

Research Article

The Human Cell Surfaceome of Breast Tumors

**Júlia Pinheiro Chagas da Cunha,^{1,2} Pedro Alexandre Favoretto Galante,^{1,3}
Jorge Estefano Santana de Souza,^{1,4} Martin Pieprzyk,⁵ Dirce Maria Carraro,⁶
Lloyd J. Old,⁷ Anamaria Aranha Camargo,^{1,3} and Sandro José de Souza^{1,8}**

¹ Ludwig Institute for Cancer Research, São Paulo 01308-050, SP, Brazil

² Center for Applied Toxinology, Butantan Institute, São Paulo 05503-900, SP, Brazil

³ Centro de Oncologia Molecular Hospital Sírio-Libanês, São Paulo 01308-050, SP, Brazil

⁴ Instituto de Bioinformática e Biotecnologia, Natal 59064-560, RN, Brazil

⁵ Fluidigm Inc., South San Francisco 94080, CA, USA

⁶ Hospital AC Camargo, São Paulo 01509, SP, Brazil

⁷ Ludwig Institute for Cancer Research, New York 10017, NY, USA

⁸ Brain Institute, Federal University of Rio Grande do Norte, Natal 59064-560, RN, Brazil

Correspondence should be addressed to Júlia Pinheiro Chagas da Cunha; julia.cunha@butantan.gov.br

Received 21 March 2013; Accepted 22 July 2013

Academic Editor: Lubna Nasir

Copyright © 2013 Júlia Pinheiro Chagas da Cunha et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction. Cell surface proteins are ideal targets for cancer therapy and diagnosis. We have identified a set of more than 3700 genes that code for transmembrane proteins believed to be at human cell surface. *Methods.* We used a high-throughput qPCR system for the analysis of 573 cell surface protein-coding genes in 12 primary breast tumors, 8 breast cell lines, and 21 normal human tissues including breast. To better understand the role of these genes in breast tumors, we used a series of bioinformatics strategies to integrate different type, of the datasets, such as KEGG, protein-protein interaction databases, ONCOMINE, and data from literature. *Results.* We found that at least 77 genes are overexpressed in breast primary tumors while at least 2 of them have also a restricted expression pattern in normal tissues. We found common signaling pathways that may be regulated in breast tumors through the overexpression of these cell surface protein-coding genes. Furthermore, a comparison was made between the genes found in this report and other genes associated with features clinically relevant for breast tumorigenesis. *Conclusions.* The expression profiling generated in this study, together with an integrative bioinformatics analysis, allowed us to identify putative targets for breast tumors.

This paper is in the memory of Lloyd J. Old

1. Introduction

Breast tumors are among the most common fatal cancers, accounting for half of the total cancer deaths in women [1]. It has been shown that breast cancer is a very heterogeneous disease with four different molecular profiles based on the expression pattern of ERBB2, estrogen and progesterone receptors, and histological grade. These classes are also associated with distinct clinical outcomes and responses to therapy [2].

According to the American Society of Clinical Oncology [3], breast cancers express some additional markers that have been shown to be useful in clinic: CA15-3, CA27.29, carcinoembryonic antigen (CEA), urokinase plasminogen activator, and plasminogen activator inhibitor 1, among others. Several of these markers are cell surface transmembrane proteins, including ERBB2 and CEA. In addition, gene signatures identified by gene-expression profiling of breast tumors, such as the MammaPrint [4], which calculates a

prognostic score for node-negative patients at stage I or II based on the expression of a pool of 70 genes, contain more than 10% of trans-membrane (TM) proteins believed to be at the cell surface [5].

Some of these cell surface markers are used as targets for therapies against breast cancer. Targeted therapies, such as lapatinib (Tykerb) [6], bevacizumab (Avastin) [7], and trastuzumab (Herceptin), are gaining special attention in cancer treatment as they act directly to cancer cells avoiding the destruction of healthy cells. Two of them are monoclonal antibodies that block directly or indirectly the activation of a trans-membrane receptor. Herceptin binds to the extracellular domain of ERBB2 and is used routinely in the treatment of patients with ERBB2 positive breast tumors. However, ERBB2 is just over expressed in 20–25% of invasive breast tumors [8, 9]. These facts emphasize the importance of exploring cell surface proteins in tumors to identify new tumor targets. Besides being targets for monoclonal antibodies that block their function, trans-membrane proteins can also be used to allow the entrance of drugs or radioactive material directly into cancer cells.

A catalog of 3,702 genes coding for TM proteins located at the surface of human cells (human cell surfaceome) was generated by us through the use of bioinformatics approaches [5]. New targets for colorectal and glioblastoma (GBM) tumors have been identified using this collection of cell surface proteins. Here we explored the human cell surfaceome searching for breast tumors targets using a high-throughput quantitative real-time PCR (qPCR) in tumor samples and tumor cell lines. Furthermore, we compared our data with other expression profiles publically available taking into consideration some clinical features important for breast tumors.

2. Materials and Methods

2.1. Public Data. Breast cancer somatic mutations were obtained from Sjöblom et al. [10] and Wood et al. [11]. Homozygotic deletions and genomic amplifications in breast cancer were retrieved from Leary et al. [12]. The genomic coordinates for these alterations were determined through a local mapping using BLAT and crossed against surfaceome genes coordinates. Gene pathways were downloaded from KEGG (<http://www.genome.jp/kegg/>). To build the protein protein interaction (PPI) database, data from MINT (December 2007 version), BIOGRID (2.0.37), INTACT (January 2008 version), HPRD (September 2007 version), BIND (May 2006 version), and DIP (January 2008 version) were used. Data obtained exclusively from functional interactions (e.g., genetic interactions present in BIOGRID) and mass spectrometry-based methods were excluded.

2.2. Samples. The twelve breast tumor samples (ERBB2 negative) were obtained from the Tumor Tissue Biobank of the Medical and Research Center, A.C. Camargo Hospital, São Paulo, after local ethics committee approval. Human tumor breast cell lines (HCC1954, MCF7, MDA231, MDA435, MDA436, MDA468, and SK-BR 3) were obtained from American Type Culture Collection (ATCC) and cultured

following ATCC instructions. A nontumorigenic and non-transformed breast cell line (HB4A) was obtained from reduction mammoplasty tissue [13]. Total RNA derived from 21 normal human tissues (thymus, prostate, fetal brain, trachea, skeletal muscle, fetal liver, uterus, small intestine, heart, bone marrow, kidney, stomach, liver, spleen, spinal cord, lung, testis, placenta, brain whole, breast, and colon) was purchased from Clontech.

2.3. Primer Design. Primers and the specific UPL probes were designed using the Universal ProbeLibrary Assay Design Center (<http://qpcr.probefinder.com/organism.jsp>) by automatically selecting an intron spanning assay. Rpl27 (NM.021574.2—ribosomal protein l27) and bcr (NM.021574.2—breakpoint cluster region) were used as the references genes. All primers and probes used are available in [5].

2.4. RNA Extraction, cDNA Synthesis, and Preamplification Reaction. RNA extraction, cDNA synthesis, and pre-amplification reaction were performed as previously described in [5]. Briefly, total RNA was extracted by using Trizol (Invitrogen), and its integrity, checked by Bioanalyser (Agilent). SuperScript III First-Strand Synthesis SuperMix (Invitrogen) was used to synthesize cDNA from 200 ng of DNA-free RNA. For the pre-amplification reaction, one fourth of cDNA (synthesized as described above) was pooled with all primers (50 nM), 10 μ L of 2 X PreAmp Master Mix (Applied Biosystems) in a cycling program of 14 cycles of 95°C for 15 sec and 60°C for 4 min.

2.5. qPCR. The qPCR reactions for 573 genes were performed in 96.96 dynamic array chips (Fluidigm) as previously described [5] following the manufacturer's instructions. In summary, the assay inlets of the M96 Dynamic Array were loaded with 2 μ M forward primer, 2 μ M reverse primer, 1 μ M UPL probe, and 0.025% Tween 20 while each sample inlets was loaded with 2.5 μ L of PreAmp sample, 3.25 μ L of 2 X AB Universal TaqMan Master Mix (Applied Biosystems), and 0.32 μ L of 20 X DA Sample Loading Solution (Fluidigm). The cycling program used consisted of 2 min at 50°C, 10 min at 95°C followed by 35 cycles of 95°C for 15 sec, 70°C for 5 sec, and 1 min at 60°C. All reactions were performed in triplicates.

2.6. Data Processing of qPCR and Selection of Overexpressed Genes. The BioMark Gene Expression Data Analysis was used to obtain the CT values and to discriminate between high-quality and low-quality reactions. Low-quality reactions and CTs higher than 30 were excluded from these analyses and named as not available (NA). The $2^{-\Delta\Delta CT}$ [14] method was used to obtain the relative gene expression values; normal whole breast (from Clontech) was used as a reference, and Rpl27 and bcr were used as reference genes. Genes were considered overexpressed in breast tumors if 3 or more primary breast tumors had a fold change higher than three times the standard deviation in all three replicates. Genes expressed in 5 or less normal tissues were considered restrictedly expressed. In order to be considered expressed,

the gene/sample must have a fold change value higher than 0.1 in two qPCR experiments. Expression in fetal brain, fetal liver, brain, placenta, and testis was not considered in this analysis.

2.7. Comparative Analysis Using ONCOMINE. For the analysis of overexpression in other tumor types, the ONCOMINE website (<http://www.oncomine.org>) was used for searching individually (“Gene search”) the presence of each of the 77 overexpressed genes in the studies that compare “Cancer & Normal” with P value < 0.01 . For the correlation analysis of over expression and prognosis, tumor grade and stage, we searched (“Profile Search”) genes over expressed (P -value < 0.01) on breast studies using the filters “Prognosis,” “Grade” and “Stage” existing at the ONCOMINE website (<http://www.oncomine.org>). Manual inspections of the studies were performed in others to select the most concordant (regarding samples types) in each category.

3. Results and Discussion

3.1. Genes Differentially Expressed in Breast Tumors. Recently, we have generated a catalog of 3,702 genes that codes for trans-membrane proteins located at cell surface [5]. In that report, we further selected 573 cell surface genes as the most promising candidates for cancer diagnostics and therapy based on their expression profiling through the use of Massively Parallel Signature Sequencing (MPSS) and Serial Analysis of Gene Expression (SAGE) libraries. We validated this strategy by finding potential new human targets for GBM and colorectal tumors. Here, we analyzed the expression of these 573 genes in 12 primary breast tumors, 1 normal breast tissue, 7 breast tumor cell lines, and 1 normal breast cell line using a high-throughput qPCR platform. We also analyzed these 573 genes in a panel of 21 normal samples obtained from our previous publication [5].

An ideal candidate for breast cancer diagnostics and therapy would be a gene product that is over expressed in breast tumors and expressed in a restricted pattern in normal tissues. Based on that, our qPCR data was analyzed for genes expressed in less than 6 normal tissues and overexpressed in at least 3 primary breast tumors when compared to normal breast tissue. We found two genes showing this expression pattern: C3orf57 and LRRC26. According to our data, the former is expressed in brain, prostate, and stomach while the latter in colon and prostate. C3orf57 (also known as ADMP), a putative G-protein coupled receptor, is also expressed in prostate epithelium and was shown to be downregulated by androgens in mice cells [15]. LRRC26, a leucine-rich repeat protein, has been previously shown to be increased in breast, prostate, colon, and pancreatic cancers. However, its subcellular localization is still a matter of debate in spite of the presence of a trans-membrane domain [16].

Irrespective of the expression pattern in normal tissues, we found 77 genes that were over expressed (in relation to normal breast) in primary breast tumors (Figure 1). Interestingly, these genes are enriched in specific genomic locations. Eight (CADM4, FXYD5, GPR77, PSENEN, SPINT2,

TMEM147, TYROBP, and UPK1A) out of 77 genes are localized at chromosome 19q (P -value 0.006). In addition, an enrichment at the chromosome cytoband 1p35 (IFI6, PTPRU, SERINC2, and SLC2A1) is also observed (P -value 4.5E-4). Interestingly, this cytoband is associated with sporadic colorectal cancer, an Epstein-Barr virus integration site, and with some hormone-dependent syndromes such as serkal syndrome, mullerian aplasia, and hyperandrogenism (OMIM—<http://www.ncbi.nlm.nih.gov/Omim/getmap.cgi?l132850>). SLC2A1, also known as GLUT-1, is a glucose transporter, that is mutated in GBM and pancreatic tumor. Furthermore, it is expressed in various tumors, like oral squamous cell carcinoma [17], primary renal tumors [18], ovarian carcinoma [19], colorectal tumors [20], thyroid carcinomas [21], and gallbladder carcinomas [22]. Differential expression of this gene in breast tumors emphasizes the importance of glucose transporters in the glucose requirements in cancer cells. As we used only ERBB2 negative samples in this study, this gene accordingly was not present in our list of over expressed genes.

Next, we looked at protein-protein interaction (PPI) networks, KEGG, and the literature to find common interaction partners among differentially expressed surfaceome genes, as we did in [5] (Figure 2). Four genes (TNFRSF19, TNFRSF12A, TMEM109, and TMBIM6) interact with members of the TRAF family. TRAF proteins are TNF receptor-associated factors that serve as adapter proteins for a wide variety of cell surface receptors. They regulated the activation of many signal transduction pathways, such as JNK, ERK, and NF- κ B culminating in regulation of cell survival, apoptosis, and differentiation [23].

Furthermore, we found that GIPC1, a member of PDZ domain containing family that has already been detected as overexpressed in breast tumors [24], interacts with four genes of our list of overexpressed surfaceome genes (SDC3, SDC4, SLC2A1, and FDZ3). GIPC1 leads to an increase of TGF beta receptors and may interfere with TGF beta signaling [25]. SDC3 and SDC4 are syndecans that contain heparan and chondroitin sulfates chains that play a role in cell-cell and cell-matrix interactions [26]. Orend et al. [27] proposes that SDC4 may play an important role in the control of anchorage-dependent cellular proliferation of fibroblasts. The FZD3, a member of frizzled gene family, contributes to the Wnt signaling pathway which leads to an increase of the beta-catenin in the nucleus promoting specific gene expression [28]. MPSS analysis also confirms that this gene is over expressed in breast tumors (data not shown).

3.2. Further Validation of Overexpressed Genes. To gain more information on the surfaceome genes over-expressed in breast tumors, a comparison with expression data from other sources was performed. We first compared our list of genes with our previous study [5] in which the expression of the same set of genes was evaluated in colorectal tumors and GBM. Seven (GPRI72A, IFI6, NPFFR2, REEP6, SLC2A1, SQLE, and TNFRSF12A) and 13 (CD300LE, CD72, CMTM3, GPRI72A, MMP14, P2RY11, SERINC2, TMEM134, TNFRSF12A, TNFRSF19, TYROBP, VSIG4, and ZDHHC12) genes are also over expressed in colorectal tumors and GBM,



FIGURE 1: Expression profile of surfaceome genes in normal tissues, breast cancer cell lines, and breast tumor samples. Expression data for normal tissues was obtained from [5]. Colors represent the fold changes values obtained by qPCR (average of three qPCRs experiments) using normal breast as a reference sample and Rpl.27 as a reference gene. Red, green and black squares represent, respectively, genes over-expressed (fold changes ranging from 3 to 5.99, 6 to 11.99, and higher than 12 standard deviation values), genes downregulated (fold changes ranging from -3 to -5.99, -6 to -11.99, and lower than -12 standard deviation values), or genes equally expressed between the respective sample and reference. White squares represent noninformative data points.

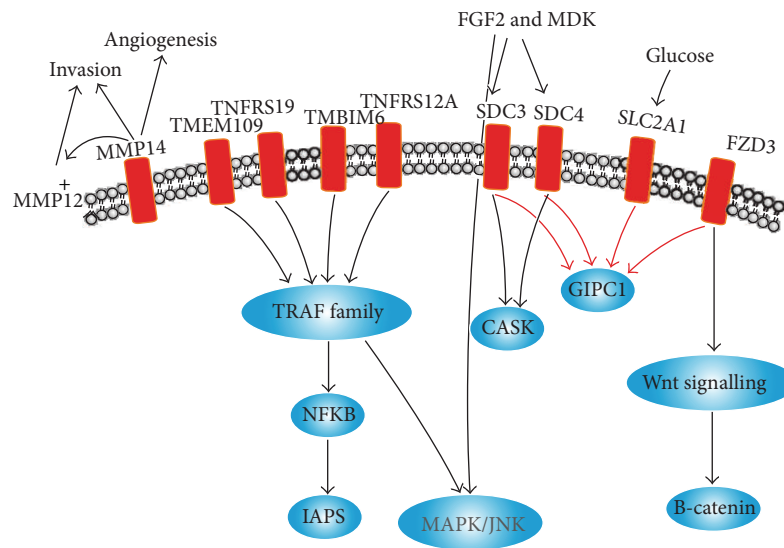


FIGURE 2: Schematic view of key pathways represented in the set of surfaceome genes overexpressed in breast tumors (see text for more details). Interactions represented by black arrows were derived from protein-protein interaction networks while interactions represented by red arrows were derived from either the literature or KEGG database.

respectively. Among them, MMP14, a membrane-bound matrix metalloprotease (MMP), seems one of the most interesting. It has been associated with tumorigenesis, metastasis, and angiogenesis [29–31] and activates pro-MMP2 [32, 33], a MMP known to be critically involved in tumor spreading. Recently, Devy et al. [34] identified a human antibody that is a potent and highly selective inhibitor of MMP-14. It reduces the incidence of metastases in orthotopic mouse model, displayed anti-invasive activity, and inhibited angiogenesis. The authors propose that this inhibitor could be used for the treatment of solid tumors.

Moreover, we compared the expression of our set of overexpressed surfaceome genes in breast tumors to 16 other tumor types, including breast, using microarray data from studies available at ONCOMINE (<http://www.oncomine.org>). As shown in Table 1, we confirmed that more than a half (57,1%) of our genes are overexpressed in at least another study using breast tumors. The other part, however, most probably reflects different experimental sensitivities between qPCR and microarray as well as different breast tumor types. In addition, 53 out of 77 are over expressed in more than 4 tumor types.

It is also important to know whether any of our over expressed surfaceome genes is associated with a clinically relevant characteristic. Genes that are overexpressed (P values < 0.01) in studies that compared samples with distinct prognosis, grades, or stages of breast tumors were selected by looking at the ONCOMINE dataset (Table 2). Thirty-seven out of 77 genes were expressed (in at least one study) in breast tumors of higher grades. Concerning the staging of breast cancer using the TNM classification, six genes (FXVD5, MMP15, PTPRU, SQLE, TM9SF1, and TNFRS12A) are overexpressed in samples with increased spread to lymph nodes, while one gene (TYROBP) is over expressed when breast tumors samples without and with metastasis were

compared. Twenty-five out of 77 genes are over expressed in samples derived from patients with high mortality rate within 5 years while 24 out of 77 genes are over expressed in samples derived from patients that relapsed within 5 years.

The result of these comparisons strongly suggests that the set of surfaceome genes generated in this report is enriched with genes important for breast tumor.

3.3. Surfaceome Genes and Genetic Alterations in Breast Tumors. We also analyzed whether the human surfaceome is associated with genes commonly altered in breast tumors based on previously published data [35] that investigated the presence of homozygous deletion, chromosome amplification, and somatic mutation in breast tumor samples. Eighteen surfaceome genes (BVES, C9orf11, CDH3, FAT, LINGO2, LRRC19, MTNR1A, OR4C6, OR4S2, PCDH8, POPDC3, PTCHD3, PTPLAD2, ROBO2, SGPP1, SLITRK1, TEK, and TLR3) were found previously to be homozygously deleted in breast tumors. Not surprisingly, none of them are present in our list of over expressed genes. In contrast, 116 surfaceome genes are localized in genomic regions that contain focal amplifications in cancer (Figure 3 and See Table S1 in Supplementary Material available online at <http://dx.doi.org/10.1155/2013/976816>). GPR172A, ORMDL3, PERLD1, PRPH2, and TREM2 are presented in our list of over expressed genes. Among them, PERLD1 is located in a chromosome 17 locus together with well known oncogenes (MYC, ERBB2, MET, and FGFR2 among others) that are frequently amplified in gastric and breast cancers [36]. However, the subcellular localization of some of these genes is still a matter of debate.

Two hundred and thirty-five surfaceome genes (6,35% of the total surfaceome collection) have been previously identified as having somatic mutations in breast tumor [10, 11]. Four of them (C19orf28, HM13, ITGAL, and MMP15) are

TABLE 1: Evaluation of breast tumor overexpressed genes in other tumor types according to ONCOMINE.

Gene	Tumors													
	Bladder	Bone marrow	Brain	Breast	Colon	Esophagus	Head and neck	Kidney	Leukemia	Liver	Lung	Ovary	Pancreas	Prostate
ANKH	—	—	—	—	•	—	•	•	—	—	—	•	—	—
ATP13A2	•	—	—	—	—	•	•	—	—	—	—	—	•	•
BRI3	—	—	•	•	—	—	•	—	—	—	—	—	—	—
C19orf28	•	•	—	•	—	—	—	—	—	—	•	•	•	•
C20orf103	—	—	—	•	•	—	•	—	—	—	—	•	•	—
C6orf129	—	—	•	•	—	—	•	—	—	—	•	—	•	—
C9orf7	•	—	•	•	—	—	—	—	—	—	•	—	•	•
C3orf57	—	—	—	—	—	—	—	—	—	—	—	—	—	—
CADM4	—	—	•	—	—	—	—	—	—	•	•	—	—	•
CD300LF	—	—	—	•	—	—	—	—	—	—	—	—	—	—
CD72	•	—	—	•	—	—	•	—	—	•	—	—	—	—
CMTM3	—	—	•	•	—	—	—	•	—	—	—	—	•	—
CNOT1	•	—	•	•	•	—	—	—	•	—	•	—	—	•
CYB561D1	—	•	—	—	—	—	—	—	—	•	—	—	•	—
DEGS2	—	—	—	—	—	—	—	—	•	—	—	—	—	—
FLJ20674	•	—	•	•	—	—	—	—	—	—	—	—	—	—
FXYD5	•	—	•	•	•	—	—	•	—	—	—	•	•	—
FZD3	—	—	—	—	—	—	•	—	—	—	•	—	—	•
GPR172A	•	—	•	•	•	•	•	—	—	•	•	•	•	•
GPR77	—	—	—	—	—	—	•	—	—	—	—	—	—	•
HM13	—	—	•	•	•	—	—	—	—	—	—	•	•	—
IFI6	•	•	•	•	—	•	•	•	—	•	—	—	•	•
ITGAL	—	—	•	—	—	—	—	—	—	—	—	—	—	—
LRRC26	—	—	—	—	—	—	—	—	—	—	—	—	—	—
MMP14	•	—	•	•	•	•	•	—	—	•	•	•	•	•
MMP15	•	—	•	—	—	•	—	—	—	—	•	•	—	•
NINJ1	—	—	•	—	•	—	•	•	—	—	—	—	—	•
NPF2FR2	—	—	—	•	—	—	—	—	—	—	—	—	—	—
NRM	—	—	•	•	—	—	—	—	•	•	•	—	•	•
ORAI1	—	—	•	—	•	•	—	—	—	—	•	—	•	—
ORMDL3	—	•	•	—	•	—	—	•	—	—	—	—	—	—
P2RY11	—	—	•	—	—	—	—	—	—	—	—	•	—	•
PAQR4	•	—	—	•	—	—	•	—	•	•	•	•	—	•
PERLD1	•	—	—	•	—	—	—	•	•	—	—	•	—	—
PRPH2	—	—	—	—	—	—	—	—	—	—	—	—	•	—
PSENEN	—	—	—	—	—	—	—	—	—	•	—	—	—	—
PTDSS1	•	•	•	•	•	•	•	—	•	•	•	—	•	•
PTPRA	•	•	•	•	—	•	—	•	•	—	•	—	•	•
PTPRU	•	—	—	—	—	•	—	—	•	—	•	—	•	•
PYCARD	•	—	•	•	—	—	•	•	—	—	—	—	—	—
REEP6	—	•	—	•	—	—	—	•	•	—	—	—	—	—
RNF167	—	•	—	—	—	—	—	—	—	—	•	—	—	•
RPN2	•	—	•	•	•	•	•	—	•	—	•	•	—	•

TABLE 1: Continued.

Gene	Tumors														
	Bladder	Bone marrow	Brain	Breast	Colon	Esophagus	Head and neck	Kidney	Leukemia	Liver	Lung	Ovary	Pancreas	Prostate	Skin
SCARB1	•	•	—	—	—	•	•	•	—	—	•	—	—	•	•
SDC3	—	—	•	•	—	•	•	•	—	—	•	—	—	•	•
SDC4	•	—	—	—	—	•	•	—	—	—	—	•	—	•	—
SERINC2	—	—	•	•	—	—	—	—	—	•	•	•	•	•	—
SLC10A3	•	•	•	•	—	—	—	—	—	•	—	—	—	—	—
SLC15A3	•	—	•	•	—	•	•	—	—	—	—	—	—	—	—
SLC1A4	—	—	•	•	•	—	•	•	•	•	•	•	—	•	—
SLC2A1	•	•	—	•	•	—	•	•	•	—	•	•	•	—	—
SMPD2	•	—	•	—	—	—	—	•	—	•	•	•	—	—	•
SPINT2	•	—	—	•	—	—	•	•	—	—	•	•	•	•	—
SQLE	•	—	•	•	—	—	•	—	—	•	•	•	—	•	•
SYNGR2	•	—	—	•	•	—	—	•	—	—	•	•	•	•	—
TM9SF1	•	—	•	—	—	—	•	—	•	•	•	•	—	•	—
TMBIM6	•	—	•	—	—	—	—	—	—	—	•	•	•	•	—
TMED3	•	•	•	—	•	—	—	•	•	—	•	—	—	•	•
TMEM109	—	•	—	•	—	—	—	—	•	—	—	—	—	—	•
TMEM132A	•	—	—	•	—	—	•	—	—	—	•	•	•	•	—
TMEM134	•	—	•	—	—	—	—	—	—	—	•	•	—	•	—
TMEM147	—	—	•	—	•	—	—	—	—	•	•	•	—	•	—
TMEM184B	•	—	—	—	—	•	•	•	•	•	•	•	•	•	•
TMEM205	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
TMPRSS4	—	—	—	—	•	—	—	—	—	—	•	•	—	—	—
TNFRSF12A	—	•	•	—	•	•	•	•	—	—	•	•	—	—	•
TNFRSF19	—	—	•	—	—	—	—	—	—	•	—	—	—	•	—
TOR1AIP2	—	•	•	•	—	—	—	—	—	•	—	—	—	—	—
TP53I13	—	—	•	•	—	—	—	—	—	—	—	—	—	—	—
TREM2	•	—	•	•	—	•	•	•	—	•	—	—	•	—	•
TRPV2	—	—	—	•	—	•	—	—	—	—	—	—	—	•	—
TSPAN1	•	—	—	•	—	•	•	—	—	—	—	•	•	•	—
TYROBP	—	—	•	•	•	—	•	•	—	—	—	—	—	—	—
UNC93B1	•	•	•	•	—	—	—	—	•	—	—	•	—	—	—
UPK1A	—	—	—	•	—	—	—	—	—	—	—	—	—	—	—
VSIG4	—	—	•	•	—	•	•	•	•	—	—	—	—	—	—
ZDHHC12	—	—	•	•	—	—	—	—	—	—	—	•	—	—	—

•: means overexpression in a given tumor in at least one ONCOMINE study ($P < 0.01$).

also over expressed in breast tumors according to our analysis (Table S1).

4. Conclusions

The expression profiling generated in this work, together with an integrative analysis using other genomics information, allows some considerations. First, like in GBM [5], signaling through members of the TRAF family seems to be important in breast tumorigenesis. In particular the gene TNFRSF12A (also known as TWEAK receptor) seems to be critical as it was also detected as overexpressed in GBM and

colorectal tumors samples [5]. We confirmed that this gene is significantly overexpressed in breast tumors and seems to be associated with local invasiveness [37]. In addition, MMP15, a metalloproteinase that was found over expressed in our samples, is also over expressed in samples with important clinical features such as higher grades, positive lymph nodes, and poor prognosis (Table 2). It is important to note that this study focus on expression of cell surface gene products thus, well-known cancer genes products localized at cytoplasm or nucleus were not detected in this dataset. In addition, we focused our analysis and discussion on genes that have been previously associated with tumorigenesis and

TABLE 2: Continued.

	Grade ¹										Stage ²										Prognosis																
	Bittner	Chin	Desmedt	Farmer	Finak	Ginestier	Hess	Kreike	Ma	Miller	Pollack	Sortie	Sortie	Sortie	Hess	Sortie	Sortie	Sortie	Turashvili	Yu	M	Bild	Desmedt	Pawitan	Sortie	VandeVijver	Ma	Wang	Wang	VandeVijver	Van de Vijver	VandeVijver					
SYNGR2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
TMPRSS4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
ORAI1	X	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
ZDHC12	X	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
SDC4	-	X	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
SDC3	-	-	-	-	-	X	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
TMEM184B	-	-	-	-	-	X	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
TOR1AIP2	-	-	-	-	-	-	-	-	X	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
TMEM147	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
TMED3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	X	-	-	-	-	-	-	-	-	-	-	-	-	-	-
TREM2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	X	-	-	-	-	-	-	-	-	-	-	-	-	-
SLC15A3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
VSIG4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ANKH	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ATP13A2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
DEGS2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
TNFRSF19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
TMEM134	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

¹ Compares lower with higher grade samples.

² Compares samples derived from Tn + x (T); Nx with Nx + 1 (N); and Mx with Mx + 1 (M).

³ Compares samples derived from live and dead patients within 5 years.

⁴ Compares samples derived from patients without and with relapse/recurrence in 5 years.

X: means overexpression in a given study ($P < 0.01$).

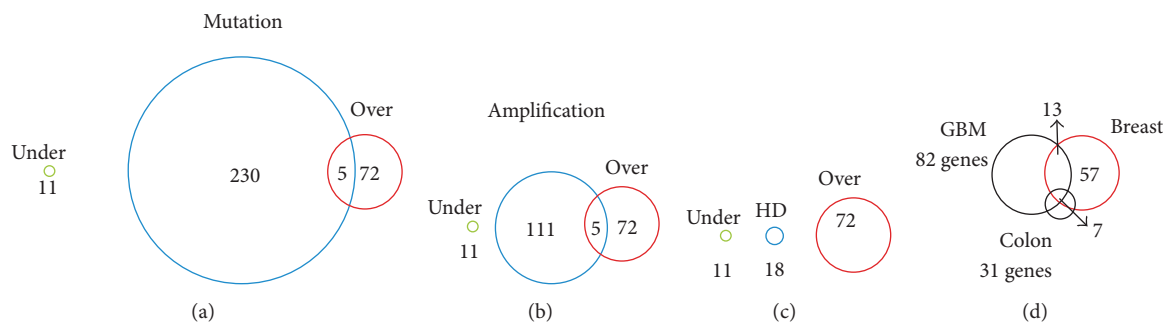


FIGURE 3: Comparison of over- and underexpressed genes in breast tumors with previous studies that identified genes with somatic mutation (a), chromosome amplification (b), or homozygous deletions (c) in breast tumors. (d) Comparison of surfaceome genes over-expressed in breast tumors with surfaceome genes over-expressed in GBM and colorectal tumors [5].

seem to have clinical relevance. Although presented evidence suggests that the overexpressed set of genes is enriched with genes important for breast tumorigenesis, additional studies are needed to establish any clinical applicability. Likewise, additional analyses are also required to uncover the function of the remaining genes that were not found to be associated with cancer and that were not explored in this work.

Abbreviations

GBM: Glioblastoma
 MPSS: Massively Parallel Signature Sequencing
 qPCR: Quantitative Real-Time Polymerase Chain Reaction
 SAGE: Serial Analysis of Gene Expression
 TM: Transmembrane.

Conflict of Interests

The authors declare that they have no conflict of interests.

Authors' Contribution

Lloyd J. Old and Sandro José de Souza designed research; Júlia Pinheiro Chagas da Cunha, Pedro Alexandre Favoretto Galante, Jorge Estefano Santana de Souza, and Martin Piepzyk performed research; Dirce Carraro contributed new reagents/analytic tools; Júlia Pinheiro Chagas da Cunha, Pedro Alexandre Favoretto Galante, Anamaria Aranha Camargo, and Sandro José de Souza analyzed data; and Júlia Pinheiro Chagas da Cunha, Anamaria Aranha Camargo, Lloyd J. Old, and Sandro José de Souza wrote the paper.

Acknowledgments

The authors are grateful to Paula Asprino, Gladis Wilner, Bruna Quevedo, Valéria Paixão, Ken Livak, and Min Lin for technical assistance. Júlia Pinheiro Chagas da Cunha was supported by a fellowship from FAPESP. This work was supported by intramural funds from the Ludwig Institute for Cancer Research and National Institutes of Health Fogarty International Center Grant 5D43TW007015-02.

References

- [1] A. Jemal, R. Siegel, E. Ward et al., "Cancer statistics, 2008," *CA: A Cancer Journal for Clinicians*, vol. 58, no. 2, pp. 71–96, 2008.
- [2] C. M. Perou, T. Sørile, M. B. Eisen et al., "Molecular portraits of human breast tumours," *Nature*, vol. 406, no. 6797, pp. 747–752, 2000.
- [3] L. Harris, H. Fritsche, R. Mennel et al., "American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer," *Journal of Clinical Oncology*, vol. 25, no. 33, pp. 5287–5312, 2007.
- [4] L. J. van't Veer, H. Dai, M. J. van de Vijver et al., "Gene expression profiling predicts clinical outcome of breast cancer," *Nature*, vol. 415, no. 6871, pp. 530–536, 2002.
- [5] J. P. C. da Cunha, P. A. F. Galante, J. E. de Souza et al., "Bioinformatics construction of the human cell surfaceome," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 39, pp. 16752–16757, 2009.
- [6] W. Xia, R. J. Mullin, B. R. Keith et al., "Anti-tumor activity of GW572016: a dual tyrosine kinase inhibitor blocks EGF activation of EGFR/erbB2 and downstream Erk1/2 and AKT pathways," *Oncogene*, vol. 21, no. 41, pp. 6255–6263, 2002.
- [7] H. X. Chen, R. E. Gore-Langton, and B. D. Cheson, "Clinical trials referral resource: current clinical trials of the anti-VEGF monoclonal antibody bevacizumab," *Oncology*, vol. 15, no. 8, pp. 1017–1026, 2001.
- [8] D. J. Slamon, G. M. Clark, S. G. Wong, W. J. Levin, A. Ullrich, and W. L. McGuire, "Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene," *Science*, vol. 235, no. 4785, pp. 182–191, 1987.
- [9] M. A. Owens, B. C. Horten, and M. M. Da Silva, "HER2 amplification ratios by fluorescence in situ hybridization and correlation with immunohistochemistry in a cohort of 6556 breast cancer tissues," *Clinical Breast Cancer*, vol. 5, no. 1, pp. 63–69, 2004.
- [10] T. Sjöblom, S. Jones, L. D. Wood et al., "The consensus coding sequences of human breast and colorectal cancers," *Science*, vol. 314, no. 5797, pp. 268–274, 2006.
- [11] L. D. Wood, D. W. Parsons, S. Jones et al., "The genomic landscapes of human breast and colorectal cancers," *Science*, vol. 318, no. 5853, pp. 1108–1113, 2007.
- [12] R. J. Leary, J. C. Lin, J. Cummins et al., "Integrated analysis of homozygous deletions, focal amplifications, and sequence alterations in breast and colorectal cancers," *Proceedings of the*

- National Academy of Sciences of the United States of America*, vol. 105, no. 42, pp. 16224–16229, 2008.
- [13] A. C. Stamps, S. C. Davies, J. Burman, and M. J. O'Hare, "Analysis of proviral integration in human mammary epithelial cell lines immortalized by retroviral infection with a temperature-sensitive SV40 T- antigen construct," *International Journal of Cancer*, vol. 57, no. 6, pp. 865–874, 1994.
- [14] K. J. Livak and T. D. Schmittgen, "Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method," *Methods*, vol. 25, no. 4, pp. 402–408, 2001.
- [15] J. Singh, L. Young, D. J. Handelsman, and Q. Dong, "Molecular cloning and characterization of a novel androgen repressible gene expressed in the prostate epithelium," *Gene*, vol. 348, no. 1-2, pp. 55–63, 2005.
- [16] K. A. Egland, X. F. Liu, S. Squires et al., "High expression of a cytokeratin-associated protein in many cancers," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 15, pp. 5929–5934, 2006.
- [17] A. W. Eckert, M. H. W. Lautner, H. Taubert, J. Schubert, and U. Bilkenroth, "Expression of Glut-1 is a prognostic marker for oral squamous cell carcinoma patients," *Oncology Reports*, vol. 20, no. 6, pp. 1381–1385, 2008.
- [18] A. Ozcan, S. S. Shen, Q. Zhai, and L. D. Truong, "Expression of GLUT1 in primary renal tumors: morphologic and biologic implications," *American Journal of Clinical Pathology*, vol. 128, no. 2, pp. 245–254, 2007.
- [19] M. Tsukioka, Y. Matsumoto, M. Noriyuki et al., "Expression of glucose transporters in epithelial ovarian carcinoma: correlation with clinical characteristics and tumor angiogenesis," *Oncology Reports*, vol. 18, no. 2, pp. 361–367, 2007.
- [20] R. S. Haber, A. Rathan, K. R. Weiser et al., "GLUT1 glucose transporter expression in colorectal carcinoma: a marker for poor prognosis," *Cancer*, vol. 83, pp. 34–40, 1998.
- [21] R. S. Haber, K. R. Weiser, A. Pritsker, I. Reeder, and D. E. Burstein, "Glut1 glucose transporter expression in benign and malignant thyroid nodules," *Thyroid*, vol. 7, no. 3, pp. 363–667, 1997.
- [22] W. K. Youn, Y.-K. Park, Y. Y. Tae, and M. L. Sang, "Expression of the GLUT1 glucose transporter in gallbladder carcinomas," *Hepato-Gastroenterology*, vol. 49, no. 46, pp. 907–911, 2002.
- [23] P. W. Dempsey, S. E. Doyle, J. Q. He, and G. Cheng, "The signaling adaptors and pathways activated by TNF superfamily," *Cytokine and Growth Factor Reviews*, vol. 14, no. 3-4, pp. 193–209, 2003.
- [24] S. Rudchenko, M. Scanlan, G. Kalantarov et al., "A human monoclonal autoantibody to breast cancer identifies the PDZ domain containing protein GIPCI as a novel breast cancer-associated antigen," *BMC Cancer*, vol. 8, article 248, 2008.
- [25] H. Kirikoshi and M. Katoh, "Expression of human GIPCI in normal tissues, cancer cell lines, and primary tumors," *International Journal of Molecular Medicine*, vol. 9, no. 5, pp. 509–513, 2002.
- [26] K. Lambaerts, S. A. Wilcox-Adelman, and P. Zimmermann, "The signaling mechanisms of syndecan heparan sulfate proteoglycans," *Current Opinion in Cell Biology*, vol. 21, no. 5, pp. 662–669, 2009.
- [27] G. Orend, W. Huang, M. A. Olayioye, N. E. Hynes, and R. Chiquet-Ehrismann, "Tenascin-C blocks cell-cycle progression of anchorage-dependent fibroblasts on fibronectin through inhibition of syndecan-4," *Oncogene*, vol. 22, no. 25, pp. 3917–3926, 2003.
- [28] B. T. MacDonald, K. Tamai, and X. He, "Wnt/ β -catenin signaling: components, mechanisms, and diseases," *Developmental Cell*, vol. 17, no. 1, pp. 9–26, 2009.
- [29] D. V. Rozanov, A. Y. Savinov, R. Williams et al., "Molecular signature of MT1-MMP: transactivation of the downstream universal gene network in cancer," *Cancer Research*, vol. 68, no. 11, pp. 4086–4096, 2008.
- [30] Y. Tsunozuka, H. Kinoh, T. Takino et al., "Expression of membrane-type matrix metalloproteinase 1 (MT1-MMP) in tumor cells enhances pulmonary metastasis in an experimental metastasis assay," *Cancer Research*, vol. 56, no. 24, pp. 5678–5683, 1996.
- [31] T.-H. Chun, F. Sabeh, I. Ota et al., "MT1-MMP-dependent neovessel formation within the confines of the three-dimensional extracellular matrix," *Journal of Cell Biology*, vol. 167, no. 4, pp. 757–767, 2004.
- [32] H. Sato, T. Kinoshita, T. Takino, K. Nakayama, and M. Seiki, "Activation of a recombinant membrane type 1-matrix metalloproteinase (MT1-MMP) by furin and its interaction with tissue inhibitor of metalloproteinases (TIMP)-2," *FEBS Letters*, vol. 393, no. 1, pp. 101–104, 1996.
- [33] K. Holmbeck, P. Bianco, J. Caterina et al., "MT1-MMP-deficient mice develop dwarfism, osteopenia, arthritis, and connective tissue disease due to inadequate collagen turnover," *Cell*, vol. 99, no. 1, pp. 81–92, 1999.
- [34] L. Devy, L. Huang, L. Naa et al., "Selective inhibition of matrix metalloproteinase-14 blocks tumor growth, invasion, and angiogenesis," *Cancer Research*, vol. 69, no. 4, pp. 1517–1526, 2009.
- [35] R. J. Leary, J. C. Lin, J. Cummins et al., "Integrated analysis of homozygous deletions, focal amplifications, and sequence alterations in breast and colorectal cancers," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 42, pp. 16224–16229, 2008.
- [36] M. Katoh and M. Katoh, "MGC9753 gene, located within PPP1R1B-STAR3-ERBB2-GRB7 amplicon on human chromosome 17q12, encodes the seven-transmembrane receptor with extracellular six-cystein domain," *International Journal of Oncology*, vol. 22, no. 6, pp. 1369–1374, 2003.
- [37] A. L. Willis, N. L. Tran, J. M. Chatigny et al., "The fibroblast growth factor-inducible 14 receptor is highly expressed in HER2-positive breast tumors and regulates breast cancer cell invasive capacity," *Molecular Cancer Research*, vol. 6, no. 5, pp. 725–734, 2008.