CANCER GENOMICS & PROTEOMICS 5: 253-262 (2008)

Comparative Clinical and Transcriptomal Profiles of Breast Cancer Between French and South Mediterranean Patients Show Minor but Significative Biological Differences

N. CHALABI[1,2], D.J. BERNARD-GALLON[1,2], Y.-J. BIGNON[1,2,3] and THE BREAST MED CONSORTIUM: F. KWIATKOWSKI[1], M. AGIER[1,6], V. VIDAL[1,6], V. LAPLACE-CHABAUD[1,6], V. SYLVAIN-VIDAL[1,6], V. BERTHOLET[7], F. DE LONGUEVILLE[7], M. LACROIX[7], G. LECLERCQ[7], J. REMACLE[7], C. SIBILLE[8], N. ZAMMATEO[8], N. BEN JAAFAR[9], A. SEFIANI[9], K. OULDIM[9], A. MÉGARBANÉ[10], N. JALKH[10], W. MAHFOUDH[11], W. TROUDI[11], A. BEN AMMAR-EL GAÏED[11] and L. CHOUCHANE[12]

[1]Département d'Oncogénétique, Centre Jean Perrin, 63011 Clermont-Ferrand Cedex 01;
[2]Centre de Recherche en Nutrition Humaine (CRNH), 63009 Clermont-Ferrand Cedex 01;
[3]Université d'Auvergne, 63001 Clermont-Ferrand 1;
[4]Diagnogène, Division d'Imaxio, Biopôle Clermont-Limagne, 63360 Saint-Beauzire, France;
[5]Université de Namur, URBC, 5000 Namur;
[6]Eppendorf Array Technologies, 5000 Namur, Belgium;
[7]Département de Génétique et de Biologie Moléculaire, Institut National d'Hygiène, Rabat, Maroc;
[8]Unité de Génétique Médicale, Université Saint-Joseph, Bayrout, Liban;
[9]Laboratoire d'Immunologie, Institut Pasteur de Tunis;
[10]Laboratoire d'Immuno-Oncologie Moléculaire,
Faculté de Médecine de Monastir, Université de Monastir, Tunis, Tunisie

Abstract. *Background: In Western countries, breast cancer incidence and mortality are higher than in Mediterranean countries. These differences have been ascribed to environmental factors but also to late-stage diagnostic and biological specific characteristics. Patients and Methods: Between September 2002 and September 2005, we collected clinical data by phone counselling 180 French and Mediterranean breast cancer patients and performed microarray experiments. Results: Characteristics of breast cancer in patients from Lebanon, Tunisia and Morocco were more aggressive (more SBR grade III and positive node invasion) and patients were 10 years younger at diagnosis. Sixteen differentially expressed genes such as MMP9, VEGF, PHB1, BRCA1, TFAP2C, GJA1 and TFF1 were also found. Additionally, an up-regulation of cytokeratins KRT8 and KRT18 may indicate a luminal B subtype in "South" (Lebanon, Tunisia and Morocco) tumors while "North" (France) tumors may more frequently be luminal A type. Conclusion: This study allowed the identification of specific clinical and transcriptomic parameters in patients from South Mediterranean countries.*

Correspondence to: Professor Yves-Jean Bignon, Département d'Oncogénétique, Centre Jean Perrin, 58 rue Montalembert BP 392, 63001 Clermont-Ferrand, France. e-mail: Yves-Jean.Bignon@cjp.fr

Key Words: Breast cancer, microarrays, European and Mediterranean countries, clinical parameters.

Breast cancer is not only the most common cause of cancer death in women in Western industrialized countries but also in Mediterranean countries (1). Breast carcinoma incidence has increased dramatically in Europe over the last 20 years while mortality has remained unchanged (2). This phenomenon could be attributed to earlier detection, better diagnostic methods and more effective therapy. In Mediterranean countries, breast cancer incidence has also increased dramatically over the last 20 years and is now a public health problem, with poor prognosis mainly due to late-stage diagnosis, but also to different risk factors linked to biologically specific characteristics.

In Europe, some risk factors have been suspected, including age, hereditary predisposition, hormonal exposure (early age at menarche, late menopause, late age at first pregnancy, or prolonged use of oral contraceptive) and other factors (diet, alcohol, obesity) (3). All these factors are different between North European and South Mediterranean patients and may influence breast cancer risk in a protective or aggravating way. Indeed, in Mediterranean countries, there are fewer breast cancer cases than in European countries. This phenomenon may be attributed to a larger number of

Table I. *Characteristics of patients between the "North" and the "South" populations.*

	North (France)	South (Lebanon, Tunisia and Morocco)
Country of residence	123	57
Height (cm)[a]	162.7±5.4 (148-177)	162.7±5.7 (152-175)
Weight (kg)[a]	66.9±11.6 (48-96)	71.0±11.8 (50-112)
Body mass index (BMI)[a]	25.3 (21.9-30.6)	26.8 (21.6-36.6)
Smoker[b]		
Yes	41 (33.3%)	16 (35.6%)
No	82 (66.7%)	29 (64.4%)
Duration of smoking (months)[a]	70.3±125.7 (0-480)	88.9±146.3 (0-600)
Consumption of alcohol[b]		
Yes	46 (37.7%)	8 (17.8%)
No	76 (62.3%)	37 (82.2%)
Duration of alcohol consumption (months)[a]	1.3±2.5 (0-14)	0.2±0.5 (0-2)

[a]Average value±standard deviation (range); [b]number of patients (percentage of patients).

Table II. *Baseline characteristics of breast cancer cases in the two studied populations.*

Characteristic (Total of cases)	North (France)	South (Lebanon, Tunisia and Morocco)
Age at menarche (years)[a]	13.1±1.6 (9-17)	12.7±1.3 (10-17)
Number of pregnancies[a]	2.4±1.7 (0-10)	3.2±2.9 (0-15)
Number of full-term pregnancies[a]	2.0±1.5 (0-8)	2.4±1.9 (0-7)
Age at first pregnancy (years)[a]	23.2±4.4 (15-42)	24.1±4.6 (18-33)
Oral contraceptive use[b]		
Yes	43 (35%)	8 (14%)
No	67 (54.5%)	37 (64.9%)
Not known	13 (10.6%)	12 (21.1%)
Duration of use (months)[a]	139.1±111.2 (2-360)	37.6±34.1 (1-84)
Breastfeeding[b]		
Yes	34 (27.6%)	37 (64.9%)
No	53 (43.1%)	4 (7%)
Not known	36 (29.3%)	16 (28.1%)
Duration of breastfeeding (Months)[a]	2.0±2.7 (0-12)	5.5±10.3 (0-50)
Age at menopause (years)[a]	50.5±3.7 (41-58)	49.6±4.8 (39-62)
Previous carcinoma[b]		
Breast	11 (8.9%)	45 (78.9%)
Ovary	2 (1.6%)	0

[a]Average value±standard deviation (range); [b]number of patients (percentage of patients).

pregnancies and breastfeeding, which are protective factors. But in these countries, there is also an early age at menarche and an obesity problem (body mass index >26) which are aggravating breast cancer risk factors.

The Breast Med Consortium, financed by the European Union, was established between five countries [France, Belgium (data not shown here), Lebanon, Morocco and Tunisia] in order to evaluate molecular characteristics of breast cancer tumors in these countries. Our goal was to provide a better breast tumor sub-classification and to discover new prognostic markers, in particular for the evaluation of early anticancer treatment. Multiparametric analyses were performed, including histopathological information, life habits, main reproductive life events, personal history of breast cancer and transcriptomic analysis.

Patients and Methods

We collected population-based tumors of breast cancer cases from France, Lebanon, Tunisia and Morocco. Eligible study patients were women who developed breast cancer between 1999 and 2003 with an initial surgery without neoadjuvant chemotherapy. All tumors were then collected and cryoconserved in nitrogen until use. Each site recruited cases through their regional patient list and a total of 398 patients were selected. Among them, 244 tumors were extracted and finally 180 RNA were hybridized through quantity and quality criteria.

Clinical data were collected by phone counselling through a questionnaire regrouping 39 questions about characteristics (Table I), baseline characteristics of breast cancer cases (Table II) and breast cancer history (Table III). All this information was transcribed in internally developed software: Breast Med Consortium Director Software version 1.2 alpha (Soluscience, Biopôle Saint-Beauzire, France), which permitted us to extract data of interest.

Microarray slides were manufactured by Diagnogene™ Society, (Division Imaxio, Biopôle Saint-Beauzire, France). Oligonucleotides (25 nM) (MWG Array Technology) were spotted in triplicate by using the Microgrid arrayer (Biorobotic, Cambridge, UK) with a complexity of 3,456 spots per glass slide. Among these 3,456 spots, 2,094 corresponded to breast cancer genes for which a patent was submitted (Agreement number 0409192). Moreover, 444 spots corresponding to 148 genes (Table IV) were chosen specifically by the Breast Med Consortium according to their suspected role in breast carcinogenesis in literature (4).

Frozen tumors were first pulverized using a French-Press and total RNA was purified using RNeasy Mini Kit as *per* manufacturer's protocol (Qiagen, Courtaboeuf, France). Integrity of the RNA samples was verified using a 2100 Bioanalyzer with RNA 6000 Nano LabChip® and BioSizing A.02.11 software (Agilent Technologies, Massy, France). Ten μg of total RNA were reverse transcribed with FairPlay™ Microarray Labeling Kit (Stratagene®, Amsterdam, The Netherlands). Target patient cDNA and a cDNA generated from a Human Universal Reference RNA (Stratagene®, Amsterdam, the Netherlands) originating from a pool of 10 cell lines were simultaneous hybridized. After purification, the two cDNA were labeled with cyanine 3 (reference cDNA) or cyanine 5 (tumor cDNA) (Amersham Biosciences, Saclay, France). Labeled cDNA were then purified through columns using a DNA-binding solution (Stratagene®, Amsterdam, The Netherlands). Cy-3 and Cy-5 labeled cDNA were mixed and concentrated using a Speed-Vacuum concentrator (45 minutes at room temperature). The samples were dissolved in hybridization buffer (Corning, New York,

Table III. *Characteristics of breast tumors*.

Characteristic (Total of cases)	North (France)	South (Lebanon, Tunisia and Morocco)
Number of tumors[a]	1.2±0.4 (1-3)	0.9±0.6 (0-3)
Age at breast cancer diagnosis (years)[a]	60.1±12.7 (23-87)	49.6±13.1 (28-79)
Tumor diameter (cm)[a]	2.55±1.6 (0.4-10)	2.94±1.77 (1-8)
Tumor side: right/ left[b]	55 (44.7%)/68 (55.3%)	29 (52.7%)/26 (47.3%)
Histological type[b]		
Invasive ductal carcinoma	98 (79.7%)	37 (64.9%)
Invasive lobular carcinoma	13 (10.6%)	8 (14%)
In situ ductal carcinoma	3 (2.4%)	2 (3.5%)
Medullar carcinoma	3 (2.4%)	0
Inflammatory carcinoma	0	1 (1.8%)
Other	4 (3.2%)	1 (1.8%)
Hormonal receptor status	118	47
Positive (estrogen/progesterone)	97 (82.2%)/79 (67%)	34 (72.3%)/31 (66%)
Negative (estrogen/progesterone)	21 (17.8%)/39 (33.1%)	13 (27.7%)/16 (34%)
ERBB2 Status (100)	100	0
Positive	8	0
Negative	43	0
Not Known	49	0

[a]Average value±standard deviation (range); [b]number of patients (percentage of patients).

USA) and hybridized on 15-mer oligonucleotide arrays for 16 h at 37°C in humidified chambers. After washing microarrays with several buffers (saline sodium citrate 2X – sodium dodecyl sulfate 0.1%; saline sodium citrate 0.1X – sodium dodecyl sulfate 0.1%; saline sodium citrate 0.1X), hybridization was visualized with a 418 GMS scanner (Affymetrix®, Santa Clara, CA, USA). GenePix Pro 6 (Molecular Device, Axon®, Saint-Grégoire, France) software was used to quantify the intensity of each spot.

To normalize signal intensities, data were submitted to R2.2.0 program and LimmaGUI Version 1.7.0 software (5). For each gene, Cy-3 and Cy-5 signal intensities were normalized using an Edwards background correction (6) and a global lowess normalization. Positive and negative controls were removed from normalized data. All clinical and trasnscriptomic data were included in a database using SEM software (7, 8). We compared Mediterranean "South" patients (Lebanon, Tunisia, Morocco) *versus* European "North" patients (France). Seven clinical parameters were analyzed and crossed with transcriptomic data to establish correlation between gene expression and patient clinical characteristics. These clinical parameters were as follows: country of residence, age at menarche, number of pregnancies, duration of breastfeeding, age at breast cancer diagnosis, estrogen receptor status and Scarff-Bloom-Richardson (SBR) grade. Statistical analyses were performed using a Bonferroni correction ($p<0.001$). Correlations were first found using global univariate analysis. Significant genes observed were then analyzed through a multivariate logistic regression.

Results

A large difference in the age at breast cancer diagnosis (±10 year, $p<0.001$) was found between North and South patients. North patients were 60.1 years old when diagnosed with breast cancer whereas South patients were 49.6 years old (Table III). Tunisian patients were in fact more than 13 years younger. We also found disparities such as tumor diameter, which was smaller in the North patients ($p=0.076$); and a predominance of SBR grade III in the South population ($p=0.00002$) which represents 63% of invasive ductal carcinomas, whereas there were only 25% in the North population. Moreover, positive node invasion (N+) was more frequent in South cases (62.5%) whereas only 40% of the North cases were N+ ($p=0.0087$).

Using SEM software, tumors were classified according to an analysis which regroups tumors by similar characteristics (Figure 1). Tumors were classified according to pathology degree (axis 1) and sociocultural (axis 2) characteristics. Axis 1 was analyzed by 3 clinical parameters: SBR grade (I, II or III), estrogen receptor status (ER+ *versus* ER–) and age at menarche (<13 or ≥13). In axis 2, tumors were assessed by 4 parameters: country of residence (North *versus* South), age at breast cancer diagnosis (<50, 50-65, ≥65 years), number of pregnancies (<3 or ≥3) and breast-feeding (Yes *versus* No)

Secondly, by global univariate analysis, we found 59 differentially expressed genes according to these seven clinical parameters (Table V). Among them, after multivariate logistic regression, only 16 genes were highly significant and correlated with these parameters ($p<0.05$). Three differentially expressed genes were found related to angiogenesis: *MMP9* (matrix metalloproteinase 9), *VEGF* (vascular endothelial growth factor) and *PHB1* (prohibitin). According to the "country of residence" criterion, *MMP9* was up-regulated in the South

Table IV. *List of the 148 genes chosen by the Breast Med Consortium.*

Gene symbol	Gene name	GeneBank number
ACTR1A	ARP1 actin-related protein 1 homolog A	NM_005736
APEX1	APEX nuclease (multifunctional DNA repair enzyme) 1	NM_001641
ARVCF	Armadillo repeat gene deletes in velocardiofacial syndrome	NM_001670
ATM	Ataxia telangiectasia mutated	U26455
BAG1	BCL-2 associated athanagene	AF022224
BAX	BCL2-associated X protein	NM_004324
BCAR1	Breast cancer anti-estrogen resistance 1	NM_014567
BCL2	B-cell lymphoma 2	NM_000633
BRCA1	Breast cancer 1,	Y08864
BRCA2	Breast cancer 2	NM_000059
BRF1	BRF1 homolog	NM_001519
CAD	Carbamoyl-phosphate synthetase 2, aspartate transcarbamylase, and dihydroorotase	NM_004341
CAV1	Caveoline1	NM_001753
CBX3	Chromobox homolog 3 (HECH)	NM_007276
CCND1	Cyclin D1	NM_053056
CCNE1	Cyclin E1	NM_001238
CD36	CD36 antigen (collagen type I receptor, thrombospondin receptor)	NM_000072
CDC42BPA	CDC42 binding protein kinase alpha (DMPK-like)	NM_003607
CDH1	Cadherin1/E-cadherin	NM004360
CDH13	Cadherin13	U59289
CDK4	Cyclin-dependent kinase 4	NM_000075
CDKN1A	Cyclin-dependent kinase inhibitor 1A (p21, Cip1)	XM_011458
CDKN1B	Cyclin-dependent kinase inhibitor 1B (p27, Kip1)	XM_006894
CDKN1C	Cyclin dependent kinase inhibitor 1C	NM_000076
CDKN2A	Cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4)	XM_005656
CEACAM5	Carcinoembryonic antigen-related cell adhesion molecule 5	NM_004363
COX6C	Cytochrome *c* oxidase subunit vic	NM_004374
CP	Ceruloplasmin (ferroxidase)	NM_000096
CRABP2	Cellular retinoic acid binding protein 2	NM_001878
CSDA	Cold shock domain protein A	NM_003651
CSF1	Colony stimulating factor 1	M37435
CSF1R	Colony stimulating factor 1 receptor, formerly McDonough feline sarcoma viral (v-fms) oncogene homolog	XM_003789
CST6	Cystatin M	U62800
CTPS	CTP synthase	NM_001905
CTSD	Cathepsin D	NM_001904
CX3CL1	Chemokine (C-X3-C motif) ligand 1	NM_002996
D123	D123 gene product	NM_006023
DLX4 (BP1)	Distal-less homeobox 4	NM_138281
EGFR	Epidermal growth factor receptor	NM_005228
ERBB2	V-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma-derived oncogene homolog	NM_004448
ESR1	Estrogen receptor alpha	NM_000125
ESR2	Estrogen receptor beta	X99101
FABP4	Fatty acid binding protein 4, adipocyte	NM_001442
FGF2	Fibroblast growth factor 2 (basic)	XM_003306
FGF8	Fibroblast growth factor 8	U36223
FGFR	Fibroblast growth factor receptor 1	NM_000604
FHIT	Fragile histidine triad gene	NM_002012
FOXA1	Forkhead box A1	NM_004496
FOXM1	Forkhead box M1	NM_021953
G22P1	Thyroid autoantigen 70 kDa	NM_001469
GATA3	GATA binding protein 3	NM_002051
GDI2	GDP dissociation inhibitor 2	NM_001494
GJA1	Connexine 43	NM_000165
GNAI3	Guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 3	NM_006496
GPX4	Glutathione peroxidase 4	NM_002085
GSTP1	Glutathione S-transferase *pi*	NM_000852
HADHA	Hydroxyacyl-coenzyme A dehydrogenase	NM_000182
HGF	Hepatocyte growth factor	M29145
HSPC195	Hypothetical protein HSPC195	NM_016463

Table IV. *continued*

Table IV. *continued*

Gene symbol	Gene name	GeneBank number
IGF2	Insulin-like growth factor 2 (somatomedin A)	NM_000612
IGFBP2	Insulin growth factor binding protein 2	NM_000597
IGFBP5	Insulin growth factor binding protein 5	NM_000599
IGFBP6	Insulin-like growth factor binding protein 6	NM_002178
IL1A	Interleukin 1, alpha	XM_017768
IL1B	Interleukin 1, beta	XM_010760
IL6	Interleukin 6 (interferon, beta 2)	XM_004777
IL8	Interleukin 8	XM_003501
ILF2	Interleukin enhancer binding factor 2	NM_004515
ING1	Inhibitor of growth family, member 1	NM_005537
ITGA6	Integrin, alpha 6	XM_002335
ITGB1	Integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29 includes MDF2, MSK12)	XM_005799
ITGB3	Integrin, beta 3 (platelet glycoprotein iiia, antigen CD61)	XM_012636
ITGB8	Integrin, beta 8	NM_002214
KAI1	Kangai 1	NM_002231
KIAA0601	KIAA0601 protein	mRNA: XM_031267
KISS1	Kiss1 metastasis suprressor	NM_002256
KLK3	Kallikrein 3	NM_001648
KRT17	Keratin 17	NM_000422
KRT18	Cytokeratin 18	NM_000224
KRT19	Cytokeratin 19	NM_002276
KRT5	Keratin 5	NM_000424
KRT8	Cytokeratin 8	NM_002273
LAMC2	Laminin, gamma 2	NM_005562
LIV-1	LIV-1 protein, estrogen-regulated	NM_012319
LRP1	Low density lipoprotein-related protein 1	NM_002332
MCM7	Minichromosome maintenance deficient 7	NM_005916
MGB1	Mammaglobin 1	NM_002411
MKI67	Ki-67	NM_002417
MMP11	Matrix metalloproteinase 11	NM_005940
MMP13	Matrix metalloproteinase 13	NM_002427
MMP14	Matrix metalloproteinase 14	NM_004995
MMP9	Matrix metalloproteinase 9	NM_004994
MTMR4	Myotubularin-related protein 4	NM_004687
MUC1	Mucine1, transmembranaire	NM_002456
MX2	Myxovirus (influenza virus) resistance 2	NM_002463
MYBL2	b-myb	NM_002466
MYC	c-myc	NM012333
NCOA3	Nuclear receptor coactivator 3	NM_006534
NIFU	Nitrogen fixation cluster-like	NM_014301
NME1	Non-metastatic cells 1	NM_000269
NSEP1	Nuclease sensitive element binding protein 1	NM_004559
ODC	Ornithine decarboxylase 1	NM_002539
OXCT	3-oxoacid coa transferase	NM_000436
P53	Tumor protein p53	AF307851
PAI-RBP1	PAI-1 mRNA-binding protein	NM_015640
PCNA	Proliferating cell nuclear antigen	NM002592
PDGFB	Platelet-derived growth factor beta polypeptide	NM_002608
PFKP	Phosphofructokinase, platelet	NM_002627
PHB1	Prohibitin	
PGR	Progesterone receptor	NM_000926
PHYH	Phytanoyl-CoA hydroxylase	NM_006214
PIP	Prolactin-induced protein (gross cystic disease fluid protein)	NM_002652
PLAT	Plasminogen activator tissue	NM_000930
PPP1CB	Protein phosphatase 1, catalytic subunit, beta isoform	NM_002709
PRAME	Preferentially expressed antigen in melanoma	NM_006115
PRLR	Prolactin receptor	NM_000949
RB1	Retinoblastoma 1 (including osteosarcoma)	XM_007211
RBL2	Retinoblastoma-like 2	NM_005611

Table IV. *continued*

Table IV. *continued*

Gene symbol	Gene name	GeneBank number
S100A4	S100 calcium binding protein A4	NM_002961
SELENBP1	Selenium binding protein 1	NM_003944
SERPINB5	Maspin / serine proteinase inhibitor clade b member 5	NM_002639
SERPINE1	Plasminogen activator inhibitor type1	M14083
SLPI	Secretory leukocyte protease inhibitor (antileukoproteinase)	NM_003064
SOX9	SRY (sex determining region Y)-box 9 (campomelic dysplasia, autosomal sex-reversal)	NM_000346
SPHAR	S-phase response	NM_006542
SPS	Selenophosphate synthetase	NM_012247
ST13	Suppression of tumorigenicity 13	NM_003932
STAB1	Stabilin 1	NM_015136
STAT1	Signal transducer and activator of transcription 1, 91 kDa	NM_007315
STC2	Stanniocalcin 2	NM_003714
TFAP2C	Transcription factor AP-2 gamma	NM_003222
TFF1	Trefoil factor 1	NM_003225
THBS1	Thrombospondin 1	NM_003246
TIMP1	Tissue inhibitor of metalloproteinase 1	NM_003254
TIMP2	Tissue inhibitor of metalloproteinase 2	XM_012690
TOB1	Transducer of ERBB2, 1	NM_005749
TOP2A	Topoisomerase2-alpha	NM_001067
TP53BP2	Tumor protein p53 binding protein, 2	NM_005426
UGTREL1	UDP-galactose transporter related	NM_005827
VEGF	Vascular endothelial growth factor	XM_004512
VEGFR1	Vascular endothelial growth factor receptor 1	NM_002019
VIM	Vimentine	NM_003380
VLDLR	Very low density lipoprotein receptor	NM_003383
VWF	Factor von willebrand	NM_000552
XBP1	X-box binding protein 1	NM_005080
ZNF161	Zinc finger protein 161	NM_007146
ZNF22	Zinc finger protein 22	BC010642

Table V. *Differentially expressed genes in "South" tumors according to global univariate analysis and multivariate logistic regression (p<0.05). To distinguish down-regulated from up-regulated genes, down-regulated genes are represented in bold.*

Clinical parameters	Univariate analysis	Logistic regression
Country of residence	*CDH1, CDKN1A, CSF1R, KIAA0601, MMP9,* ***PIP****, PRAME,* ***VEGF***	*MMP9,* ***VEGF****, CDKN1*
Age at menarche	***IL1A, KIAA0601, LRP1****, PHB1,* ***PLAT****, PRLR, SERPINE1, ST13, TFAP2C,* ***THSB1***	***KIAA0601****, TFAP2C*
Nursing	*APEX1, BRCA1,* ***CRABP2****, FGF8,* ***MMP9****, PHYP*	*BRCA1,* ***MMP9***
Number of pregnancies	*ACTR1A,* ***CCNE1****, CDC42BPA,* ***GJA1****, GSTP1,* ***LIV_1***	*ACTR1A,* ***GJA1***
Age at breast cancer diagnosis	***BRCA2****, CEACAM5, GPX4, ING1,* ***KRT18****, MYBL2, PAI_RBP1,* ***S100A4, SERPINE1, ZNF161***	*CEACAM5,* ***KRT18***
Positive estrogen receptor	***BAG1, COX6C, CRABP2****, CSF1, GNAI3,* ***IGF2****, ITGB3, KRT18, KRT8, LAMC2,* ***MMP9****, MYBL2,* ***NCOA3, ODC, PHB1, SERPINB5****, SELENBP1, SPHAR,* ***STC2, TFF1,*** *TOB1,* ***VEGF, XBP1, ZNF161***	*KRT18, KRT8,* ***MMP9, PHB1, STC2, TFF1***
SBR grade (1 *vs*. 2+3)	*APEX1,* ***BAG1****, CCNE1, CP,* ***CTPS, GNAI3, GSTP1****, HGF,* ***MMP11****, MYBL2,* ***OXCT****, PHB, TFAP2C,* ***VEGF***	*CP,* ***OXCT****, TFAP2C,* ***VEGF***

population whereas *VEGF* was rather up-regulated in the North tumors. *MMP9* was also found to be down-regulated in the South group according to the "breastfeeding" parameter. Down-regulation of *MMP9* and *PHB1* was linked to the presence of estrogen receptor in South tumors and a down-regulation of *VEGF* in South tumors was associated with SBR grade. Six genes tightly linked with breast cancer were also altered: *BRCA1* (breast cancer 1), *TFAP2C* (transcription factor

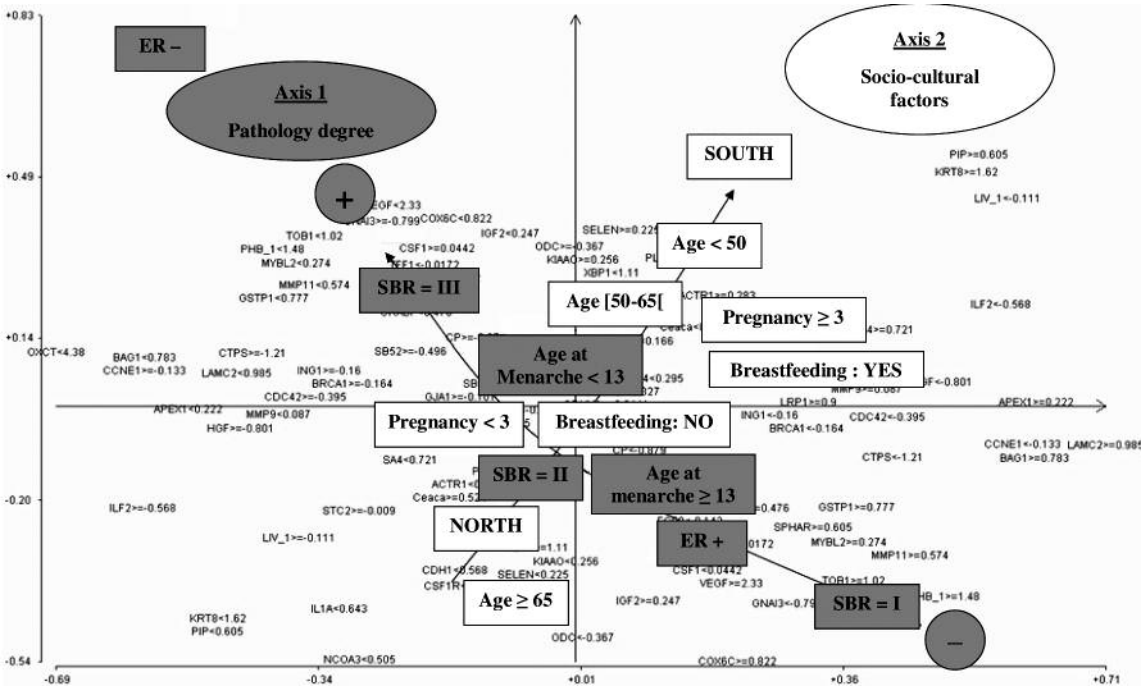

Figure 1. *Corresponding analysis using SEM software. Tumors had been ranged according to 2 axes: socio-cultural parameters* □ *and pathology degree* ■.

AP2-gamma), *GJA1* (gap junction protein alpha 1), *TFF1* (trefoil factor 1), *KRT8* (keratin 8) and *KRT18* (keratin 18). *BRCA1* was up-regulated in South tumors according to the "breastfeeding" criterion. *TFAP2C* was also up-regulated in regard to "the age at menarche" and "SBR grade" parameters whereas *GJA1* and *TFF1* were down-regulated respectively in association with the number of pregnancies and in ER-positive South tumors. An up-regulation of cytokeratins *KRT8* and *KRT18* was observed in ER-positive South tumors.

Other genes involved in various pathways were also differentially expressed according to clinical data. We noted up-regulation of *CDKN1A* (cyclin-dependent kinase inhibitor 1A) in South tumors, *ACTR1A* (actin-related protein 1A) according to the number of pregnancies, *CEACAM5* (carcinoembryonic antigen-related cell adhesion molecule 5) in relation to the age at breast cancer diagnostic and *CP* (ceruloplasmin) with SBR grade; whereas *KIAA0601* (amine oxidase, flavin-containing, 2) according to the age at menarche, *STC2* (stanniocalcin 2) in relation with ER-positive tumors and *OXCT* (3-oxoacid CoA transferase 1) with SBR grade were found to be down-regulated.

Discussion

The aim of the Breast Med Consortium was to compare clinical data and transcriptomic profiling expression between French and South Mediterranean breast cancer patients. We first noticed a 10-year age difference at breast cancer diagnosis between these two populations. This is in accordance with the data of the GLOBOCAN (2002): the average age at breast cancer diagnosis in Maghreb is of 45-50 years whereas in France it is around 60 years (9). Therefore, the population of postmenopausal women (55-65 years) is under-represented in the Maghreb hence relatively young women having breast cancer could be recruited by the health system. This difference in age breast cancer diagnosis led to the hypothesis that tumor characteristics should be different between the North and the South patients. Indeed, among the South patients, we found tumor diameter to be more important, and more tumors with SBR grade III and positive lymph nodes.

Tumors were classified according to seven discriminant clinical characteristics: country of residence, age at menarche, breastfeeding, number of pregnancies, age at breast cancer diagnosis, positive estrogen receptor status and SBR grade. We found 59 differentially expressed genes linked to these clinical parameters. Among them, 16 seemed to be more significant. We observed differential expression of *MMP9, VEGF* and *PHB1* involved in tumor invasion and angiogenesis. *MMP9* was found to be up-regulated in South tumors and down-regulated according to the "breastfeeding" criteria. MMP9 expression was related to aggressive tumor behavior. However, in the South group, there were more SBR grade III tumors than in the North group. Moreover, 64.9% of the South population reported breast-feeding for a period of 5.5 months

versus 27.6% during 2 months in the North population. It has been demonstrated that breastfeeding for more than 3 months could reduce breast cancer risk by 16% (10). This phenomenon was explained by a highly significant linear relationship between breast cancer risk and the cumulative number of cycles before a first full-term pregnancy. During pregnancy and breastfeeding, mammary gland differentiation may inhibit carcinogenesis initiation (11).

We also found down-regulation of *VEGF* in South tumors according to SBR grade. *VEGF* mediates angiogenic activity in a variety of estrogen target tissues. Studies demonstrated that estradiol (E2)-regulated *VEGF* gene transcription requires a variant estrogen response element (ERE) (12). However, in our population, clinical data showed that there were fewer ER-positive tumors in the South population, which could explain the underexpression of VEGF in this population. The *PHB1* gene was also found down-regulated in estrogen-positive South tumors. PHB1 is a multifunctional membrane protein established as a vascular marker of adipose tissue (13). In our data, we found that South women exhibited an average BMI of 26.8, with an upper value of 36.5 which expresses severe obesity. However, proliferation of tumor cells depends on new blood vessel formation that accompanies malignant progression. Although white fat is a nonmalignant tissue but highly vascularized, it could nevertheless quickly proliferate and expand (13). These authors also showed that targeting a proapoptotic peptide to PHB1 in the adipose vasculature caused ablation of white fat in mice. Because PHB1 is also expressed in blood vessels of human white fat, the work suggested the development of targeted drugs for treatment of obese patients.

Transcriptomic analysis revealed differentially expressed genes related to breast cancer. Indeed, we observed an up-regulation of *BRCA1* in relation with breastfeeding in South tumors. Women carrying deleterious *BRCA1* mutation had a reduced risk of breast cancer of 45% if they breastfed for one year or more (14). Furthermore, individuals with low levels of *BRCA1* may have increased breast epithelial cell proliferation in response to the increased estrogen exposure of pregnancy. But the same authors demonstrated that women who carry a *BRCA1* mutation had poor milk production. Nevertheless, breast cancer risk had been found to increase in women with nonhereditary breast cancer who tried to breastfeed but could not (15). In our data, South patients breastfed more than North patients. *BRCA1* is implicated in proliferation and differentiation of the mammary gland by among other functions, suppressing estrogen-mediated breast cell proliferation (16). Moreover, breastfeeding is also related to breast estrogen levels. Human milk production may also allow carcinogen excretion such as organochlorine xenoestrogen compounds (17). Indeed, accumulation of these molecules in the mammary gland has been associated with a potential role in breast cancer etiology (18).

We also observed an up-regulation of *TFAP2C* in relation with the "age at menarche" and the "SBR grade" parameters. This gene is involved in the overexpression of *HER2* in human breast cancer cells (19). A positive expression of HER2 is associated with the luminal B tumor subtype (20). This is generally associated with a positive-estrogen receptor expression, while the negative-estrogen receptor tumors are rather classified as basal-like subtype. Moreover, luminal tumors also expressed KRT8 and KRT18 cytokeratins (21). In our data, we also noted an up-regulation of these two cytokeratins in South tumors related to the expression of estrogen receptor. These results could indicate that there were more luminal B subtype tumors in the South population. In North patients, we observed luminal A subtype tumors, which expressed both estrogen and progesterone receptors but not HER2. Nevertheless, HER2 expression has not been systematically evaluated for all the tumors. Moreover, we observed a down-regulation of *TFF1* in South tumors according to the presence of estrogen receptor. *TFF1* is a breast cancer estrogen-inducible gene of which mRNA and protein expressions are detected in approximately 50% of human breast tumors (22). It has been demonstrated that the *TFF1* gene is transcriptionally induced in MCF-7 human breast cancer cells by estrogen (23).

We also noted a down-regulation of *GJA1* in South tumors in relation to the number of pregnancies. GJA1, also known as connexin 43 (Cx43), is a protein involved in gap junctions which facilitate cell-to-cell adhesion and provide pathways for direct intercellular communication. It has been determined that normal mammary epithelial cells that expressed Cx26 and Cx43, but not tumor cells that did not express them, contained functional gap junctions (24). Cx43 is constitutively expressed at a uniform low level throughout the cell cycle. In our data, we also found an up-regulation of *CDKN1*, an effector of *p53* which controls the G1- to S-phase transition in mammals; p53 may mediate its biological role as a tumor suppressor. CDKN1 protein mediates p53 suppression of tumor cell growth by participating in the regulation of a G2/M checkpoint (25). There was also an up-regulation of *CEACAM5* in South tumors in relation to the age at diagnosis. The CEA immunoassay is useful in the diagnosis and serial monitoring of cancer patients for recurrent disease or response to therapy. Highest CEA levels are usually associated with patients with liver metastases. Thus, South tumors may be more aggressive than North tumors. This observation could be linked to the fact that in South tumors, we observed a bigger tumor diameter, more frequent SBR grade III and greater positive node invasion. To corroborate this hypothesize, it may be useful to complete clinical data with a 5-year follow-up which may inform about metastasis development.

In conclusion, this study is the first report which compares transcriptomic and clinical data between French and South Mediterranean breast cancer patients. This work allowed specific clinical parameters in Mediterranean countries to be

identified such as precocity in the age at breast cancer development and tumor aggressiveness. Moreover, microarray data permitted various genes to be distinguished which could be potential biomarkers for breast cancer in the South countries. Nevertheless, the major differences observed between these two populations were highly correlated with social factors such as breastfeeding and high number of pregnancies in the South population.

Acknowledgements

The "Breast Med Consortium" is a European project linked by a contract n° ICA3-CT-2002-10005. We thank T. Jalabert and J. Fayolle from the "Departement d'Oncogénétique" of Centre Jean Perrin for technical assistance. This work was supported by grants from 'La Ligue Nationale Française de Lutte Contre le Cancer' and 'Les Comité du Puy-de-Dôme et du Cantal'.

References

1 Parkin DM: International variation. Oncogene *23(38)*: 6329-6340, 2004.
2 Marshall E: The politics of breast cancer. Science *259(5095)*: 616-617, 1993.
3 Hulka BS and Stark AT: Breast cancer: cause and prevention. Lancet *346(8979)*: 883-887, 1995.
4 Hedenfalk I, Duggan D, Chen Y, Radmacher M, Bittner M, Simon R, Meltzer P, Gusterson B, Esteller M, Kallioniemi OP, Wilfond B, Borg A, Trent J, Raffeld M, Yakhini Z, Ben-Dor A, Dougherty E, Kononen J, Bubendorf L, Fehrle W, Pittaluga S, Gruvberger S, Loman N, Johannsson O, Olsson H and Sauter G: Gene-expression profiles in hereditary breast cancer. N Engl J Med *344*: 539-548, 2001.
5 Wettenhall JM and Smyth GK: LimmaGUI: a graphical user interface for linear modeling of microarray data. Bioinformatics *20*: 3705-3706, 2004.
6 Edwards D: Non-linear normalization and background correction in one-channel cDNA microarray studies. Bioinformatics *19*: 825-833, 2003.
7 Kwiatkowski F, Girard M, Hacene K and Berlie J: Sem: a suitable statistical software adaptated for research in oncology. Bull Cancer *87*: 715-721, 2000.
8 Glehen O, Kwiatkowski F, Sugarbaker PH, Elias D, Levine EA, De Simone M, Barone R, Yonemura Y, Cavaliere F, Quenet F, Gutman M, Tentes AA, Lorimier G, Bernard JL, Bereder JM, Porcheron J, Gomez-Portilla A, Shen P, Deraco M and Rat P: Cytoreductive surgery combined with perioperative intraperitoneal chemotherapy for the management of peritoneal carcinomatosis from colorectal cancer: a multi-institutional study. J Clin Oncol. *22*: 3284-2392, 2004.
9 Ferlay J, Bray F, Pisani P and Parkin D M: Cancer Incidence, Mortality and Prevalence Worldwide GLOBOCAN 2002. IARC CancerBase No. 5, version 2.0. Lyon, IARCPress: 2004.
10 Clavel-Chapelon F: Cumulative number of menstrual cycles and breast cancer risk: results from the E3N cohort study of French women. Cancer Causes Control *13*: 831-838, 2002.
11 Russo J, Hu YF, Silva I D and Russo IH: Cancer risk related to mammary gland structure and development. Microsc Res Tech *52*: 204-223, 2001.
12 Mueller MD, Vigne JL, Minchenko A, Lebovic DI, Leitman DC and Taylor RN: Regulation of vascular endothelial growth factor (VEGF) gene transcription by estrogen receptors alpha and beta. Proc Natl Acad Sci USA *97*: 10972-10977, 2000.
13 Kolonin MG, Saha PK, Chan L, Pasqualini R and Arap W: Reversal of obesity by targeted ablation of adipose tissue. Nat Med *10*: 625-632, 2004.
14 Jernström H, Lubinski J, Lynch HT, Ghadirian P, Neuhausen S, Isaacs C, Weber BL, Horsman D, Rosen B, Foulkes WD, Friedman E, Gershoni-Baruch R, Ainsworth P, Daly M, Garber J, Olsson H, Sun P and Narod SA: Breast-feeding and the risk of breast cancer in *BRCA1* and *BRCA2* mutation carriers. J Natl Cancer Inst *96*: 1094-1098, 2004.
15 Yang CP, Weiss NS, Band PR, Gallagher RP, White E and Daling JR: History of lactation and breast cancer risk. Am J Epidemiol *138*: 1050-1056, 1993.
16 Fan S, Wang J, Yuan R, Ma Y, Meng Q, Erdos MR, Pestell RG, Yuan F, Auborn KJ, Goldberg ID and Rosen EM: BRCA1 inhibition of estrogen receptor signaling in transfected cells. Science *284*: 1354-1356, 1999.
17 Harris JR, Lippman ME, Veronesi U and Willett W: Breast cancer (1). N Engl J Med *327*: 319-328, 1992.
18 Moysich KB, Ambrosone CB, Vena JE, Shields PG, Mendola P, Kostyniak P, Greizerstein H, Graham S, Marshall JR, Schisterman EF and Freudenheim JL: Environmental organochlorine exposure and postmenopausal breast cancer risk. Cancer Epidemiol Biomarkers Prev *7*: 181-188, 1998.
19 Bosher J M, Williams T and Hurst HC: The developmentally regulated transcription factor AP-2 is involved in c-erbB-2 overexpression in human mammary carcinoma. Proc Natl Acad Sci USA *92*: 744-747, 1995.
20 Sørlie T, Wang Y, Xiao C, Johnsen H, Naume B, Samaha RR and Børresen-Dale AL: BMC Genomics *7*: 127-142, 2006.
21 Millikan RC, Newman B, Tse CK, Moorman PG, Conway K, Smith LV, Labbok MH, Geradts J, Bensen JT, Jackson S, Nyante S, Livasy C, Carey L, Earp HS and Perou CM: Epidemiology of basal-like breast cancer. Breast Cancer Res Treat *109*: 123-139, 2008.
22 Tomasetto C, Rio MC, Gautier C, Wolf C, Hareuveni M, Chambon P and Lathe R: hSP, the domain-duplicated homolog of pS2 protein, is co-expressed with pS2 in stomach but not in breast carcinoma. Embo J *9*: 407-414, 1990.
23 Roberts M, Wallace J, Jeltsch J M and Berry M: The 5' flanking region of the human *pS2* gene mediates its transcriptional activation by estrogen in MCF-7 cells. Biochem Biophys Res Commun *151*: 306-313, 1988.
24 Lee SW, Tomasetto C, Paul D, Keyomarsi K and Sager R: Transcriptional downregulation of gap-junction proteins blocks junctional communication in human mammary tumor cell lines. J Cell Biol *118*: 1213-1221, 1992.
25 Yu D, Jing T, Liu B, Yao J, Tan M, McDonnell TJ, Hung MC: Overexpression of ErbB2 blocks Taxol-induced apoptosis by upregulation of $p21^{Cip1}$, which inhibits $p34^{Cdc2}$ kinase. Mol Cell *2*: 581-591, 1998.

Received May 6, 2008
Revised July 15, 2008
Accepted August 6, 2008