Gene Expression Profiling of Epithelial–Mesenchymal Transition in Primary Breast Cancer Cell Culture

LUIGI MINAFRA[1*], VALENTINA BRAVATÀ[1*], GIUSI IRMA FORTE[1], FRANCESCO PAOLO CAMMARATA[1],
MARIA CARLA GILARDI[1,2,3] and CRISTINA MESSA[1,2,4]

[1]*Institute for Bioimaging and Molecular Physiology, National Research Council, -LATO, Cefalù, Italy;*
[2]*Nuclear Medicine, San Raffaele Scientific Institute, Milan, Italy;*
[3]*Department of Health Sciences, Tecnomed Foundation, University of Milano-Bicocca, Milan, Italy;*
[4]*Nuclear Medicine Center, San Gerardo Hospital, Monza, Italy*

Abstract. *Background/Aim: Epithelial–mesenchymal transition (EMT) is a process co-opted by cancer cells to invade and form metastases. In the present study we analyzed gene expression profiles of primary breast cancer cells in culture in order to highlight genes related to EMT. Materials and Methods: Microarray expression analysis of primary cells isolated from a specimen of a patient with an infiltrating ductal carcinoma of the breast was performed. Real-Time Quantitative Reverse Transcription PCR (qRT-PCR) analyses validated microarray gene expression trends. Results: Thirty-six candidate genes were selected and used to generate a molecular network displaying the tight relationship among them. The most significant Gene Ontology biological processes characterizing this network were involved in cell migration and motility. Conclusion: Our data revealed the involvement of new genes which displayed tight relationships among them, suggesting a molecular network in which they could contribute to control of EMT in breast cancer. This study may offer a basis for understanding complex mechanisms which regulate breast cancer progression and for designing individualized anticancer therapies.*

Breast cancer (BC) represents a highly complex group of tumors with distinct biological behaviour, prognosis and response to therapy (1-4). The application of gene expression profiling technology has provided molecular classification of BC into clinically relevant subtypes, allowing for new

*These Authors contributed equally to this work.

Correspondence to: Dr. Luigi Minafra and Dr. Valentina Bravatà, Contrada Pietrapollastra-Pisciotto, 90015-Cefalù, PA, Italy. E-mail: luigi.minafra@ibfm.cnr.it, valentina.bravata@ibfm.cnr.it

approaches to predict disease recurrence and response to different treatments, and the likelihood of metastatic progression, and new insights into related signaling pathways (5, 6). Indeed, microarray-based expression studies have showed that BC is an heterogeneous disease at both the molecular and clinical level, which comprises of at least five subtypes with distinct gene expression patterns, associated with different prognoses and outcomes, namely luminal A, luminal B, HER2-overexpressing, basal-like, and normal-like types (7-9). However, histological and molecular parameters currently identified are still insufficient to give a reliable prediction of the clinical outcome for individual patients. Prediction of the metastatic potential of cancer cells is a critical point in the treatment of BC, as well as for other types of tumors. Nowadays, despite progress in our understanding of cellular processes which control the metastatic process (10-13), further characterization of the molecular mechanisms that regulate BC progression and invasion is required in order to improve recovery from BC.

The metastatic program of a carcinoma starts when the neoplastic epithelial cells begin to detach from the tissue of origin, lose their stationary phenotype and acquire motile capability with a mesenchymal phenotype. This process, known as epithelial-to-mesenchymal transition (EMT) which was recognized in the past in morphogenetic steps during development, has more recently been described as a strategy for neoplastic cells, including BC cells, to induce invasion and metastasis (14-18). Findings that reveal the appearance of EMT *in vivo* are sometimes controversial, probably due to histological preparation of *ex vivo* tissues that are not sufficiently adequate to reveal the phenotypic and molecular signs of EMT. To overcome these limitations, primary cell cultivation from individual BC tissues is a relevant procedure for the *in vitro* study of an ample variety of neoplastic cell features, including metastatic traits. Recently, we described an analytical investigation of a neoplastic primary cell population isolated from an infiltrating ductal carcinoma (IDC) of the breast which

revealed phenotypic traits and expression of several mesenchymal biomarkers, such as vimentin and alpha-smooth muscle actin (α-SMA), typical of EMT (19). Real-Time Quantitative Reverse Transcription PCR (qRT-PCR) analysis showed a high level of expression of multiple genes associated with EMT, particularly TGF-β, the major inducer of the EMT in BC (20) and some of the most representative targets of the TGF-β signal transduction pathway, supporting the occurrence of the process in cells of our study (19). Although some reports describe microarray analysis investigating EMT gene signatures *in vitro* by using immortalized BC cell lines (21-23), to our knowledge only one other study has examined EMT transition in BC primary culture by microarray analysis (24). For this reason, this field needs to be further explored. Thus, the purpose of this study was to analyze gene expression profile by cDNA microarray of BC primary culture in order to highlight genes related to EMT in BC.

Materials and Methods

Tissue and cell cultures. An extra portion of breast IDC tissue, unnecessary for histological diagnosis after surgery, was obtained for this study. The patient gave her written informed consent for research participation according to the Helsinki Declaration and the study was approved by the Ethical Committee of the San Raffaele G. Giglio Hospital, Cefalù (number of protocol: C.E.2012/16). The tissue was disaggregated mechanically immediately after surgery, subjected to enzymatic digestion with collagenase/hyaluronidase and the BC primary cells, hereafter named simply BC cells, were isolated by applying differential centrifugation method as previously described (19). The BC cells were grown at 37°C under 5% CO_2-95% air atmosphere in a humidified incubator in a selective medium of DMEM with antibiotics supplemented with 10 mM Hepes (Invitrogen, Carlsbad, CA, USA), 1 mg/ml bovine serum albumin (BSA), 10 ng/ml cholera toxin, 0.5 μg/ml hydrocortisone, 5 μg/ml insulin, 5 ng/ml epidermal growth factor (EGF), 5 μg/ml apo-transferrin (all from Sigma-Aldrich, St. Louis, MO, USA). All gene expression experiments were conducted when BC cells were at the fourth growth passage. The non-tumor MCF-10A cell line was used as mammary epithelial reference sample in line with others *in vitro* BC studies (30-31) and cultured according to instructions (American Type Culture Collection, Manassas, VA, USA). The follow-up of the patient was not part of this study.

Whole-genome cDNA microarray expression analysis. Total RNA was extracted from BC and MCF-10A cells using Trizol and the RNeasy mini kit (Invitrogen) according to the manufacturer's guidelines. RNA concentration and purity were determined spectrophotometrically using the Nanodrop ND-1000 (Thermo Scientific Open Biosystems, Lafayette, CO, USA) and RNA integrity, measured as RNA integrity number (RIN) values, was assessed using a Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA). Only samples with a maximum RIN of 10 were used for further microarray analysis. Five hundred nanogrammes of total RNA from BC primary culture and MCF-10A cell line were used for cRNA synthesis and labelling according to Agilent Two-Color Microarray-Based Gene Expression Analysis protocol. Samples were labelled with Cy5 and Cy3 dye (Agilent Technologies). Fluorescent complementary cRNA samples

(825 ng) were hybridized onto Whole Human Genome 4×44K microarray (Agilent Technologies) GeneChips containing all known genes and transcripts of an entire human genome. Four replicates were performed. Array hybridization was conducted for 17 h at 65°C. Images were achieved by an Agilent's DNA Microarray Scanner with Sure Scan high-Resolution Technology (Agilent Technologies) and analyzed using Feature Extraction expression software (Agilent Technologies, Santa Clara, CA, USA) that found and placed microarray grids, rejected outlier pixels, accurately determined feature intensities and ratios, flagged outlier features, and calculated statistical confidences.

Statistical data analysis, background correction, normalization and summary of expression measure were conducted with GeneSpring GX 10.0.2 software (Agilent Technologies). Data were filtered using two-step procedure: first the entities were filtered based on their flag values P (present) and M (marginal) and then filtered based on their signal intensity values, this enables very low signal values or those that have reached saturation to be removed. Statistically significant differences were computed by the Student's *t*-test and the significance level was set at $p<0.05$. The false discovery rate (FDR) was applied as a multiple test correction method. Average gene expression values between experimental groups were compared (on log scale) by means of a modified ANOVA ($p<0.05$). Genes were identified as being differentially expressed compared to MCF-10A cells if they had a fold-change (FC) in expression of at least 3 and a *p*-value <0.05.

The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (25) and are accessible through GEO Series accession number GSE53175 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE53175).

Microarray data are available in compliance with Minimum Information About a Microarray Experiment (MIAME) standards.

MetaCore network analyses. Microarray data sets were also analyzed by pathway analysis using the network building tool MetaCore (Thomson Reuters, Philadelphia, PA, USA) consisting of millions of relationships between proteins derived from literature publications about proteins and small molecules (including direct protein interaction, transcriptional regulation, binding, enzyme-substrate, and other structural or functional relationships). Results, *i.e.* maps of protein lists from the uploaded dataset, were then compared with all the possible pathway maps for all the proteins in the database, and the p-value was calculated based on hypergeometric distribution probability test. The most representative networks significantly changed were selected and analysed.

qRT-PCR analysis. Candidate genes for qRT-PCR analysis were chosen based on microarray results. One microgramme of total RNA was reverse-transcribed into cDNA with SuperScriptII reverse transcriptase according to the manufacturer's specifications (Invitrogen). One microlitre of cDNA (50 ng RNA equivalent) was analyzed by real-time PCR (1 cycle 95°C for 20 s and 40 cycles of 95°C for 3 s, 60°C for 30 s) in triplicate using Fast 7500 Real-Time PCR System (Applied Biosystems). Amplification reactions were performed in a 20 μl reaction volume containing 10 pmoles of each primer and the Fast SYBR Green Master Mix according to the manufacturer's specifications (Applied Biosystems). Reaction specificity was controlled by post-amplification melting curve analysis. The oligonucleotide primers were selected with the Primer3 software (26-27) and tested for human specificity using NCBI database. Primer sequences (forward and reverse) used are listed in
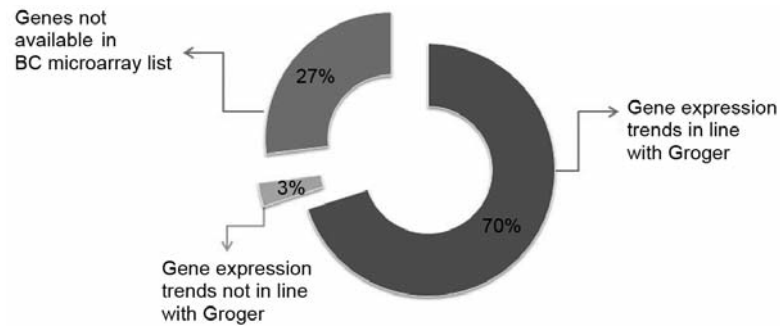
Figure 1. *Expression of genes in primary breast cancer cells compared to the Epithelial–Mesenchymal Transition core gene list of Gröger et al. (29).*

Table I. Quantitative data, normalized *versus* rRNA 18S gene, were analyzed by average of triplicates Ct (cycle threshold) according to the $2^{-\Delta\Delta ct}$ method using SDS software (Applied Biosystems). The data shown were generated from three independent experiments and the values are expressed relative to mRNA levels in the non-tumuorigenic MCF-10A control cells as the mean±SD.

STRING interaction analyses. In order to analyse potential interactions among selected genes, STRING database of known and predicted protein interactions was used (http://string-db.org). Direct (physical) and indirect (functional) associations derived from four sources: genomic context, high-throughput experiments, conserved coexpression, and previous bibliographical knowledge, were analysed as previously described (3).

Results

Overview of cDNA microarray gene expression. In the present study, a two-color microarray-based gene expression analysis (Agilent Technologies) was conducted on BC primary culture and on MCF-10A cells used as a reference sample. Comparative differential gene expression analysis revealed that 2,982 genes in BC cells had expression levels significantly altered by 3-fold or greater compared to the MCF-10A reference group: 1,553 genes were down-regulated and 1,429 were up-regulated (GSE53175). Altered transcripts were selected and grouped according to their involvement in specific biological pathways using integrated pathway enrichment analysis with GeneGo MetaCore. Datasets were loaded into the MetaCore software and the top enriched canonical metabolic pathways were analyzed. The result of this mapping revealed the involvement of a set of molecules controlling specific process networks such as cytoskeleton remodelling, cell-cycle regulation and cell adhesion in BC primary culture in comparison with the reference sample, processes known to have a role in EMT. In addition, lists of differentially expressed genes were analyzed in order to search for factors potentially involved in EMT or in BC, and new ones to propose as new markers of interest. Thirty-six candidate genes were selected, validated and analyzed using

PubMatrix tool (28). In this way, lists of terms such as gene names can be assigned to a genetic, biological, or clinical relevance in a rapid flexible systematic fashion in order to confirm our assumptions.

EMT core list validation. Our microarray data were compared with the 'EMT core gene list' recently proposed by Gröger *et al*. (29). Based on cluster analysis of the gene expression studies (GES) derived from a multitude of cell types and EMT initiation methods, Gröger and colleagues conducted a meta-analysis of gene expression signature defining the EMT during cancer progression. They proposed a meaningful EMT core gene list which describes the majority of the involved genes across the analyzed GES. Cluster analysis of genes shared among at least 10 datasets containing 130 genes, out of which 67 were up- and 63 were down-regulated, were further classified into five categories: (i) cell adhesion and migration, (ii) development, cell differentiation and proliferation, (iii) angiogenesis and wound healing, (iv) metabolism, (v) other or unclassified, according to single enrichment analysis. As shown in Table II and in Figure 1, 91 (~70%) out of these 130 genes were found to be de-regulated in BC primary culture, with the same trend proposed by Gröger *et al*. (44 down-regulated genes and 47 up-regulated). Thirty five (~27%) genes were not included in our BC microarray dataset, while four genes (~3%) showed an opposite expression trend with respect to Gröger *et al*.'s list: Fibulin 1 (FBLN1), Latent Transforming growth factor Beta binding Protein 1 (LTBP1) and Tissue Factor Pathway Inhibitor (TFPI) were down-regulated, while Follistatin (FST) was up-regulated in BC cells.

Microarray validation experiments. Firstly, we assayed mRNA levels of the four genes showing an opposite gene expression trend with respect to Gröger's list: *FBLN1*, *LTBP*, *TFPI* and *FST* genes. qRT-PCR experiments confirmed the microarray expression trend (Tables III and IV).

Moreover, based on the microarray data set, Pubmatrix results and MetaCore analyses, we chose 32 candidate genes

Table I. *Primer sequences used for qRT-PCR analyses*

Gene symbol	Full name	Forward primer 5'>3'	Reverse primer 5'>3'	Template size (base pairs)
ACTA2	Homo sapiens actin, alpha 2, smooth muscle, aorta	gctgttttcccatccattgtg	ttggtgatgatgccatgttct	148
AREG	Homo sapiens amphiregulin	tctccactcgctcttccaac	attgaggtccaatccagcagc	190
CD24	Homo sapiens CD24 signal transducer mRNA	acatgggaacaaacagatcg	tgttcaatgcaaatcaattcc	165
CDH1	Homo sapiens cadherin 1, type 1, E-cadherin (epithelial)	ccaagtgcctgcttttgatga	cccctacccctcaactaac	166
CDH2	Homo sapiens cadherin 2, type 1, N-cadherin (neuronal)	gacaatgcccctcaagtgtt	ccattaagccgagtgatggt	119
COL6A3	Homo sapiens collagen, type VI, alpha 3	caggtttgctcaggggttcata	actggaccagagcaaaatggaaa	287
COL8A1	Homo sapiens collagen, type VIII, alpha 1	gccacctcaaattcctcctca	cttgttcccctcgtaaactgg	252
CTGF	Homo sapiens connective tissue growth factor	catcttcggtggtacggtgt	gccaaacgtgtcttccagtc	259
CTSK	Homo sapiens cathepsin K	cccgcagtaatgacaccctttt	aaggcattggtcatgtagccc	258
ESRP2	Homo sapiens epithelial splicing regulatory protein 2	ccctacatgctctgcactgat	aacaattgactgctgggctctt	292
FAP	Homo sapiens fibroblast activation protein, alpha	tacccaaaggctggagctaag	ctcctgggtctttggacaatc	257
FBLN1	Homo sapiens fibulin 1	cgggcatacatgcatcaacac	gaacagctgcagaggtagga	318
FGF2	Homo sapiens fibroblast growth factor 2 (basic)	ctgtgatggaatggtcaggaat	ttccatttgttccgactgagtc	288
FGFBP1	Homo sapiens fibroblast growth factor binding protein 1	gtgctcagaacaaggtgaacg	ttgagagagatgccctcctc	276
FN1	Homo sapiens fibronectin 1	ggaaagcatatgcagccaac	ctacagtattgcgggccaga	121
FST	Homo sapiens follistatin	cccaactgcatcccctgtaaa	gagctgcctggacagaaaaca	275
ID2	Homo sapiens inhibitor of DNA binding 2, dominant negative helix-loop-helix protein	ccgtgaggtccgttaggaaaa	ggtgatgcaggctgacaatag	254
IL6	Homo sapiens interleukin 6 (interferon, beta 2)	gaggagacttgcctggtgaaa	aaagctgcgcagaatgagatga	272
KRT18	Homo sapiens keratin 18	ccagtctgtggagaacgaca	atctgggcttgtaggcctttt	159
KRT19	Homo sapiens keratin 19	tcgagggacaggaagatcac	attggcaggtcaggagaaga	158
KRT5	Homo sapiens keratin 5	tttgtctccaccacctcctc	cctgggaaccaaagaatgtgt	195
KRT6A	Homo sapiens keratin 6A	caagctcaccttccaggact	gagtgtgagaggctggagga	114
KRT8	Homo sapiens keratin 8	gggaagctggtgtctgagtc	ctcctgttcccagtgctacc	135
LOXL2	Homo sapiens lysyl oxidase-like 2	ctcctgctgcttctgtacc	ctcgctgtgcttcctcttct	286
LTBP1	Homo sapiens latent transforming growth factor beta binding protein 1	tcgtaccaagggcttcctgt	cccggtaggcattccatcatt	280
MMP3	Homo sapiens matrix metallopeptidase 3 (stromelysin 1, progelatinase)	cgatgcagccatttctgataag	aagtgcccatattgtgccttct	302
MXRA5	Homo sapiens matrix-remodelling associated 5	ttcagggtgtctgtgctctg	gggcaagctggagttttgaag	252
PDGFRB	Homo sapiens platelet-derived growth factor receptor, beta polypeptide	gggtaccccaagaaggatgt	gtaaggtgccaacctgcaatg	277
POSTN	Homo sapiens periostin, osteoblast specific factor	gacacacccgtgaggaagtt	accacaggaggctaactcca	281
SNAI1	Homo sapiens Snail family zinc finger 1	gcgagctgcaggactctaat	ggacagagtcccagatgagc	135
TFPI	Homo sapiens tissue factor pathway inhibitor (lipoprotein-associated coagulation inhibitor)	gcagtctggcatctcaacaac	gcctagggattgggtgaact	343
TGFB1	Homo sapiens transforming growth factor, beta 1	ccctggcacccaactattgc	aggcagaagttggcatggta	133
THBS1	Homo sapiens thrombospondin 1	gtggaagagcatcaccctgt	ggtgaggaggacactggta	269
TP63	Homo sapiens tumor protein p63	ttgcccctcctagtcatttg	ctggggtggctcataaggta	110
TWIST	Homo sapiens twist homolog 1 (Drosophila)	gtccgcagtcttacgaggag	ccagcttgagggtctgaatc	159
VIM	Homo sapiens vimentin	aacaaccgacactcctacaaga	tggttggatacttgctggaaa	142
WNT5A	Homo sapiens wingless-type MMTV integration site family, member 5A	gggaggaaagtccccagaaat	tgaatggcacgcaattaccttgt	275

and performed qRT-PCR validation experiments (Table III and IV). Genes for validation were chosen based on two considerations: i) factors known to be involved in EMT transition or in BC; ii) lesser known genes involved in these processes to propose as new molecular markers. qRT-PCR analyses were performed in parallel with MCF-10A cells. Genes selected were some of the most representative targets of TGF-β signal transduction pathway, as well as cytoskeleton

and adhesion proteins, and downstream targets of this pathway (32, 33). qRT-PCR analysis revealed an increase of mRNA levels of genes known to be involved in mesenchymal progression, EMT transition and in matrix modelling such as Transforming Growth Factor beta 1 (*TGFB1*), Twist homolog 1 (Drosophila) (*TWIST1*), Vimentin (*VIM*), Actin, alpha 2, smooth muscle, aorta (*ACTA2*), Cadherin 2, type 1, N-cadherin (*CDH2*), Wingless-type MMTV integration site family,

Table II. *Microarray data for gene expression of cultured primary breast cancer cells compared with to Groger's EMT core gene list.*

	Down-regulated genes in accordance with Groger *et al.* (29) (n=44)	Up-regulated genes in accordance with Groger *et al.* (29) (n=47)	Unchanged (n=35)	Discordant (n=4)
Cell adhesion and migration	CD24, CDH1, CXADR, DSG3, EPHA, JUP, MPZL2, PLXNB1, S100P, SLC7A5, SYK	ADAM12. CDH2. COL1A1, COL3A1, COL5A1, COL6A1, COL6A3, CTGF, CYP1B1, DLC1, FBLN5, FGF2, FN1, MMP2, MYL9, NR2F1, NRP1, PPAP2B, PRKCA, RECK, SERPINE1, SERPINE2, SPOCK1, TPM1, VCAN, WNT5A	CDH11, CXCL16, ELF3, EPCAM, FGFR1, HAS2, LUM, NID2, PLAT, OVOL2, TGM2, TNFAIP6	FBLN1
Development, cell differentiation and proliferation	ABLIM1, ADRB2, ALDH1A3, ANK3, BIK, CTSL2, FGFR3, GJB3, IFI30, IL18, KLK7, KRT15, KRT17, LSR, MBP, PKP2, PPL, PRSS8, SPINT1	CDKN2C, EMP3, FBN1, IGFBP3, IL1R1, MME, PMP22, PTX3, SYNE1, TAGLN, VIM	CA2, FGFR2, MAP7, OCLN, PTGER2, RAPGEF5, SRGN, SULF1, TUBA1A, ZEB1	FST, LTBP1
Angiogenesis and wound healing	–	DCN, LOX	–	TFPI
Metabolism	SORL1, ST6GALNAC2	GALNT10	ABCA1, GPX3, SLC22A4, SLC27A2, SMPDL3B	–
Others or unclassified	CDS1, FAM169A, FXYD3, KLK10, LAD1, MTUS1, RHOD, SERPINB1, SLPI, TPD52L1, TSPAN1, ZNF165	C5orf13, CDK14, EML1, FSTL1, LTBP2, RGS4, TMEM158	AGR2, C10orf10, MAP1B, PLS1, PRRG4, SYT11, TMEM30B, ZHX2	–

member 5A (*WNT5A*), Matrix MetalloPeptidase 3 (MMP3), Lysyl Oxidase-Like 2 (*LOXL2*), Matrix-remodelling Associated 5 (*MXRA5*), Collagen type VIII alpha 1 (*COL8A1*), Collagen type VI alpha 3 (*COL6A3*), Fibronectin 1 (*FN1*) and connective tissue growth factor (*CTGF*). Other less well-known genes involved in these processes were also up-regulated: Platelet-derived Growth Factor Receptor, beta polypeptide (*PDGFRB*), Thrombospondin 1 (*THBS1*), Fibroblast Activation Protein, alpha (*FAP*), interleukin-6 (IL6), Keratin-19 (*KRT19*), Periostin (*POSTN*), Fibroblast Growth Factor 2 (*FGF2*), Inhibitor of DNA binding 2 (ID2) and cathepsin K (*CTSK*) (Table III).

On the other hand, down-regulation was observed for some epithelial genes and factors including also several keratins, such as E-cadherin (*CDH1*), Amphiregulin (*AREG*), Epithelial Splicing Regulatory Protein 2 (*ESRP2*), Keratin-6A (*KRT6A*), Keratin-8 (*KRT8*), Keratin-5 (*KRT5*) and Keratin-18 (*KRT18*). In addition, down-regulation of Fibroblast Growth Factor Binding Protein 1 (*FGFBP1*), Tumor Protein p63 (*TP63*) and CD24 signal transducer (*CD24*) mRNA levels was observed (Table IV). All qRT-PCR validation experiments confirmed Microarray gene expression trends as shown in Table III and IV.

Thirty-six candidate gene interactions. In order to verify potential interactions among markers identified, we generated a molecular network by introducing these 36 genes into STRING, a molecular tool able to elaborate physical and functional associations among proteins (34). Despite data heterogeneity, Figure 2 clearly displays the tight relationships existing among molecules, showing that 34 out of 36 genes are closely connected in a single network. Moreover, the STRING analysis reveals that the most significant GO biological processes characterizing this network (Figure 3) are involved in cell migration and motility, in agreement with behaviour featured in tumor cells wring EMT.

Discussion

EMT is a recognized process involved in the morphogenetic steps during development, but which is also co-opted by various cancer cells to enable them to invade and form metastases at distant sites.

Over the past decade, a number of GES regarding EMT were conducted on a variety of cell types and different modes of EMT induction. Recently, Gröger and colleagues, using a meta-analysis of GES performed on different cell types and treatment modalities, defined a transcriptome signature named EMT core gene list, containing 130 well-known genes involved in EMT, as well as new genes selected among the multitude of GES so far described (29). All BC samples used in those studies were immortalized BC cell lines, which are considered well-established procedures in biomedicine for a more complete understanding of cellular cancer process.

A further contribution to these studies is offered by the primary cell cultures which more tightly maintain tumor

Table III. *Up-regulated genes in breast cancer primary culture.*

Gene symbol	Entrez Gene ID	Description	Microarray Fold Change	qRT-PCR Fold Change
ACTA2	59	Homo sapiens actin, alpha 2, smooth muscle, aorta (ACTA2)	15.91	28*
CDH2	1000	Homo sapiens cadherin 2, type 1, N-cadherin (neuronal) (CDH2)	26.70	8.2*
COL6A3	1293	Homo sapiens collagen, type VI, alpha 3 (COL6A3)	80.06	>45
COL8A1	1295	collagen, type VIII, alpha 1	102.02	33
CTGF	1490	Homo sapiens connective tissue growth factor (CTGF)	79.57	42
CTSK	1513	Homo sapiens cathepsin K (CTSK)	113.77	7.4
FAP	2191	Homo sapiens fibroblast activation protein, alpha (FAP)	86.99	5.4
FGF2	2247	Homo sapiens fibroblast growth factor 2 (basic) (FGF2)	43.33	42
FN1	2335	Homo sapiens fibronectin 1 (FN1)	35.63	11*
FST	10468	Homo sapiens follistatin (FST)	42.98	7.9
ID2	3398	Homo sapiens inhibitor of DNA binding 2, dominant negative helix-loop-helix protein (ID2)	16.42	44
IL6	3569	Homo sapiens interleukin 6 (interferon, beta 2) (IL6)	89.42	9.2
KRT19	3880	Homo sapiens keratin 19 (KRT19)	16.21	19.8
LOXL2	4017	Homo sapiens lysyl oxidase-like 2 (LOXL2)	24.29	11.44
MMP3	4314	Homo sapiens matrix metallopeptidase 3 (stromelysin 1, progelatinase) (MMP3)	117.57	9.4
MXRA5	25878	Homo sapiens matrix-remodelling associated 5 (MXRA5)	40.23	15.5
PDGFRB	5159	Homo sapiens platelet-derived growth factor receptor, beta polypeptide (PDGFRB)	118.88	2.2
POSTN	10631	Homo sapiens periostin, osteoblast specific factor (POSTN)	281.80	35
TGFB1	7040	Homo sapiens transforming growth factor, beta 1 (TGFB1)	3.33	6.8*
THBS1	7057	Homo sapiens thrombospondin 1 (THBS1)	347.62	2.7
TWIST	7291	Homo sapiens twist homolog 1 (Drosophila) (TWIST1)	3.03	20.9*
VIM	7431	Homo sapiens vimentin (VIM)	9.68	36.5*
WNT5A	7474	Homo sapiens wingless-type MMTV integration site family, member 5A (WNT5A)	53.08	2.9

*Minafra *et al*. (19).

Table IV. *Down-regulated genes in breast cancer primary culture*

| Gene symbol	Entrez Gene ID	Description	Microarray Fold Change	qRT-PCR Fold Change
AREG	374	Homo sapiens amphiregulin (AREG)	−100.90	0.0039
CD24	100133941	Homo sapiens CD24 signal transducer Mrna	−8.68	<0.0039
CDH1	999	Homo sapiens cadherin 1, type 1, E-cadherin (epithelial) (CDH1)	−72.34	0.0026*
ESRP2	80004	Homo sapiens epithelial splicing regulatory protein 2 (ESRP2)	−33.05	0.0071
FBLN1	2192	Homo sapiens fibulin 1 (FBLN1)	−42.27	0.032
FGFBP1	9982	Homo sapiens fibroblast growth factor binding protein 1 (FGFBP1)	−84.19	<0.0039
KRT18	3875	Homo sapiens keratin 18 (KRT18)	−4.71	0.15*
KRT5	3852	Homo sapiens keratin 5 (KRT5)	−6.03	0.0023*
KRT6A	3853	Homo sapiens keratin 6A (KRT6A)	−7.82	0.0031*
KRT8	3856	Homo sapiens keratin 8 (KRT8)	−6.84	0.0051*
LTBP1	4052	Homo sapiens latent transforming growth factor beta binding protein 1 (LTBP1)	−2.85	0.91
TFPI	7035	Homo sapiens tissue factor pathway inhibitor (lipoprotein-associated coagulation inhibitor) (TFPI)	−2.80	0.25
TP63	8626	Homo sapiens tumor protein p63 (TP63)	−208.88	0.32

*Minafra *et al*. (19).
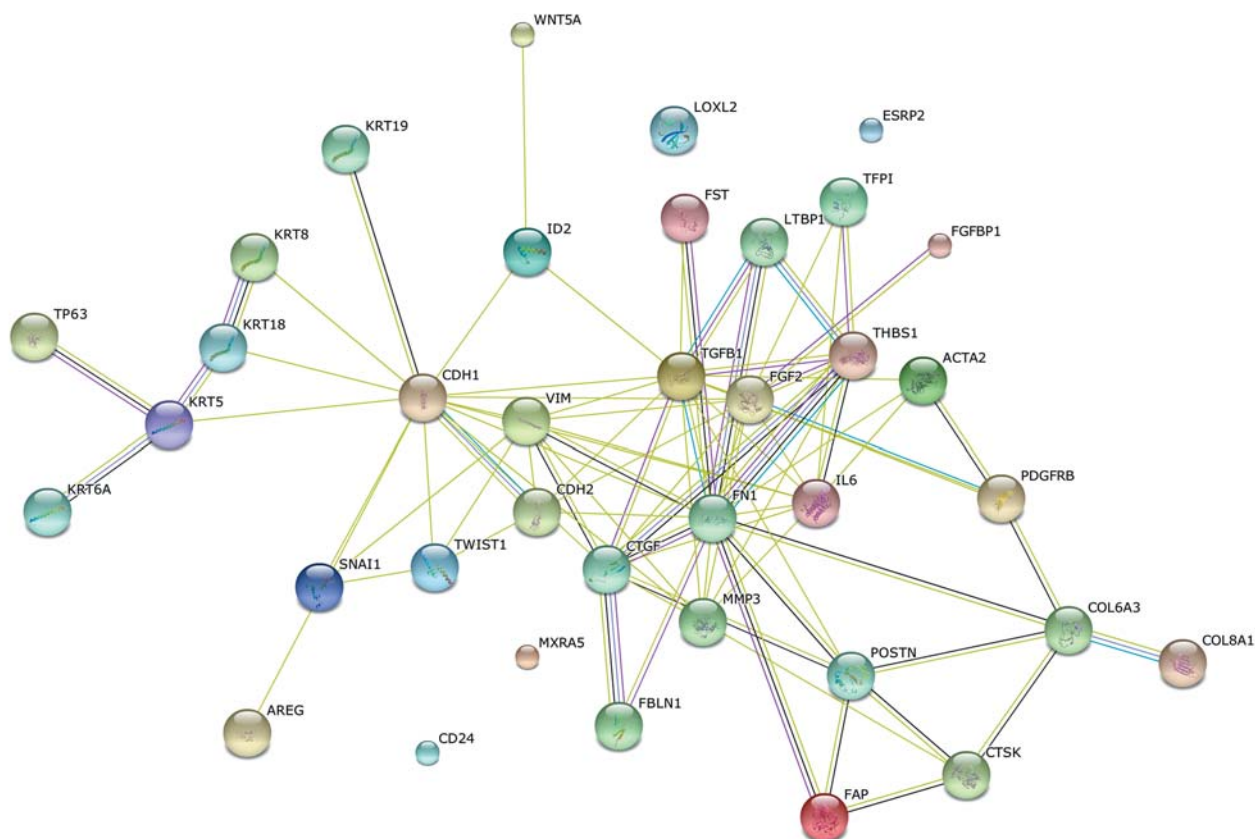
Figure 2. *STRING analysis of known and predicted protein interactions existing among the 36 selected genes closely connected in a single network.*

characteristics and behaviour *in vivo*. Indeed, their application in clinical research could represent a powerful system for studying molecular mechanisms of human disease such as BC in order to define individualized anticancer therapies in the future. Recently, we described EMT in a primary cell culture derived from a patient with breast IDC which revealed phenotypic traits and expression of several mesenchymal biomarkers. The EMT phenomenon is difficult to predict due to its dynamic evolution inside tissues. Thus, the aim of the present study was to analyze the gene expression profile by cDNA microarray of BC cells from primary culture in order to identify known and less well-known putative EMT markers of BC. Indeed, a very limited number of reports describe gene expression profile of EMT in BC cell lines (21-23, 35-36), and to our knowledge, only one other study has examined EMT transition in BC primary culture by microarray approach (24).

Several bibliographical data are available concerning the role of some of the selected genes in EMT transition such as: *ACTA2, CD24, CDH1, CDH2, CTSK, CTGF, FGF2, FN1, ID2, IL6, KRT18, KRT8, LOXL2, MMP3, TP63, POSTN, TGFB1, THBS1, TWIST, VIM* and *WNT5A*. For this reason, their functional role will be not discussed here.

Conversely, no data or few are available on *COL6A3, COL8A1, AREG, ESRP2, FAP, PDGFRB, FST, LTBP1, KRT19, KRT5, KRT6A, FBLN1, FGFBP1, MXRA5* and *TFPI* genes in EMT.

In the tumor microenvironment, both stromal and cancer cells contribute along with various types of extracellular matrix (ECM) proteins to actively remodel the microenvironment and favour tumor growth and metastasis. EMT provides a mechanism where cells adopt a motile phenotype and, regarding this process, ECM proteins are often described: for example increased collagen production is linked with malignancy. Nevertheless, no one has described the role of *COL8A1* in EMT and only recent work describes COL6A3 protein as being actively involved in mammary tumour progression through enhancing EMT, fibrosis and chemokine activity, thereby recruiting stromal cells to the tumor microenvironment (37, 38). Both *COL6A3* and *COL8A1* were up-regulated in our study. Notably, all these activities are associated with acquired drug resistance: Park *et al*. suggest that this protein is a predictor of chemoresistance to cisplatin in patients with solid tumors, including BC, being able to enhance EMT. Thus, *COL6A3*
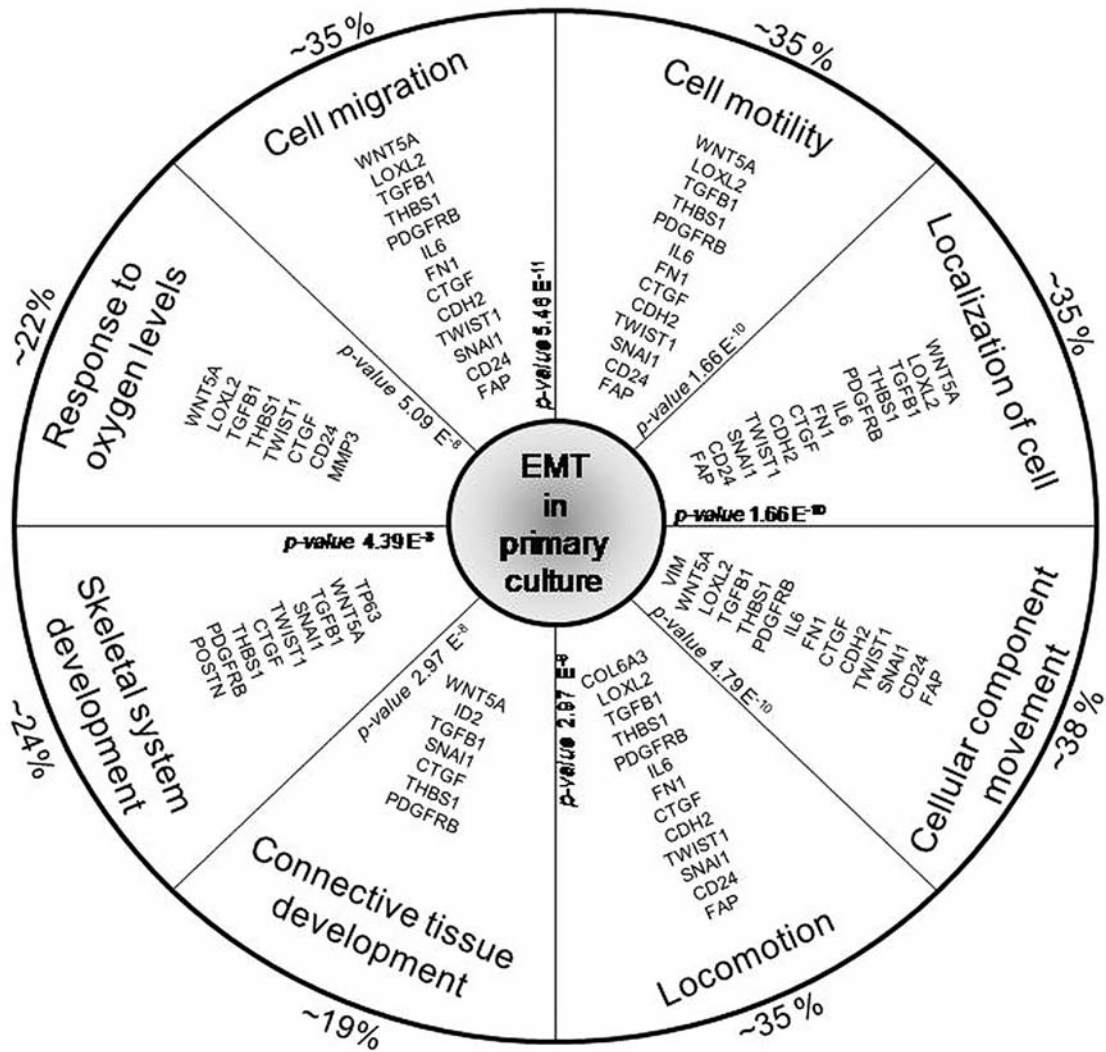
Figure 3. *Analysis of the most significant Gene Ontology biological processes highlighted in the STRING network.*

has been highlighted as a promising candidate triggering drug resistance against platinum-based therapeutics, since its levels are vastly increased in cisplatin-resistant cancer cells *in vitro*. However, the detailed molecular mechanisms underlying these correlations remain elusive (39).

AREG, or amphiregulin protein, is an EGF receptor ligand, essential for *in vitro* proliferation of human keratinocytes under autocrine and growth factor-stimulated conditions (40). AREG was recently described by Yang *et al*. to be able to mediate EGF-independent growth and malignant behaviour of mammary epithelial cells in a non-cell-autonomous manner. It is a target for a transcriptional co-activator named TAZ factor having a PDZ-binding motif, which in human mammary epithelial cells promotes EMT, but no additional data are available on this topic. Positive

correlation between TAZ and AREG proteins in a patient with BC was revealed *in vivo*. Of particular note, AREG can be secreted into the blood system, and its detection would be non-invasive and quite simple (41, 42).

To date, the epithelial splicing regulatory proteins (ESRPs) are the only known splicing factors that exhibit epithelial cell-type-specific expression and that undergo pronounced changes in expression during the EMT. A loss of ESRP2 expression is a generalized event during the EMT, and was shown by studies in which the induction of EMT by TGFB1, SNAIL, ZEB1, ZEB2, and E-cadherin knockdown was accompanied by a nearly complete abrogation of ESRP expression (22, 43, 44). Cells lacking ESRP1 and ESRP2 lost cell–cell adhesion, acquired a stellate-like shape, and became more motile. Changes in cell morphology and

behaviour through abrogation of an epithelial splicing program suggests that the ESRPs regulate transcripts that, while expressed in both epithelial and mesenchymal cells, encode protein isoforms with unique functions specific to each cell type. For these reasons Warzecha and Carstens suggest a model whereby the ESRPs are master regulators of epithelial cell function and morphology (45). Interestly, only two articles describe the relation between ESRPs and BC, showing their down-regulation after EMT (TGFB-induced) during BC progression (46-49), in line with our data.

During EMT, tumor cells interact with activated stromal fibroblasts, also known as cancer-associated fibroblasts, in order to define tumour phenotype development. Quiescent fibroblasts become activated through mutual paracrine effects with tumour cells. Some differentially expressed markers of cancer-associated fibroblasts, including fibroblast activation protein (FAP or seprase) and PDGFRB, were up regulated in BC primary culture cells. Although an influence of these markers on EMT is postulated, their functional importance for modulation of cancer cell phenotype is not fully understood (49, 50).

Currently, no data regarding EMT and *FST, LTBP1, KRT19, KRT5, KRT6A, FGFBP1* and *TFPI*, are available, but a relation between these factors and BC was described. Thus, they could represent new markers of EMT in BC.

In addition, scarce information exists for *FBLN1* or *MXRA5* genes and EMT or BC. FBLN1 is a secreted glycoprotein that becomes incorporated into a fibrillar extracellular matrix and has been indicated as a prognostic marker for near-term events (2.5 years) for both patients with lymph node-positive and those with lymph node-negative BC (51). Moreover, *MXRA5* gene, which encodes for one of the matrix-remodelling associated proteins, has been recently described as a new cancer gene, frequently mutated in non-small cell lung carcinoma (52) and by us described for the first time in EMT and BC. Further analysis should clarify their role on this matter.

Concluding, gene expression data for BC cells of primary culture are in line with those reported in literature concerning EMT in cancer cells, with special focus on BC. In the present work we describe the involvement of novel genes able to guide this process, highlighting a molecular network that could contribute to the control of EMT. EMT markers identified here should act as prognostic indicators, for example in response to drug therapies or to ionizing radiation treatments, in which this transition may play a key role not yet well-known (53, 54).

Conclusion

Our gene expression data show the involvement of novel genes which display tight relationships among them, suggesting a molecular network in which they could contribute to regulation of EMT in BC. We trust that this study may offer a further basis for understanding the complex mechanisms which regulate BC progression, such as the EMT, and for designing individualized anticancer therapies in the future.

Competing Interests

The Authors declare that they have no competing interests'.

Authors' Contributions

All Authors participated to the conception, design, interpretation, elaboration of the findings of the study, drafting and revising the final elaborate. In particular, VB and LM designed the study: LM isolated and maintained BCpc primary culture and VB performed microarray experiments and gene expression profile analyses. VB, LM, FPC and GIF performed qRT-PCR assays. Professor M.C. Gilardi and Professor C. Messa participated in the elaboration of the findings of the study, drafting and revising the final elaborate. All Authors read and approved the final content of the manuscript.

Acknowledgements

References

1   Voduc KD, Cheang MC, Tyldesley S, Gelmon K, Nielsen TO and Kennecke H: Breast cancer subtypes and the risk of local and regional relapse. J Clin Oncol *28*: 1684-1691, 2010.

2   Harbeck N, Sotlar K, Wuerstlein R and Doisneau-Sixou S: Molecular and protein markers for clinical decision making in breast cancer: Today and tomorrow. Cancer Treat Rev *40(3)*: 434-44, 2014.

3   Bravatà V, Cammarata FP, Forte GI and Minafra L: "Omics" of HER2-positive breast cancer. OMICS *17(3)*: 119-129, 2013.

4   Bravatà V, Stefano A, Cammarata FP, Minafra L, Russo G, Nicolosi S, Pulizzi S, Gelfi C, Gilardi MC and Messa C: Genotyping analysis and [18]FDG uptake in breast cancer patients: a preliminary research. J Exp Clin Cancer Res *32*: 23, 2013.

5   Sotiriou C and Pusztai L: Gene-expression signatures in breast cancer. N Engl J Med *60*: 790-800, 2009.

6   Sotiriou C and Piccart MJ: Taking gene-expression profiling to the clinic: When will molecular signatures become relevant to patient care? Nat Rev Cancer *7*: 545-553, 2007.

7   Reis-Filho JS, Westbury C and Pierga JY: The impact of expression profiling on prognostic and predictive testing in breast cancer. J Clin Pathol *59*: 225-231, 2006.

8   Wirapati P, Sotiriou C, Kunkel S, Farmer P, Pradervand S, Haibe-Kains B, Desmedt C, Ignatiadis M, Sengstag T, Schütz F, Goldstein DR, Piccart M and Delorenzi M: 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.

9   Weigelt B, Baehner FL and Reis-Filho JS: The contribution of gene expression profiling to breast cancer classification, prognostication and prediction: a retrospective of the last decade. J Pathol *220*: 263-280, 2010.

10 Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, Talantov D, Timmermans M, Meijer-van Gelder ME, Yu J, Jatkoe T, Berns EM, Atkins D and Foekens JA: Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet *365*: 671-679, 2005.

11 Smid M, Wang Y, Klijn JG, Sieuwerts AM, Zhang Y, Atkins D, Martens JW and Foekens JA: Genes associated with breast cancer metastatic to bone. J Clin Oncol *24*: 2261-2267, 2006.

12 Smid M, Wang Y, Zhang Y, Sieuwerts AM, Yu J, Klijn JG, Foekens JA and Martens JW: Subtypes of breast cancer show preferential site of relapse. Cancer Res *68*: 3108-3114, 2008.

13 Kennecke H, Yerushalmi R, Woods R, Cheang MC, Voduc D, Speers CH, Nielsen TO and Gelmon K: Metastatic behavior of breast cancer subtypes. J Clin Oncol *28*: 3271-3277, 2010.

14 Baum B, Settleman J and Quinlan MP: Transitions between epithelial and mesenchymal states in development and disease. Semin Cell Dev Biol *19*: 294-308, 2008.

15 Hugo H, Ackland ML, Blick T, Lawrence MG, Clements JA, Williams ED and Thompson EW: Epithelial-Mesenchymal and Mesenchymal-Epithelial Transitions in Carcinoma Progression. J Cell Physiol *213*: 374-383, 2007.

16 Yang J and Weinberg RA: Epithelial–mesenchymal transition: at the crossroads of development and tumour metastasis. Dev Cell *14*: 818-829, 2008.

17 Tanos B and Rodriguez-Boulan E: The epithelial polarity program: machineries involved and their hijacking by cancer. Oncogene *27(55)*: 6939-6957, 2008.

18 Tomaskovic-Crook E, Thompson EW and Thiery JP: Epithelial to mesenchymal transition and breast cancer. Breast Cancer Res *11(6)*: 213, 2009.

19 Minafra L, Norata R, Bravatà V, Viola M, Lupo C, Gelfi C, Messa C: Unmasking epithelial–mesenchymal transition in a breast cancer primary culture: A study report. BMC Res Notes *5*: 343, 2012.

20 Wendt MK, Jason A, Smith BS and Schiemann WP: Transforming growth factor-β-induced epithelial–mesenchymal transition facilitates epidermal growth factor-dependent breast cancer progression. Oncogene *29*: 6485-6498, 2010.

21 Mani SA, Guo W, Liao MJ, Eaton EN, Ayyanan A, Zhou AY, Brooks M, Reinhard F, Zhang CC, Shipitsin M, Campbell LL, Polyak K, Brisken C, Yang J and Weinberg RA: The epithelial–mesenchymal transition generates cells with properties of stem cells. Cell *133(4)*: 704-715, 2008.

22 Taube JH, Herschkowitz JI, Komurov K, Zhou AY, Gupta S, Yang J, Hartwell K, Onder TT, Gupta PB, Evans KW, Hollier BG, Ram PT, Lander ES, Rosen JM, Weinberg RA and Mani SA: Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. Proc Natl Acad Sci USA *107(35)*: 15449-15454, 2010.

23 Sims, AH, Larionov AA, Harrison DJ and Katz E: Use of microarray analysis to investigate EMT gene signatures. Methods Mol Biol *1046*: 85-95, 2013.

24 Nishikata T, Ishikawa M, Matsuyama T, Takamatsu K, Fukuhara T and Konishi Y: Primary culture of BC: a model system for epithelial–mesenchymal transition and cancer stem cells. Anticancer Res *33(7)*: 2867-2873, 2013.

25 Edgar R, Domrachev M and Lash AE: Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res *30(1)*: 207-210, 2002.

26 Rozen S and Skaletsky HJ: Primer3 on the WWW for general users and for biologist programmers. *In*: Bioinformatics Methods and Protocols: Methods in Molecular Biology. Krawetz S and Misener S (eds.). Humana Press, Totowa, NJ, pp. 365-386, 2000.

27 Primer3 tool. http//fokker.wi.mit.edu/primer3.

28 Becker KG, Hosack DA, Dennis G Jr., Lempicki RA, Bright TJ, Cheadle C and Engel J: PubMatrix: a tool for multiplex literature mining. BMC Bioinformatics *10*: 4: 61, 2003.

29 Gröger CJ, Grubinger M, Waldhör T, Vierlinger K and Mikulits W: Meta-analysis of gene expression signatures defining the epithelial to mesenchymal transition during cancer progression. PLoS One *7(12)*: e51136, 2012.

30 Zientek-Targosz H1, Kunnev D, Hawthorn L, Venkov M, Matsui S, Cheney RT and Ionov Y: Transformation of MCF-10A cells by random mutagenesis with frameshift mutagen ICR191: A model for identifying candidate breast-tumour suppressors. Molecular Cancer *7*: 51, 2008.

31 Zhang Q, Fan H, Shen J, Hoffman RM, Xing HR: Human breast cancer cell lines co-express neuronal, epithelial, and melanocytic differentiation markers *in vitro* and *in vivo*. PLoS One *5*: e9712, 2010.

32 Moustakas A and Heldin CH: Signaling networks guiding epithelial-mesenchymal transitions during embryogenesis and cancer progression. Cancer Sci *98*: 1512-1520, 2007.

33 Wendt MK, Allington TM and Schiemann WP: Mechanisms of the epithelial-mesenchymal transition by TGF-beta. Future Oncol *5*: 1145-1168, 2009.

34 Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguez P, Doerks T, Stark M, Muller J, Bork P, Jensen LJ and von Mering C: The STRING database in 2011: Functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Res *39*: D561-568, 2011.

35 Lobert S, Graichen ME and Morris K: Coordinated regulation of β-tubulin isotypes and epithelial-to-mesenchymal transition protein ZEB1 in BC cells. Biochemistry *52(32)*: 5482-90, 2013.

36 Bundy LM and Sealy L: CCAAT/enhancer binding protein beta (C/EBPbeta)-2 transforms normal mammary epithelial cells and induces epithelial to mesenchymal transition in culture. Oncogene *22(6)*: 869-83, 2003.

37 Park J and Scherer PE: Adipocyte-derived endotrophin promotes malignant tumour progression. J Clin Invest *22*: 4243-4256, 2012.

38 Park J and Scherer PE: Endotrophin a novel factor linking obesity with aggressive tumour growth. Oncotarget *2*: 1487-1488, 2012.

39 Park J, Morley TS and Scherer PE: Inhibition of endotrophin, a cleavage prouct of collagen VI, confers cisplatin sensitivity to tumours. EMBO Mol Med *5(6)*: 935-948, 2013.

40 Stoll SW, Rittié L, Johnson JL and Elder JT: Heparin-binding EGF-like growth factor promotes epithelial–mesenchymal transition in human keratinocytes. J Invest Dermatol *132(9)*: 2148-2157, 2012.

41 Yang N, Morrison CD, Liu P, Miecznikowski J, Bshara W, Han S, Zhu Q, Omilian AR, Li X and Zhang J: TAZ induces growth factor-independent proliferation through activation of EGFR ligand amphiregulin. Cell Cycle *11(15)*: 2922-2930, 2012.

42 Huang Y, Fernandez SV, Goodwin S, Russo PA, Russo IH, Sutter TR and Russo J: Epithelial to mesenchymal transition in human breast epithelial cells transformed by 17beta-estradiol. Cancer Res 67: 11147-11157, 2007.

43 Onder TT, Gupta PB, Mani SA, Yang J, Lander ES and Weinberg RA: Loss of E-cadherin promotes metastasis *via* multiple downstream transcriptional pathways. Cancer Res *68*: 3645-3654, 2008.

44 Thomson S, Petti F, Sujka-Kwok I, Mercado P, Bean J, Monaghan M, Seymour SL, Argast GM, Epstein DM and Haley JD: A systems view of epithelial–mesenchymal transition signaling states. Clin Exp Metastasis *28*: 137-155, 2011.

45 Horiguchi K, Sakamoto K, Koinuma D, Semba K, Inoue A, Inoue S, Fujii H, Yamaguchi A, Miyazawa K, Miyazono K and Saitoh M: TGF-β drives epithelial-mesenchymal transition through δEF1-mediated down-regulation of ESRP. Oncogene *31(26)*: 3190-3201, 2012.

46 Shapiro IM, Cheng AW, Flytzanis NC, Balsamo M, Condeelis JS, Oktay MH, Burge CB and Gertler FB: An EMT-driven alternative splicing program occurs in human breast cancer and modulates cellular phenotype. PLoS Genet *7*: e1002218, 2011.

47 Warzecha CC and Carstens RP: Complex changes in alternative pre-mRNA splicing play a central role in the epithelial-to-mesenchymal transition (EMT). Semin Cancer Biol *2(5-6)*: 417-27, 2012.

48 Horvath A, Pakala SB, Mudvari P, Reddy SD, Ohshiro K, Casimiro S, Pires R, Fuqua SA, Toi M, Costa L, Nair SS, Sukumar S and Kumar R: Novel insights into BC genetic variance through RNA sequencing. Sci Rep *3*: 2256, 2013.

49 Ostman A and Augsten M: Cancer-associated fibroblasts and tumour growth–bystanders turning into key players. Curr Opin Genet Dev *19*: 67-73, 2009.

50 Schulte J, Weidig M, Balzer P, Richter P, Franz M, Junker K, Gajda M, Friedrich K, Wunderlich H, Östman A, Petersen I and Berndt A: Expression of the E-cadherin repressors Snail, Slug and Zeb1 in urothelial carcinoma of the urinary bladder: relation to stromal fibroblast activation and invasive behaviour of carcinoma cells. Histochem Cell Biol *138(6)*: 847-60, 2012.

51 Mefford D and Mefford J: Stromal genes add prognostic information to proliferation and histoclinical markers: a basis for the next generation of BC gene signatures. PLoS One *7(6)*: e37646, 2012.

52 Xiong D, Li G, Li K, Xu Q, Pan Z, Ding F, Vedell P, Liu P, Cui P, Hua X, Jiang H, Yin Y, Zhu Z, Li X, Zhang B, Ma D, Wang Y and You M: Exome sequencing identifies MXRA5 as a novel cancer gene frequently mutated in non-small cell lung carcinoma from Chinese patients. Carcinogenesis *33(9)*: 1797-1805, 2012.

53 Minafra L, Bravatà V, Russo G, Ripamonti M and Gilardi MC: Gene expression profiles in irradiated cancer cells. AIP Conf. Proc. 2013, 1546: 93. doi: 10.1063/1.4816614

54 Chen W, Wu S, Zhang G, Wang W and Shi Y: Effect of AKT inhibition on epithelial-mesenchymal transition and ZEB1-potentiated radiotherapy in nasopharyngeal carcinoma. Oncol Lett *6(5)*: 1234-1240, 2013.