Genes selected after application modified logistic regression in the microarrays gene expression for breast cancer.

Francielly Morais-Rodrigues[1*%], Diego Lucas Neres Rodrigues[1*], Rita Silvério-Machado[1], Lucas Gabriel Rodrigues Gomes[1], Rodrigo Bentes Kato[1], Roselane Gonçalves dos Santos[1], Vasco A. C. Azevedo[1], Marcos A dos Santos[2].

1 Institute of Biological Sciences, Federal University of Minas Gerais, Brazil. Av. Antônio Carlos, 6627, Belo Horizonte, MG 31270-901,Brazil.

2 Department of Computer Science, Federal University of Minas Gerais, Brazil Av Antônio Carlos, 6627 Belo Horizonte, MG 31270-901,Brazil.

Abstract:

The modified logistic regression (Morais-rodrigues et al., 2020) was used to classify breast cancer subtypes using all microarray database samples. A stabilizing term, that allows the assignment of values to $\alpha i*$ parameters, was inserted in this methodology and with that, during the derivation there is the insertion of the identity matrix (positive defined) which added to the other semi-defined part, results in a positive defined matrix, which has auto values > 0 and determinant > 0, square matrix is full rank if it is reversible (determinant > 0), which results in a single solution. In the results it was observed that some genes we relocated topologically in the extremities after plotting the parameters $\alpha i*$, these parameters are related to the expression of genes with a suppressor or oncogenic profile in breast cancer, and with genes not studied yet. Some of these genes were found in gene regulatory networks from the search of Iglesias-Martinez et al.(2016), and S-score values were associated with these genes, negative value S-score is indicative of tumor suppressing or reduced gene activity and the positive value S-score is indicative of oncogene or increased gene activity (de Souza et al., 2014). In view of the importance of these genes, this article provides a literary review of them.

Keywords: Modified logistic regression, breast cancer subtypes, microarray, stabilizing term, *αi** parameters, genes, suppressor, oncogenic, oncogene

Introduction

The modified logistic regression method presented in the research of Morais-rodrigues *et al*., (2020) showed that the modified methodology allowed for the efficient classification using of all samples of breast cancer tumors. The classification was possible without samples reduction using all tested data sets, showing an efficacy of at least 80% and exploring all possible data combinations. Thus, it proved to be a good model for the breast cancer subtypes classification (systems: TNBC, HER2 Luminal A-LumA, Luminal B-LumB, Basal-like). This methodology also highlighted genes tumor suppressors roles and oncogens. The research then searched for their participation in the Gene Regulatory Networks (GRNs), according to the search of Iglesias-Martinez, *et al*., (2016), it by measuring the gene expression profile of MCF-7 cells stimulated with heregulin(HRG) and epidermal growth factor (EGF) over time, and also associated S-score values to these genes, which according to de Souza et al., (2014) the negative indicative of tumor suppressing or reduced gene activity and the positive indicative of oncogene or increased gene activity. In light to the importance of research on breast cancer suppressors and oncogenes, we decided to make these literary reviews of the genes highlighted in study by Morais-rodrigues et al., (2020).

Table 1, adapted from the research of Morais-rodrigues *et al*., (2020), shows these genes are located

inside EGF and HRG, and predicted their roles as either oncogenes or tumor suppressor genes involved breast cancer, through the S-score system, which integrates genome-wide data (de Souza *et al*., 2014). Stimulation of BC cells with HRG leads to variations in landscape topography in the GRN reconstructed using this data. As such, genes holding the most positive and negative values associated with the α_i^* parameters and selected for each system, were also researched for its involvement in the HGF induced GRN. Some of these genes can be seen located topologically in the extremities Figure 1, and others because they are a little lower, were not indicated in the figure, but are still close to the genes of the extremities.

The MVK gene (positive S-score, oncogene activity), was found to be regulated by GATA2, a transcription factor known to play a crucial role in hematologic malignancies and tumor progression. In breast cancer (BC), it can promote proliferation, and in hepatocellular carcinoma it can lead to poor prognosis (Liu, *et al*., 2015). The WIPI1 gene was found to be regulated by FOXA1, this TF is directly related to the expression of Androgen receptors (AR), and this AR and Estrogen Receptors (ER) regulate cell proliferation and differentiation in BC (Rangel, *et al*., 2018). Moreover, the WIPI1 gene is upregulated in a variety of tumors, such as breast cancer, having gathered attention in the context of breast cancer metastasis understanding (D'Arcangelo, *et al*., 2018). The ADORA2B, which has positive S-score, characterizing it as having oncogene activity, appears in the genetic network regulated by ESRRA, which was implicated in breast cancer progression and it is regulated in prostate, ovary, colon and oral cancers. It is known for having a role in the migration and invasion of cancer cells throughout the body (Tiwari, *et al*., 2015). In the research of Zhang *et al*., 2018, it has been shown that the dysregulation of miRNAs is implicated in the progression of various cancers. They also showed that the expression of miR-135a leads a decrease in the progression of breast cancer, and its inhibition leads to the progression of the disease. This miRNA influenced the targeting of the ESR1, ESRRA and NCOA1 genes by inhibiting the estrogen receptor alpha (ERα), and in the regulatory process between miR-135a and ERα, the miRNA regulated the response to tamoxifen. They could note in tamoxifen resistance cells there was a decrease in the expression of miR-135 in ERα + breast cancer cells, and in tamoxifen sensitive cells there was an increase in miR135a expression.

WWOX is a known tumor suppressor gene in BC and no expression of this gene directly influences the metastasis of triple-negative breast cancer (Chang *et al*.,2018; Pospiech *et al*., 2018) and it is regulated in this EGF induced GRN by GATA3, a gene that is referred to in the literature as being particularly useful as a marker for metastatic breast carcinoma, especially in TNBC subtype.
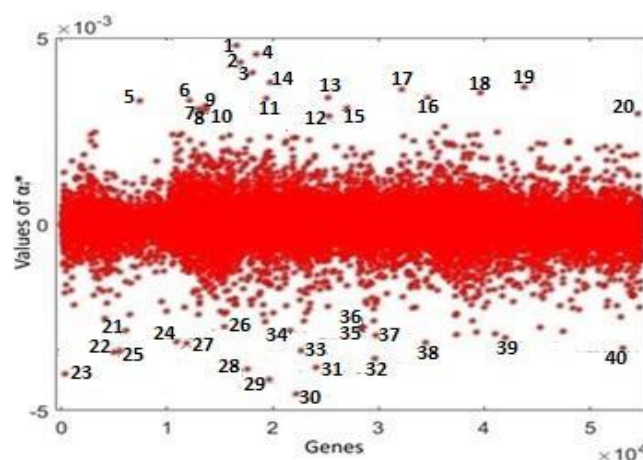

Figure 1: Genes associated with features that represent the 20 most positive and 20 most negative *αi** parameters after applying modified logistic regression from research by Morais-rodrigues *et al*., (2020): 1-ADORA2B, 2-CD33, 3-PDE2A, 4-KCNJ12, 5-SHC1, 6- KIAA0355, 7-WIPI1, 8-MVK, 9-ADCY9, 10-TOX, 11-KIAA1009, 12-ZNF174, 13-WWOX, 14-EYA, 15-SLC25A17, 16-DET1, 17-NUDT7, 18-TCHP, 19-STARD9,20-WIZ, 21-PMVK, 22-SKIV2L2, 23-HDAC9, 24-PDGFRL, 25-FLJ35024, 26-GTF2E1, 27-ALDH1L1, 28-VDR-Cdx2, 29-SPAG11B, 30-PTGDS, 31-CAMTA2,32-SEMA6A, 33-LPHN3, 34-SOCS5, 35-ZNHIT2, 36-MED31, 37-KLHL7, 38- FCRL2, 39-CCDC126, 40-TNFRSF10A Morais-rodrigues *et al*., (2020).

The genes MED23, USP42, ITGA2B, CPEB3, MTR and TELO2 distinguishes between TNBC subtype samples against all other breast cancer subtypes. The first two were found to be regulated by SIX5 that in a recent study was shown to be correlated with the clinico-pathological parameters of BC patients. Iglesias-Martinez and collaborators as having pointed the gene out as having a potential major role in breast cancer (Iglesias-Martinez, *et al*., 2016). In the work of Lin, et al., 2017, it was proven that the inhibition of the MED23 gene implied in a stop in the growth and proliferation of breast cancer. This inhibition also controls distal relapse and metastasis, but its negative S-score (Table 1) puts it with genes with either suppressive activity or reduced activity. The mutual regulation of HER2 and MED23 induces the formation of the everolimus drug resistance mechanism (Lin, et al., 2017). The ubiquitin proteasome system (USP) is a key regulator of fundamental cellular processes on which cancer cells depend. In particular, USP42 was described as being involved in gastric cancer (D'Arcy, *et al*., 2015; Hou, et al., 2016) and it also has a positive S-score (Table 1) putting it together with genes oncogene activity. Likewise, the transcription factor BHLHE40 regulates the ITGA2B and CPEB3 genes, which is believed to be involved in cell differentiation and metastasis, as well as being positively associated with the malignant phenotype of invasive breast cancer (Sethuraman, *et al*., 2018). However its negative S-score (Table 1) puts it together genes with suppressive activity, which in turn regulates CASP7, according to Chaudhary *et al*., (2016), and they claim that this gene induces the development and differentiation of dental tissue and encodes a cysteine protease enzyme linked to apoptosis and inflammation. When CASP7 expression occurs, it induces the growth and proliferation of cancer cells in BC, but did not show apoptosis in primary breast cancer. Although this research has proven that CASP7 can be overexpression in breast cancer and that it is related to the ERα status of the disease, when this happens causes this gene to regulate negatively p21Cip together with the inhibition of CASP7 by peptide inhibitors inducing a low proliferation and growth of BC cells (Chaudhary, et al., 2016). MTR, which has a positive S-score (Table 1) putting it among the genes with oncogene activity, is regulated by ELF-1, as well as by NF-YA, among others. ELF-1 is thought to modulate breast cancer progression to some extent without having an impact on the survival of breast cancer patients (Budka, et al., 2018; Gerloff, et al., 2011). NF-Y was shown to be a heterotrimer composed of NF-YA, NF-YB and NF-YC that turns on cancer related genes and that specific cancer-driving nodes are generally under NF-YA/B control (Benatti, *et al*., 2016). NF-YA was detected in this GRN as a regulator for the IGFBP-3 apoptosis gene. IGFBP-3 is known as insulin-like growth factor binding protein-3 and is overexpressed along with EGFR (Epidemal Growth Factor Receptor) in TNBC, because of this interaction between EGFR and IGFBP-3 promotes resistant treatments. In some case this protein is tumor suppressor in other promoter. For these reasons, and because this subtype of breast cancer has a poor prognosis due to its lack of markers, researches on this topic is very important. (Julovi, *et al*.,2018; Marzec, *et al*., 2015). The TELO2 gene, which has a positive S-score (Table 1) putting it among the gebe with oncogene activity, is regulated by BDP1(protein tyrosine phosphatase). In the work of Thaler *et al*. (2017), this TF protein (expressed in the Her2 subtype) has been shown to be stabilized by carfilzomib and bortezomib, slow down the growth of this subtype. This slowing is due the weakening of the activation of the PI3K / Akt-pathway and blocking of the ERα expression. BDP1 also regulates IGFBP-3 in this EGF induced GRN.

For the HER2 subtype samples against all other samples presented in the results from research by Morais-rodrigues *et al*., (2020), the genes ZKSCAN4 and CNN2 are regulated by SIX5. Both genes have a negative S-score, indicating they may have suppressive activity, but the involvement of CNN2 in cancer, metastasis and the patient survival is not yet clear, but because it plays an important role in cell division and migration, this gene is being studied in breast, prostate cancers, among others, as a way to helps regulate

tumor growth (Hu, *et al*.,2017; Qiu *et al*., 2017). According to Huang *et.al*, 2019, the ZKSCAN4 gene is expressed in the basal-like subtype and has carcinogenic promoter activity.

The SEC61A1 (positive S-score) gene is regulated by CHD2, the MRPL9 gene is regulated by NFE2 and the METTL1 gene by STAT2-STAT6. Although CHD2 was not previously studied in the context of BC, some have mentioned that CHD2 may play a crucial role not only in the proliferation of BC cells but also as having a potential clinical relevance in designing new BC treatments. The MRPL9 gene, which has a positive S-score (Table 1), putting it with genes with oncogene activity, encodes the mitochondrial ribosomal protein L9. It is worth remembering the Mitochondria is responsible for powering breast cancer metabolism as mitochondrial respiration was proven to be required for tumor growth and development. In recent research metabolic alterations are being studied for the possibility causing resistance to anticancer drugs (Villanueva, 2015; Zaal and Berkers, 2018). NF-E2-related factor-2 (NRF2) is one of the major regulators of antioxidant enzyme expression, causing vitamin D to induce these enzymes to create antioxidant defense mechanisms. Sometimes, the Vitamin D increases the expression of some genes that may be linked to DNA repair or play important roles in the proliferation and/or progression of breast carcinoma, and still to be associated with poor prognosis in several different cancers and regulates MRPL9 (Jeon and Shin, 2018). The MERRL1 gene, which has a positive S-score (Table 1), putting it with the gene with oncogene activity, is regulated by STAT2-STAT6: the STAT family of proteins is frequently implicated in breast tumorigenesis in both solid and hematopoietic cancers. STAT6 and STAT3 are important for the regulation of mammary cell differentiation with a substantial body of evidence indicating the involvement of STAT6 and other STATs in breast cancer formation, progression, prognosis and prediction, although STAT2 has not yet been associated with breast cancer (Laudisi, *et al.* 2018; Shahmarvand, *et al.* 2018*;* Terrassa, 2018). The INTS4 and SCN1B genes (both positive S-score, oncogene activity,Table 1) are featured as being regulated by SIX5 and FOXA1, respectively. The SCN1B enforced expression in breast cancer is associated with increases in tumor growth and metastasis in this carcinoma, and could promote metastasis to both lung and liver (Gong, *et al*., 2018). The INTS4 gene goes through mutations implying the progression of adrenocortical carcinoma, bladder cancer, breast cancer, squamous cell carcinoma of the cervix and endocervical adenocarcinoma, colon adenocarcinoma, esophageal carcinoma, glioblastoma, head and neck squamous cell carcinoma, renal clear cell carcinoma, hepatic adenocarcinoma, hepatic carcinoma pulmonary disease, squamous cell carcinoma of the lung, pancreatic adenocarcinoma, pheochromocytoma and paraganglioma, prostate adenocarcinoma, rectal adenocarcinoma, sarcoma, cutaneous skin melanoma, stomach adenocarcinoma, testicular germ cell tumors, thymoma, endometrial carcinoma of the uterine body,uterine carcinosarcoma and uveal melanoma(Federico, *et al*., 2017).

For the Luminal B subtype samples against all other samples, TSEN2 (positive S-score Table 1), GSTM3 (Glutathione S-transferases-GSTs, negative S-score, Table 1) contribute to the individual breast cancer risk and with GSTP1 are still involved in the progression of cervical cancer (Checa-Rojas, *et al*., 2018; Jaramillo-Rangel, *et al*., 2015). The ZNF516 gene, which has a negative S-score (Table 1), putting it with gene with suppressive or reduced activity, has been revealed as being regulated by FOXA1, GATA3 and ZEB1. This last TF is reported as ZEB1 as required for breast tumor initiation and maintenance (Zhou, *et al*., 2017), because of that, the S-score of the gene should be positive, indicating an oncogene or gene with increased activity

The genes FAM102B and ZNF514, expressed in TNBC subtype samples, both genes have positive S-scores (Table 1), that put them with oncogene activity, showing in this GRN as being regulated by NFE2 and MXI1, respectively, which are TFs with a known role in cell differentiation and found to be two of the

largest junctions in this network (Iglesias-Martinez, *et al*., 2016). The MKL1 and PKN2 genes, both of which have negative S-scores (Table 1) that put them with suppressive or reduced gene activity, also were show to be regulated CHD2 for the former and by both this TF and SIX5 for the latter.

Table 1: Genes located inside EGF and HRG with value S-score negative indicative of tumor suppressing or reduced gene activity and positive indicative of oncogene or increased gene activity. Adapted from search of Morais-rodrigues *et al*., (2020).

Genes located inside EGF and HRG both induced GRN	
Genes / S-score / Brast Cancer subtype	
Tumor suppressing:	Oncogene:
SKIV2L2 / −1.54 / (HER2, LumA, LumB)	PMVK / 4.11 / (HER2, LumA, LumB) (S-score needed to be negative)
PDGFRL / −2.52 / (HER2, LumA)	ADORA2B / 1.54 / (TBNC, Lum A, Lum B)
GTF2E1 / −0.95 / (TNBC, LumA, LumB)	KIAA0355 / 0.41 / (TNBC, LumB)
CAMTA2 / −0.99 / (TNBC, LumA)	MVK / 0.60 / (TNBC, LumA)
SEMA6A / 0.73 / (HER2, LumA) (S-score needed to be negative)	WWOX / 0.75 / (TNBC, HER2, LumA, LumB)
MED31 / −2.08 / (LumA, LumB)	WIPI1 / 2.06 / (TNBC, LumA)
CCDC126 / 1.31 / (HER2, LumA, LumB) (S-score needed to be negative)	ADCY9 / 2.17 / (HER2, LumA)
SLC25A17 / −1.56 / (HER2, LumA) (S-score needed to be positive)	EYA4 / 0.77 / (HER, LumA)
TNFRSF10A / −3.20 / (HER2, LumB)	ZNF174 / 2.46 / (TNBC, LumA)
GSTM3/ -1.01 /(LumB) (S-score needed to be positive)	TCHP / 0.48 / (HER2, LumA)

Genes located inside HRG induced GRN	
Tumor suppressing:	Oncogene :
MED23 / -3.58 (TBNC)	CD33 / 0.89 / (TNBC, HER2, LumA)
HDAC9 / 1.95 / (HER2, LumA)(S-score needed to be negative)	SHC1 / 3.44 / (HER2, LumB)
PTGDS / -1.34 / (TNBC, LumA)	FAM102B / 1.07 / (TNBC)
DET1 / -1.52 / (TNBC, LumA, LumB) (S-score needed to be positive)	ZNF514 / 0.45 / (TNBC)
MKL1 / -2.30 / (TNBC)	ME2 / 1.15 / (HER2)
PKN2 / -2.40 / (TNBC)	FOLR1/2.20/ (LumB)
LRRC8E / -1.64 / (HER2)	MACROD1/1.24/ (LumB)
HECTD1/ -1.31/ (LumA)	PI4KB/3.92/ (LumB)
CCDC92/ -1.34/ (LumA)	
CUL5/ -2.10/ (LumA)	
CHD9/ -0.82/ (LumA)	
HDAC11/ -1.22/ (LumA)	
HPS1/ -0.97/ (LumA)	
VIPR2/ -0.39/ (LumA)	
SPRED2/-0.62/ (LumB)	
FANCA/-1.23/ (LumB)	
GLI3/-3.43/ (LumB)	

Genes located inside EGF induced GRN	
Tumor suppressing:	Oncogene :
ZKSCAN4 / -0.19/ (HER2) (According to Huang *et.al*, 2019 the S-score needed to be positive)	USP42/0.90/ (TNBC)
CNN2/ -1.58/ (HER2) (S-score needed to be positive)	ITGA2B/2.12/(TNBC)
ZNF516/ -1.98 /(LumB) (S-score needed to be positive)	CPEB3/-1.54/(TNBC) (S-score needed to be positive)
	MTR/3.39/ (TNBC)
	TELO2/2.51/(TNBC)
	WIZ / 0.51 / (HER 2, LumA)
	SEC61A1 / 1.32 / (HER2)
	MRPL9/ 4.49 / (HER2)
	MERRL1/ 2.07 / (HER2)
	INTS4/ 1.82 / (LumA)
	SCN1B/ 1.10 / (LumA)
	TSEN2 / 0.83 /(LumB)

The LRRC8E gene in the HER2 subtype is regulated by MXI1 and still in this system,. The LRRC8E has negative S-score (Table 1) puts it with suppressive activity or reduced gene activity, unlike it, the ME2 has a positive S-score. The ME2 is found to be regulated by RXRA-VDR complex, found to be another largest junction in this network and also with a known role in cell differentiation

HECTD1, CCDC92, CUL5, CHD9, HPS1, VIPR2 and HDAC11, have negative S-scores (Table 1) put them with suppressive or reduced gene activity, and KLHL20 has positive S-score (Table 1) that puts it with oncogene activity. The first three are shown to be regulated by RFX5, whose expression was found to be predictive of BC patient survival in addition to SIX and CHD2 (Iglesias-Martinez, et al., 2016). CUL5 has been suggested to be a tumor suppressor in in breast and gastric cancers (Fay, et al., 2003; Daniela Mennerich, Kubaichuk, and Kietzmann, 2019), and is also regulated by MXI1, similarly to the CHD9 gene. HPS1 is regulated by NFE2. The VIPR2, HDAC11 and KLHL20 genes are regulated by RXRA complexes: RXRA-VDR for the first and RXRA-NR1H3 for the last two. The RXRA-NR1H3 complex has been previously described as a master regulator of lipid synthesis in mammary epithelial cells and as one of the largest transcriptional hubs in this network (Iglesias-Martinez, *et al.*, 2016). These RXRA complexes also regulate SPRED2 (negative S-score, suppressive activity or reduced gene activity, Table 3) and FOLR1 (positive S-score, oncogene activity, Table 1) revealed in the Luminal B subtype, with the latter being also regulated by NFE2, along with FANCA (negative S-score, suppressive activity or reduced gene activity, Table 1). The GSTM3(Glutathione S-transferases-GSTs, negative S-score) and MACROD1 (positive S-score, oncogene activity, Table 1) genes are regulated by RAD21, that like the RXRA-VDR complex, was found to be on of the largest junctions in this network and also with a known role in cell differentiation. The gene PI4KB (positive S-score, oncogene activity, Table 1) was studied by Waugh (2014) and he noticed that the observed mutations did not contribute to breast cancer, but this gene is regulated by CTBP2 that if it has overexpression it may be involved in making the cells become malignant and breast cancer proliferation and migration. (Yang, *et al.*, 2017). Finally, GLI3 (negative S-score, suppressive activity, Table 1) is featured in this system and regulated by MXI1.

The CD33 and SHC1 genes have positive S-score (Table 1) and put them with oncogene activity, were found to be regulated by the RXRA complexes RXRA-VDR and RXRA-NR1H3, respectively. In myeloid cells from human breast cancer, the CD33 gene is expressed in some cells type, such as: colorectal cells, head and neck squamous cell carcinoma, ovarian (high grade), pancreatic and gastrointestinal stromal tumor cells. This gene is more expressed with high levels of CD33 being associated with reduced overall patient survival, accelerated tumor progression and can promote metastasis. The mechanism in which Myeloid Cells can influence the modulation of the tumor phenotype is currently begin studied (Elliott, *et al.*, 2017; Peng, *et al.*, 2016). In the context of polymorphisms, a study of the association of the condition with breast cancer risk, the single nucleotide polymorphism (SNP) Met300Val in SHC1 gene shows a protective effect in breast cancer, while the common form of the protein is associated with vascularization and metastatic spread of breast tumors (Northey, *et al.*, 2013; Wagner, *et al.*, 2004).

But some genes are not inside the EGF and HRG networks. The gene Tox (expressed in the subtypes TNBC and Lum B) has S-score 2.77. It is aberrantly expressed or mutated in several different kinds of malignancies such as breast cancer and proved to be potential diagnostic or prognostic markers in this type of cancer (Yu and Li 2015). Genes KCNJ12 (expressed in the subtypes HER2 and LumA) with S-score 1.8, KIAA1009 (expressed in the subtypes TNBC, LumA and LumB) with Sscore −1.34, NUDT7 (expressed in the subtypes HER2 and LumA) with S-score 0 and STADARD9 no S-score. ZNHIT2 with S-score −0.02 (TNBC, LumA, LumB), FLJ35024 S-score 0.36 (LumA, LumB), ALDH1L1 S-score 1.52 (HER2, LumA, LumB), LPHN3 S-score 1.50 (HER2, LumA), SOCS5 S-score 0.27 (TNBC,LumA), KLHL7 S-score 1.35

(TNBC, HER2, Lum B), The next genes do not have S-score FCRL2 (LumA), VDR-Cdx2 (LumA), SPAG11B(HER2, LumA, LumB).

The VDR gene is responsible for encoding a vitamin D-associated receptor which is expressed in most body tissues including neoplastic cells. Fact is that the growth of skin, breast, prostate and colon cancers is regulated through vitamin D on signaling pathways (Chauhan and Sakharkar, 2017). Over 470 SNPs have been discovered in the VDR gene in different individuals (Osman, *et al*.,2015), and recently the correlation between the effect of vitamin D on the regulation of the VDR gene focusing on Cdx2 polymorphism has been established: there is a close association between the VDR-Cdx2 polymorphism and breast cancer, in which patients with the polymorphism show a more aggressive phenotype (Kurucu, *et al*.,2019; Rai, *et al*.,2017).

In the suppressor of cytokine signaling (SOCS) family of proteins, the SOCS5 (S-score needed to be positive, but isn't), whose expression in breast cancer tissue was found to reducing tumor growth significantly. The group formed by the genes SOCS 4 to SOCS 6 Regulate epidermal growth factor (EGF) signaling and all of these genes are disease modulators (autoimmune, inflammatory, allergic, bacterial and cancer)(Duncan, *et al*., 2017; Sasi, *et al*., 2014).

In the research of Beniaminov, et a.,(2018), they showed that there is no relationship between the expression levels of ALDH1L1 (considered a tumor suppressor) or promoter methylation levels with yhe stage of breast cancer. However, silencing the ALDH1L1 gene may be a mechanism for tumor development.

The transcripts of the LPHN3 genes are associated with the axillary node status in breast cancer, and are considered potential clinical marker for accurately predicting axillary node status and tumor aggressiveness. Commonly, the first route of spread for breast carcinoma is through the axillary lymph nodes (Aust, *et al*.,2016; kotepui, *et al*.,2012).

The ZNHIT2, STARD9, PDE3A and WIPI1 genes are in a way related to breast cancer in literature, although this relation is not very clear. In a study characterizing the mutational pattern of several Basal-like breast cancer model cell lines, aiming to improve our understanding of Basal-like breast cancer biology and for the development of drug targets for this aggressive subtype, and according to Olsson, *et al*.,(2015) the gene ZNHIT2 shows mutations when found in the Basal-like subtype. The STARD9 gene product is associated with mitotic microtubule formation and cell division and was associated with breast cancer in an in vitro study performed by Torres, *et al*.,(2011) and collaborators where depletion of STARD9 in MCF-7 (breast adenocarcinoma) cell lines caused the pericentriolar material to fragment, and a START Domain Protein Overexpressed in Breast Cancer (Floris, *et al*.,2019). The PDE3A gene is expressed in hundreds of cancer cell lines, including breast cancer cell lines and is associated with cancer maintenance (de Waal, *et al*., 2015).

The CAMTA2 gene, also known as KIAA1009 , may act as tumor suppressor, however, when mutated, allows for malignant cell growth (UniProt, 2019). Despite there begin no direct reference in the literature to the ties between this gene and breast cancer, the CAMTA2 gene maps to human chromosome 17 to which two key tumor suppressor genes are associated, namely, p53 and BRCA1 (a genetic determinant of early onset breast cancer). Like CAMTA2, also MED31 and DET1 genes are not described as being explicitly related to breast cancer. However, the MED31 gene was validated as a target of microRNA-1 (miR-1) in osteosarcoma (Jiang, *et al*.,2014; Liu, *et al*.,2015; Ryota, *et al*.,2019) and DET1 was identified as one of the 54 miR-155-specific tar- get genes in B-cell lymphoma (Slezak, *et al*.,2015, Solé, *et al*.,2018), with miR-1 working as a critical regulator in both osteosarcoma and breast cancer and miR-155 having a known

oncogene role in breast cancer (Bacci,2016; Howard and Yang, 2018).

Still in the context of cancer suppression, PDGFRL is a known tumor suppressor gene in breast cancer (Kawata, *et al.*, 2017). The HDAC9 gene has an essential role in the regulation of breast cancer cell proliferation and patient survival (Huang, *et al.*, 2018).

The FLJ35024 gene encodes a very low density lipoprotein receptor (VLDLR) that constitutes an apolipoprotein E-VLDLR ligand receptor system has been shown to be overexpressed in human Triple-Negative Breast Cancer *in vitro* (Chen *et al.*, 2015; Shiang *et al.*, 2012).

ADCY9 gene encodes an enzyme that acts in the transmission of cellular signals. Studies in the literature indicate that this gene can be used as a biomarker for colon cancer considering that its overexpression is strongly related to the development of this pathology (Yi et al., 2018). Another important synthesized protein is a peroxisomal transporter of coenzyme A, FAD and NAD + encoded by the gene SLC25A17 sequence. The expression of this protein associated with others is related to recurrence and survival and patients with neuroblastoma (Agrimi et al., 2012; Khan et al., 2015).

The KIAA1009 gene, also known as CEP162, is considered to be an important cytoskeletal agent and one of those responsible for the formation of the mitotic spindle. This role is also associated with the development of glioma and is suggested as a putative target for antimitotic drugs for cancer treatment (Leon et al., 2006). Likewise, the KIAA0355 gene (acronym for GARRE) is associated with genes in the RAC family, especially RAC1. Their joint action seems to be related to cellular metabolic regulation and is part of the processes of cellular movement (Bagci et al., 2020; Hanahan and Weinberg, 2011). Specific reports of the RAC1 gene relate it to processes of malignant cell proliferation using in vivo experiments involving Kaposi's sarcoma as a model (Ma et al., 2009).

Another factor capable of regulating cell activity through cytoskeleton modulation is semaphorin 6A (SEMA6A). This protein is able to regulate the expression of TUBB3, an important betatubulin that had its overexpression related to cancers of ovary, breast, lung and gastrointestinal tract (Prislei et al., 2008). Its expression in central nervous system tissue and its correlation with glioblastoma have made semaphorin 6A a study target as a putative biomarker for glioblastoma (Zhao et al., 2015).

From another perspective, molecular mechanisms related to accessibility to chromatin may also be related to cancer pathways. This is the case of the proteins WIZ (or ZNF803) and ZNF174. They are zinc-fingers, proteins with the activity of reversibly binding to the DNA molecule and changing its accessibility to proteins in the replissome (Nygaard et al., 2016; Simon et al., 2015). The importance of types of zinc fingers in the development and progression of cancer variants is already known and their role varies according to the type of protein and the reported gene. However, in this case, zinc fingers can regulate genes related to the production of apoptotic factors and involved in cell proliferation, which culminates in neoplastic processes (Jen and Wang, 2016).

The MVK gene is responsible for coding the enzyme mevalonate kinase, which is involved in the production of cholesterol and other steroids (Buhaescu and Izzedine, 2007) For this reason, the low expression of MVK is one of the main factors responsible for mevalonic aciduria and hyperimmunoglobulinemia D syndrome (Haas and Hoffmann, 2006). On the other hand, pathways related to mevalonic acid are also related to cancer progression, therefore, overexpression of MVK may be related to the development of neoplasms (Thurnher et al., 2012).

In contrast, the EYA4 gene is a regulatory element that acts by regulating molecular mechanisms of

transcription and translation. Its action is essential for the healthy development of kidneys, eyes, ears and muscle, in addition to regulating cardiac action (Liu et al., 2015; Williams et al., 2015).

Still dealing with the regulation of cellular metabolism, TCHP is a gene that encodes the trichoplein keratin filament binding protein. This molecular element is associated with mitochondria, and is usually subregulated in cases of epithelial and breast neoplasms (Cerqua et al., 2010; Vecchione et al., 2009).

NUDT7 is a gene that encodes the Nudix hydrolase protein, an enzyme that eliminates toxic metabolites from the cell's nucleotide degradation. Organisms with low expression of NUDT7 tend to have an increase in the deposition of lipids (Song et al., 2018). On the other hand, the overexpression of this gene reduces the concentration of circulating lipids by activation of cascading coenzyme A metabolic pathways (Shumar et al., 2019).

In another way, Sperm Associated Antigen 11B (SPAG11B) is a protein with antimicrobial action that is found in sperm and very expressed in the epididymis (Narmadha and Yenugu, 2015; von Horsten et al., 2002), and because of this property it is also related to anticancer agents (Gaspar et al., 2013).

In dealing with other neoplasms, it is important to highlight the role of the FCRL2 gene. This gene encodes the FC receptor found especially in cells and tissues related to the immune response such as B cells, lymph nodes and spleen (Jackson et al., 2010). Reports indicate that the underexpression of this gene is related to a worse prognosis in patients with chronic lymphocytic leukemia (Nückel et al., 2009; Shea et al., 2019).

As for the protein Kelch-like family member 7 (KLHL7), it is a protein related to the ubiquitination process, which is linked to protein degradation and regulation of the proteasome activity (Kigoshi et al., 2011). Differences in KLHL7 expressions were previously related to different subtypes of breast cancer and lesion severity, considering that it may be related to the degradation of oncogenic factors (Kurozumi et al., 2018).

On the other hand, the Potassium Inwardly Rectifying Channel Subfamily J Member 12 (KCNJ12) protein is related to the control of potassium influx, and has been associated with diseases such as squamous cell carcinoma and colon cancer (Yuan et al., 2017). The literature also points out that ion channels are in fact more expressed in cancer development, as they are related to the fact that electrical disorders stimulate cell differentiation and proliferation (Biasiotta et al., 2016).

The CCDC126 gene was poorly related to disease. However, studies already indicate that its underexpression is linked to the development of eye diseases associated with the development of photoreceptor structures (Omori et al., 2011). In contrast, the expression of the GTF2E1 gene (general transcription factor IIE subunit 1) was directly related to a poor prognosis in cases of development of colorectal cancer regulating activities of key proteins and functions of transcription and cell permeability (Chang et al., 2020).

The SKIV2L2 gene, also known as MTREX, is responsible for encoding the Mtr4 exosome RNA helicase protein. Its underexpression was related to low mitotic rate due to its role in cell differentiation. Its performance has been proven in murine models of neoplasms (Onderak and Anderson, 2017).

Finally, the overexpression of TNFRSF10A, a member of the tumor necrosis factor receptor superfamily, has been widely studied as a good prognostic biomarker in cases of neoplasms developed in the prostate and pancreas (Che et al., 2018; Gundlach et al., 2018).

2 Conclusions

In this text was presented a brief history of some genes for your association to parameters αi∗that reveal extreme positive and negative values and located topologically at the ends of the graph presented in Figure 1 in the research by Morais-rodrigues et al., 2020. These parameters are related to the expression of each one of the genes presented. It could be noted that these genes are present in different types of cancer and also in breast cancer. But we do not fail to notice that some of these genes are regulated by the same transcription factor. In this case, a more in-depth study would be interesting to answer what would happen to these genes if the activity of these transcription factors or genes was decreased. It is a next step to be thought for the continuation of this research.

Competing interests: The authors declare that they have no competing interests.

Consent to publish: Not Applicable.

Ethics approval and consent to participate: Not Applicable.

References

[1] Agrimi, G., et al. (2012). The human gene SLC25A17 encodes a peroxisomal transporter of coenzyme A, FAD and NAD+. Biochem. J. 443, 241–247. https://doi.org/10.1042/BJ20111420.

[2] Aust,G. et al. (2016) Adhesion GPCRs in Tumorigenesis. Handb Exp Pharmacol, 234: 369–396

[3] Bacci, M., et al. (2016). mir-155 drives metabolic reprogramming of er+ breast cancer cells following long-term estrogendeprivation and predicts clinical response to aromatase inhibitors. Cancer Research 76, 1615–26.

[4] Bagci, H., et al. (2020). Mapping the proximity interaction network of the Rho-family GTPases reveals signalling pathways and regulatory mechanisms. Nat. Cell Biol. 22, 120–134. https://doi.org/10.1038/s41556-019-0438-7

[5] Benatti, P et al., (2016). Direct non transcriptional role of NF-Y in DNA replication. Biochimica et Biophysica Acta 1863(4).

[6] Beniaminov, et al. (2018) Deep Sequencing Revealed a CpG Methylation Pattern Associated With ALDH1L1 Suppression in Breast Cancer. Frontiers in Genetics, (9) Article 169.

[7] Biasiotta, A., et al. (2016). Ion channels expression and function are strongly modified in solid tumors and vascular malformations. J. Transl. Med. 14. https://doi.org/10.1186/s12967-016-1038-y

[8] Budka, J.A. et al. (2018) Common ELF1 deletion in prostate cancer bolsters oncogenic ETS function, inhibits senescence and promotes docetaxel resistance. Genes & Cancer, Vol. 9 (5-6).

[9] Buhaescu, I. and Izzedine, H., 2007. Mevalonate pathway: A review of clinical and therapeutical implications. Clin. Biochem. 40, 575–584. https://doi.org/10.1016/j.clinbiochem.2007.03.016

[10] Cerqua, C., et al. (2010). Trichoplein/mitostatin regulates endoplasmic reticulum–mitochondria juxtaposition. EMBO Rep. 11, 854–860. https://doi.org/10.1038/embor.2010.151

[11] Chang,R. et al. (2018) Loss of Wwox drives metastasis in triple-negative breast cancer by JAK2/STAT3 axis. Nature Communications, 9:3486.

[12] Chang, Z., et al. (2020). Identification of Prognostic Dosage-Sensitive Genes in Colorectal Cancer Based on Multi-Omics. Front. Genet. 10. https://doi.org/10.3389/fgene.2019.01310

[13] Chaudhary, S. et al. (2016) Overexpression of caspase 7 is ERalpha dependent to affect proliferation and cell growth in breast cancer cells by targeting p21(Cip). Oncogenesis, 5, e219.

[14] Chauhan,B. and Sakharkar,P (2017) Role of vitamin D receptor (VDR) gene polymorphism. WORLD JOURNAL OF PHARMACY AND PHARMACEUTICAL SCIENCES,6(7).

[15] Che, Z., et al. (2018). The impact of TRAIL on proliferation of secretory prostate cancer PC-3 cell and LMO2 gene expression. Eur. Rev. Med. Pharmacol. Sci. 22, 7172–7177. https://doi.org/10.26355/eurrev_201811_16249

[16] Checa-Rojas,A. et al. (2018) GSTM3 and GSTP1: novel players driving tumor progression in cervical cancer. Oncotarget,, Vol. 9, (31)

[17] Chen, C. et al. (2015) Genome-Wide Loss of Heterozygosity and DNA Copy Number Aberration in HPVNegative Oral Squamous Cell Carcinoma and Their Associations with Disease-Specific Survival. PLoS ONE 10(8) e0135074.

[18] D'Arcangelo, D. et al. (2018) WIPI1, BAG1, and PEX3 Autophagy-Related Genes Are Relevant Melanoma Markers. Oxidative Medicine and Cellular Longevity, Volume 2018, Article ID 1471682, 12 pages.