Abstract Number: 841 **Presentation Number:** PS4-02-24

Identifying Prognostic Epithelial-Mesenchymal Transition and Tumor Stem Cell Markers in Breast Cancer Subtypes

Vanessa D.Bertoni [1,2], Bruno R.Cavalcante [2,3], Gisele V. Rocha [2], Clarissa Gurgel[2,3]*

[1]-Oncology Multidisciplinary Assistence (AMO / DASA, BA, Brazil);[2]- Gonçalo Moniz Institute, Oswaldo Cruz Foundation (IGM-FIOCRUZ/BA, Brazil);[3]- Department of Pathology and Forensic Medicine, School of Medicine, Federal University of Bahia (UFBA, Brazil).

* vanessadybal@gmail.com

INTRODUCTION: Breast cancer (BC) remains the most prevalent malignancy among women, excluding skin tumors. Despite significant therapeutic advancements, a high relapse rate, affecting approximately 20% of all cases, persists. To improve clinical outcomes and develop novel therapeutic strategies, a deeper understanding of mechanisms driving tumor resistance, such as epithelial-mesenchymal transition (EMT) and tumor stem cell (TSC) biology, is crucial.

OBJECTIVE: This study aimed to identify immunohistochemically targetable markers related to EMT and TSC in different breast cancer subtypes with prognostic significance.

METHODS: We analyzed a public The Cancer Genomic Atlas (TCGA) breast cancer patient database encompassing clinical, histopathological, transcriptomic, and prognostic data of 989 patients. Bioinformatics tools were employed to identify differentially expressed genes (DEGs) impacting survival, analyze associated biological processes, construct co-expression networks, and investigate interactions with EMT and TSC-related pathways. Patients were categorized by their clinical immunohistochemical subtypes: triple-negative (TNBC), HER2-positive, and hormone receptor-positive (HR+) tumors. Protein-protein interaction (PPI) network analysis was integrated to identify the most relevant proteins within each subtype that correlated with survival.

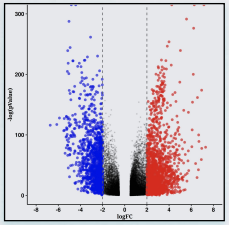


Figure 2. Volcano graph showing the 1033 DEGs found. In blue are the down-regulated genes, and in red are the up-regulated ones.

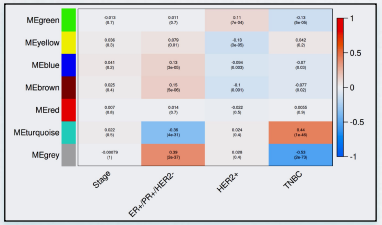


Figure 3. Detailing the correlation strength between the gene co-expression modules and the analysis categories.

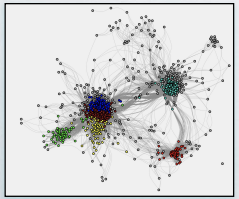


Figure 4 Overview of DEGs in their gene co-expression communities and networks. DEGs: differentially expressed genes.

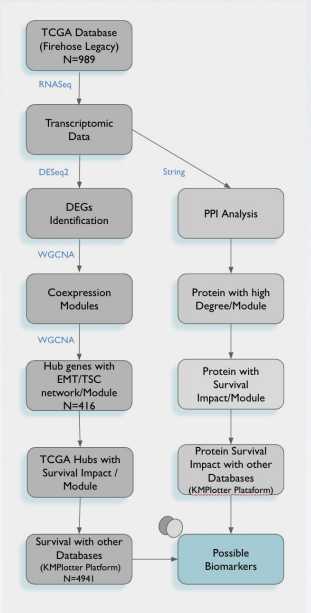


Figure 1. Trial schematic workflow. 989 patients with BC were analysed from the TCGA database to identify breast cancer (BC) differentially expressed genes (DEGs), their coexpression modules; relationship with molecular subtypes; and EMT/TSC genes networks by module. The survival DEGs association was studied both in TCGA data, and KMPlotter platform. Protein network (PPI) analysis was coupled to the study, allowing the identification of the most relevant proteins in each group with an impact on survival. The analysis of both genes and proteins with survival relevance lead to possible prognostic biomarkers to be validated.

Module	EMT/TSC genes	Genes	Genes with Survival Impact (TCGA)	Genes with Survival Impact (KMPlotter)	Proteins	Proteins with Survival Impact (KMPlotter)
Turquoise TCGA n=104; KMPlotter n=126	*DLL3, CCNE1, EZH2, PDX1, CCNA2, TERT, CDK1*	AURKB, CDCA8, CENPA, NDC80, KIF2C, KIFC1, FOXM1, HJURP, RAD54L, GTSE1, PLK1, TTK		*CENPA, CDCA8, NDC80, PLK1*	CDK1, CCNB1, TOPA2, BUB1, BUB1B, NCAPG, TTK, KIF11, CCNB2, CDC20, ASPM, AURKB, CDC45, UBE2C, CDCA8	CCNB1, CDCA8
Green TCGA n=104; KMPlotter n=379	*FN1, GJB2*	*SYNDIG1, COL10A1, SLC24A2, LINC1614, LINC00922, KLKP1, ECM2, P4HA3, INHBA, COL11A1, LRRC15, MMP11, LINC02544*	*SYNDIG1, COL10A1, SLC24A2, LINC1614, LINC00922, KLKP1, ECM2, INHBA, COL11A1*	*LINC00922, ECM2, P4HA3, LRRC15*	COMP, FN1, BGN, COL11A1, MMP13, COL10A1, ADAMTS14, P4HA3, MFAP2, WISP1	COMP, P4HA3
Blue TCGA n=104 KMPlotter n=2218	*ACVR1C, PPARG, FZD4*	*GLYAT, LEP, PLIN1, FABP4, KCNIP2, SLC19A3, AQP7, PLIN4, TMEM132C, ITIH5, GPD1, MRAP, LIPE*	*ITIH5*	*GLYAT, LEP, PLIN1, FABP4, SLC19A3, AQP7, PLIN4, TMEM132C, ITIH5, GPD1, MRAP, LIPE*	PPARG, ADIPOQ, LIPE, LEP, FABP4, LPL, CD36, PLIN1, SLC2A4, DGAT2	ADIPOQ, LIPE, LEP, FABP4, DGAT2
Brown TCGA n=104 KMPlotter n=2218	*IGF1, DTX1, KLF4, ALDH1A1*	*PDE2A, TNS1, MFAP4, ABCA9, SVEP1, CAVIN2, CCL14, GPIHBP1, INMT, LRRN4CL, ABCA10, ABCA8*	*LRRN4CL*	*PDE2A, TNS1, MFAP4, ABCA9, SVEP1, INMT, LRRN4CL, ABCA10, ABCA8*	IL6, CAV1, KLF4, FOS, IGF1, CLDN5, FOSB, ALDH1A1, EGR1, PGM5, AKAP12, CAVIN2, CAV2, ALDH1A2	CAV1, KLF4, FOS, IGF1, CLDN5, FOSB, ALDH1A1, EGR1, PGM5, AKAP12, CAV2,

RESULTS: Analysis of 989 BC patients from the Firehose Legacy TCGA database revealed DEGs and co-expression gene communities correlated with breast cancer clinical subtypes. Within these communities, DEGs related to EMT and TSC were identified, and those with a significant impact on survival were selected. PPI analysis pinpointed proteins within each community exhibiting a high degree of interaction and clinical endpoint impact. After integration with genomic findings, the following potential protein markers were identified for further investigation in paraffin-embedded patient material: CCNB1 and CDCA8 (TNBC); COMP, P4HA3, and ECM2 (HER2-positive tumors); PPARG, ADIPOQ, LIPE, LEP, FABP4, DGAT2; CAV1,KLF4, FOS, IGF1, CLDN5, FOB, ALDH1A1, EGR1,PGM5, AKAP12, CAV2 (HR+ tumors).

Figure 5. Table showing the genes identified in each module related to EMT and TSC, the main differentially expressed genes (DEGs), and their survival impact. Connection networks between module DEGs and EMT/TSC genes were constructed, and survival analyses were performed using TCGA and validated with KMPlotter. For each module, genes with validated survival impact are listed. A subsequent TCGA-based protein–protein interaction (PPI) network analysis identified hub proteins and those associated with survival.

CONCLUSION: We identified subtype-specific prognostic protein markers associated with EMT and TSC in breast cancer. These findings are currently undergoing validation through immunohistochemical analysis of paraffin-embedded material from Brazilian patients and in commercial breast cancer tissue microarray. A Brazilian patients RNA analysis will be performed.

GURGEL LAB

FIOCRUZ