.G. analyzed data; C.G. wrote the paper.

<sup>1</sup> The abbreviations used are: ACN, acetonitrile; Akt, protéine kinase B; AUC, area under the curve; CHCA, a-Cyano-4-hydroxycinnamic acid; DP, desmoplakin (protein); DSP, desmoplakin (gene); eNOS, endothelial nitric oxide synthase; ERK, mitogen-activated pro-

8). In this context, taking into account the tumor microenvironment, we investigated a cohort of 83 TNBC samples without microdissection by a quantitative proteomic approach using iTRAQ labeling. Based on clinical data, we established a protein signature of the most aggressive tumors. From these differentially expressed proteins, some of them appeared to be potential therapeutic targets.

#### PATIENTS AND METHODS

**Patients**—The study involved 83 patients diagnosed and treated at the Institut of Cancérologie de l'Ouest (ICO) for a TNBC, between early 1998 and 2007. The primary inclusion criterion was an adequate fresh tumor obtained from a resected tumor sample (see below). Patients were included if they fulfilled the following criteria: (a) female primary unilateral invasive ER/PR and HER2 negative—breast carcinoma without previous or concomitant malignancies; (b) T1T2, N0N1 N2 N3-, M0 staging according to UICC criteria; (c) older than 18 years old; and (d) surgical first-line treatment. All patients 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. All the patients received the same adjuvant chemotherapy (FEC100) and radiotherapy treatments.

Patients were followed up for disease evolution. The 83 tumors were divided in two cohorts: a training cohort ( $n = 42$ ) corresponding to patients diagnosed at ICO Paul Papin (Angers) and a test cohort ( $n = 41$ ) corresponding to patients diagnosed at ICO René Gauducheau (Nantes). The clinicopathological characteristics of these TNBC cohorts are listed in [Supplemental data S1](#). Follow-up data were collected for all patients, including, disease-free survival (DFS; time from diagnosis to first recurrence of the disease or contralateral breast cancer or second primary other cancer) and overall survival (OS, time from diagnosis to death from any cause). Recurrences were defined as locoregional (breast, mammary region, or regional lymph nodes) or metastatic (visceral or not). 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. This study did not need ethical approval.

**Tumor Characteristics**—Pathological data were reviewed by two pathologists. Tumor size (pT) was measured on fresh resection specimens, as the maximum diameter (mm) of the tumor. Histological type was determined according to the WHO classification and histological grade according to the Elston and Ellis methods.

ER and PR status were accessed by immunohistochemistry on representative formalin-fixed tumors blocks at a 4- $\mu$ m thickness. Tumors were determined as negative when < 10% cells stained positive. All patients where HER-2 negative, that means an immunostaining 1+, score according to the Her-

cepTest scoring system or 2+ without HER-2 gene amplification investigated by *in situ* fluorescence.

**Sample Collection**—All specimens were collected immediately after surgery, snap frozen, and stored in liquid nitrogen until the time of analysis. The time between the resection of the breast tumor and its freezing is less than 1 h. We also selected four normal macroscopic areas for our control pool. Frozen sections (12  $\mu$ m thick) of either TNBC or normal areas were embedded in OCT and cut on a cryostat (Bright Instrument Co. Ltd., St. Margarets Way, UK). Specific sections were stained with toluidine blue for visual reference, and each tissue section from all specimens was evaluated by experienced pathologists for cancer cell proportion determination. Samples containing less than 75% of tumor cells were removed.

**Protein Extraction from Frozen Tissues**—Ten frozen sections per tumor were lysed in a buffer consisting of 0.1 M Tris-HCl, pH 8.0; 0.1 M DTT; and 4% SDS at 95 °C for 90 min. Detergent was removed from the lysates, and the proteins were digested with trypsin using the FASP protocol (9) using spin ultrafiltration units of nominal molecular weight cut of 30,000. Using YM-30 microcon filter units (Cat. No. MRCF0R030, Millipore) containing protein concentrates, 200  $\mu$ l of 8 M urea in 0.1 M Tris/HCl, pH 8.5 (UA), was added, and samples were centrifuged at 14,000  $g$  at 20 °C for 8 min. This step was repeated three times. Then 6  $\mu$ l of 200 mM MMTS in 8 M urea was added to the filters, and the samples were incubated for 20 min. Filters were washed three times with 200  $\mu$ l of 8 M UA followed by six washes with 100  $\mu$ l 0.5 M TEAB. Finally, trypsin (AB sciex, Carlsbad, CA) was added in 100  $\mu$ l 0.5 M TEAB to each filter. The protein to enzyme ratio was 100:1. Samples were incubated overnight at 37 °C, and released peptides were collected by centrifugation. Samples were then dried completely using a SpeedVac and resuspended in 100  $\mu$ l of 0.5% trifluoroacetic acid (TFA) in 5% acetonitrile and were desalted via PepClean C-18 spin columns (Pierce Biotechnology, Rockford, IL). Peptide content was determined using Micro BCA Protein Assay Kit (Pierce-Thermo Scientific).

**Peptide Labeling with iTRAQ Reagents**—Each peptide solution was labeled at room temperature for 2 h with one iTRAQ reagent vial previously reconstituted with 70  $\mu$ l of ethanol for 4plex iTRAQ reagent and reconstituted with 50  $\mu$ l of isopropanol for 8plex iTRAQ reagent. A mixture containing small aliquots from each labeled sample was analyzed by MS/MS to determine a proper mixing ratio to correct for unevenness in peptide yield. Labeled peptides were then mixed in 1:1:1:1 (or 1:1:1:1:1:1:1:1) ratio. Peptide mixture was then dried completely using a SpeedVac.

**Peptide OFFGEL Fractionation**—For pl-based peptide separation, we used the 3100 OFFGEL Fractionator (Agilent Technologies, Böblingen, Germany) with a 24-well setup using our protocol (10). Briefly, prior to electrofocusing, samples were desalted onto a Sep-Pak C18 cartridge (Waters). For the

24-well setup, peptide samples were diluted to a final volume of, respectively, 3.6 ml using OFFGEL peptide sample solution. To start, the IPG gel strip of 24 cm-long (GE Healthcare, München, Germany) with a 3–10 linear pH range was rehydrated with the Peptide IPG Strip Rehydration Solution according to the manufacturer protocol for 15 min. Then, 150  $\mu$ l of sample was loaded in each well. Electrofocusing of the peptides was performed at 20 °C and 50  $\mu$ A until the 50 kVh level was reached. After focusing, the 24 peptide fractions were withdrawn, and the wells were washed with 200  $\mu$ l of a solution of water/methanol/formic acid (49/50/1). After 15 min, the washing solutions were pooled with their corresponding peptide fraction. All fractions were evaporated by centrifugation under vacuum and maintained at –20 °C. Just prior nano-LC, the fractions were resuspended in 20  $\mu$ l of H<sub>2</sub>O with 0.1% (v/v) TFA.

**Capillary LC Separation**—The samples were separated on an Ultimate 3,000 nano-LC system (Dionex, Sunnyvale, USA) using a C18 column (PepMap100, 3  $\mu$ m, 100 A, 75  $\mu$ m id  $\times$  15 cm, Dionex) at 300 nL/min flow rate. Buffer A was 2% ACN in water with 0.05% TFA, and buffer B was 80% ACN in water with 0.04% TFA. Peptides were desalted for 3 min using only buffer A on the precolumn, followed by a separation for 105 min using the following gradient: 0 to 20% B in 10 min, 20% to 45% B in 85 min, and 45% to 100% B in 10 min. Chromatograms were recorded at the wavelength of 214 nm. Peptide fractions were collected using a Probot microfraction collector (Dionex). We used CHCA (LaserBioLabs, Sophia-Antipolis, France) as MALDI matrix. The matrix (concentration of 2 mg/ml in 70% ACN in water with 0.1% TFA) was continuously added to the column effluent via a micro “T” mixing piece at 1.2  $\mu$ l/min flow rate. After 12 min run, a start signal was sent to the Probot to initiate fractionation. Fractions were collected for 10 s and spotted on a MALDI sample plate (1,664 spots per plate, Applied Biosystems, Foster City, CA).

**MALDI-MS/MS**—MS and MS/MS analyses of offline spotted peptide samples were performed using the 5800 MALDI-TOF/TOF Analyzer (AB sciex) and 4000 Series Explorer software, version 4.0. The instrument was operated in a positive ion mode and externally calibrated using a mass calibration standard kit (AB sciex). The laser power was set between 2,800 and 3,400 for MS and between 3,600 and 4,200 for MS/MS acquisition. After screening all LC-MALDI sample positions in MS-positive reflector mode using 1,500 laser shots, the fragmentation of automatically selected precursors was performed at a collision energy of 1 kV using air as collision gas (pressure of  $\sim 2 \times 10^{-6}$  Torr) with an accumulation of 2,000 shots for each spectrum. MS spectra were acquired between  $m/z$  1,000 and 4,000. For internal calibration, we used the parent ion of Glu1-fibrinopeptide at  $m/z$  1,570.677 diluted in the matrix (30 femtomoles per spot). Up to 12 of the most intense ion signals per spot position having an S/N > 12 were selected as precursors for MS/MS acquisition. Peptide and protein identification were performed by the Protein-

Pilot™ Software V 4.0 (AB Sciex) using the Paragon algorithm as the search engine (11). Each MS/MS spectrum was searched for *Homo sapiens* species against the Uniprot/swissprot database (UniProtKB/Sprot 20,120,208 release 01, with 525,997 sequence entries). The searches were run using the fixed modification of methylmethanethiosulfate labeled cysteine parameter enabled. Other parameters such as tryptic cleavage specificity, precursor ion mass accuracy, and fragment ion mass accuracy are MALDI 5800 built-in functions of ProteinPilot software. The detected protein threshold (unused protscore (confidence) in the software was set to 1.3 to achieve 95% confidence, and identified proteins were grouped by the ProGroup algorithm (AB sciex) to minimize redundancy. The bias correction option was executed.

A decoy database search strategy was also used to estimate the false discovery rate (FDR), defined as the percentage of decoy proteins identified against the total protein identification. The FDR was calculated by searching the spectral against the Uniprot *H. sapiens* decoy database. The estimated low FDR of 0.9% indicated a high reliability in the identified proteins.

**Quantification of Relative Protein Expression**—We employed a customized software package, iQuantitator (12–14) to infer the magnitude of change in protein expression. The software infers sample-dependent changes in protein expression using Markov chain, Monte Carlo, and Bayesian statistical methods. Basically, this approach was used to generate means and 95% credible intervals (upper and lower) for each protein expression of each tumor of the training set and the test set by using peptide-level data for each component peptide. For proteins whose iTRAQ ratios were down-regulated in tissues, the extent of down-regulation was considered further significant if the higher limit of the credible interval had a value lower than 1. Conversely, for proteins whose iTRAQ ratios were up-regulated in tumors, the extent of up-regulation was considered further significant if the lower limit of the credible interval had a value greater than 1. The width of these credible intervals depends on the data available for a given protein. Since the number of peptides observed and the number of spectra used to quantify the change in expression for a given protein are taken into consideration, it is possible to detect small but significant changes in up- or down-regulation when many peptides are available. The peptide selection criteria for relative quantification were performed as follows. Only peptides unique for a given protein were considered for relative quantification, excluding those common to other proteins. In cases where a peptide could be assigned to more than one protein, it is eliminated from consideration prior to analysis. Proteins were identified on the basis of having at least two peptides with an ion score above 95% confidence. The protein sequence coverage (95%) was estimated for specific proteins by the percentage of matching amino acids from the identified peptides having confidence greater than or equal to



95% divided by the total number of amino acids in the sequence.

**Functional Annotation and Network Analysis**—Gene ontology (GO) terms for identified proteins were extracted, and overrepresented functional categories for differentially abundant proteins were determined by the high throughput GOMiner tool (National Cancer Institute, <http://discover.nci.nih.gov/gate2.inist.fr/gominer/>) (15). All proteins that were subjected to iQuantitor analysis served as the background list, and GO terms with at least five proteins were used for statistical calculations. A *p* value for each term was calculated via the one-sided Fisher's exact test, and FDR was estimated by permutation analysis using 1,000 randomly selected sets of proteins sampled from the background list. Statistically significant (FDR<25%) GO terms were clustered based on the correlation of associated proteins to minimize potential redundancy in significant GO terms.

**Statistical Analysis**—To visualize clustering of groups, a two-way (by protein and tumor ID) hierarchical clustering was performed on log<sub>2</sub>-transformed data. Further multivariate statistics and modeling was performed with SIMCA (SIMCA 13.0, Umetrics, Sweden) (16). The analysis was performed on mean-centered, unit-variance-scaled data, assuming equal importance of each protein regardless of relative abundance and magnitude of variance between samples. Principal component analysis (PCA) (17, 18), was performed to get an overview of the data, detect clustering of the data, and pick up outliers if any. The PCA summarizes the variation of the data matrix (*i.e.* protein ratios) and shows the relationship between the observations. For classification and identification of proteins differentiating relapse from relapse-free tumors/patients, we used orthogonal partial least square analysis (OPLS) (19). The OPLS analysis detects the protein expression data that covaries with the defined clinical groups. For optimization of the OPLS models, we used the variable importance in the projection value to judge protein influence (including prediction performance) on the model. The OPLS models were validated by sevenfold full cross-validation. Proteins with high variable importance in the projection throughout the cross-validation of the model (95% confidence interval) were selected for the optimized model. We used the plots of the scores predicted in the cross-validation and analysis of variance (CV-ANOVA) to evaluate the model validity. Fisher exact test were calculated for the training set *versus* the test set and for the relapse *versus* no-relapse cohorts.

Survival rates were calculated using the nonparametric Kaplan–Meier method, and log-rank tests were performed to evaluate the difference in the time between recurrence and nonrecurrence groups. Multivariate Cox models were used to assess the prognostic value of each variable.

**Immunohistochemistry and Scoring**—The 42 tumors from the training set (ICO René Gauducheau) were studied by immunohistochemistry. The immunohistochemistry was carried out on 4- $\mu$ m thick paraffin embedded sections of forma-

lin-fixed tumor samples. Details of the antigen retrieval technique and dilution of primary antibodies (TrpRS, DP, and TPS1) are described in [Supplemental data S2](#). The immunolabeling technique was performed by a benchmark automatized tissue staining system (Ventana Medical System, Tucson, AZ). The immunohistochemistry was evaluated semiquantitatively by the percentage of cytoplasmic staining cells, the intensity, and the presence or not of secretory granules. To exclude subjectivity, all slides were evaluated by two pathologists who had no knowledge of the patients' identities or clinical status. In discrepant cases, the two pathologists reviewed the slides together and reached a consensus. The percentage of immunopositive stained cells (A) was divided into four grades as: <10% (0); 10–30% (1); 30–50% (2); 50–70% (3); and >70% (4). Second, the intensity of staining was scored by evaluating the average staining intensity (B) of the positive cells (0, none; 1, weak; 2, intermediate; and 3, strong). The score for each section was measured as A  $\times$  B, and the result was defined as negative (–, 0), weakly positive (+, 1–3), positive (++, 4–7), and strongly positive (+++, 8–12). The immunohistochemical data were subjected to statistical analysis. All quantitative data were recorded as mean  $\pm$  S.D. Comparison between multiple groups was performed by one-way ANOVA and Wilcoxon rank test (*p* value<.05).

**Receiver Operating Characteristic Curves**—Individual and combined biomarker performances were investigated on the receiver operating characteristic curves with linear discriminant analysis. Linear discriminant analysis was used to find a linear combination of features that characterizes or separates two or more classes of objects or events. The resulting combination was then used as a linear classifier. To determine how accurately the learning algorithm was able to predict data, cross-validation and bootstrapping methods were used. In leave-one-out cross-validation, one sample was removed from the dataset, and a classifier was generated using the remaining samples to predict the status of the removed sample. In 10-fold cross-validation, the data were divided into 10 subsets of approximately equal size, and 10 iterations of training and validation were performed. The 0.632+ bootstrap cross-validation uses resampling technique. The 10-fold cross-validation and bootstrapping procedures were replicated 100 times. Statistical analyses were performed using TANAGRA (v1.4.49).

## RESULTS

The training tumors were profiled by iTRAQ-LC-MS/MS approach ([Supplemental data S3](#)). The baseline clinical features of patients were similar between the ICO Paul Papin training set and the ICO René Gauducheau test cohorts, although patient tumor size were slightly bigger in the training set (Table I). The median follow-up for the good prognosis patients in the training and test sets was 168 (range = 68–279) and 203 (range = 51–413) months, respectively. In the training cohort, 14 patients experienced a relapse (13 distant

TABLE I  
Clinicopathological characteristics of patients for tissue proteomic studies

Patient characteristics	Training set (n = 42)		Test set (n = 41)	
Age (years) median [min-max]	55[28–71]		57[29–84]	
< 50 (%)	15 (34.8)		10 (24.4)	
> = 50 (%)	28 (65.2)		31 (75.6)	
	No-recurrence	Recurrence	No-recurrence	Recurrence
	28 (67.4)	14 (32.6)	24 (58.5)	17 (41.5)
Grade 1	0		1	
Grade 2	2		5	
Grade 3	26		18	
Lymph node status				
Positive (%)	4		9	
Negative (%)	24		15	
pT (mm)				
8–10	1		1	
11–20	22		13	
21–50	4		9	
>50	0		1	
Type of surgery				
mastectomy	4		8	
tumorectomy	24		16	

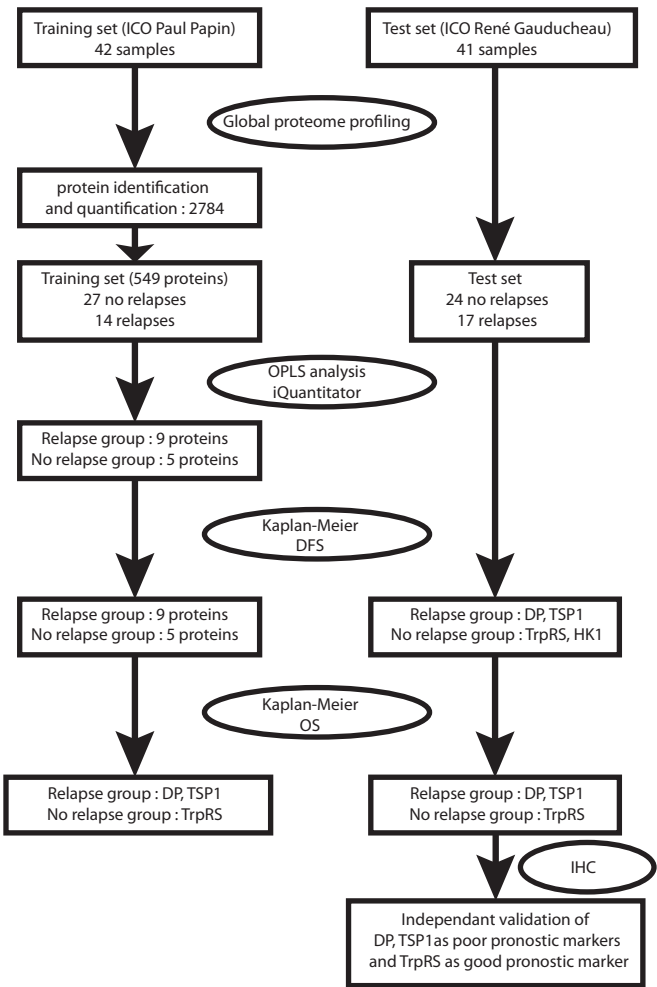


FIG. 1. Flow chart of experimental design for the development and validation of the biomarkers.

metastasis and 1 contralateral), and among these patients, we recorded 11 deaths. In the test set, 17 patients experienced a relapse (15 distant metastasis and 2 contralateral), and among these patients, we recorded 13 deaths. Clinical data used for data analysis were updated until January 2013.

**Identification of Expressed Proteins in the Training Samples: Proteomic Coverage of 42 Triple-Negative Breast Tumors**—As we considered the microenvironment was an integral, essential part of the tumor, the samples were not microdissected, but each tumor section was validated as containing more than 75% tumor cells by pathologists. Using Protein Pilot and iQuantitor software, we identified and quantified a total of 2,784 nonredundant proteins with at least two peptides, according to the schematic workflow of the experimental design presented Fig. 1. By taking into consideration both the peptide and spectra numbers, this approach allowed us to detect small but significant expression changes, provided that several peptides are detected. Using this analysis, we were able to obtain a list of quantified proteins from the 20 iTRAQ experiments. Following Metacore analysis using the “Enrichment of protein function” function, (Supplemental data S4), we identified 690 enzymes, 58 phosphatases, 122 proteases, 105 kinases, 73 ligands, 82 transcription factors, and 83 receptors. This analysis showed that the best enrichment score and *p* value were assigned to the GO Process “Metabolic Process” and to the “Cytoskeleton Remodeling” pathway (Supplemental data S4). Among these 2,784 proteins, 220 proteins met our definition for differential expression in a comparison between tumor and normal tissues: 126 were overexpressed, and 93 were underexpressed (Supplemental data S3).

**A Proteomic Coverage of Different Status**—We used the iQuantitor software to quantify protein expression between

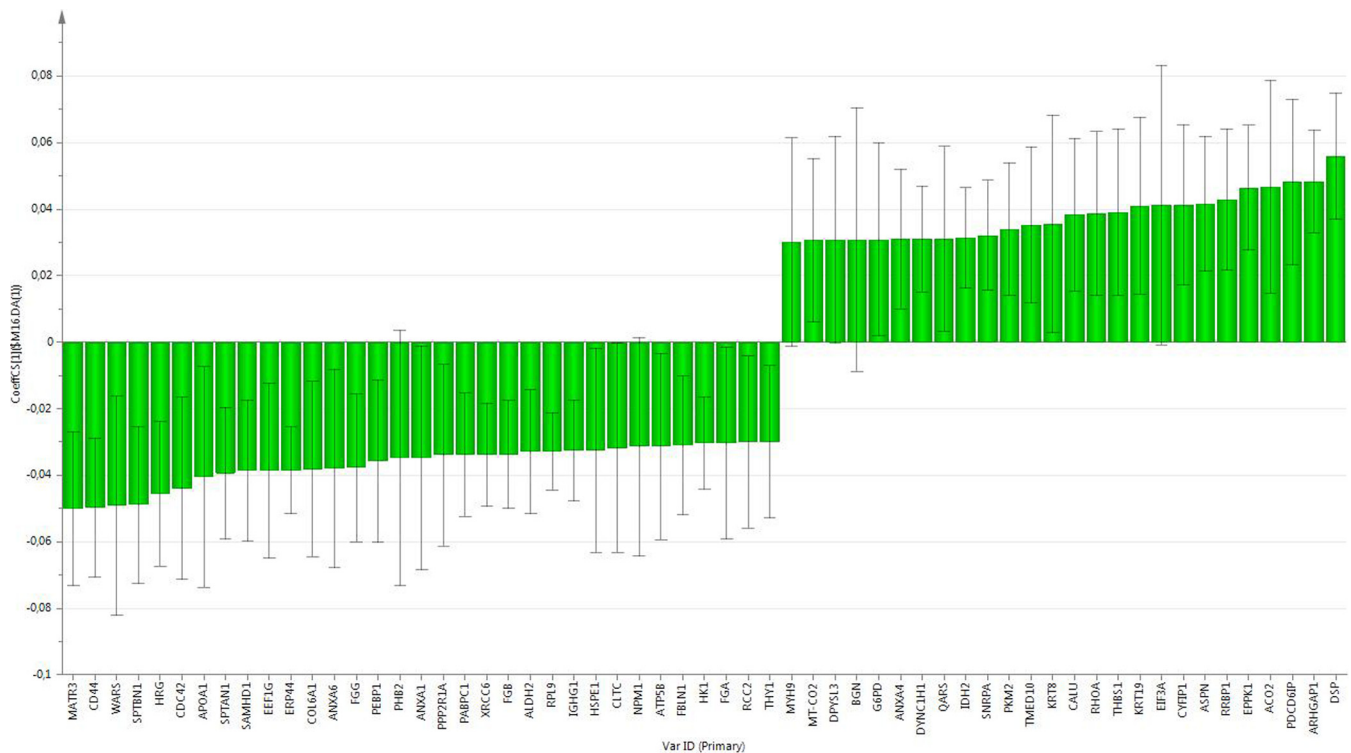


FIG. 2. Supervised orthogonal partial least squares discrimination analysis (OPLS-DA) to distinguish between no-relapse and relapse groups. OPLS-DA analysis was performed on 549 proteins for which quantitative value was available in all the tumors. The optimized OPLS model included 59 proteins. Among these proteins, 33 were assigned to the group without recurrence and 26 to the group with recurrence. Error bars indicate the confidence intervals of the coefficients.

TABLE II

Gene	Protein	Expression	iTRAQ ratio
<b>No relapse</b>			
WARS	Tryptophanyl-tRNA synthetase	(TrpRS)	overexpressed
<b>Relapse</b>			
DSP	Desmoplakin	(DP)	overexpressed
THBS1	Thrombospondin-1	(TSP1)	overexpressed
IDH2	Isocitrate dehydrogenase [NADP]	(IDH)	overexpressed

the different status “relapse” (Supplemental data S5) and “no relapse” (Supplemental data S6). For the relapse group, 295 proteins were significantly differentially expressed: 165 were overexpressed, and 130 were underexpressed. The Metacore analysis of this list of proteins indicated a cytoskeleton re-modeling with a  $p$  value =  $9.2 \times 10^{-12}$  for the Process Network “Regulation of Cytoskeleton Rearrangement” and a best enrichment score and  $p$  value for “Binding” ( $p = 9.4 \times 10^{-26}$ ) in the GO Molecular Functions term. It should be noted that 26 secreted proteins were found in this list characterizing the relapse group (Supplemental data S5). For the no-relapse group, 189 proteins were significantly differentially expressed: 98 were overexpressed, and 91 were underexpressed. For this group, the best score for the Process Network was obtained for “Cell adhesion\_Integrin-mediated cell-matrix adhesion” ( $p = 7.5 \times 10^{-11}$ ). The protein class “ligands” was found

to have to best z-score in the module “Enrichment for Protein Function” with 15 proteins (Supplemental data S6).

**Classification Based on Relapse Status**—We investigated if we could detect differences between the relapse and no-relapse groups in terms of protein levels in the triple-negative tumors by OPLS analysis. This analysis was performed on 549 proteins for which quantitative value was available in all the tumors. The OPLS model, initially based on all 549 proteins, was optimized by stepwise removal of proteins with small variable importance in the projection value. This was performed until the model did not improve anymore as judged by the CV-ANOVA  $p$  value, indicative of the probability that the model is the result of chance alone. The optimized OPLS model included 59 proteins ( $p = 2.1 \times 10^{-15}$ ) (Fig. 2). Among these proteins, 33 were assigned to the group without recurrence and 26 to the group with recurrence. These proteins

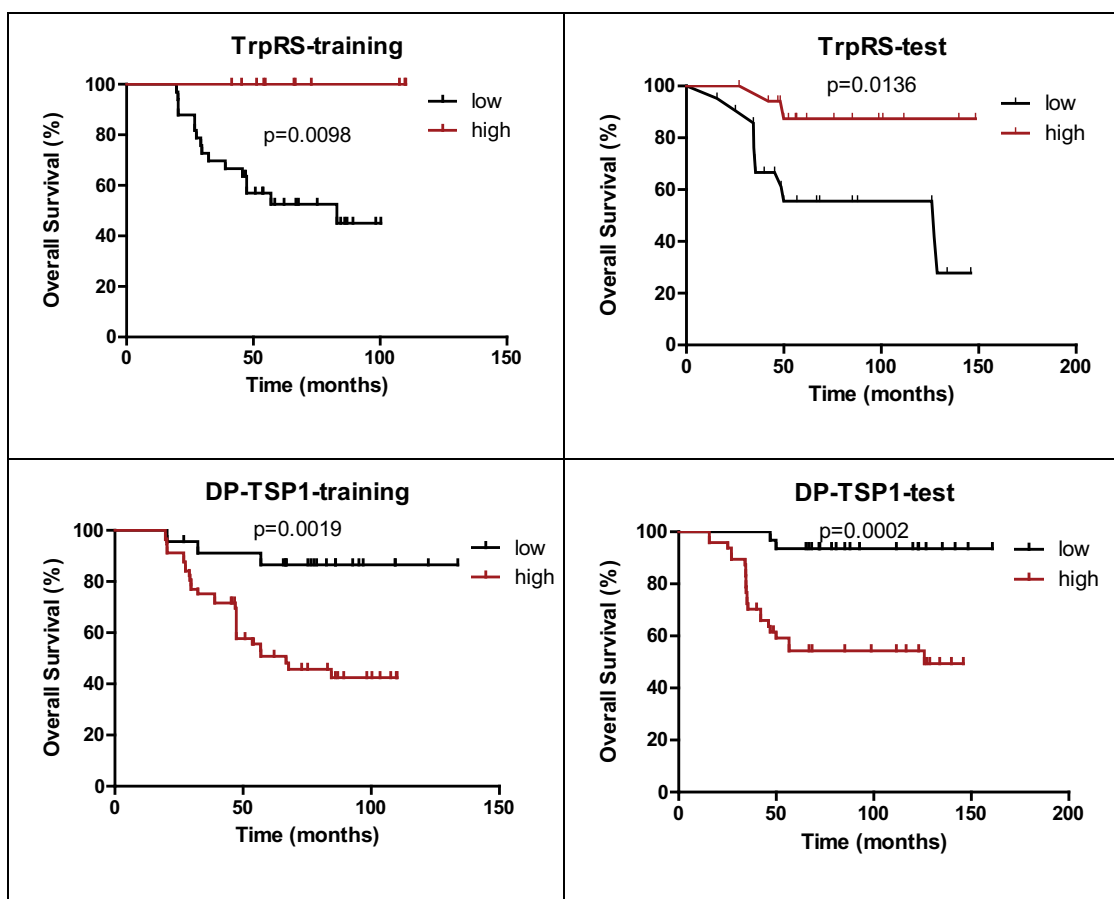


FIG. 3. **Kaplan-Meier survival estimates.** Kaplan-Meier analysis shows that high TrpRS and low DP-TSP1 are strongly predictive of overall survival.

were matched against a database consisting of known protein signaling pathways using Metacore. For the no-relapse group, two significant pathways ( $p < .05$ ) were found: Blood coagulation ( $p = 4.4 \times 10^{-6}$ ) and Chemotaxis\_Lipoxin inhibitory action on fMLP-induced neutrophil chemotaxis ( $p = .0003$ ). The relapse group was characterized by just one significant pathway: Cytoskeleton remodeling\_Keratin filaments ( $p = 7.9 \times 10^{-7}$ ) (Supplemental data S7).

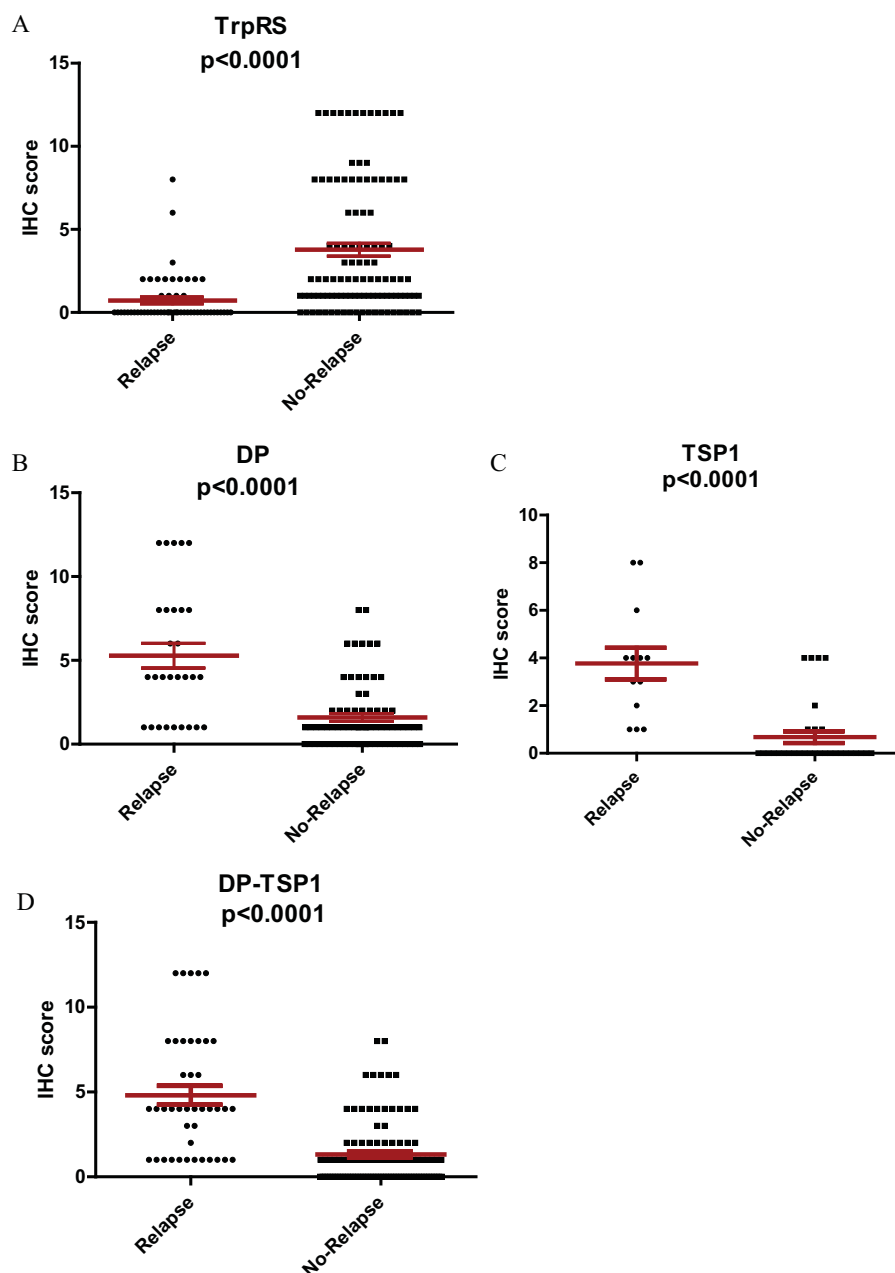
**Proteomic Signature of Triple-Negative Breast Tumors Relapse**—By combining protein lists obtained from the univariate (with iQuantitor) and the multivariate analyses (OPLS), we generated two lists of proteins that characterize the relapse (DP, TPS1, G6PD, IDH, KRT19, KRT8, EPPK1, ARH-GAP1, DPYSL3,) and no-relapse (TrpRS, HSPE1, SAMHD1, HK1, IGHG1) groups of triple-negative breast tumors. These proteins were submitted to a Fisher exact test and only TrpRS, TSP1, DP, and IDH were validated (Table II and Supplemental data S8).

**Prognostic Value of the Markers in the Training Set and in the Test Set**—The prognostic value of the markers was evaluated through estimation of disease-free survival (DFS) and overall survival (OS) using the Kaplan-Meier method. By the same iTraQ quantitative proteomic approach used in the training set,

41 tumors were analyzed. Then, in the both training and test cohorts, the patients were divided into two categories based on the median iTraQ expression data for each marker: high expression (protein levels higher than median) versus low expression (protein levels lower than median).

In the training set, for the no-relapse group, a high expression level of TrpRS was correlated with a significantly better DFS compared with a low expression ( $p = .0129$ ) (Supplemental data S9). In the test set, TrpRS ( $p = .049$ ) was validated. When we considered the expression of TrpRS with the OS rates, we showed that TrpRS was validated in the training ( $p = .098$ ) and in the test set ( $p = .0136$ ) (Supplemental data S10).

In the training set, for the relapse group, patients' tumors with high expression level of any of the three proteins experienced a significantly worse DFS compared with those with low expression ( $p < .0001$ ,  $p = .0330$ ,  $p = .0016$ ) for DP, TPS1, and IDH, respectively (Supplemental data S11). In the test set, DP ( $p = .002$ ), TPS1 ( $p < .0001$ ), and IDH ( $p = .0040$ ) were validated. Furthermore, tumors with high DP ( $p = .0209$ ), TSP1 ( $p = .0364$ ), and IDH ( $p = .0007$ ) expression were also associated with lower OS rates in the training set (Supplemental data S12) while only DP ( $p = .0256$ ) and TPS1 ( $p =$



**FIG. 4. Validation of TrpRS as good prognostic biomarker and TSP1 and DP as bad prognostic biomarkers.** Different distribution of IHC scores (staining intensity  $\times$  staining percentage) between relapse and no-relapse groups for TrpRS (A), DP (B), TSP1 (C), and combined DP-TSP1 (D) (mean  $\pm$  S.D. are shown).

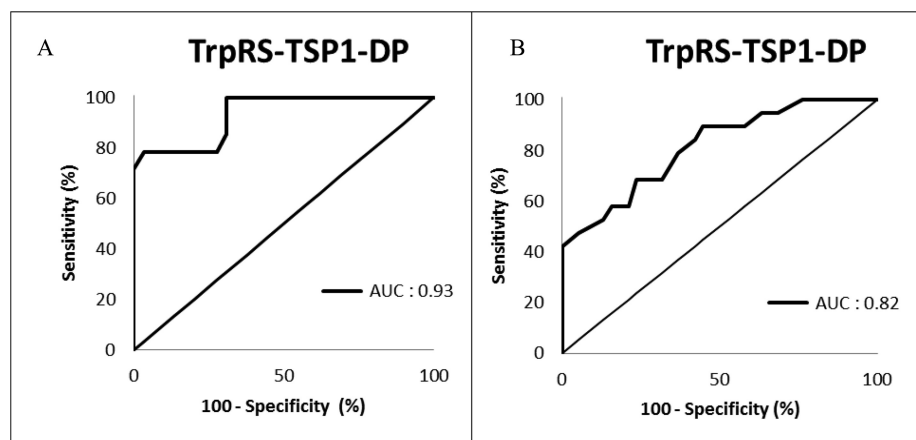
.0018) were validated in the test cohort. By combining the expression levels of DP and TSP1, we showed that patients who have a high expression of these both biomarkers in their tumors have a significantly worse DFS and OS compared with those with low expression of DP and TSP1 (Fig. 3). Finally, DP and TSP1 expression was a strong independent prognostic factor to predict risk of recurrence and TrpRS expression an independent prognostic factor to predict no recurrence. No other prognostic factors, such as age, tumor size, or nodal status of the patients, were statistically significantly associated with recurrence (Supplemental data S13).

**Validation of Dysregulated Protein Expression**—To confirm the dysregulation of the best biomarker candidates for the

relapse and no-relapse groups, the expressions of TrpRS, and TSP1 and DP were analyzed by immunohistochemistry using paraffin-embedded tissues isolated from the training and the test cohorts. Representative pictures of TrpRS, TSP1, and DP staining in no-relapse and relapse cases are shown Fig. 4 for the test set. For the no-relapse group, intense cytoplasmic staining was shown for TrpRS whereas we observed a significant ( $p < .0001$ ) decrease staining intensity in the no-relapse tumor group for this marker. Inversely, for TSP1 and DP, the no-relapse cases exhibited a moderate cytoplasmic staining whereas the cytoplasmic staining increased significantly in the relapse cases (Fig. 4 and Supplemental data S14). In order to evaluate the performance of our protein signature by IHC



FIG. 5. Receiver operating characteristics curves of the 3 protein signature in the training set (A) and in the test set (B).



approach, we investigated the receiver operating characteristic curves that discriminate the no-relapse and relapse groups. Individual biomarker and combination of two or three biomarkers were investigated. The corresponding AUC was estimated in the training and in the test cohorts ([Supplemental data S15](#)). In the training set, each individual biomarker was able to discriminate the both groups, but we obtained the best AUC for the association of TSP1-TrpRS (AUC: 0.92) and for the combination of the three biomarkers TSP1-TrpRS-DP (AUC: 0.93). In the test set, each biomarker was also able to discriminate between both groups, but the best protein combination to discriminate between no-relapse and relapse groups was obtained with the association of the three biomarkers (AUC : 0.82) (Fig. 5). Finally, to confirm the robustness of the learning algorithm, accuracy was estimated using common methods based on resampling: cross-validation (leave-one-out and repeat 10-fold cross-validation) and bootstrap. Interestingly, results show stable estimate accuracies for the linear discriminant analysis classifier ([Supplemental data S15](#)).

#### DISCUSSION

Despite the many recent advances in breast tumor treatments through targeted therapies, no specific treatment exists for the triple-negative breast tumors, and there are no prognostic molecular markers that would predict whether a tumor will behave aggressively or remain indolent. It is abundantly clear that tumor biology plays a significant role in resultant tumor behavior. Unfortunately, triple-negative breast primary tumors that are placed in the same prognostic category based on currently used parameters may behave differently. It is our hypothesis that the underlying biology of these tumors and differences in its detail will determine a particular tumor's potential for aggressiveness. In addition, we can use these biological differences to identify novel molecular markers that may be useful for diagnostic, prognostic, or predictive purposes, the success of which would pave the road for a new era of personalized medicine in breast cancer.

In this study, we performed quantitative proteomic profiling of 83 triple-negative breast tumors to identify biomarkers for

good and bad prognostic. Although the size of the training set and the test set are limited, we propose the couple DP/TPS1 for the recurrence and bad prognostic and TrpRS for a good prognostic.

Desmoplakin is the principal plaque protein of desmosomes, involved in the adhesion junctions found in various tissues. Desmosomes are intercellular junctions that provide strong adhesion among cells. These proteins are ubiquitously expressed in epithelia and play a critical role in the maintenance of epithelial tissue integrity. Recently, studies suggest that desmosomes participate in the regulation of cell motility, growth, differentiation, and apoptosis (20–22). DP, as founding member of the plakin family, is an obligate component of desmosomal plaques (23). Two isoforms of DP have been reported so far, DP I (322 kDa) and DP II (259 kDa), both encoded by the DSP gene on human chromosome 6p24.3. DP proteins interact with plakoglobin ( $\gamma$ -catenin), plakophilins, and intermediate filaments, providing the intimate link between desmosomal cadherins and the cytoskeleton (24, 25) and belong to the pathway cytoskeleton Remodeling\_Keratins-filaments and Gap junctions found in Metacore analysis. This is in agreement with the fact that this pathway is the top-ranked pathway characterizing the relapse group in our proteomic approach.

The second poor prognostic biomarker, TPS1, thrombospondin-1, is an extracellular matrix glycoprotein (26) and was initially recognized as an antiangiogenic factor (27, 28). More recently, TSP1 was shown to induce proangiogenic activity in breast cancer cells (29). Other experiments suggested that TSP1 accelerates invasion and metastasis in breast (30), pancreatic (31), thyroid (32), and prostate (33) cancers. Recent study on invasive ductal carcinoma of the breast have reported that TSP1 is highly expressed in tumors associated with lymph node metastasis (34). In this context, TSP1 is not a specific marker of triple-negative breast tumors but rather a marker for aggressive tumors.

The other interesting protein as a good prognostic is TrpRS, which is a tryptophanyl-tRNA synthetase. Several tRNA synthase proteins have been identified as secreted cytokines that

control angiogenesis and immune responses and that may have roles in the tumor microenvironment (35). This protein is 653 amino acids in length and is involved in protein synthesis, regulation of RNA transcription and translation, and cytokine activities in inflammatory and angiogenic signaling pathways (36). The native enzyme lacks angiogenic activity. Proteolysis or alternative splicing of its N-terminal 47 amino acids, generates a mature mini-TrpRS or T2-TrpRS (37, 38), which possesses the angiostatic cytokine function (39). The production of mini-TrpRS is stimulated by IFN $\gamma$ . The secretion of TrpRS is mediated by the dissociation from a ternary complex that is formed with annexin II and S100A10 in the cytosol (40). Interestingly, cytosolic S100A10 was observed to be decreased in abundance in this study. VE-cadherin, a calcium-dependent adhesion molecule, was identified as a receptor for mini-TrpRS (41) and is selectively expressed and concentrated at the intercellular junctions of endothelial cells.

TrpRS has been demonstrated to regulate ERK, Akt, and eNOS activation pathways that are associated with angiogenesis, cytoskeletal reorganization, and shear stress-responsive gene expression (40). Interestingly, a recent report indicated that low expression levels of TrpRS are related to an increased risk of disease recurrence and reduced survival of patients with colon cancer (42), although it is not known whether this finding is related to the angiostatic activity of TrpRS.

Future prospective clinical trials are needed to further consolidate the validity of this biomarker signature. Nevertheless, we propose the establishment of an IHC test to calculate the score of TrpRS, DP, and TPS1 and to evaluate the degree of aggressiveness of the triple-negative tumors.

**Acknowledgments**—Tissues used in this work were provided by: Institut Régional du Cancer Nantes-Atlantique tumor bank, funded by the Institut National du Cancer (INCa) and the Cancéropôle Grand Ouest, and Anjou Maine ICO Paul Papin tumor bank funded by the INCa.

\* This work was supported by grants from the Ligue Contre le Cancer (comité du Maine et Loire), Cancéropôle Grand Ouest and Pays de Loire region.

§ This article contains [supplemental material](#) [Supplemental data S1-S15](#).

\*\* To whom correspondence should be addressed: Institut de Cancérologie de l'Ouest Paul Papin, INSERM U892, CNRS 6299, 2 rue Moll, 49000 Angers, France. Fax: 33 2 41 35 27 30; E-mail: catherine.guette@ico.unicancer.fr

## REFERENCES

- Foulkes, W. D., Smith, I. E., and Reis-Filho, J. S. (2010) Triple-negative breast cancer. *N. Engl. J. Med.* **363**, 1938–1948
- Harris, L., Fritzsche, H., Mennel, R., Norton, L., Ravdin, P., Taube, S., Somerfield, M. R., Hayes, D. F., and Bast, R. C., Jr. (2007) American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. *J. Clin. Oncol.* **25**, 5287–5312
- Gámez-Pozo, A., Ferrer, N. I., Ciruelos, E., López-Vacas, R., Martínez, F. G., Espinosa, E., Vara, J. Á. (2013) Shotgun proteomics of archival triple-negative breast cancer samples. *Proteomics Clin. Appl.* **3–4**, 283–291
- Kim, M. J., Kim, D. H., Jung, W. H., and Koo, J. S. (2014) Expression of metabolism-related proteins in triple-negative breast cancer. *Int. J. Clin. Exp. Pathol.* **7**, 301–12
- Muñiz Lino, M. A., Palacios-Rodríguez, Y., Rodríguez-Cuevas, S., Bautista-Piña, V., Marchat, L. A., Ruiz-García, E., Astudillo de la Vega, H., González-Santiago, A. E., Flores-Pérez, A., Díaz-Chavez, J., Carlos-Reyes, Á., Álvarez-Sánchez, E., and López-Camarillo, C. (2014) Comparative proteomic profiling of triple-negative breast cancer reveals that up-regulation of RhoGDI-2 is associated to the inhibition of caspase 3 and caspase 9. *J. Proteomics*. **111**, 198–211
- Liu, N. Q., Stingl, C., Look, M. P., Smid, M., Braakman, R. B., De Marchi, T., Sieuwerts, A. M., Span, P. N., Sweep, F. C., Linderholm, B. K., Mangia, A., Paradiso, A., Dirix, L. Y., Van Laere, S. J., Luiders, T. M., Martens, J. W., Foekens, J. A., Umar, A. (2014) Comparative proteome analysis revealing an 11-protein signature for aggressive triple-negative breast cancer. *J. Natl. Cancer Inst.* **106**, 376
- Albini, A., and Sporn, M. B. (2007) The tumour microenvironment as a target for chemoprevention. *Nat. Rev. Cancer*. **7**, 139–147
- Bizzarri, M., Cucinal, A., and Proietti, S., (2010) The tumor microenvironment as a target for anticancer treatment. *Oncobiol. Targets*. **1**, 3–11
- Wiśniewski, J. R., Zougman, A., Nagaraj, N., and Mann, M. (2009) Universal sample preparation method for proteome analysis. *Nat. Methods* **6**, 359–362
- Ernault, E., and Guette, C. (2011) OFFGEL-Isocratic Focusing Fractionation for the analysis of complex proteome. *Neuroproteomics*, Edited by Ka Wan Li, Humana Press Inc, Springer protocols U.S. 145–158
- Shilov, I. V., Seymour, S. L., Patel, A. A., Loboda, A., Tang, W. H., Keating, S. P., Hunter, C. L., Nuwaysir, L. M., and Schaeffer, D. A. (2007) The Paragon Algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Mol. Cell. Proteomics* **6**, 1638–1655
- Schwacke, J. H., Hill, E. G., Krug, E. L., Comte-Walters, S., and Schey, K. L. (2009) iQuantitator: A tool for protein expression inference using iTRAQ. *BMC Bioinformatics* **10**, 342
- Grant, J. E., Bradshaw, A. D., Schwacke, J. H., Baicu, C. F., Zile, M. R., and Schey, K. L. (2009) Quantification of protein expression changes in the aging left ventricle of *Rattus norvegicus*. *J. Proteome Res.* **8**, 4252–4263
- Besson, D., Pavageau, A. H., Valo, I., Bélanger, A., Eymery-Morin, C., Moulière A., Chassevent, A., Boisdron-Celle M., Morel A., Solassol J., Campone M., Gamelin E., Barré B., Coqueret O., Guette C. (2011) A quantitative proteomic approach of the different stages of colorectal cancer establishes OLFM4 as a new nonmetastatic tumor marker. *Mol. Cell. Proteomics* **10**, M111.009712
- Zeeberg, B. R., Qin, H., Narasimhan, S., Sunshine, M., Cao, H., Kane, D. W., Reimers, M., Stephens, R. M., Bryant, D., Burt, S. K., Elnekave, E., Hari, D. M., Wynn, T. A., Cunningham-Rundles, C., Stewart, D. M., Nelson, D., and Weinstein, J. N. (2005) High-throughput GoMiner, an “industrial-strength” integrative gene ontology tool for interpretation of multiple-microarray experiments, with application to studies of common variable immune deficiency (CVID). *BMC Bioinformatics* **6**, 168
- Eriksson, L., Antti, H., Gottfries, J., Holmes, E., Johansson, E., Lindgren, F., Long, I., Lundstedt, T., Trygg, J., and Wold, S. (2004) Using chemometrics for navigating in the large data sets of genomics, proteomics, and metabolomics (gpm). *Anal. Bioanal. Chem.* **380**, 419–429
- Geladi P, and Esbensen K. (1991) Regression on Multivariate Images -Principal Component Regression for Modeling, Prediction and Visual Diagnostic-Tools. *J. Chemometr.* **5**, 97–111
- Wold, S., Jonsson, J., Sjostrom, M., Sandberg, M., Rännar, S. (1993) DNA and peptide sequences and chemical processes multivariately modeled by principal component analysis and partial least-squares projections to latent structures. *Anal. Chim. Acta* **277**, 239–253
- Trygg, J., and Wold, S. (2002) Orthogonal projections to latent structures (O-PLS). *J. Chemometr.* **16**, 119–128
- Allen, E., Yu, Q. C., and Fuchs, E., (1996) Mice expressing a mutant desmosomal cadherin exhibit abnormalities in desmosomes, proliferation, and epidermal differentiation. *J. Cell Biol.* **133**, 1367–1382
- Wan, H., South, A. P., and Hart, I. R., (2007) Increased keratinocyte proliferation initiated through downregulation of desmoplakin by RNA interference. *Exp. Cell Res.* **313**, 2336–2344
- Rickelt, S., Winter-Simanowski, S., Noffz, E., Kuhn C., Franke, W. W. (2009) Upregulation of plakophilin-2 and its acquisition to adherens junctions identifies a novel molecular ensemble of cell-cell-attachment character-

- istic for transformed mesenchymal cells. *Int. J. Cancer*, **125**, 2036–2048
23. Leung, C. L., Green, K. J., and Liem, R. K., (2002) Plakins: A family of versatile cytolinker proteins. *Trends Cell Biol.*, **12**, 37–45
  24. Jonkman, M. F., Pasmooij, A. M., Pasmans, S. G., van den Berg, M. P., Ter Horst H. J., Timmer A., Pas H. H., (2005) Loss of desmoplakin tail causes lethal acantholytic epidermolysis bullosa. *Am. J. Hum. Genet.* **77**, 653–660
  25. Kowalczyk, A. P., Bornslaeger, E. A., Borgwardt, J. E., Palka, H. L., Dhaliwal, A. S., Corcoran, C. M., Denning, M. F., and Green, K. J. (1997) The amino-terminal domain of desmoplakin binds to plakoglobin and clusters desmosomal cadherin-plakoglobin complexes. *J. Cell Biol.* **139**, 773–784
  26. Lawler, J. W., Slayter, H. S., and Coligan, J. E. (1978) Isolation and characterization of a high molecular weight glycoprotein from human blood platelets. *J. Biol. Chem.* **253**, 8609–8616
  27. Iruela-Arispe, M. L., Lombardo, M., Kruttsch, H. C., Lawler, J., and Roberts, D. D. (1999) Inhibition of angiogenesis by thrombospondin-1 is mediated by 2 independent regions within the type 1 repeats. *Circulation*. **100**, 1423–1431
  28. Locopo, N., Fanelli, M., and Gasparini, G., (1998) Clinical significance of angiogenic factors in breast cancer. *Breast Cancer Res. Treat.* **52**, 159–173
  29. Hyder, S. M., Liang, Y., Wu, J., and Welbern, V. (2009) Regulation of thrombospondin-1 by natural and synthetic progestins in human breast cancer cells. *Endocr. Relat. Cancer*. **16**, 809–817
  30. Fontana, A., Filleur, S., Guglielmi, J., Frappart, L., Bruno-Bossio, G., Boissier, S., Cabon, F., and Clézardin, P. (2005) Human breast tumors override the antiangiogenic effect of stromal thrombospondin-1 in vivo. *Int. J. Cancer*. **116**, 686–691
  31. McElroy, M. K., Kaushal, S., Tran Cao, H. S., Moossa, A. R., Talamini, M. A., Hoffman, R. M., and Bouvet, M. (2009) Upregulation of thrombospondin-1 and angiogenesis in an aggressive human pancreatic cancer cell line selected for high metastasis. *Mol. Cancer Ther.* **8**, 1779–1786
  32. Sid, B., Langlois, B., Sartelet, H., Bellon, G., Dedieu, S., and Martiny, L. (2008) Thrombospondin-1 enhances human thyroid carcinoma cell invasion through urokinase activity. *Int. J. Biochem. Cell Biol.* **40**, 1890–1900
  33. Trojan, L., Schaaf, A., Steidler, A., Haak, M., Thalmann, G., Knoll, T., Gretz, N., Alken, P., and Michel, M. S. (2005) Identification of metastasis-associated genes in prostate cancer by genetic profiling of human prostate cancer cell lines. *Anticancer Res.* **25**, 183–191
  34. Horiguchi, H., Yamagata, S., Rong Qian, Z., Kagawa, S., and Sakashita, N. (2013) Thrombospondin-1 is highly expressed in desmoplastic components of invasive ductal carcinoma of the breast and associated with lymph node metastasis. *J. Med. Invest.* **60**, 91–96
  35. Kim, S., You, S., and Hwang, D. (2011) Aminoacyl-tRNA synthetases and tumorigenesis: more than housekeeping. *Nat. Rev. Cancer* **11**, 708–718
  36. Wakasugi, K., and Schimmel P. (1999) Two distinct cytokines released from a human aminoacyl-tRNA synthetase. *Science* **284**, 147–151
  37. Tolstrup, A. B., Bejder, A., Fleckner, J., and Justesen, J. (1995) Transcriptional regulation of the interferon- $\gamma$ -inducible tryptophanyl-tRNA synthetase includes alternative splicing. *J. Biol. Chem.* **270**, 397–403
  38. Turpaev, K. T., Zakhariev, V. M., Sokolova, I. V., Narovlyansky, A. N., Amchenkova, A. M., Justesen, J., and Frolova, L. Y. (1996) Alternative processing of the tryptophanyl-tRNA synthetase mRNA from interferon-treated human cells. *Eur. J. Biochem.* **240**, 732–737
  39. Wakasugi, K., Slike, B. M., Hood, J., Otani, A., Ewalt, K. L., Friedlander, M., Cheres, D. A., and Schimmel, P. (2002) A human aminoacyl-tRNA synthetase as a regulator of angiogenesis. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 173–177,
  40. Kapoor, M., Zhou, Q., Otero, F., Myers, C. A., Bates, A., Belani, R., Liu, J., Luo, J. K., Tzima, E., Zhang, D. E., Yang, X. L., and Schimmel, P. (2008) Evidence for annexin II-S100A10 complex and plasmin in mobilization of cytokine activity of human TrpRS. *J. Biol. Chem.* **283**, 2070–2077
  41. Tzima, E., Reader, J. S., Irani-Tehrani, M., Ewalt, K. L., Schwartz, M. A., and Schimmel, P. (2005) VE-cadherin links tRNA synthetase cytokine to anti-angiogenic function. *J. Biol. Chem.* **280**, 2405–2408
  42. Ghanipour, A., Jirstrom, K., Pontén, F., Glimelius, B., Pahlman, L., and Birgisson, H. (2009) The prognostic significance of tryptophanyl-tRNA synthetase in colorectal cancer. *Cancer Epidemiol. Biomarkers Prev.* **18**, 2949–2956