



OPEN

Epigenetic regulation of triple negative breast cancer (TNBC) by TGF- β signaling

Radhakrishnan Vishnubalaji¹ & Nehad M. Alajez^{1,2}✉

TGF β signaling plays crucial role during development and cancer, however the role for TGF β signaling in regulating the noncoding part of the human genome in triple negative breast cancer (TNBC) is still being unraveled. Herein, we provide the transcriptional landscape of TNBC in response to TGF β activation and subsequent inhibition employing SB431542, selective TGF β 1 Receptor ALK5 Inhibitor. Our data revealed 72 commonly upregulated [fold change (FC) \geq 2.0], including PLAU, TPM1, TAGLN, COL1A1, TGFBI, and SNAI1, and 53 downregulated (FC \leq 2.0) protein coding genes in BT-549 and MDA-MB-231 models in response to TGF β 1 activation. Alignment to the geocode (V33) identified 41 upregulated (FC \geq 2.0) and 22 downregulated (FC \leq 2.0) long non-coding RNA (lncRNA) in response to TGF β 1 activation, which were inhibited by concurrent treatment with SB431542. To place our data from the in vitro models into their clinical context, we identified AC015909.1, AC013451.1, CYP1B1-AS1, AC004862.1, LINC01824, AL138828.1, B4GALT1-AS1, AL353751.1, AC090826.3, AC104695.4, ADORA2A-AS1, PTPRG-AS1, LINC01943, AC026954.3, TPM1-AS, ZFPM2-AS1, AC007362.1, AC112721.2, MALAT1, AL513314.2, AC112721.1, AC010343.3, LINC01711, and MAP3K2-DT lncRNA expression to positively correlate with TGF β 1 expression in a cohort of 360 TNBC patients. To provide mechanistic insight into lncRNA regulation by TGF β signaling, SMAD2/3 ChIP-Seq data from BT-549 TNBC model retrieved from Gene Expression Omnibus (GEO) revealed direct binding of SMAD2/SMAD3 to the promoter of AC112721.1, AC112721.2, MALAT1, HHIP-AS1, LINC00472, and SLC7A11, suggesting their direct regulation by TGF β 1/SMAD2/SMAD3 pathway. Interestingly, AC112721.1, AC112721.2 exhibited higher expression in TNBC compared to normal breast tissue suggesting a possible role for those lncRNA in TNBC biology. Our miRNA analysis in the BT-549 model in response to exogenous TGF β 1 revealed several affected miRNAs (2.0 \leq FC \leq 2.0), whose expression pattern was reversed in the presence of SB431542, suggesting those miRNA as plausible targets for TGF β regulation. In particular, we observed hsa-miR-1275 to be downregulated in response to TGF β 1 which was highly predicted to regulate PCDH1, FIBCD1, FXYP7, GDNF, STC1, EDN1, ZSWIM4, FGF1, PPP1R9B, NUA1, PALM2AKAP2, IGFL3, and SPOCK1 whose expression were upregulated in response to TGF β 1 stimulus. On the other hand, hsa-miR-181b-5p was among the top upregulated miRNAs in response to TGF β 1, which is also predicted to regulate CDKN1B, TNFRSF11B, SIM1, and ARSJ in the BT-549 model. Taken together, our data is the first to provide such in depth analysis of lncRNA and miRNA epigenetic changes in response to TGF β signaling in TNBC.

Tumorigenesis is a complex process driven by numerous genetic alterations, altered signaling pathways, and activation of transcription factors over an extended period of time ultimately transforming normal into a cancerous cell. Predominantly, multidimensional signal transduction cascades regulate normal cellular homeostasis, including the transforming growth factor-beta (TGF- β) pathway. TGF β appears to play dual role as a tumor suppressor in early stage of cancer progression and tumor promotor by enhancing tumorigenesis that undergo epithelial–mesenchymal transition (EMT) and eventually leads to chemo-resistant metastatic cancer at the advanced stage^{1–3}. TGF- β is the prototypic super family members of over 30 secreted large members comprising an extended structurally associated group of cytokines/regulatory polypeptides. Including, TGF- β s, bone morphogenetic proteins (BMPs), growth and differentiation factors (GDFs), activins (ACTs), inhibins (INHs),

¹Translational Cancer and Immunity Center (TCIC), Cancer Research Center, Qatar Biomedical Research Institute (QBRI), Hamad Bin Khalifa University (HBKU), Qatar Foundation (QF), P.O. Box 34110, Doha 00000, Qatar. ²College of Health and Life Sciences, Hamad Bin Khalifa University (HBKU), Qatar Foundation (QF), Doha, Qatar. ✉email: nalajez@hbku.edu.qa

glial-derived neurotrophic factors (GDNFs) and nodal growth differentiation factor (Nodal). TGF- β is one among the major regulators of cellular homeostasis in multicellular organisms, which binds to two different membranous serine/threonine kinase receptors TGF- β type II and TGF- β type I (activin-like kinase 5, ALK5). Upon ligand binding, TGF- β -type II receptor phosphorylate TGF- β -type I receptor, subsequently translocate SMAD complex into the nucleus and ultimately regulates various target genes including FOXH1, MIXER, RUNX-proteins and E2F⁴⁻⁶.

TGF signaling has well documented role in tumorigenesis. TGF- β and BMP2 signaling plays significant roles throughout breast cancer (BC) development and promote metastasis by triggering bone metastasis genes IL11 and CTGF in vivo^{7,8}. Furthermore, TGF- β was found to prime tumor cells for pulmonary metastasis through initiation of angiopoietin-like 4 (ANGPTL4) via SMAD pathway in BC microenvironment⁹.

Despite the conspicuous role of TGF- β in cellular development and pathogenesis, the role of TGF β signaling in regulating non-coding RNAs is still being unraveled. Data in the literature suggested an interplay between miRNA and TGF β signaling including miR-34 and miR-203/SNAI1, miR-200 and miR-205/ZEB, miR-1 and miR-200/SLUG and the autocrine TGF- β /miR-200 negative feedback loops identified at the core of TGF- β induced EMT¹⁰.

Similarly, long non-coding RNAs (lncRNAs) are also involved in TGF- β mediated EMT. lncRNAs can exert their function as chromatin modifiers and through regulation of various mRNAs by competitively binding to microRNAs (miRNAs). Limited studies have shown the functional effects of lncRNAs in the context of TGF- β signaling in various cancers. Amongst, few studies have identified lncRNAs involved in TGF- β induced EMT processes, such as lncRNA-ATB, AC026904.1, UCA1, HIT, ROR, HOXA-AS2, CCAT2, which are highly expressed, whereas NKILA, ANCR, lnc-Spry1 and MEG3 are oftentimes downregulated in BC¹¹.

Recently, we have shown the association of novel LINC01614 with TGF- β and FAK signaling in HR⁺/HER2⁺ BC molecular subtype¹². However, the role of TGF- β signaling in regulating lncRNA expression in triple negative breast cancer (TNBC) remains elusive. Herein, we provide complete picture of the cross regulation between miRNA, lncRNA, and mRNA in the context of TGF β signaling employing transcriptome and bioinformatics analyses. We also provide direct correlation between the expression of several lncRNAs and TGF β signaling in TNBC patients' specimens.

Results

Comparative analysis of the protein coding transcriptome in response to TGF β signaling in TNBC.

To characterize TGF- β -induced transcriptional alterations in TNBC, BT-549 and MDA-MB-231 models were exposed to rhTGF- β (10 ng/ml) as single agent or in combination with SB431542 (10 μ M/ml), a well-known TGF- β inhibitor. Heatmap depicting differentially expressed protein coding genes under the aforementioned experimental conditions is presented in Fig. 1a where each column represents one experimental condition, and each row represents an mRNA. Expression level of each gene (log₂) in a single treatment condition is depicted according to the color scale. Enriched functional annotations and pathways are illustrated on the left side of the heatmap. When comparing rhTGF- β and rhTGF- β + SB431542 to DMSO control, we observed three major common clusters in both BT-549 and MDA-MB-231. In Fig. 1a, cluster C1 in hierarchical clustering illustrate G-protein signaling pathway, blood circulation, positive regulation of synaptic transmission GABAergic, mammary gland development and defense response to virus functions were activated by rhTGF- β in BT-549 and MDA-MB-231, and on the contrary, these functional transcripts were reversed by SB431542 treatment to a level comparable to DMSO control. In addition, SB431542 reversed the effect of rhTGF- β which induced pathways of proteinaceous extracellular matrix, extracellular space and angiogenesis in cluster C2. Similar reversible effects by SB431542 were observed in regulation of tyrosine phosphorylation of stat3 protein, regulation of activin receptor signaling pathway, regulation of body fluid levels, organ morphogenesis, positive regulation of fibroblast proliferation and cell adhesion and TGF- β receptor signaling pathways (Fig. 1a). The gene expression principal component analysis (PCA) plot delivers insights into the relationship between samples; mainly we could observe the reversed effects of SB431542 through adjacent falling into DMSO condition, based on PC1, 56% and PC2, 36% of the variation attribution (Fig. 1b). The expression of selected number of differentially expressed genes (MMP14, FAP, SNAI1, BHLHE40, TGFBI, IL1R1 and PDE7B) was validated using qRT-PCR in BT-549, MDA-MB-231, and HCC70 TNBC models (Fig. 1c).

In order to identify commonly regulated genes by TGF β signaling, we compared the mRNA profile of BT-549 and MDA-MB-231 cells under different experimental conditions. Venn diagram depicting the overlap in upregulated (FC \geq 2.0) and downregulated (FC \leq 2.0) genes in both TNBC models and identified commonly 72 upregulated (FC \geq 2.0) and 53 downregulated (FC \leq 2.0) genes, which were reversed by SB431542 treatment (Fig. 2a,b), implying their regulation by TGF β signaling. Furthermore, heatmap depicting commonly deregulated protein-coding genes under different experimental conditions, respectively (Fig. 2c, supplementary table 1). Not surprisingly, several of the identified gene targets are associated with TGF- β /SMAD2/SMAD3 pathway, such as PLA2, TPM1, TAGLN, COL1A1, TGFBI, and SNAI1.

Effects of TGF β signaling on lncRNA transcriptome in TNBC.

We subsequently sought to characterize alterations in the lncRNA transcriptome in response to TGF β signaling. Figure 3a provides hierarchical clustering of differentially expressed lncRNAs in BT-549 and MDA-MB-231 models where each column represents specific treatment condition, and each row represents an lncRNA expression (log₂) where the expression is depicted according to the color scale (supplementary table 2). The lncRNAs expression PCA confirmed the relationship between different treatment showing close proximity between rhTGF- β + SB431542 and DMSO, which are distantly displaced from the rhTGF- β group on PC1, 43% and PC2, 30% of the variation attribution (Fig. 3b). Subsequently, we sought to identify commonly dysregulated lncRNAs in response to TGF β 1 stimula-

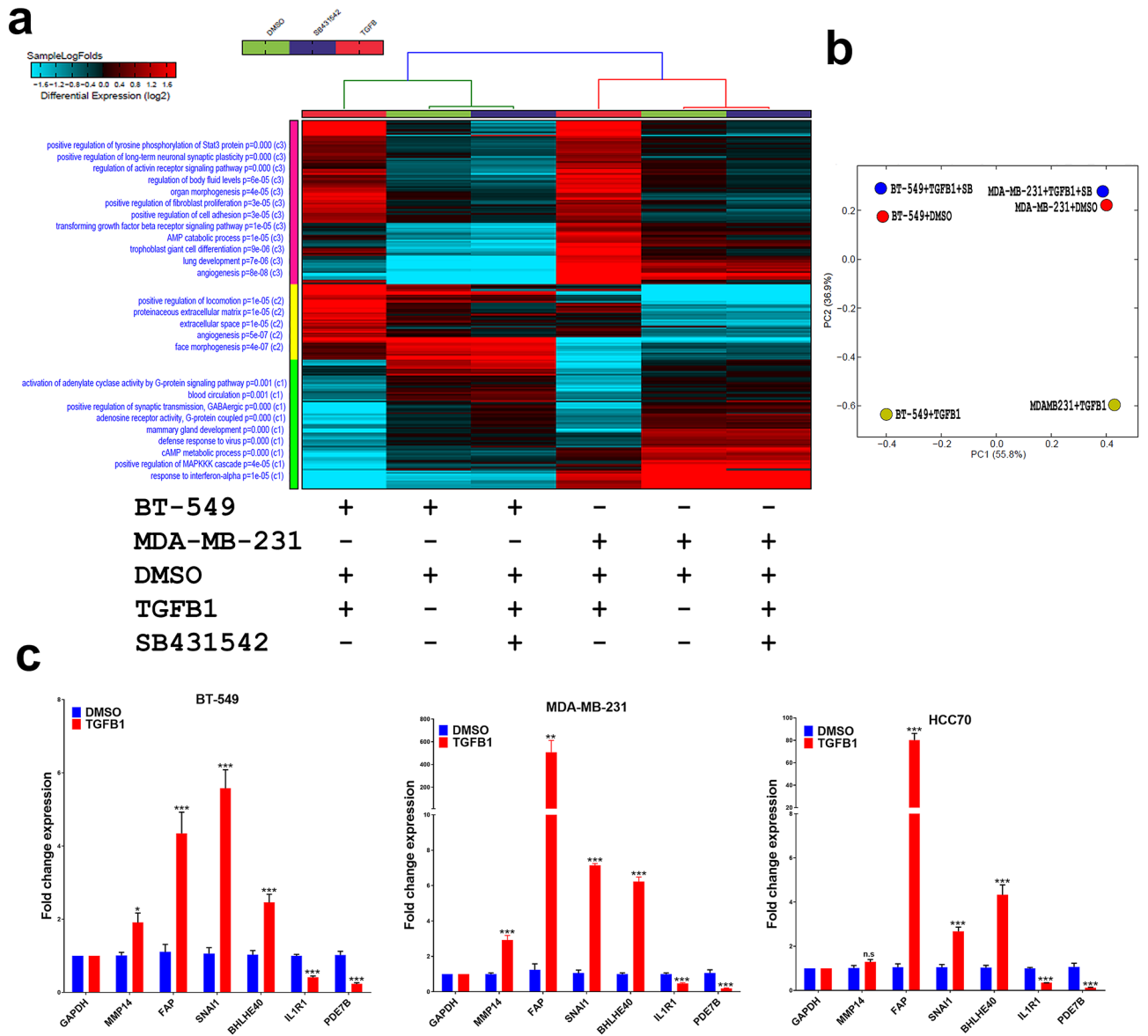


Figure 1. (a) Heatmap depicting changes in protein coding gene expression in BT-549 and MDA-MB-231 under different experimental conditions. Enriched gene ontology (GO) associations/pathways are indicated on the y axis. (b) Principal component analysis (PCA) for the mRNA transcriptome of BT-549 and MDA-MB-231 under different experimental conditions. (c) The expression of selected number of differentially expressed genes was validated using qRT-PCR in BT-549, MDA-MB-231, and HCC70 TNBC models. Data are presented as mean ± SEM (n = 6). *P < 0.05; **P < 0.005; ***P < 0.0005, n.s. not significant.

tion, which also were reversed by concurrent treatment with SB431542 in both TNBC models. Venn diagram depicting the overlap in up and downregulated lncRNAs in both TNBC models identified 41 commonly upregulated (FC ≥ 2.0) and 22 downregulated (FC ≤ 2.0) lncRNAs (Fig. 4a,b). Illustration of the differentially expressed lncRNAs in BT-549 and MDA-MB-231 TNBC models is shown as heatmap (Fig. 4c, supplementary table 3).

Correlation between lncRNA expression and TGFβ signaling in a large cohort of TNBC patients. In order to place our findings from the in vitro models into their clinical context, we investigated the correlation between TGFβ signaling and the expression of the differentially expressed lncRNAs employing a cohort of 360 TNBC patients previously described¹³. Our analysis revealed positive correlations (R² ≥ 0.2) between TGFβ1 expression and the expression of AC015909.1, AC013451.1, CYP1B1-AS1, AC004862.1, LINC01824, AL138828.1, B4GALT1-AS1, AL353751.1, AC090826.3, AC104695.4, ADORA2A-AS1, PTPRG-AS1, LINC01943, AC026954.3, TPM1-AS, ZFPM2-AS1, AC007362.1, AC112721.2, MALAT1, AL513314.2, AC112721.1, AC010343.3, LINC01711 and MAP3K2-DT. Therefore, our analysis underlining a plausible correlation between TGFβ signaling the aforementioned lncRNAs in TNBC (Fig. 5).

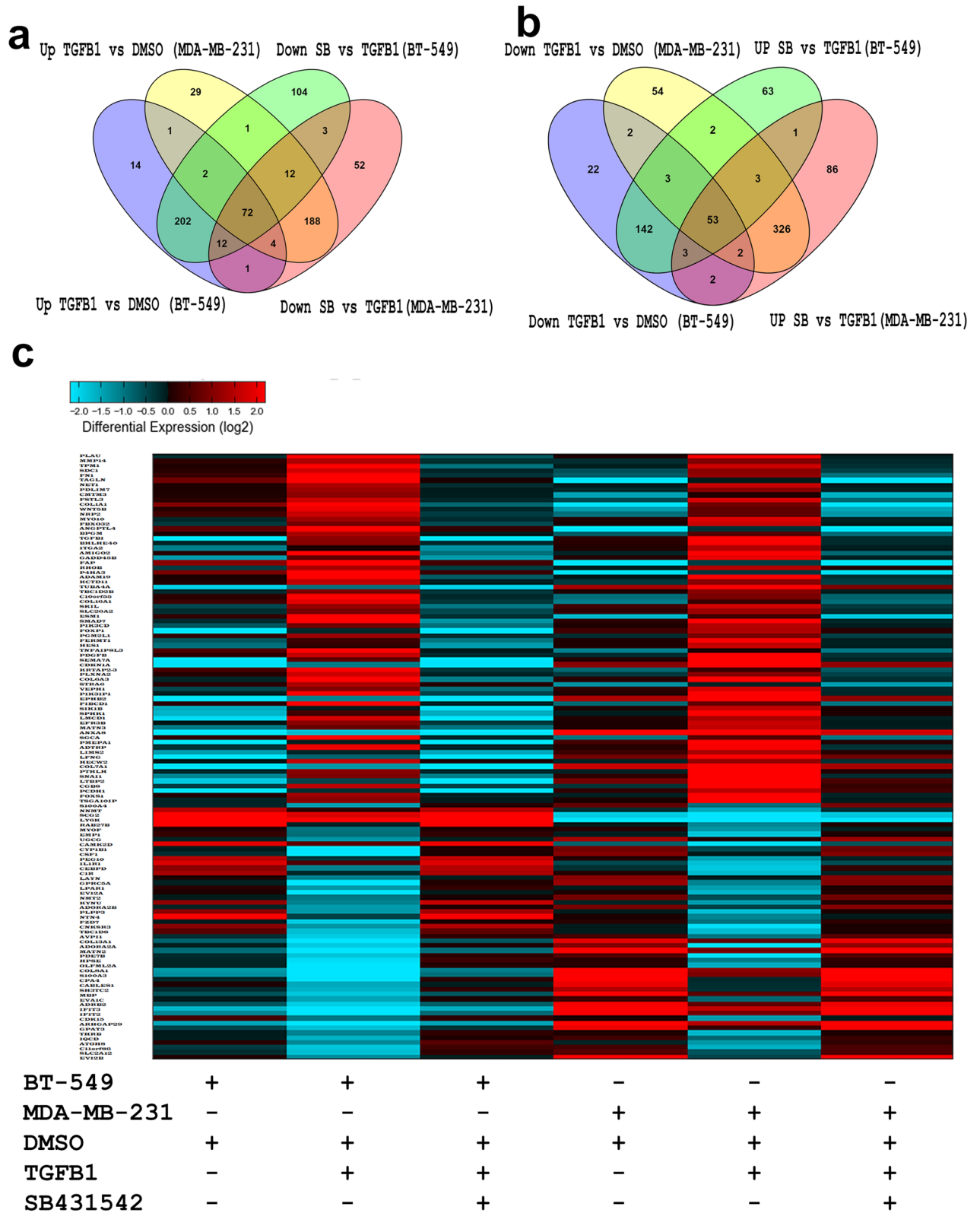


Figure 2. Venn diagram depicting the overlap in upregulated (a) or downregulated (b) genes in BT-549 and MDA-MB-231 under the indicated experimental conditions. Upregulated genes are defined as the genes upregulated ($FC \geq 2.0$) by TGFB and those were downregulated ($FC \leq 2.0$) by SBSB431542. Downregulated genes were defined as the genes downregulated by TGFB1 ($FC \leq 2.0$) and those were induced by SB431542 ($FC \geq 2.0$). (c) Heatmap depicting commonly altered (up and down) protein coding genes under different experimental conditions.

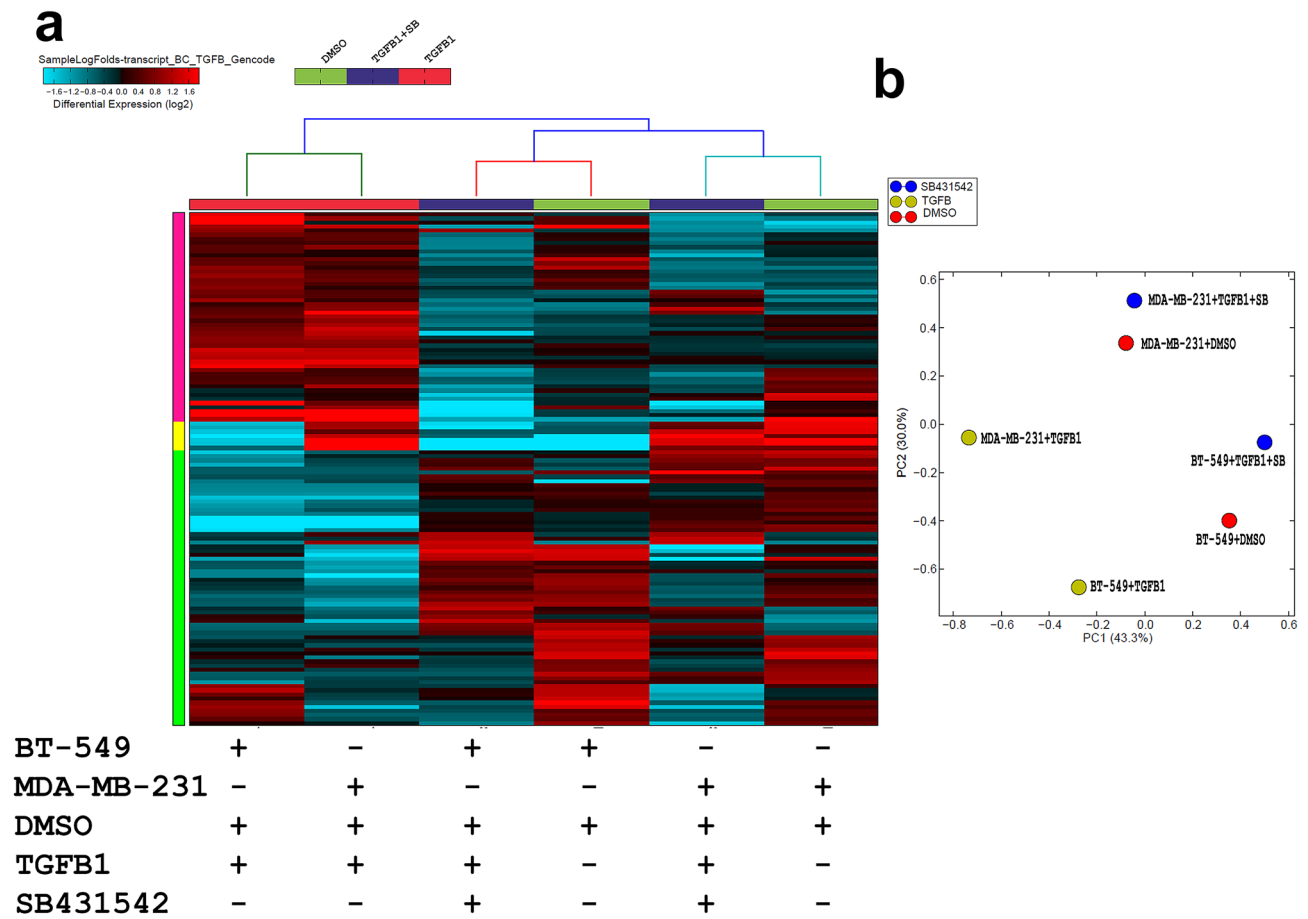


Figure 3. (a) Heatmap depicting changes in lncRNA expression under different experimental conditions. (b) PCA analysis.

ChIP-seq analysis provides support of direct binding of SMAD2/SMAD3 to the promoter region of several TGF- β -dependent lncRNAs. In order to provide direct evidence showing regulation of the identified lncRNAs from current study by SMAD2/SMAD3 signaling, we analyzed chromatin immunoprecipitation (using SMAD2/SMAD3 ab) and sequencing (ChIP-seq) data from the BT-549 cell model previously described¹⁴. Data presented in Fig. 6a revealed direct binding of SMAD2/SMAD3 to the promoter region of the AC112721.1, AC112721.2, MALAT1, HHIP-AS1, LINC00472, SLC7A11, along with TGF β I as positive control, suggesting their direct regulation by TGF β I/SMAD2/SMAD3 pathway in BT-549. AC112721.1, AC112721.2 exhibited highest peak score and higher expression levels of both lncRNAs was observed in TNBC patients (n = 360) compared to controls (n = 88, Fig. 6b).

Integrated analysis of miRNA and mRNA regulatory networks in response to TGF β signaling in TNBC. To identify potential miRNA-mRNA regulatory networks triggered by TGF β in TNBC, we performed miRNA sequencing in BT-549 model in response to rhTGF- β (10 ng/ml) as single agent or in combination with SB431542 (10uM/ml). Our data revealed 37 upregulated (FC \geq 2.0) and 19 downregulated (FC \leq 2.0) miRNAs whose expression could be reversed by SB431542 treatment (Fig. 7a,b). Data from miRNA and mRNA expression were subsequently integrated using the IPA microRNA target filter, which utilizes complex target prediction and validation databases, to provide deeper insight into the miRNA-mRNA regulatory networks. Our analysis revealed the pairing relationship of downregulated hsa-miR-1275, hsa-miR-30c-1-3p, hsa-miR-335-5p, hsa-miR-128-1-5p, hsa-miR-142-3p, hsa-miR-140-3p, hsa-miR-141-5p, hsa-miR-503-3p, hsa-miR-769-3p, hsa-miR-30a-3p and hsa-let-7a-2-3p and potential binding to several upregulated mRNAs in rhTGF- β treated BT-549 cells (Fig. 7c, supplementary table 4). For instance, downregulated hsa-miR-1275 exhibited pairing with 30 upregulated mRNAs associated with apelin cardiac fibroblast signaling pathway, actin cytoskeleton signaling, dendritic cell maturation, adipogenesis pathway, HOTAIR regulatory, cardiac hypertrophy signaling (enhanced), breast cancer regulation by Stathmin1, axonal guidance signaling, cAMP-mediated signaling, cardiac β -adrenergic signaling and protein kinase A Signaling. Likewise, hsa-141-5p also has binding relationship with 11 mRNAs associated with human embryonic stem cell pluripotency, TGF- β signaling, complement system and some of pathways which are similar to has-miR-1275 (supplementary table 4). Likewise, IPA analysis of high-confidence interaction revealed the pairing relationship of upregulated hsa-miR-181b-5p, hsa-miR-582-5p, hsa-miR-377-3p, hsa-miR-154-5p, hsa-miR-379-5p, hsa-miR-409-5p, hsa-miR-329-3p, hsa-miR-134-5p, hsa-miR-654-3p, hsa-miR-485-3p, hsa-miR-1185-1-3p, hsa-miR-376a-3p, hsa-miR-376c-3p, hsa-miR-487b-3p,

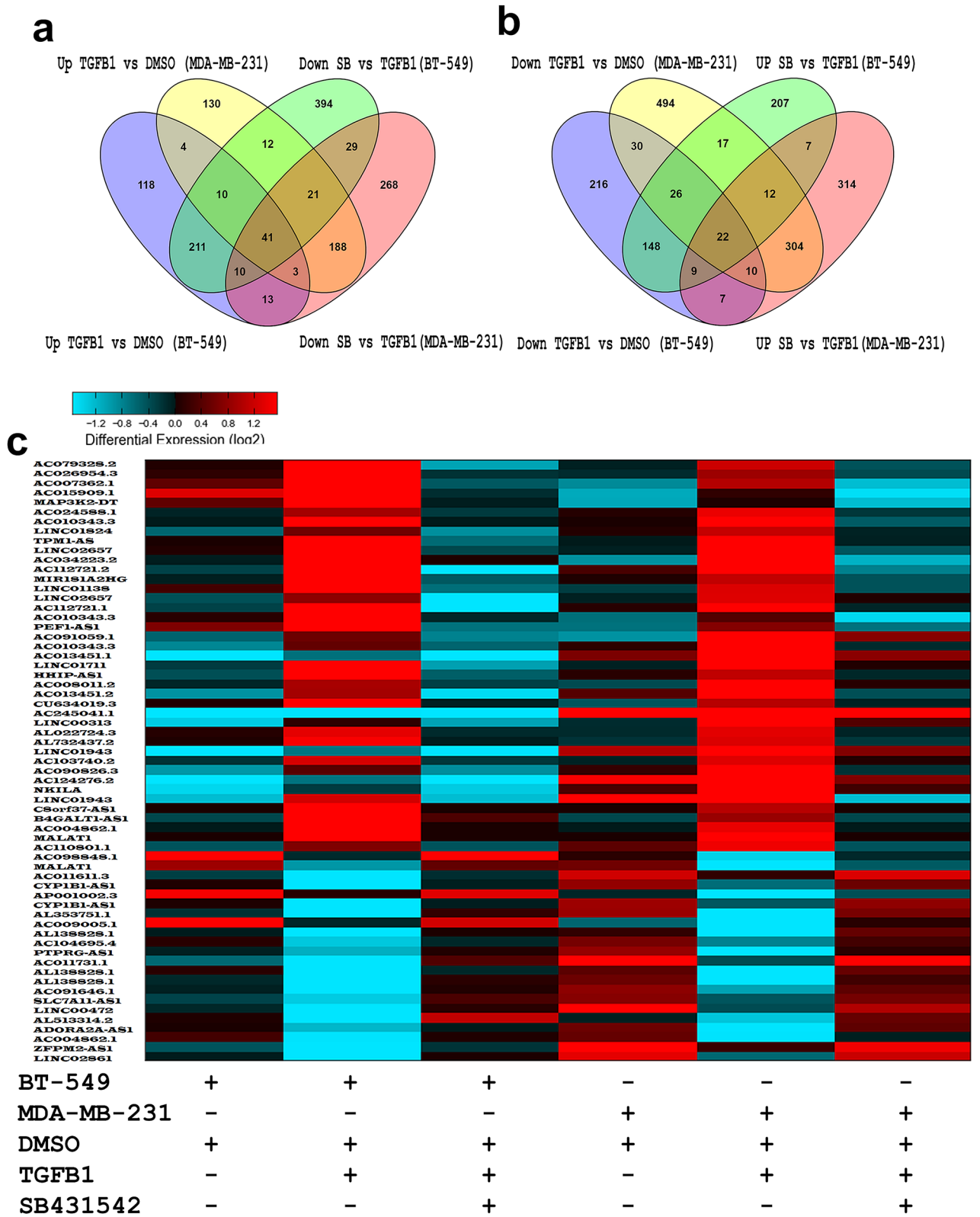


Figure 4. Venn diagram depicting the overlap in upregulated (a) or downregulated (b) lncRNAs in BT-549 and MDA-MB-231 under the indicated experimental conditions. Upregulated genes are defined as the genes upregulated ($FC \geq 2.0$) by TGFB1 and those were downregulated ($FC \leq 2.0$) by SBSB431542. Downregulated genes were defined as the genes downregulated ($FC \leq 2.0$) by TGFB1 and those were induced ($FC \geq 2.0$) by SB431542. (c) Heatmap depicting commonly altered (up and down) lncRNAs under different experimental conditions.

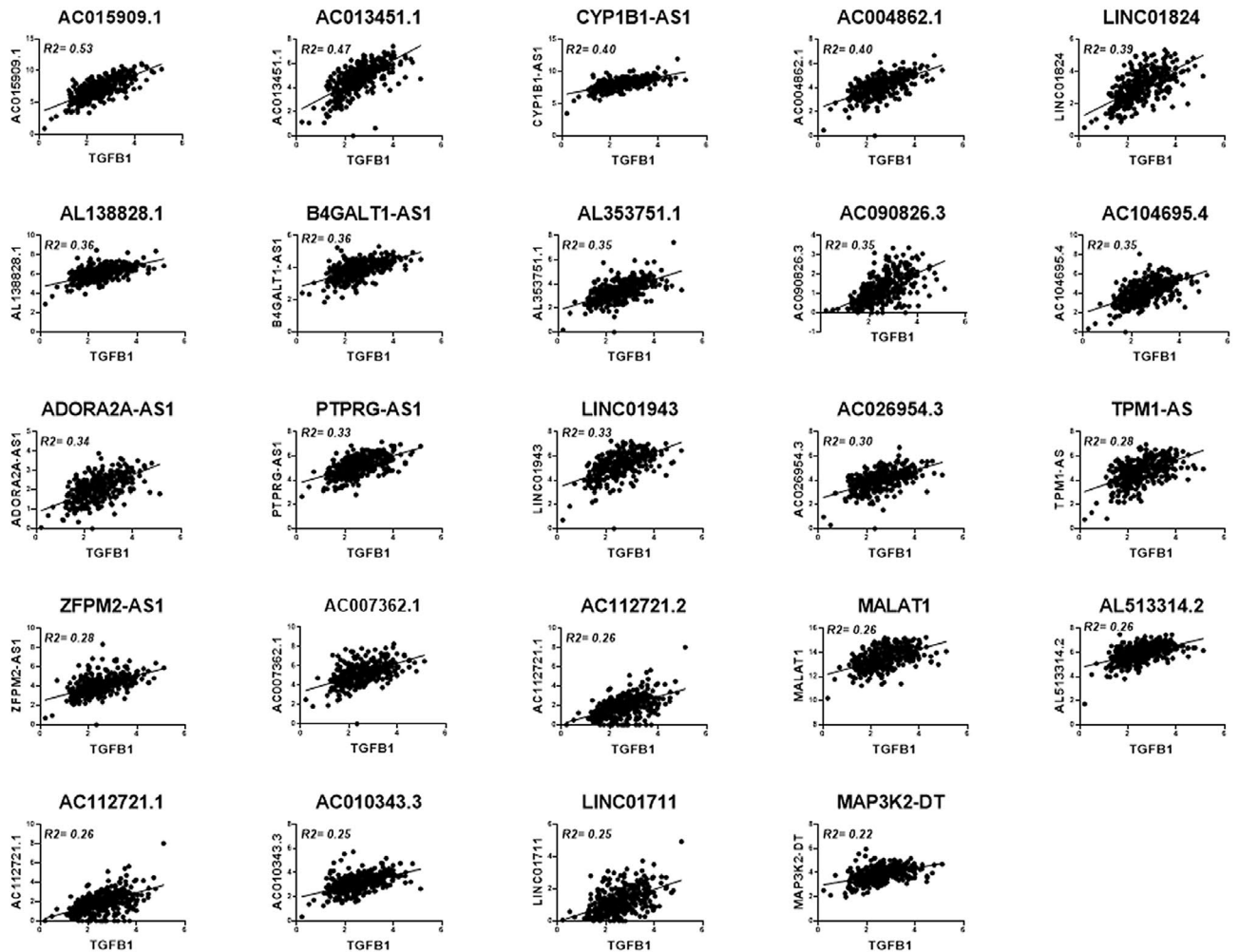


Figure 5. Dot plot depicting the correlation between the expression of TGFB1 and the indicated lncRNAs in a cohort of 360 TNBC patients.

hsa-miR-299-3p, hsa-miR-889-3p, hsa-miR-299-3p, hsa-miR-582-5p, hsa-miR-190a-5p, hsa-miR-2355-3p, hsa-miR-3129-3p and hsa-miR-382-5p and numerous downregulated mRNA targets (Fig. 7d, supplementary table 4). Most of the mRNAs are associated with integrin signaling, leukocyte extravasation signaling, NF- κ B activation by viruses, role of tissue factor in cancer, breast cancer regulation by Stathmin1, eicosanoid signaling and androgen biosynthesis signaling pathway (supplementary table 4). miRNA-mRNA network analyses in the MDA-MB-231 model revealed common networks (hsa-miR-212-3p, hsa-miR-181a-2-3p, and hsa-miR-582-3p) however each cell model also displayed their networks (Supplementary table 5).

Discussion

TNBC represent an aggressive subtype of BC which lack expression of common targets such as estrogen, progesterone and human epidermal growth factor receptors for targeted therapies¹⁵. In complex disease, tumor heterogeneity and reciprocal connection between tumor cells and adjacent cell populations like stromal cells, cancer-associated fibroblasts (CAFs) and extracellular matrix are major factors that regulate tumorigenesis and treatment response. Particularly, CAF in axillary lymph nodes drives metastases in BC through corresponding mechanisms, especially EMT that involves TGF- β pathway¹⁶, although the role of TGF- β in TNBC is not fully understood. The relationship between TGF- β pathways, especially downregulation of TGF- β receptors, indicates the aggressive BC phenotype and stipulates contradictory correlation with patient's outcomes¹⁷. A previous study showed that intratumoral high expression of TGF- β 1 was observed in TNBC (52.5%) compared with non-TNBC (27.5%), implying plausible role for TGF β signaling in TNBC biology. TGF- β 1 was found to enhance the tumorigenic property of MDA-MB-231 cells through activating SMAD2 and P38 signaling pathways. Overall, these studies indicate the association of TGF- β 1 with poor disease-free survival, aggressiveness, and metastasis status of BC^{18,19}.

In current study we provide the transcriptional portrait of mRNA, lncRNA, and miRNA in the context of TGFB1 activation and reversal by SB431542 in two TNBC models. Our data identified 71 commonly upregulated and 53 downregulated mRNAs in response to TGFB1 activation, which were also reversed by concurrent treatment with TGFB1 and SB431542, implying their regulation by TGFB pathway. Several of the identified

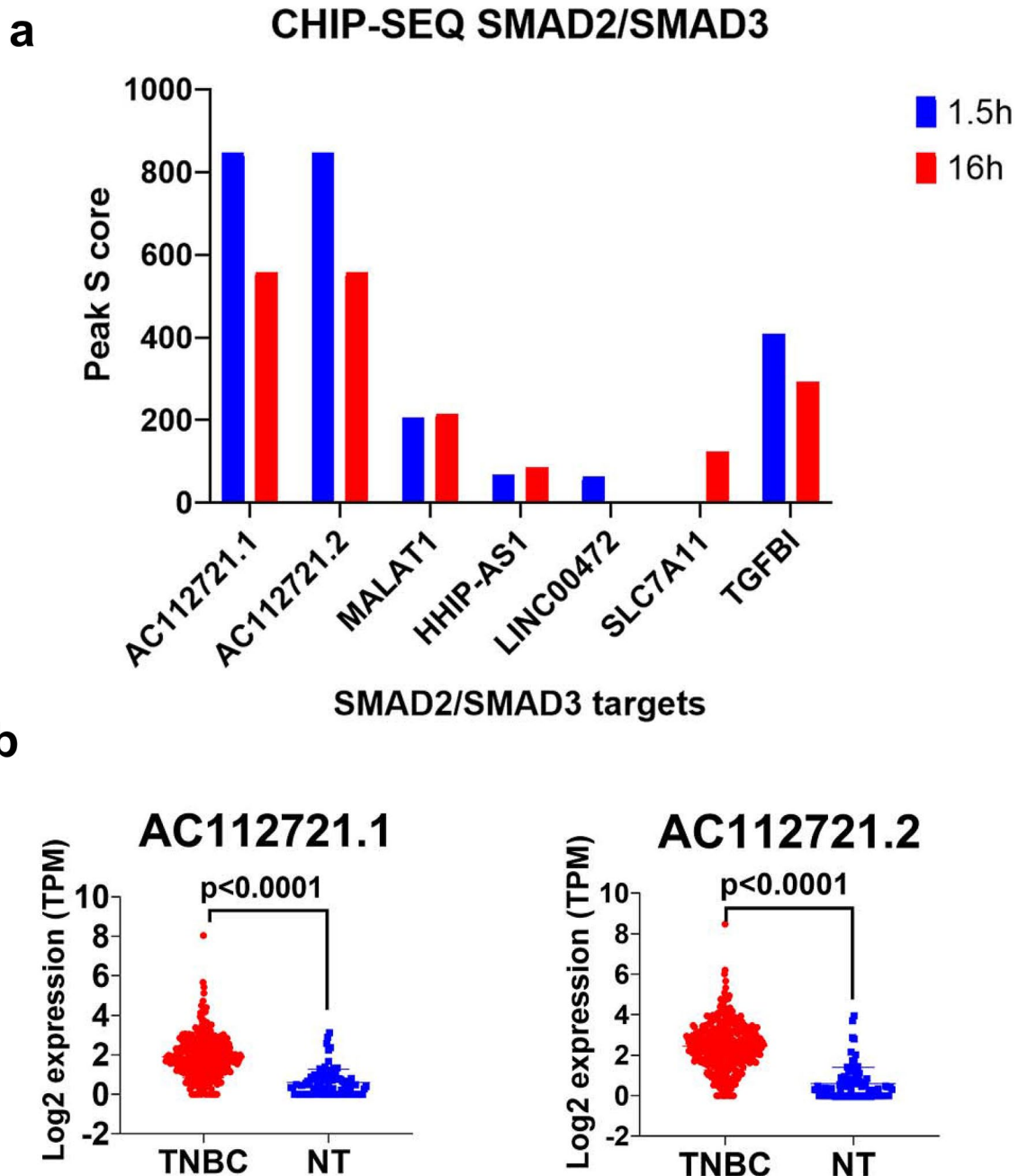


Figure 6. (a) ChIP-Seq analysis demonstrating direct binding of SMAD2/SMAD3 to the promoter region of the indicated lncRNAs. TGFBI was used as positive control. (b) Expression of AC112721.1 and AC112721.2 in a cohort of 360 TNBC and 88 normal breast tissue.

genes are known TGF β /SMAD2/SMAD3 targets, such as PLA1, TPM1, TAGLN, COL1A1, TGF- β I, and SNAI1 thus validating our experimental system^{1,20}. A number of functional categories were enriched in response to TGF- β treatment in both TNBC models, including proteinaceous extracellular matrix, extracellular space and angiogenesis, tyrosine phosphorylation of stat3 protein, activin and TGF β receptor signaling pathway, regulation of body fluid levels, organ morphogenesis, positive regulation of fibroblast proliferation and cell adhesion pathways. Activin belongs to TGF- β super family and Activin-A correlates with poor survival rate in advanced BC. Furthermore, Activin-A enhance anchorage-independent growth, angiogenesis, invasion, stemness and EMT of BC through SMAD signaling pathway²¹. The reversible effects of SB431542 confirmed the association of these functional categories to TGF- β signaling.

While the functional role of TGF- β in BC progression has been investigated rigorously through gene regulation and phenotypic changes, the role of TGF- β pathway in regulating the noncoding part of the human genome is still being unraveled. our data revealed aberrant expression of TGF- β -regulated lncRNAs and miRNAs. Herein we identified 41 upregulated and 22 downregulated lncRNAs in response to TGF β signaling. Our in vitro findings were subsequently validated in a cohort of 360 publicly-available TNBC patient dataset. Twenty-four

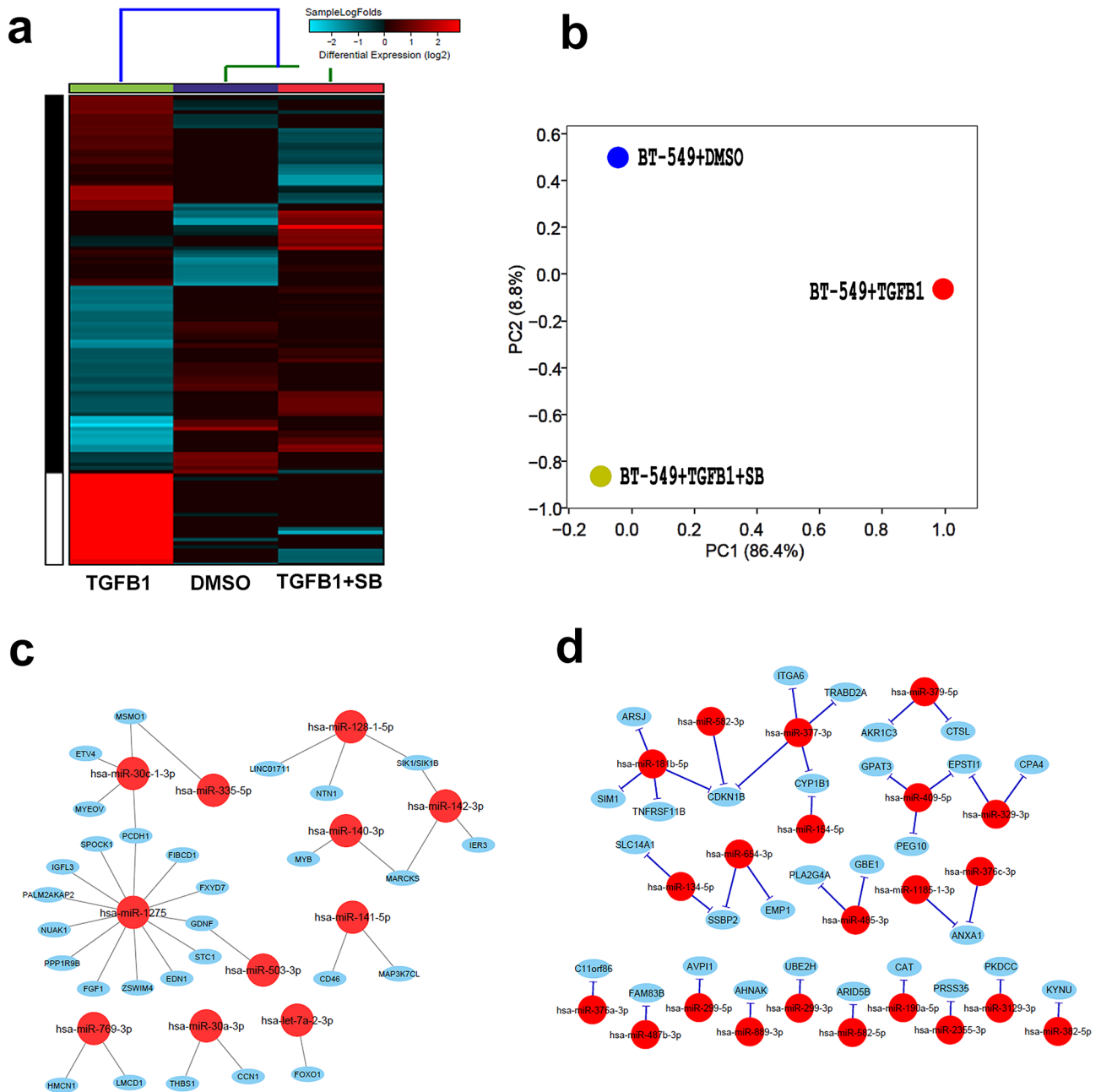


Figure 7. (a) Heatmap depicting changes in miRNA expression under different experimental conditions. (b) PCA analysis. (c) Highly predicted or validated mRNA targets of up (left) and down (right) based on Targetscan and IPA perdition that are affected in BT-549 under different experimental conditions.

lncRNAs strongly correlated with TGFβ1 expression in TNBC tissue. AC015909.1, AC013451.1, CYP1B1-AS1, AC004862.1, LINC01824, AL138828.1, B4GALT1-AS1, AL353751.1, AC090826.3, AC104695.4, ADORA2A-AS1, PTPRG-AS1, LINC01943, AC026954.3, TPM1-AS, ZFPM2-AS1, AC007362.1, AC112721.2, MALAT1, AL513314.2, AC112721.1, AC010343.3, LINC01711, and MAP3K2-DT showed highest correlation with TGFβ signaling. To provide direct evidence of the regulation of those lncRNAs by TGFβ signaling, we explored GEO ChIP-seq dataset and identified direct binding of AC112721.1, AC112721.2, MALAT1, HHIP-AS1, LINC00472 and SLC7A11 promoters by SMAD2/SMAD3, thus providing evidence of direct regulation of those lncRNAs by TGFβ through SMAD2/SMAD3 binding.

In addition to lncRNA, we also characterized the miRNA transcriptome regulated by TGFβ signaling in TNBC. hsa-miR-382-5p, hsa-miR-376c-3p, and hsa-miR-379-5p were the top three miRNAs that highly upregulated by rhTGF-β and those were reversed by simultaneous treatment with rhTGF-β and SB431542. Interestingly, hsa-miR-382-5p was previously reported to exert oncogenic function augmenting the tumorigenic and metastatic potential of BC through diminish the tumor suppressive role of RERG (Ras-related and estrogen-regulated growth inhibitor), a Ras GTPase superfamily member on the oncogenic Ras/ERK pathway²². We observed

hsa-miR-181a-5p to be upregulated in BT-549 and MDA-MB-231 cells in response to TGF β stimulation. hsa-miR-181a-5p was previously shown to hinder cell cycle inhibitor CDKN1B and other genes such as ARHGEF3, SLC2A3, SSBP2, TNFRSF11B, PHLDA1, SIM1, ARSJ, EVI2A and PLPP3. Oncogenic role of hsa-miR-181 was also reported in BC through the targeting SPRY4²³. Furthermore, knockdown of SOX2 decreased the level of cell cycle protein CCND1, CDK4 and CDK6, and reduced the expression of hsa-miR-181a-5p and hsa-miR-30e-5p in MDA-MB-231²⁴. Previous study has shown four miRNAs including hsa-miR-148b, hsa-miR-376c, hsa-miR-409-3p and hsa-miR-801, which were upregulated in the plasma of BC patients²⁵. In current study we found hsa-miR-376c and hsa-miR-409-3p to be highly induced by rhTGF- β treatment in TNBC.

In addition to the upregulated miRNAs, our data also identified several downregulated miRNAs in response to TGF β activation, including hsa-miR-140-3p, hsa-miR-128-1-5p and hsa-miR-141-5p. Loss of those miRNAs upregulates INHBA (inhibin; activin A) and VDR (vitamin D receptor), which are associated with TGF- β /SMAD signaling pathways.

In current study, we observed the downregulation of hsa-miR-128-1-5p in response to TGF- β signaling. Previous study reported loss of hsa-miR-128 as one mechanism through which breast tumor-initiating cells (BT-IC) can resist doxorubicin, through regulation of Bmi-1 and ABCC5 multidrug resistance protein²⁶. hsa-miR-128a was also found to negatively target TGF- β R1 in letrozole-resistant BC cells²⁷. Likewise, hsa-miR-34 family inhibits apoptotic functions through regulation of a wide-ranging of pro-apoptotic and oncogenes controllers including SMAD4, NOTCH1 and TP53 in metastatic BC²⁸. Our revealed downregulation of hsa-miR-34c-3p in response to TGF β stimulation and predicted KRTAP2-3/KRTAP2-4 and MSMO1 (cholesterol biosynthesis) as potential target. It remains however to be investigated if TGF β regulates those miRNAs directly or indirectly. While data presented in current study was focused on TNBC, the effects of TGF β signaling on miRNA and lncRNA expression in breast cancer with other molecule subtypes remains to be investigated.

In closing, our data provide the first insight into epigenetic regulation of lncRNA and miRNA by TGF β signaling in TNBC and suggest their potential utilization as disease biomarkers and therapeutic targets.

Materials and methodology

Cell culture, recombinant TGF- β treatment, and small molecule inhibition. Human BT-549 and MDA-MB-231 TNBC models were cultured in Dulbecco's modified Eagle's medium supplemented with D-glucose 4500 mg/l, 4 mM L-glutamine and 110 mg/l sodium pyruvate, 10% fetal bovine serum and 1 \times penicillin-streptomycin (Pen-Strep) (all purchased from vGibco-Invitrogen, Waltham, MA, USA). Both TNBC models were cultured in 6 well plates (duplicate), 0.2×10^6 cells/well and treated with rhTGF- β 1 (10 ng/ml, Peprotec, London, UK), TGF- β inhibitor (SB431542; 10 μ M, Selleckchem Inc., Houston, TX, USA), and combination of rhTGF- β 1 and TGF- β inhibitor for 48 h and subsequently RNA was isolated. Experiment were carried out with DMSO control¹².

RNA isolation and quality assessment. RNA was isolated using the RNA/DNA/Protein Purification Plus Kit (Norgen Biotek Corp, Ontario, Canada) as per the manufacturer's instructions. Briefly, cells pellets were resuspended in the lysis buffer followed by RNA was extracted under appropriate environment. The concentration and purity of extracted RNA were measured using NanoDrop 2000c (Thermo scientific, MA, USA) and RNA were stored at -80 $^{\circ}$ C. The concentration and quality of RNA was analyzed using capillary electrophoresis using Agilent RNA 6000 Nano Kit (Agilent Technologies, CA, USA) and Agilent 2100 Bioanalyzer (Agilent Technologies) as per the manufacturer's instructions. Samples with RNA Integrity Number (RIN) >8 were considered for library preparation²⁹.

Total RNA library preparation and RNA sequencing. Total RNA samples with a RIN higher than 8 were used as input for the library preparation by using TruSeq Stranded Total RNA Library Prep Gold kit (Cat #: 20,020,598) from Illumina following the manufacturer's protocol. Briefly, 500 ng of total RNA was subjected to rRNA depletion and then to fragmentation. The first-strand cDNA synthesis was performed with random hexamers and SuperScript II Reverse Transcriptase (Cat#: 18064014) from ThermoFisher Scientific. The second cDNA strand synthesis was performed by substitution of dTTP with dUTP. The double-stranded cDNA is then end-repaired and adenylated. Barcoded DNA adapters were ligated to both ends of the double-stranded cDNA and then amplified. The libraries quality was checked on an Agilent 2100 Bioanalyzer system and quantified using Qubit 2.0 fluorometer (Invitrogen). The libraries were pooled, clustered on a cBot platform, and sequenced on an Illumina HiSeq 4000 at a minimum of 50 million paired end reads (2×75 bp) per sample²⁹.

RNA-Seq data analysis of TGF β 1 and SB431542 treated TNBC models. FASTQ files were subsequently pseudo aligned to the Gencode release 33 index (mRNA and lncRNA) and reads were subsequently counted using KALLISTO 0.42.1³⁰ as we described before^{31,32}. Normalized TPM (Transcripts Per Million) expression values were subsequently subjected to differential analysis and hierarchical clustering using cosine for columns and cosine for rows, while gene set and pathway enrichment were performed using Gene Ontology (GO)-Elite algorithm in AltAnalyze v.2.1.3 as described before³³. MicroRNA analysis was carried out in CLC genomics workbench 20.0 using built-in small RNA analysis workflow. miRNA count reads were normalized using the TMM (trimmed mean of M values) normalization method and log₂ CPM (Counts per Million) values were subsequently subjected to differential analysis. The microRNA Target Filter in IPA was subsequently employed to identify potential miRNA-mRNA networks. This bioinformatics approach provides insights into the biological effects of microRNAs, using corresponding mRNA expression data as well as experimentally validated interactions from TarBase and miRecords, and the predicted microRNA-mRNA interactions from TargetScan. Only highly predicted and experimentally predicted networks were included in final analysis. DIANA-LncBase

v3 was utilized to identify miRNA–lncRNA interactions. Cytoscape 3.8.1 was used to construct miRNA–mRNA and miRNA–lncRNA networks.

Quantitative reverse transcription PCR (qRT-PCR). Five-hundred ng of RNA was reverse transcribed to cDNA using High Capacity cDNA Reverse Transcription kit (Applied Biosystems, Foster City, CA, USA). Real time PCR was performed using PowerUp SYBR Green Master Mix (Applied Biosystems) on QuantStudio 7/6 Flex qPCR (Applied Biosystems) using the following primer pairs: MMP14 F: CGAGGTGCCCTATGCCTAC, MMP14 R: CTCGGCAGAGTCAAAGTGG; FAP F: TGTGCATTGTCTTACGCCCT, FAP R: CCGATCAGGTGATAAGCCGT; SNAI1 F: CCTCCCTGTCAGATGAGGAC, SNAI1 R: CCAGGCTGAGGTATTCTCTG; IL1R1 F: GAGCGGCAGGAATGTGACAA, IL1R1 R: GAGGGTGCCTTACCTGGA, BHLHE40 F: GACGGGAATAAAGCGGAGC, BHLHE40 R: CCGGTCACGTCTCTTTTCTC; PDE7B F: GGTTGAGAGGTGTGGCGAAAT, PDE7B R: CCTTAGTCGTATATCTCCCAGCA, GAPDH F: GGAGCGAGATCCCTCCAAAT, GAPDH R: GGCTGTTGTCATACTTCTCATGG. Relative levels of transcripts were determined using the $2^{-\Delta\Delta CT}$ Method relative to GAPDH reference gene.

miRNA sequence library preparation. MiRNA library was prepared using QIAseq miRNA Library Kit (Qiagen). Before proceeding to cDNA conversion, NGS adaptor containing unique identifiers were ligated to the each 3' and 5' end of miRNA separately. Subsequently, NGS RT initiator, universal sequence and reverse transcription (RT) primer contains a combined unique molecular index (UMI) were added to the reaction, RT primer binds to a region of the 3' adapter and enables conversion of the 3'/5' ligated miRNA to cDNA. After RT, cDNA products were undergone cleanup procedure using a QIAseq miRNA NGS (QMN) beads. Next, library amplification was performed with HotStarTaq DNA Polymerase and wet universal forward primers that assigned with unique index (IDP). Finally, final products were cleaned with QMN bead. The reaction setup, components, concentrations and technical points were followed as per manufacturer's instructions (QIAseq miRNA Library Kit Handbook for Illumina NGS Systems).

TNBC RNA-Seq data retrieval and analysis. The transcriptome data were retrieved from 360 TNBC and 88 normal from the Sequence Read Archive (SRA) database (<https://www.ncbi.nlm.nih.gov/sra/SRP157974>) using the SRA toolkit version 2.9.2 as previously described. Clinical characteristics of the patients and controls were described before¹³. Paired end RNASeq FASTQ files were subsequently pseudo aligned to the human genome and reads were counted using KALLISTO 0.42.1³⁰. The R-squared goodness-of-fit measure for linear regression models was used to identify the correlation between the expression of TGFB1 and the indicated lncRNA using Graphpad prism 8.0.

ChIP-seq analysis. Chip-Seq peak score for SMAD2/SMAD3 from the BT-549 model was retrieved from the GSE104352 study as described by Sundqvist et al.¹⁴. In brief, SMAD2/3 ChIP-seq was conducted on the BT-549 TNBC cell model at 1.5 h and 15 h. ChIP-seq data sets were then aligned using Bowtie (version 1.1.0), while binding regions were identified using MACS software Model based analysis of ChIP-seq (version 1.4.2). Chip-Seq peak score peaks were retrieved from GEO.

Statistical analysis. Correlation analysis were conducted in Graphpad Prism 8.0 software and an $R^2 \geq 0.2$ was considered significant (Graphpad Software, San Diego, CA). Two-tiled t test was used to compare analysis groups.

Received: 28 February 2021; Accepted: 12 July 2021

Published online: 29 July 2021

References

- Wakefield, L. M. & Roberts, A. B. TGF-beta signaling: Positive and negative effects on tumorigenesis. *Curr. Opin. Genet. Dev.* **12**(1), 22–29 (2002).
- Bierie, B. & Moses, H. L. Tumour microenvironment: TGFbeta: The molecular Jekyll and Hyde of cancer. *Nat. Rev. Cancer* **6**(7), 506–520 (2006).
- Ikushima, H. & Miyazono, K. TGFbeta signalling: A complex web in cancer progression. *Nat. Rev. Cancer* **10**(6), 415–424 (2010).
- Feng, X. H. & Derynck, R. Specificity and versatility in TGF-beta signaling through Smads. *Annu. Rev. Cell Dev. Biol.* **21**, 659–693 (2005).
- Javelaud, D. & Mauviel, A. Mammalian transforming growth factor-betas: Smad signaling and physio-pathological roles. *Int. J. Biochem. Cell Biol.* **36**(7), 1161–1165 (2004).
- Imamura, T., Hikita, A. & Inoue, Y. The roles of TGF-beta signaling in carcinogenesis and breast cancer metastasis. *Breast Cancer* **19**(2), 118–124 (2012).
- Kang, Y. et al. A multigenic program mediating breast cancer metastasis to bone. *Cancer Cell* **3**(6), 537–549 (2003).
- Katsuno, Y. et al. Bone morphogenetic protein signaling enhances invasion and bone metastasis of breast cancer cells through Smad pathway. *Oncogene* **27**(49), 6322–6333 (2008).
- Padua, D. et al. TGFbeta primes breast tumors for lung metastasis seeding through angiopoietin-like 4. *Cell* **133**(1), 66–77 (2008).
- Lamouille, S., Subramanyam, D., Bellochio, R. & Derynck, R. Regulation of epithelial–mesenchymal and mesenchymal–epithelial transitions by microRNAs. *Curr. Opin. Cell Biol.* **25**(2), 200–207 (2013).
- Hao, Y., Baker, D. & Ten Dijke, P. TGF-beta-mediated epithelial–mesenchymal transition and cancer metastasis. *Int. J. Mol. Sci.* **20**(11), 2767 (2019).

12. Vishnubalaji, R., Shaath, H., Elkord, E. & Alajez, N. M. Long non-coding RNA (lncRNA) transcriptional landscape in breast cancer identifies LINC01614 as non-favorable prognostic biomarker regulated by TGFbeta and focal adhesion kinase (FAK) signaling. *Cell Death Discov.* **5**, 109 (2019).
13. Jiang, Y. Z. *et al.* Genomic and transcriptomic landscape of triple-negative breast cancers: Subtypes and treatment strategies. *Cancer Cell* **35**(3), 428–440 (2019).
14. Sundqvist, A. *et al.* JUNB governs a feed-forward network of TGFbeta signaling that aggravates breast cancer invasion. *Nucleic Acids Res.* **46**(3), 1180–1195 (2018).
15. Bianchini, G., Balko, J. M., Mayer, I. A., Sanders, M. E. & Gianni, L. Triple-negative breast cancer: Challenges and opportunities of a heterogeneous disease. *Nat. Rev. Clin. Oncol.* **13**(11), 674–690 (2016).
16. Pelon, F. *et al.* Cancer-associated fibroblast heterogeneity in axillary lymph nodes drives metastases in breast cancer through complementary mechanisms. *Nat. Commun.* **11**(1), 404 (2020).
17. Hachim, I. Y., Hachim, M. Y., Lopez-Ozuna, V. M., Ali, S. & Lebrun, J. J. A dual prognostic role for the TGFbeta receptors in human breast cancer. *Hum. Pathol.* **57**, 140–151 (2016).
18. Ding, M. J. *et al.* Association between transforming growth factor-beta1 expression and the clinical features of triple negative breast cancer. *Oncol. Lett.* **11**(6), 4040–4044 (2016).
19. Bahhnassy, A. *et al.* Transforming growth factor-beta, insulin-like growth factor I/insulin-like growth factor I receptor and vascular endothelial growth factor-A: Prognostic and predictive markers in triple-negative and non-triple-negative breast cancer. *Mol. Med. Rep.* **12**(1), 851–864 (2015).
20. Ahmed, S. & Nawshad, A. Complexity in interpretation of embryonic epithelial–mesenchymal transition in response to transforming growth factor-beta signaling. *Cells Tissues Organs* **185**(1–3), 131–145 (2007).
21. Bashir, M., Damineni, S., Mukherjee, G. & Kondaiah, P. Activin-A signaling promotes epithelial–mesenchymal transition, invasion, and metastatic growth of breast cancer. *NPJ Breast Cancer* **1**, 15007 (2015).
22. Ho, J. Y. *et al.* MicroRNA-382-5p aggravates breast cancer progression by regulating the RERG/Ras/ERK signaling axis. *Oncotarget* **8**(14), 22443–22459 (2017).
23. Tian, Y. *et al.* MicroRNA181 serves an oncogenic role in breast cancer via the inhibition of SPRY4. *Mol. Med. Rep.* **18**(6), 5603–5613 (2018).
24. Liu, K. *et al.* SOX2 regulates multiple malignant processes of breast cancer development through the SOX2/miR-181a-5p, miR-30e-5p/TUSC3 axis. *Mol. Cancer* **16**(1), 62 (2017).
25. Cuk, K. *et al.* Circulating microRNAs in plasma as early detection markers for breast cancer. *Int. J. Cancer* **132**(7), 1602–1612 (2013).
26. Zhu, Y. *et al.* Reduced miR-128 in breast tumor-initiating cells induces chemotherapeutic resistance via Bmi-1 and ABCC5. *Clin. Cancer Res.* **17**(22), 7105–7115 (2011).
27. Masri, S. *et al.* The role of microRNA-128a in regulating TGFbeta signaling in letrozole-resistant breast cancer cells. *Breast Cancer Res. Treat.* **124**(1), 89–99 (2010).
28. Imani, S., Wu, R. C. & Fu, J. MicroRNA-34 family in breast cancer: From research to therapeutic potential. *J. Cancer* **9**(20), 3765–3775 (2018).
29. Vishnubalaji, R., Sasidharan Nair, V., Ouararhni, K., Elkord, E. & Alajez, N. M. Integrated transcriptome and pathway analyses revealed multiple activated pathways in breast cancer. *Front. Oncol.* **9**, 910 (2019).
30. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**(5), 525–527 (2016).
31. Elango, R. *et al.* MicroRNA expression profiling on paired primary and lymph node metastatic breast cancer revealed distinct microRNA profile associated with LNM. *Front. Oncol.* **10**, 756 (2020).
32. Shaath, H., Toor, S. M., Nair, V. S., Elkord, E. & Alajez, N. M. Transcriptomic analyses revealed systemic alterations in gene expression in circulation and tumor microenvironment of colorectal cancer patients. *Cancers (Basel)* **11**(12), 1944 (2019).
33. Emig, D. *et al.* AltAnalyze and DomainGraph: Analyzing and visualizing exon expression data. *Nucleic Acids Res.* **38**(Web Server issue), W755–762 (2010).

Acknowledgements

This work was supported by Qatar Biomedical Research Institute (QBRI) start-up fund (Grant No. QB13) for Nehad M. Alajez.

Author contributions

R.V. and R.E. performed experiments and wrote manuscript. N.M.A. conceived study, obtained funding, analyzed data, and finalized manuscript. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-94514-9>.

Correspondence and requests for materials should be addressed to N.M.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021