BMC Cancer

NSMCE2, a novel super-enhancer-regulated gene, is linked to poor prognosis and therapy resistance in breast cancer

Carolina Di Benedetto, Justin Oh, Zainab Choudhery, Weiquan Shi, Gilmer Valdes and Paola Betancur[*]

Abstract

Background: Despite today's advances in the treatment of cancer, breast cancer-related mortality remains high, in part due to the lack of effective targeted therapies against breast tumor types that do not respond to standard treatments. Therefore, identifying additional breast cancer molecular targets is urgently needed. Super-enhancers are large regions of open chromatin involved in the overactivation of oncogenes. Thus, inhibition of super-enhancers has become a focus in clinical trials for its therapeutic potential. Here, we aimed to identify novel super-enhancer dysregulated genes highly associated with breast cancer patients' poor prognosis and negative response to treatment.

Methods: Using existing datasets containing super-enhancer-associated genes identified in breast tumors and public databases comprising genomic and clinical information for breast cancer patients, we investigated whether highly expressed super-enhancer-associated genes correlate to breast cancer patients' poor prognosis and to patients' poor response to therapy. Our computational findings were experimentally confirmed in breast cancer cells by pharmacological SE disruption and gene silencing techniques.

Results: We bioinformatically identified two novel super-enhancer-associated genes – NSMCE2 and MAL2 – highly upregulated in breast tumors, for which high RNA levels significantly and specifically correlate with breast cancer patients' poor prognosis. Through in-vitro pharmacological super-enhancer disruption assays, we confirmed that super-enhancers upregulate NSMCE2 and MAL2 transcriptionally, and, through bioinformatics, we found that high levels of NSMCE2 strongly associate with patients' poor response to chemotherapy, especially for patients diagnosed with aggressive triple negative and HER2 positive tumor types. Finally, we showed that decreasing NSMCE2 gene expression increases breast cancer cells' sensitivity to chemotherapy treatment.

Conclusions: Our results indicate that moderating the transcript levels of NSMCE2 could improve patients' response to standard chemotherapy consequently improving disease outcome. Our approach offers a new avenue to identify a signature of tumor specific genes that are not frequently mutated but dysregulated by super-enhancers. As a result, this strategy can lead to the discovery of potential and novel pharmacological targets for improving targeted therapy and the treatment of breast cancer.

Keywords: Breast cancer, Super-enhancers, NSMCE2, MAL2, Therapy resistance

Background

Breast cancer is the most diagnosed cancer type in women and the second leading cause of cancer-related female deaths in the US [1]. Breast cancer is a heterogeneous disease, but the identification of shared mechanisms

*Correspondence: paola.betancur@ucsf.edu

Department of Radiation Oncology, University of California San Francisco, San Francisco, USA

Di Benedetto *et al. BMC Cancer*   (2022) 22:1056

Page 2 of 16

that drive tumor progression has significantly improved patients' treatment in the last decades. Breast tumors can be divided into subtypes using two parameters: (I) At the molecular level based on the protein expression of three receptors: estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2) [2]. Tumors that express hormone receptors ER and PR belong to the ER + PR + subtype, while hormone receptor negative tumors with elevated HER2 levels belong to the HER2 + subtype. A third subtype called Triple Negative (TN) includes tumors that do not express ER, PR and HER2. (II) At the RNA level, based on gene expression profiles, breast tumors can be divided into 5 subtypes: Luminal A, Luminal B, HER2-enriched, Basal and Normal-like [3–5]. Generally, Luminal tumors are characterized by ER gene expression. Luminal A tumors have higher ER gene expression levels than Luminal B tumors, whereas Luminal B tumors have higher gene expression levels of proliferation-related genes than Luminal A tumors. HER2-enriched, Basal and Normal-like tumors are characterized by either low or absent ER gene expression levels. In addition, HER2-enriched tumors highly express several genes in the HER2 amplicon at 17q22.24 including HER2. Basal tumors are characterized by high expression of keratins 5 and 17, laminin, and fatty acid binding protein 7. Moreover, tumors of this subtype do not express ER, PR and HER2 genes just like TN classified subtypes. Thus, Basal tumors are typically considered TN. Finally, the Normal-like classified tumors show high expression levels of many genes expressed by adipose tissue and other nonepithelial cell types and are characterized by gene expression patterns that are similar to those of normal breast tissue. Normal-like tumors are the less frequent subtype [4].

The subtype classification of breast tumors using these parameters in combination with other clinicopathological characteristics, such as age, tumor size, histological grade, and lymph node positivity, is clinically relevant for determining treatment recommendations [6, 7]. For instance, therapies aimed to block the activation of oncogenic receptors like ER and/or HER2 are available for the treatment of breast tumors expressing these receptors. While most patients with ER + breast tumors receive endocrine therapy alone (e.g., tamoxifen), patients with HER2 + tumors are usually treated with anti-HER2 therapy plus chemotherapy. As these molecular targets are absent in TN tumors, TN breast cancer patients are frequently treated with chemotherapy.

Despite today's current technology and advancements for the treatment of cancer, breast cancer-related mortality, especially for TN breast cancer, remains high [7, 8]. One of the factors contributing to this is the lack of targeted therapies against TN tumors, which maintains TN cancers as an unmet clinical need [9]. Another factor contributing to breast cancer recurrence and high mortality is tumor resistance to targeted therapies and/or chemotherapy, which favors the selection of less susceptible tumor cells over time [10]. Thus, approaches aimed at identifying novel breast cancer targets of resistance, mostly for TN subtype, are urgently needed.

Super-enhancers (SEs) are long stretches of DNA containing clusters of enhancers that drive expression of genes of cell identity [11, 12] and of other pivotal functions such as homeostasis [13–15]. Because several oncogenic pathways [16, 17] and tumor immunosuppressive pathways [18] are also regulated by SEs, studying the role of SEs on the dysregulation of tumor promoting genes has gained attention over the last decade. SEs are typically enriched in acetylation at histone 3 lysine 27 (H3K27Ac) [19], a hallmark of open chromatin that allows DNA to be accessible to the transcriptional machinery. Lysine acetylation in histones is recognized by Bromodomain and Extraterminal (BET) proteins through epigenetic reader domains called bromodomains (BRDs). Bromodomains consists of a left-handed bundle of four alpha helices that form a hydrophobic pocket, which is where the protein binds to the acetylated lysine residue [20]. BRD4 is a member of the BET protein family that binds preferentially to hyperacetylated chromatin regions, hence SEs, to facilitate rapid gene transcription by linking enhancers or promoters to the TEFb (transcription elongation factor) complex [21]. Previous studies have shown that disrupting BRD4 selectively affects the expression of genes that are associated with SEs [16]. Since SEs are known to regulate the expression of several oncogenes, targeting these regulatory regions has been proposed as an anti-cancer therapy. This approach has the potential to simultaneously target multiple dysregulated pathways at the transcriptional level across cancer types.

In this work, by using publicly available databases containing gene expression data and clinical information obtained from cancer patients, we identified two SE-associated genes – NSMCE2 and MAL2 – that are enriched in breast tumors when compared to non-cancerous tissue and that are linked to breast cancer patients' poor prognosis. We show that pharmacological disruption of SE architecture using BET inhibitors decreases expression of these genes in vitro. Moreover, through analyses of tumor microarray data linked to therapy response of cancer patients, we found that high expression of NSMCE2 predicts poor response to chemotherapeutic drugs for patients diagnosed with breast cancer, especially patients diagnosed with aggressive TN or HER2 + tumors. Our results suggest that inhibiting NSMCE2 function by pharmacological inhibition or gene expression reduction (e.g., through SE blockade), could sensitize breast cancer

Di Benedetto *et al. BMC Cancer*   (2022) 22:1056

Page 3 of 16

cells to chemotherapy in specific cohorts of breast cancer patients that do not respond to standard antitumor drugs. This combinatorial approach has the potential to improve disease outcome for patients that typically would not survive after many years of cancer treatment and thus needs further investigation.

Methods

Analyses of cancer patients' datasets

TCGA breast cancer dataset [5] was downloaded from the UCSC XenaBrowser Platform [22]. METABRIC data [23] were downloaded from cBioPortal [24]. UCSC Toil RNAseq study containing TCGA and GTEX gene expression was used to compare gene expression in tumor and normal tissues [25], and it was downloaded from the UCSC XenaBrowser. Gene expression correlation matrix for SE-associated genes was done by hierarchical clustering using the package 'corrplot' on RStudio. Genetic alterations in breast cancer patients on the TCGA dataset or in breast cancer cell lines on The Cell Line Encyclopedia [26] were analyzed using cBioPortal. Kaplan–Meier and Multivariate Cox analyses were both performed using the packages 'survival' and 'survminer' on RStudio. Survival package is required for both statistical analyses, while survminer is used to generate the graphics. Samples were divided in two groups based on the median gene expression. Kaplan–Meier curves were compared using the log-rank test. Association of gene expression and response to therapy in breast cancer patients was done using ROC Plotter [27].

Cell culture

MCF7 (ER+PR+), HCC1954 (HER2+), BT549 (TN) and 293 T cell lines were obtained from ATCC. Hs578T (TN) and MDAMB231 (TN) were kindly provided by Dr. Mary Helen Barcellos-Hoff and Dr. Denise Muñoz, respectively, at UCSF. MCF7, HCC1954, BT549 and Hs578T cells were cultured in RPMI medium (Gibco), while MDAMB231 and 293 T cells were cultured in DMEM medium (Gibco). Media were supplemented with 10% heat inactivated FBS (Atlanta Biologicals), 100 units/mL Penicillin and 100 μg/mL Streptomycin (Gibco). BT549 and Hs578T were also supplemented with 0.01 mg/mL insulin (Sigma). Cells were grown at 37 °C in a humified atmosphere at 5% CO_2.

Drugs

For BET inhibition, cells were treated with 1 μM JQ1 (Sigma) or 1 μM IBET-151 (GSK1210151A, Selleckchem) dissolved as 10 mM in DMSO; DMSO was used as vehicle control. For treatment with chemotherapeutic drugs, 10 mM Doxorubicin hydrochloride (Sigma) dissolved in water and 1 mM Paclitaxel from Taxus brevifolia (Sigma) dissolved in DMSO were added to the cells' media at the concentrations indicated in the figures. DMSO (vehicle control) was used at a concentration matching the maximum concentration of paclitaxel used in a given experiment.

MTT assay

Cells were seeded in 96-well plates (at 7000 cells/well) and cultured overnight. Then, cells were exposed to drug treatments for the indicated times. MTT (Sigma) was added to each well following the manufacturer's instruction. After incubation for 4 h at 37 °C, formazan crystals were dissolved in 100% DMSO, and absorbance was read at 570 nm on a Biotek plate reader. Combination Index (CI) was calculated for JQ1 in combination with each chemotherapeutic drug using the Response Additivity approach [28] as follows:

$$CI = \left(E_{JQ1} + E_{chemo}\right)/E_{JQ1+chemo}$$

Where E_{JQ1} is the effect of JQ1 on cell viability reduction, E_{chemo} is the effect of a particular concentration of chemotherapeutic drug on cell viability reduction, and $E_{JQ1+chemo}$ is the effect of the combined treatment on cell viability reduction for that concentration of chemotherapeutic drug. $CI=1$ indicates additive effect, $CI<1$ indicates synergistic effect, and $CI>1$ indicates antagonistic effect.

RNA extraction, cDNA synthesis and qPCR

Total RNA was extracted using RNeasy Plus kits (Qiagen). cDNA was reversed transcribed using SuperScript III First-Strand Synthesis SuperMix (Invitrogen) and then amplified on the QuantStudio 5 Real Time PCR System (Applied Biosystems). Specific primers designed to amplify the gene of interest were combined with cDNA and Platinum® SYBR® Green qPCR SuperMix-UDG (Invitrogen) following the manufacturer's instructions. qPCR was carried out using the following method: an initial incubation at 50 °C for 2 min followed by incubation at 95 °C for 2 min; 40 cycles at 95 °C for 10 s followed by incubation at 60 °C for 30 s; and a final step for melting curve generation. Results were analyzed using the comparative Ct method [29]. Values were normalized to β-Actin expression. Primers used in this study are:

NSMCE2: AGGACGCCATTGTTCGCAT, GCTACAGCCAATTTGAGGGCA
MAL2: TCCGTGACAGCGTTTTTCTTTTC, TGCTTCCAATAAAAAGGCTCCAA
β-Actin: TCCCTGGAGAAGAGCTACG, GTAGTTTCGTGGATGCCACA

Di Benedetto *et al. BMC Cancer*   (2022) 22:1056

Page 4 of 16

Apoptosis assay by flow cytometry

Cells were collected, washed with HBSS and stained with APC Annexin V (Biolegend, Catalogue number 640919) in Annexin V binding buffer (Biolegend) for 30 min at room temperature. After washing, Annexin V positive events were quantified using a Northern Light cytometer (Cytek). Data were analyzed on SpectroFlo software (Cytek).

Lentiviral transduction for NSMCE2 knockdown

NSMCE2 was knockdown using GIPZ lentiviral shRNA constructs (Horizon Discovery). Mature antisense sequences used for gene knockdown, are as follows:

shRNA Non-silencing control #RHS4346: ATC TCGCTTGGGCGAGAGTAAG
shRNA NSMCE2 #V2LHS_179143: TACATAATG GTTTAGTTGC
shRNA NSMCE2 #V3LHS_354626: TATTGTAGA TTGAACAGCC
shRNA NSMCE2 #V3LHS_354628: ATTTTGAAA GTCTGCATCA

Lentiviral particles were generated by co-transfection of 293 T cells with the packaging plasmids psPAX2 and pMD2G, with either a mix of the three pGIPZ-shNS-MCE2 specific constructs or the pGIPZ-shControl plasmid using TurboFect Transfection Reagent (Life Technologies). Lentiviral particles were harvested 48 h after transfection in the culture media and used to infect target cells. After 24 h of incubation, transduced cells were selected with puromycin. Knockdown efficiency was quantified by qPCR.

Statistical analysis

All experiments were done in triplicates unless otherwise indicated. Results were plotted as mean values ± standard deviation using GraphPad Prism7 and statistically analyzed using ANOVA followed by Dunnet's or Bonferroni's post-test, as appropriate. *P*-values less than 0.05 were considered statistically significant.

Data availability

The dbGAP accession number for the PDX breast tumor H3K27ac ChIP-Seq datasets previously generated is phs001264.v1.p1 [18].
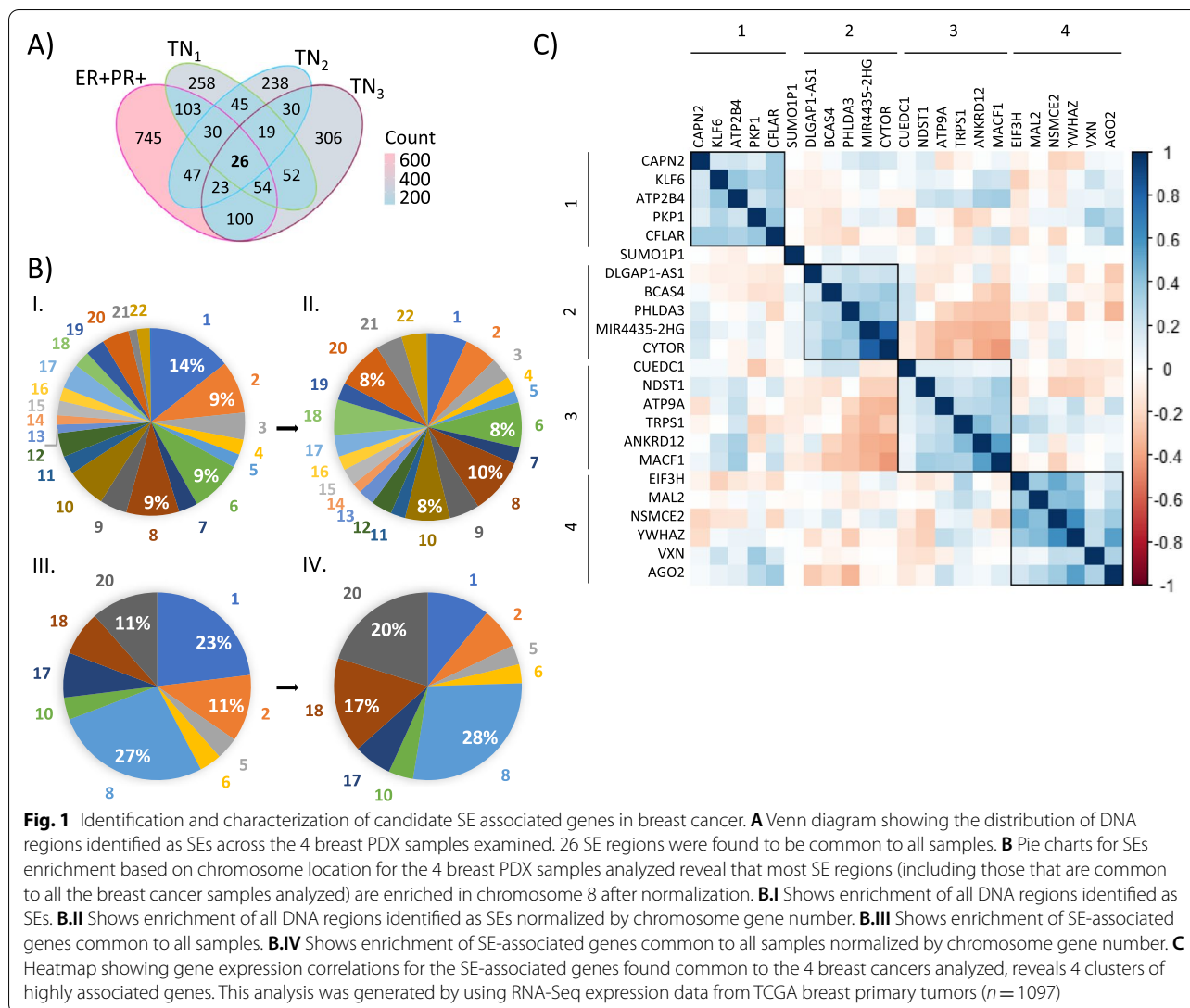
Results

Identification and characterization of candidate SE-associated genes in breast cancer

In order to find genes associated to SEs, first we identified non-coding genomic regions carrying SEs from a dataset generated by applying the Rank Ordering of Super-Enhancers (ROSE) algorithm on H3K27Ac enrichment data [12, 16, 30]. H3K27Ac data was obtained from chromatin immunoprecipitation experiments followed by high throughput sequencing analyses (ChIP-Seq) performed in 4 patient-derived Xeno transplant (PDX) breast tumors from a previous study [18]. We found that a tumor sample classified as hormone-receptor positive (ER + PR +) has 1194 DNA regions identified as SEs associated to 1128 genes, while the remaining three TN tumors contained 640, 527 and 651 SEs associated to 610, 458, and 587 genes, respectively (Fig. 1A and Additional file. 1). Our annotation analyses revealed that from the 3012 SEs identified, the majority are located either in chromosome 1 (14%), 8 (9%), 2 (9%), or 6 (9%) (Fig. 1B, panel I). Since chromosome gene number widely varies across chromosomes, SE enrichment based on chromosome location was normalized by chromosome gene number (Fig. 1B, panel II). Interestingly, when the number of genes per chromosome is considered, we found that SEs are also enriched in chromosome 8 (10%). From the total SE-associated genes, we identified 26 SE-associated genes common to all breast tumors analyzed (Fig. 1A). Again, we found that most of these 26 SE-associated genes are located in chromosomes 8 (27%) or 1 (23%) (Fig. 1B, panel III). Moreover, when gene number per chromosome is considered, common breast cancer SE-associated genes remain particularly enriched in chromosome 8 (28%) (Fig. 1B, panel IV). Given that chromosome 8 is known to contain many genes with tumor progression roles [31, 32], our findings suggest SEs may be a frequent mechanism that dysregulates expression of oncogenes in breast tumors.

In order to investigate whether within these 26 SE-associated genes there are groups of genes that share mechanisms of gene regulation or interact in similar pathways, we performed a correlation matrix using hierarchical clustering of gene expression information from breast tumors. For this, we used bulk RNA-sequencing (RNA-Seq) data from breast tumors available for 23 out of the 26 genes on The Cancer Genome Atlas (TCGA) database [5]. Our correlation analysis identified four clusters of genes whose expression levels are significantly correlated (Fig. 1C). From the hierarchical clustering we noted [1] a strong positive correlation between expression of CYTOR and MIR4435-2HG – two homolog non-coding RNA genes reported to be correlated with patients' poor prognosis in glioblastoma and in low-grade glioma [33] – and [2] an anticorrelation between expression of genes in the second and third clusters. Next, we performed a pathway enrichment analysis using Metascape [34], a web tool that combines functional

Di Benedetto *et al. BMC Cancer*   (2022) 22:1056

Page 5 of 16

Fig. 1 Identification and characterization of candidate SE associated genes in breast cancer. **A** Venn diagram showing the distribution of DNA regions identified as SEs across the 4 breast PDX samples examined. 26 SE regions were found to be common to all samples. **B** Pie charts for SEs enrichment based on chromosome location for the 4 breast PDX samples analyzed reveal that most SE regions (including those that are common to all the breast cancer samples analyzed) are enriched in chromosome 8 after normalization. **B.I** Shows enrichment of all DNA regions identified as SEs. **B.II** Shows enrichment of all DNA regions identified as SEs normalized by chromosome gene number. **B.III** Shows enrichment of SE-associated genes common to all samples. **B.IV** Shows enrichment of SE-associated genes common to all samples normalized by chromosome gene number. **C** Heatmap showing gene expression correlations for the SE-associated genes found common to the 4 breast cancers analyzed, reveals 4 clusters of highly associated genes. This analysis was generated by using RNA-Seq expression data from TCGA breast primary tumors ($n = 1097$)

enrichment, interactome (e.g., protein to protein, protein to DNA) analysis and gene annotation leveraging over 40 independent knowledgebases. Through Metascape, we found that genes grouped in the third cluster are target genes for the transcription factor NF1, a tumor suppressor gene that has been reported to be associated with an increase in breast cancer risk when lost [35].

To further investigate the causes of gene expression correlation among the SE-associated genes in breast tumors, we analyzed co-occurrence of genetic alterations of genes within each cluster. For this, we used cBioPortal [24], an open-source platform that facilitates the interactive exploratory analysis and visualization of large-scale cancer genomics datasets. Using cBioPortal on TCGA genomic data, we observed that highly co-expressed genes in cluster 1 (CAPN2, ATP2B4 and PKP1) and cluster 4 (EIF3H, MAL2, NSMCE2,

YWHAZ, VXN and AGO2) have a frequency of genetic alterations between 8 to 10% and 6 to 14%, respectively, mainly due to co-occurring DNA amplifications (Fig. S1). CAPN2, ATP2B4 and PKP1 are located on chromosome 1 between the cytogenetic bands 1q32.1 and 1q41, while genes in cluster 4 (except for VXN) are located on chromosome 8 in a region spanning cytogenetic bands 8q22.3 to 8q24.3. Importantly, these regions on chromosomes 1 and 8 are frequently amplified in cancer disease [5, 36, 37]. In line with our findings, recent reports show that SEs are frequently associated with genes located within large tandem duplications in breast cancers [38]. Overall, our gene correlation analyses suggest that genes found in clusters 1 and 4 have SE-mediated enhanced transcription that help increase transcript levels of genes that are highly amplified and potentially involved in breast cancer development.

Di Benedetto *et al. BMC Cancer*   (2022) 22:1056

Page 6 of 16

High gene expression levels of NSMCE2 and MAL2 SE-associated genes in breast tumors are linked to patients' poor prognosis in breast cancer

Previously it has been reported that SEs can enhance the expression of tumor-associated genes. For instance, in T-cell acute lymphoblastic leukemia, somatic mutations create a SE upstream of the TAL1 oncogene [17]. Recently, by comparing breast tumor samples with healthy breast tissue samples, we found that the immunosuppressive signal CD47 is transcriptionally upregulated by an SE [18]. Thus, we decided to investigate whether the gene expression levels of our 26 SE-associated candidate genes are higher in breast tumors when compared to healthy samples by using the TCGA database and the Genotype-Tissue Expression (GTEX) study [25]. TCGA compiles RNA-Seq data of tumors or matched adjacent tissue (considered "healthy" tissue due to absence of cancerous characteristics) obtained from cancer patients, for whom clinical information is also available. Such information includes histological grade and survival data. GTEX provides open access to bulk RNA-Seq gene expression data from non-diseased tissues obtained from donors. Mining these datasets, we found RNA-Seq data available for 21 out of the 26 SE-associated genes of our interest (RNA-Seq data was not available for a pseudogene nor for four non-coding RNA genes). From these 21 genes, we found that 9 showed significantly higher expression in breast primary tumors when compared to matched adjacent normal tissue or to breast normal tissue samples from the GTEX study (Fig. 2A and Table S1). These data suggest that SEs could be driving the overexpression of these 9 genes in breast tumors.

Next, to investigate whether high levels of gene expression in breast tumors for any of the 26 SE-associated genes has clinical relevance, we performed a univariate Kaplan–Meier analysis using RNA-Seq data linked to patients' survival outcomes available for 1097 breast primary tumors through TCGA. We found that high transcript levels of 3 genes – NSMCE2, MAL2 and EIF3H – significantly correlate with worse overall survival in breast cancer patients (Log-rank test, p-value < 0.05) (Fig. 2B). To validate these results on an independent dataset, we performed the same univariate Kaplan–Meier analysis using the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) dataset, which contains microarray gene expression data from 2000 breast primary tumors along with long-term clinical follow-up [23]. On the METABRIC dataset, high expression of MAL2 significantly associates with worse overall survival of breast cancer patients, while high NSMCE2 follows the same trend without reaching statistical significance (Fig. 2B). However, on the METABRIC patient

cohort, high expression levels of both genes are significantly associated with reduced probability of relapse-free survival – the length of time a cancer patient survives without any signs or symptoms of cancer after the end of primary treatment – thus, confirming that NSMCE2 and MAL2 expression are linked to poor prognosis in breast cancer. On the other hand, we did not observe a correlation between high EIF3H expression and breast cancer patients' negative prognosis on the METABRIC dataset, even though both TCGA and METABRIC datasets do not differ on the probability of overall survival for their respective patients' cohorts (Fig. S2A) nor on other demographic characteristics [39]. Since we found that high levels of NSMCE2 and MAL2 gene expression highly correlate with patients' negative prognosis, we focused our subsequent analyses on these two genes.

To further confirm that the association of high levels of gene expression with patients' negative prognosis is independent of other risk factors (covariates), we constructed a multivariate Cox model of survival using TCGA data, which includes age, stage and RNA expression subtype as covariates. We observed that, after adjusting for these clinical covariates, higher NSMCE2 and MAL2 gene expression levels remained significantly associated with patients' poor prognosis (Fig. 2C). NSMCE2 encodes a member of a family of E3 small ubiquitin-related modifier (SUMO) ligases, part of the structural maintenance of chromosomes 5/6 (SMC 5/6) complex. This protein complex plays key roles in DNA double-strand break repair by homologous recombination and in chromosome segregation during cell division [40–42]. MAL2 is a multispan transmembrane protein required for transcytosis, an intracellular transport pathway used to deliver membrane-bound proteins and exogenous cargos from the basolateral to the apical surface [43].

Next, we investigated in other cancer types (apart from breast cancer) whether expression of NSMCE2 and MAL2 is higher in tumor samples when compared to normal tissue and whether high levels of any of these two genes associate with poor survival. To approach this, we compared RNA levels from an array of cancer samples from TCGA Pan-Cancer with RNA levels from corresponding normal tissue samples from GTEX and TCGA studies as we did for breast cancer. In line with what we observed in breast cancer, NSMCE2 or MAL2 are significantly higher in the cancer samples compared to the normal corresponding samples in most of the cancer types studied (Fig. S2B). When we asked whether their overexpression is also correlated with patients' poor survival in other cancer types from the TCGA Pan-Cancer dataset, we observed that high NSMCE2 RNA expression levels associate with worse

Di Benedetto *et al. BMC Cancer*   (2022) 22:1056

Page 7 of 16

overall survival in 6 other cancer types: pancreatic ductal adenocarcinoma, uveal melanoma, kidney papillary cell carcinoma, lung adenocarcinoma, sarcoma and brain lower grade glioma (Table S2). For MAL2, we found that, while high gene expression is linked to shorter survival in 4 cancer types (pancreatic ductal adenocarcinoma, uveal melanoma, thymoma and uterine corpus endometrial carcinoma), it is linked to better survival in bladder urothelial and head and neck squamous cell carcinoma. Thus, although NSMCE2 and MAL2 are highly expressed in Pan-Cancer tumors, their association to patients' negative survival outcomes is specific to certain cancer types, including breast cancer.

NSMCE2 and MAL2 are regulated by SEs in ER + PR +, HER2 + and TN breast cancer cells

Since we found that SEs are associated to NSMCE2 and MAL2 expression in breast tumors, we hypothesized that in breast cancer cells, expression of these genes should be reduced when disrupting the SE-enhancing function by blocking the binding of BRD4 to SEs with the inhibitors JQ1 or IBET151 [44, 45]. To test this, we used 5 breast cancer cell lines representing different breast cancer subtypes: MCF7 (ER + PR +), HCC1954 (HER2 +), BT549, MDAMB231 and Hs578T (TN). We also chose these cell lines because NSMCE2 and MAL2 are amplified in all except in MDAMB231 (visualized using cBioPortal on data from the Cell Line Encyclopedia [26], Fig. S3A), which recapitulates our observation in breast cancer patients' datasets where these two genes were frequently amplified.

Treatment of ER + PR +, HER2 +, and TN cancer cell lines with non-toxic concentration of either JQ1 or IBET151 (Fig. S3B) had the following effects on gene expression levels (Fig. 3): We observed a reduction on NSMCE2 RNA in all cell lines treated with BET inhibitors, confirming regulation of NSMCE2 by SEs. MAL2 expression was only reduced by BET inhibition in cell

lines representing HER2 + and TN breast cancer subtypes, but it was curiously increased after 24 h of treatment in the TN BT549, perhaps as an indirect effect of BET inhibition on downregulating MAL2 repressors. In summary, pharmacological blockade of SEs confirmed NSMCE2 and MAL2 are associated with SEs in breast cancer cells.

High gene expression levels of NSMCE2 correlate with breast cancer patients' poor response to chemotherapy.

Given that our data shows that high NSMCE2 and MAL2 expression are associated with breast cancer patients' poor survival outcomes, we next investigated if high levels of these markers could also be negatively associated with response to standard cancer therapies. To investigate this, we used the online platform ROC Plotter [46], which integrates multiple transcriptome-level gene expression datasets available through GEO into a single database containing information for 3104 breast cancer patients linked to corresponding response data for a range of treatments, including endocrine therapies, anti-HER2 therapies and chemotherapeutic drugs [27]. The transcriptome data is obtained from patients' biopsies prior to treatment and patients are grouped into responders and non-responders for a given treatment based on clinical characteristics. Responder and non-responder patients are compared via two diverse statistical approaches: The Mann–Whitney test (non-parametric T-test) and the Receiver Operating Characteristic (ROC) test. The ROC test measures how much 'the gene expression level' model is capable of distinguishing between the two classes: responders and non-responders. ROC plots are useful to compute the Area Under the Curve value (AUC), which shows the predictive power of the gene. The higher the AUC value, the better the model is at distinguishing between patients who are responders versus non-responders. For cancer biomarkers with potential clinical utility, the AUC value should be higher

(See figure on next page.)

Fig. 2 High NSMCE2 and MAL2 RNA expression in tumors is linked to breast cancer patients' poor prognosis. **A** Box plots showing that gene expression levels are significantly higher in breast tumor samples when compared to normal samples for 9 SE-associated genes. Expression levels information for normal samples (Normal Tissue GTEX and Solid Tissue Normal TCGA) was obtained from GTEX and TCGA RNA-Seq datasets, respectively. Expression levels from primary breast tumors (Primary Tumor TCGA) were obtained from TCGA RNA-Seq datasets. Gene expression levels were compared by ANOVA. **B** Univariate Kaplan–Meier plots showing breast cancer patients' survival probability over time based on gene expression for each NSMCE2, MAL2, or EIF3H. These analyses were performed using breast primary tumors RNA-Seq data mined from two independent datasets, TCGA ($n = 1097$) or METABRIC ($n = 1904$). Red and blue curves represent samples that show high and low gene expression levels, respectively, relative to the median expression value for each gene. High levels of either NSMCE2 or MAL2 were found to be significantly correlated with breast cancer patients' poor prognosis in the two datasets. Plots were analyzed using the Log-rank test. OS = overall survival, RFS = relapse free survival. **C** Multivariate Cox model of overall survival for NSMCE2 and MAL2 RNA expression in TCGA breast primary tumors including age, stage and tumor subtype as covariates show that high levels of either NSMCE2 or MAL2 are significantly correlated with patients' poor prognosis
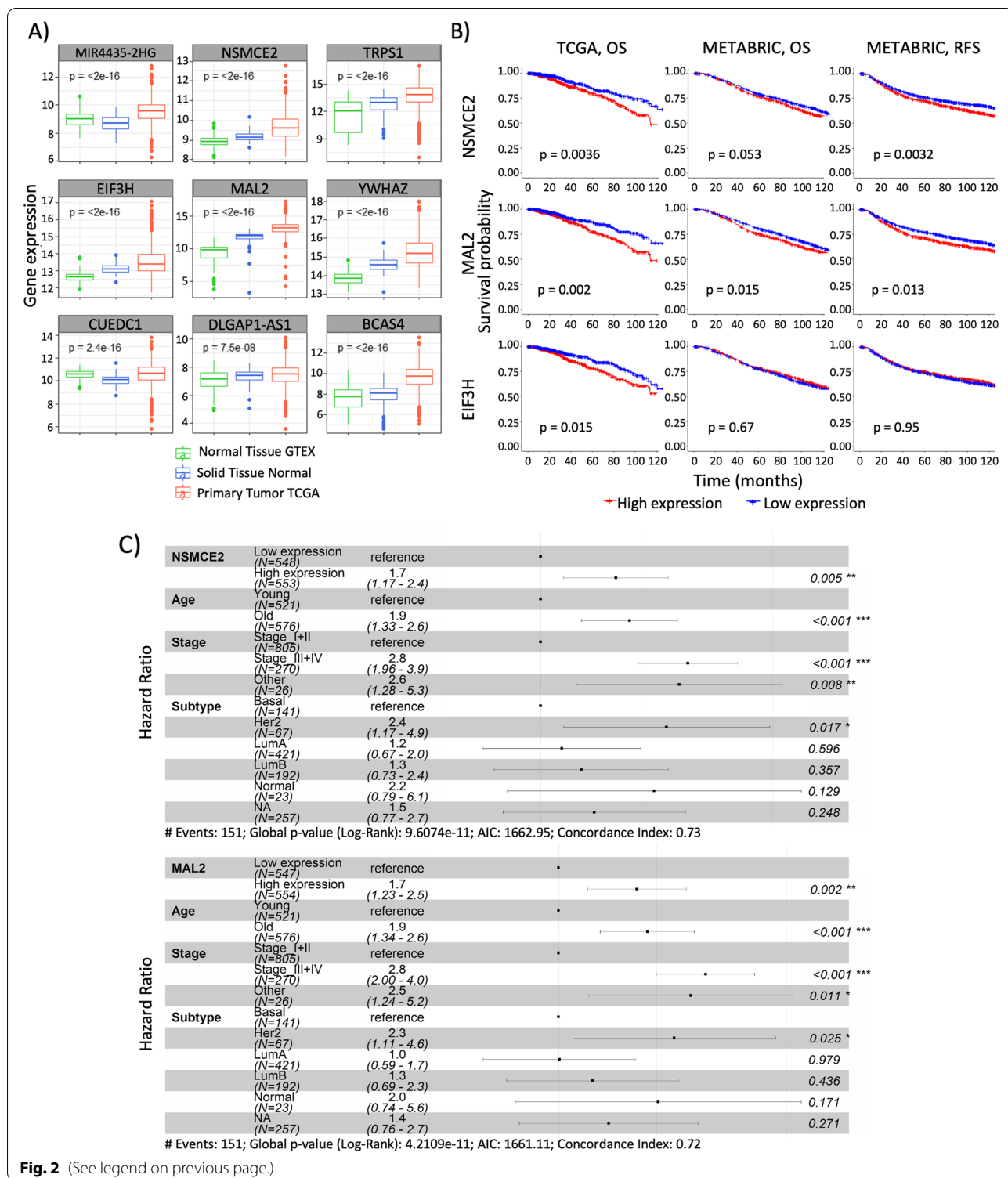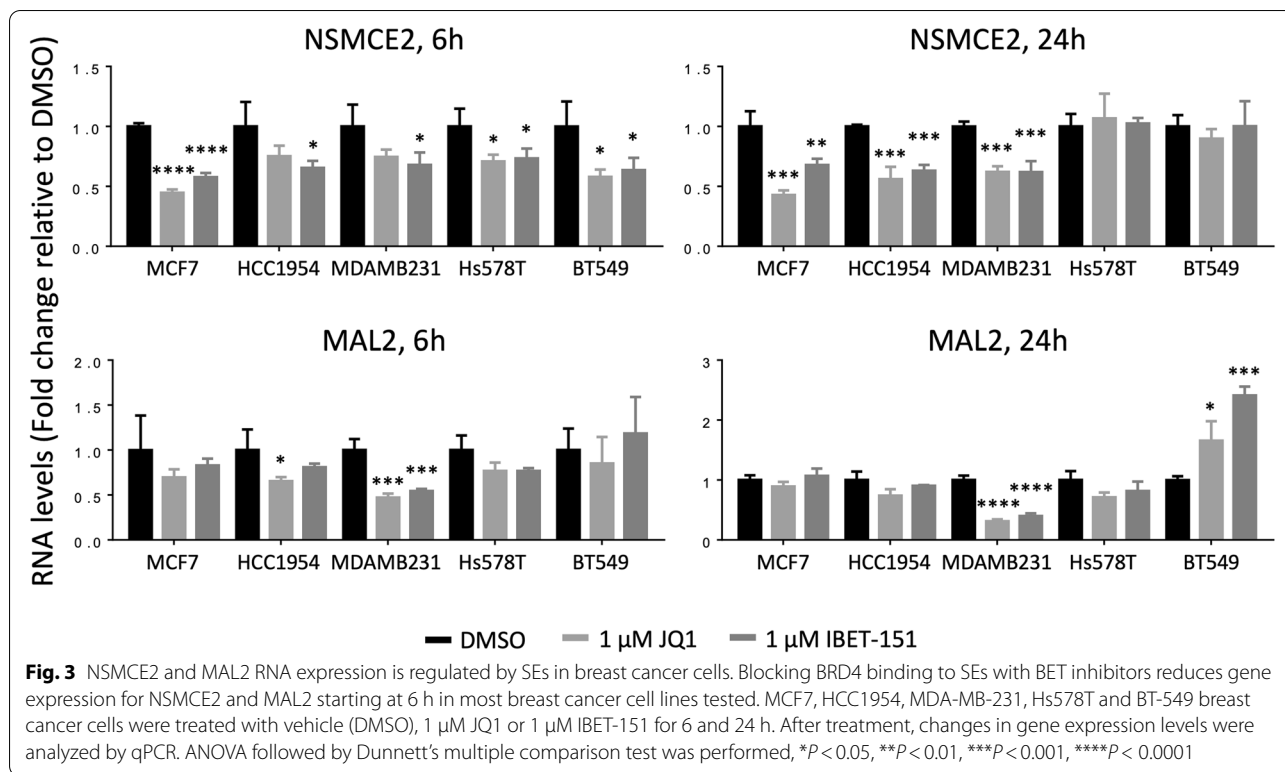
Di Benedetto *et al. BMC Cancer*   (2022) 22:1056

Page 8 of 16

Fig. 2 (See legend on previous page.)

than 0.6, while AUC values higher than 0.7 are obtained by top quality cancer biomarkers [27].

We focused our analysis on breast cancer patients who received chemotherapy due to the larger sample size in

the dataset compared to that for patients who received endocrine or anti-HER2 therapies ($n = 507$ versus $n = 160$ and $n = 151$, respectively). Using ROC Plotter, we found that high levels of NSMCE2 gene expression correlate

Fig. 3 NSMCE2 and MAL2 RNA expression is regulated by SEs in breast cancer cells. Blocking BRD4 binding to SEs with BET inhibitors reduces gene expression for NSMCE2 and MAL2 starting at 6 h in most breast cancer cell lines tested. MCF7, HCC1954, MDA-MB-231, Hs578T and BT-549 breast cancer cells were treated with vehicle (DMSO), 1 μM JQ1 or 1 μM IBET-151 for 6 and 24 h. After treatment, changes in gene expression levels were analyzed by qPCR. ANOVA followed by Dunnett's multiple comparison test was performed, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$

with lower pathological complete response to chemotherapy of breast cancer patients (Fig. 4A, Mann–Whitney test p-value $= 0.000032$ and AUC $= 0.617$). Moreover, we found that high levels of NSMCE2 gene expression strongly correlate with poor response to chemotherapy for patients diagnosed with grade III breast cancer (Fig. 4B, Mann–Whitney test p-value $= 0.00099$ and AUC $= 0.655$). We also performed the same analyses for the 21 SE-associated genes we found in breast cancer and for which gene expression data is available on ROC plotter. We did not see significant correlation (AUC < 0.6, Mann–Whitney test p-value < 0.05) between high levels of gene expression and breast cancer patients' poor response to chemotherapy for any of the other SE-associated genes analyzed (including MAL2) when analyzing the entire database or when focusing on Grade III breast cancer (Table S3). This indicates that the observed correlation between high NSMCE2 gene expression levels and poor response to chemotherapy is specific.
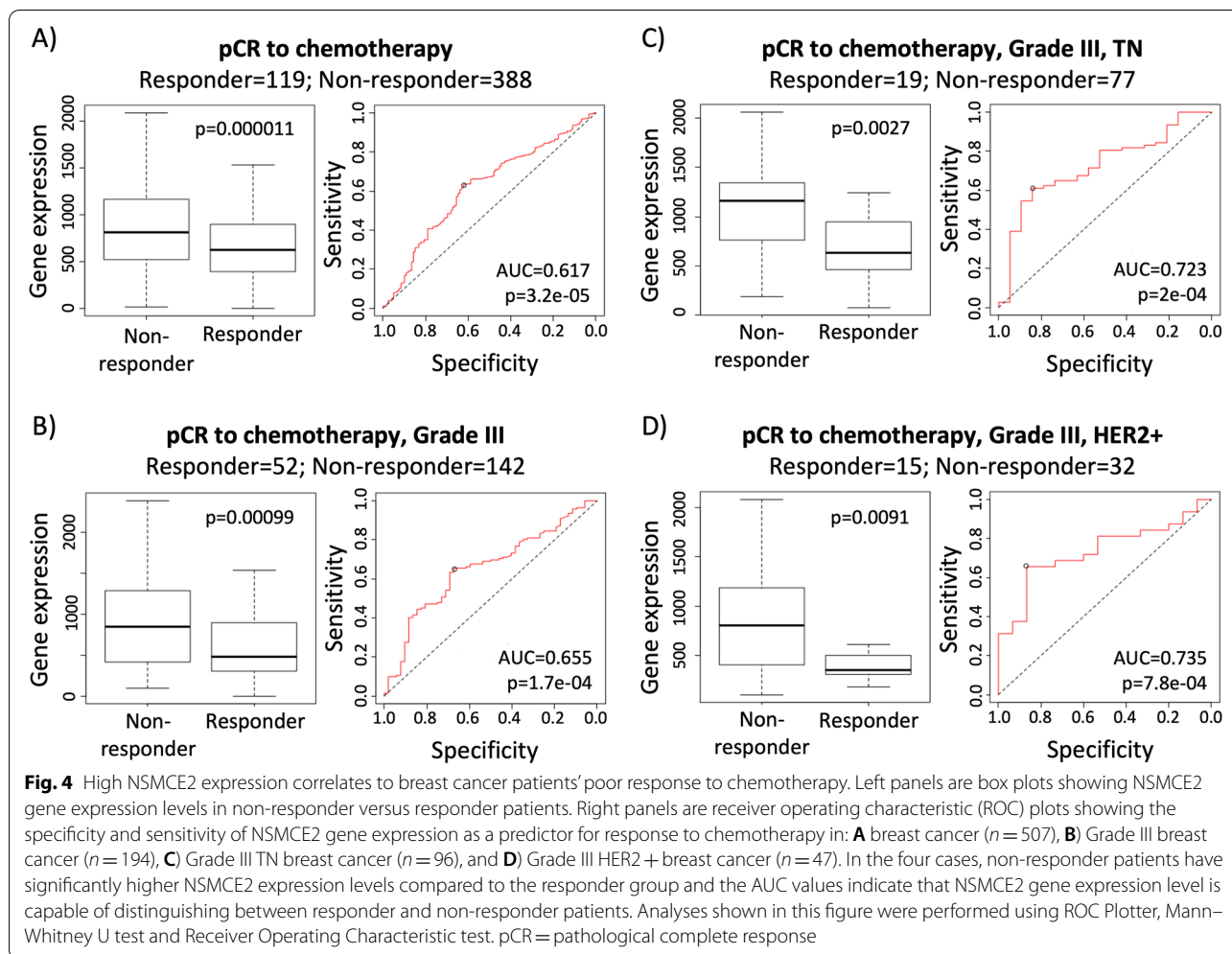
Since the drug regimen for breast cancer patients is highly determined by the tumor molecular subtype, we next analyzed the correlation between high NSMCE2 gene expression levels with response to chemotherapy for each ER + PR +, HER2 + or TN tumor subtypes. Here, we observed a stronger correlation between high levels of NSMCE2 and patients' poor response to chemotherapy when the tumors were classified as grade III TN or grade III HER2 + tumors (Fig. 4C and D, AUC $= 0.723$, Mann–Whitney test p-value $= 0.0027$ and AUC $= 0.735$, Mann–Whitney test p-value $= 0.0091$, respectively). Thus, our results indicate that breast cancer patients with high NSMCE2 RNA expression in aggressive breast tumors of either TN or HER2 + molecular subtype respond poorly to chemotherapeutic drugs. Since the AUC values obtained for NSMCE2 are usually values observed for top quality cancer biomarkers [27], our findings suggest that NSMCE2 expression could potentially be used to pinpoint a group of patients (especially those diagnosed with grade III TN or HER2 + breast cancer) that may not respond to chemotherapy.

Lastly, using the ROC Plotter tool, we found that NSMCE2 high expression levels do not correlate with patients' response to therapy for ovarian cancer, colorectal cancer, and glioblastoma [47, 48] (data not shown), thus, confirming that high levels of NSMCE2 gene expression are specifically associated with breast cancer patients' poor response to chemotherapy.

Lowering NSMCE2 transcript levels sensitizes breast cancer cells to chemotherapeutic agents.

So far, our work shows that [1] NSMCE2 and MAL2 genes are regulated by SEs at the transcript level in breast cancers; [2] the expression of these genes can be reduced by BET inhibition; and [3] high NSMCE2 expression

Fig. 4 High NSMCE2 expression correlates to breast cancer patients' poor response to chemotherapy. Left panels are box plots showing NSMCE2 gene expression levels in non-responder versus responder patients. Right panels are receiver operating characteristic (ROC) plots showing the specificity and sensitivity of NSMCE2 gene expression as a predictor for response to chemotherapy in: **A** breast cancer ($n = 507$), **B**) Grade III breast cancer ($n = 194$), **C**) Grade III TN breast cancer ($n = 96$), and **D**) Grade III HER2 + breast cancer ($n = 47$). In the four cases, non-responder patients have significantly higher NSMCE2 expression levels compared to the responder group and the AUC values indicate that NSMCE2 gene expression level is capable of distinguishing between responder and non-responder patients. Analyses shown in this figure were performed using ROC Plotter, Mann–Whitney U test and Receiver Operating Characteristic test. pCR = pathological complete response

correlates to poor response to chemotherapy in grade III TN or HER2 + breast cancer patients. The latter result indicates that our findings may have important clinical implications. Most chemotherapeutic agents work by activating the program of apoptosis in cancer cells through DNA damage induction or through cell cycle inhibition. Since NSMCE2 is known to be required for DNA damage repair at different steps of the process and for chromosomal segregation during mitosis [40–42, 49–51], NSMCE2 may inhibit chemotherapy-induced apoptosis, thus contributing to the therapy resistance we observed in breast cancer patients showing high NSMCE2 gene expression. To test whether reduction of NSMCE2 transcript levels by SE disruption can sensitize breast cancer cells to chemotherapy-induced apoptosis, we treated breast cancer cells with BET inhibitor JQ1 and the standard chemotherapeutical drugs doxorubicin and paclitaxel. Doxorubicin is part of the anthracycline group of chemotherapeutic agents that exert antitumor action by both DNA intercalation and inhibition of the enzyme

topoisomerase II, which results in DNA damage during DNA replication and the eventual induction of apoptosis. Paclitaxel, on the other hand, is a class of taxanes, an antimitotic drug that affects the stabilization of microtubules, thus, disrupting the cell's ability to divide [52]. Each of these two antitumor agents are widely used for the treatment of breast cancer, more often for the treatment of TN classified tumors. They are also used in combination with targeted therapy for the treatment of certain HER2 + subtypes. For instance, paclitaxel is used in combination with trastuzumab (anti-HER2 therapy) for the treatment of small, node-negative HER2 + breast cancer [7].

Even though we observed that NSMCE2 expression is reduced by SE blockade in all the breast cancer cell lines tested, we decided to use TN breast cancer cells, MDAMB231, BT549 and Hs578T, and the HER2 + cells HCC1954 for our experiments, since we found that high NSMCE2 gene expression levels significantly associate with poor response to chemotherapy for patients

Di Benedetto *et al. BMC Cancer*   (2022) 22:1056

Page 11 of 16

diagnosed with TN or HER2 + breast cancer. Initially, cells were pre-treated with the BET inhibitor JQ1 for 3 h to block SE transcriptional function (this is a window of time we have previously found successfully inhibits SE function). After the 3 h, cells were treated with JQ1 in combination with several concentrations of either of the chemotherapeutic drugs. We assessed cell viability by MTT assays. For all cell lines tested, we observed that the reduced cell viability resulting from the combined treatment is highly dependent on the chemotherapeutic agent used and on its concentration (Fig. S4A). For instance, a combination of JQ1 and doxorubicin has an additive effect on cell survival reduction when using doxorubicin at 1 μM. Such additive effect is gradually lost at higher concentrations of doxorubicin, where we see strong decline in cell viability regardless of SE inhibition by JQ1, probably due to high toxicity and massive DNA damage caused by doxorubicin. On the other hand, for any of the breast cancer cell lines tested, paclitaxel effectiveness at reducing the number of viable cells did not improve when given in combination with JQ1; in fact, an antagonistic effect was observed upon combination of these drugs. Next, to replicate the scenario we observe when analyzing the correlation of NSMCE2 gene expression levels with response to chemotherapy in breast cancer patients (where we see that low NSMCE2 expression levels correlate with good response to chemotherapy), we pre-treated cells with JQ1 for 24 h in order to allow enough time to reduce expression levels of SE-associated genes (in particular for NSMCE2, Fig. 3) before doxorubicin treatment. Here, we also observed that the addition of JQ1 was additive to the reduction in cell viability effect of doxorubicin (Fig. S4B). To assess whether this reduction in viable cell numbers is due to increased apoptosis, we measured apoptosis by Annexin V staining using flow cytometry. Breast cancer cells were pretreated with JQ1 for 24 h prior to the addition of JQ1 and/or chemotherapeutic drugs. As shown in Fig. 5A, doxorubicin highly increases apoptosis of BT549 cells and of Hs578T to a lesser extent. Moreover, the combination of doxorubicin with JQ1 treatment strongly synergizes to induce apoptosis in both breast cancer cell lines. On the contrary, neither paclitaxel alone nor in combination with JQ1 increase apoptosis of the treated breast cancer cell lines (Fig. 5B).

In summary, our results demonstrate that JQ1 synergizes with doxorubicin to induce apoptosis suggesting that reducing NSMCE2 expression by SE inhibition may contribute to the increase in apoptosis of the breast cancer cell lines tested. To investigate whether NSMCE2 downregulation has an effect on cell viability upon treatment with chemotherapeutic agents, we next generated BT549 or Hs578T cells stably expressing a short hairpin RNA against NSMCE2 (shNSMCE2) to knockdown gene expression. After checking NSMCE2 silencing was efficient (Fig. S4C), we performed MTT assays to measure cell viability upon chemotherapeutic drug treatments. Here we found that silencing NSMCE2 in both cell lines significantly contributed to a greater reduction in cell viability after 48 h of incubation with doxorubicin. Knocking down NSMCE2 had a significant but smaller contribution to cell viability reduction after 72 h of paclitaxel treatment (Fig. 5C). Taken together, these results show that reducing NSMCE2 levels sensitizes breast cancer cells to chemotherapy, thus implying that NSMCE2 high levels contribute resistance to chemotherapeutic agents.

Since NSMCE2 is required for DNA damage repair, we hypothesized that more NSMCE2 is produced at the transcriptional level upon chemotherapy-induced DNA damage. Indeed, we observed that NSMCE2 RNA levels significantly increase 24 h after doxorubicin treatment in BT549 and Hs578T, but not upon paclitaxel treatment (Fig. 5D). Given that NSMCE2 is associated with SEs in both cell lines, we next investigated if doxorubicin-induced NSMCE2 upregulation is mediated by SEs. Treatment of breast cancer cells with SE inhibitor JQ1 abrogates the effect of doxorubicin in NSMCE2 upregulation, indicating that NSMCE2 upregulation by doxorubicin is mediated by SEs (Fig. 5E). Collectively, our results show that NSMCE2 expression is transcriptionally upregulated upon doxorubicin treatment through SEs, and this could be a factor contributing to resistance to chemotherapy during breast cancer treatment. In sum, our experiments show that the correlation between high NSMCE2 expression levels in breast tumors and cancer patients' poor response to chemotherapy is due, in part, to increased resistance to chemotherapeutic drugs driven by NSMCE2.

Discussion
In this work, we mined publicly available cancer datasets containing cancer patients' genomic and clinical information and performed experimental perturbations to identify novel genes erroneously upregulated in breast cancer by SEs and that are associated to patients' poor outcome and resistance to cancer therapies. By using this approach, we identified two highly upregulated genes in tumors, NSMCE2 and MAL2, for which high levels of gene expression significantly correlate with breast cancer patients' negative prognosis. Through pharmacological SE disruption assays performed in breast cancer cell lines, we experimentally confirmed that SEs regulate the expression of NSMCE2 or MAL2. Importantly, by mining a database containing breast cancer treatment and response information, we found that high levels of
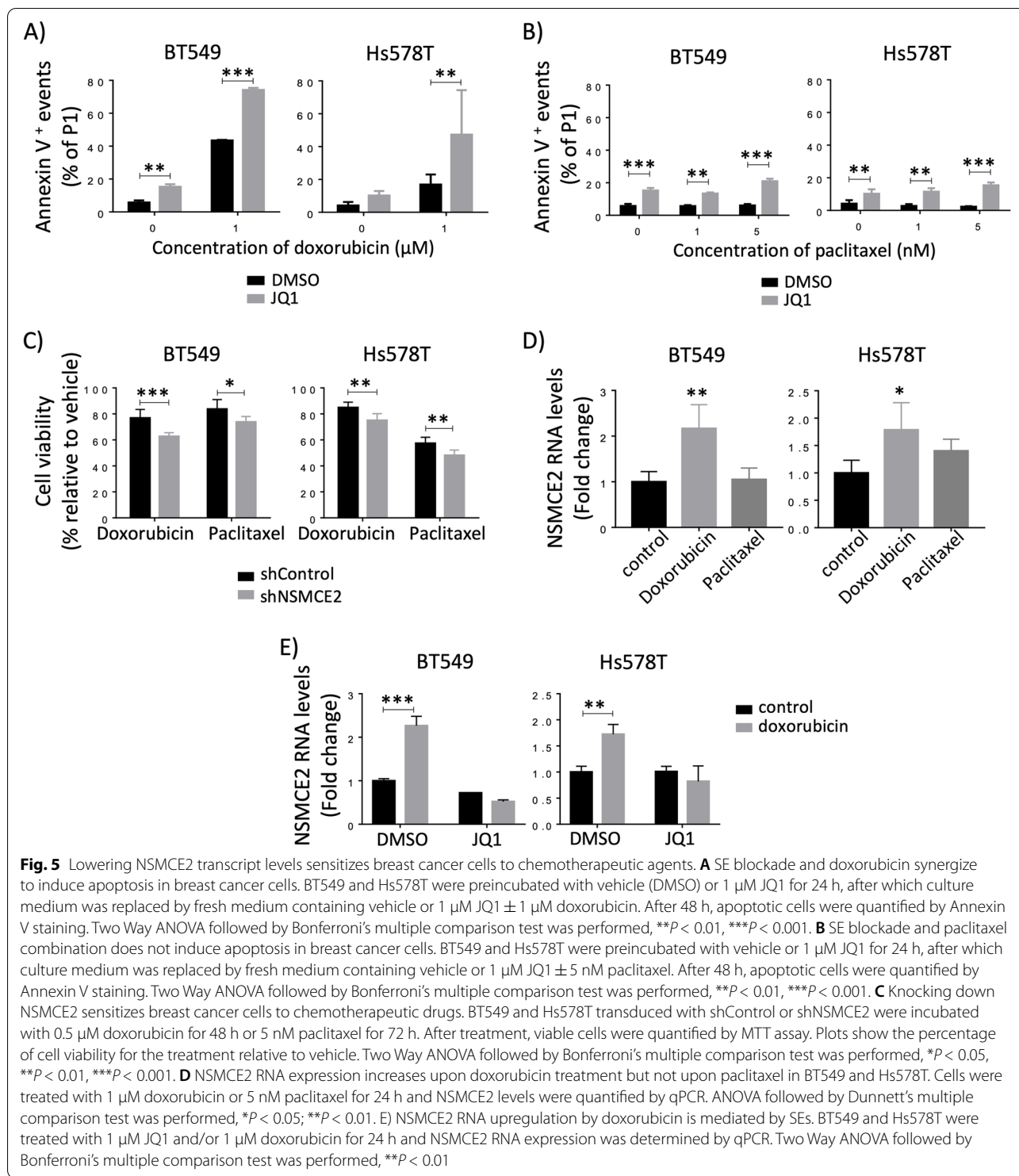
Fig. 5 Lowering NSMCE2 transcript levels sensitizes breast cancer cells to chemotherapeutic agents. **A** SE blockade and doxorubicin synergize to induce apoptosis in breast cancer cells. BT549 and Hs578T were preincubated with vehicle (DMSO) or 1 μM JQ1 for 24 h, after which culture medium was replaced by fresh medium containing vehicle or 1 μM JQ1 ± 1 μM doxorubicin. After 48 h, apoptotic cells were quantified by Annexin V staining. Two Way ANOVA followed by Bonferroni's multiple comparison test was performed, **$P < 0.01$, ***$P < 0.001$. **B** SE blockade and paclitaxel combination does not induce apoptosis in breast cancer cells. BT549 and Hs578T were preincubated with vehicle or 1 μM JQ1 for 24 h, after which culture medium was replaced by fresh medium containing vehicle or 1 μM JQ1 ± 5 nM paclitaxel. After 48 h, apoptotic cells were quantified by Annexin V staining. Two Way ANOVA followed by Bonferroni's multiple comparison test was performed, **$P < 0.01$, ***$P < 0.001$. **C** Knocking down NSMCE2 sensitizes breast cancer cells to chemotherapeutic drugs. BT549 and Hs578T transduced with shControl or shNSMCE2 were incubated with 0.5 μM doxorubicin for 48 h or 5 nM paclitaxel for 72 h. After treatment, viable cells were quantified by MTT assay. Plots show the percentage of cell viability for the treatment relative to vehicle. Two Way ANOVA followed by Bonferroni's multiple comparison test was performed, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$. **D** NSMCE2 RNA expression increases upon doxorubicin treatment but not upon paclitaxel in BT549 and Hs578T. Cells were treated with 1 μM doxorubicin or 5 nM paclitaxel for 24 h and NSMCE2 levels were quantified by qPCR. ANOVA followed by Dunnett's multiple comparison test was performed, *$P < 0.05$; **$P < 0.01$. E) NSMCE2 RNA upregulation by doxorubicin is mediated by SEs. BT549 and Hs578T were treated with 1 μM JQ1 and/or 1 μM doxorubicin for 24 h and NSMCE2 RNA expression was determined by qPCR. Two Way ANOVA followed by Bonferroni's multiple comparison test was performed, **$P < 0.01$

NSMCE2, especially in grade III TN and HER2 + breast cancers, are strongly correlated with patients' poor response to chemotherapy. Accordingly, we experimentally showed that reducing NSMCE2 gene expression levels increases the effectiveness of chemotherapeutic agents.

Previously it was reported that NSMCE2 is involved in maintaining telomeres' length in cancer cell lines,

Di Benedetto *et al. BMC Cancer*   (2022) 22:1056

Page 13 of 16

in this manner preventing telomere-mediated cell cycle arrest and the activation of the senescence pathway [53]. Also in cancer, NSMCE2 depletion has been reported to prevent cell cycle progression of breast cancer cells [54]. Similar to NSMCE2, MAL2 has a role in tumorigenesis, but through modulation of the immune system. MAL2 overexpression has been previously shown to reduce tumor cells' antigen presentation by promoting the endocytosis of tumor antigens [55]. On a different study, Jeong and colleagues showed that MAL2 contributes to breast cancer by inducing accumulation of HER2 on the cell surface, thus enhancing oncogenic HER2 signaling [56]. These known tumor-promoting roles for NSMCE2 and MAL2 suggest that SE-driven overexpression of these two genes can support a variety of tumor-promoting processes in breast cancers cells. Since our results demonstrate that NSMCE2 and MAL2 gene expression levels are reduced by pharmacological SE blockade, targeting SEs could work as a therapeutic approach to reduce these tumor promoting processes.

Through mining genomic datasets, we also found that NSMCE2 and MAL2 are frequently amplified in breast cancer patients and in breast cancer cell lines. Recently it was reported that SEs are frequently found on amplified DNA regions in tumor cells [38]. DNA amplification is known to contribute to cancer progression by mediating overexpression of oncogenes or genes that promote cancer therapy resistance [57]. Thus, frequent copy number gains of the NSMCE2 and MAL2 loci in breast cancer suggest that higher gene expression levels for these genes indeed contribute to tumor progression. Since we also demonstrated that downregulation of NSMCE2 and MAL2 gene expression can be achieved by blocking SEs, collectively these data suggest that both mechanisms (i.e., SE-driven dysregulation and gene amplification) can increase expression of these oncogenes in tumor cells. Thus, reducing the expression of these genes through SE perturbation could have potential therapeutic value.

Currently, several ongoing clinical trials are studying SE disruption through BET inhibition, alone or in combination with other therapies, for the treatment of both hematological malignancies and solid tumors [58]. However, SEs are also involved in the control of processes in healthy cells, bringing concerns about toxicity when using SE inhibitors as treatment. Some common adverse effects observed in current clinical trials include thrombocytopenia (most frequent dose-limiting toxicity), nausea, diarrhea, and fatigue [59–61]. For this reason, synergistic therapies have gained interest because they lead to increased efficacy with lower drug dose, thus reducing toxicities associated with higher concentrations. Following this approach,

BET inhibitors have been shown to sensitize tumor cells to antitumor drugs when tested preclinically. For instance, it has been shown that JQ1 synergizes with daunorubicin, an anthracycline chemotherapeutic drug, to induce apoptosis in acute myeloid leukemia cells [62]. More recently, a combination of JQ1 and the CDK4/6 inhibitor palbociclib has been found to synergistically inhibit cell growth through induction of cell cycle arrest in TN breast cancer cell lines [63].

Our results show that JQ1 sensitizes breast cancer cells to the chemotherapeutic drug doxorubicin by increasing apoptosis and by exerting an additive effect that inhibits cell growth when given in combination with chemotherapy. Although an additive cytotoxic effect has been reported for JQ1 and doxorubicin-treated primary osteosarcoma cells [64], we are the first to report this effect on breast cancer. On the contrary, we did not see an increase on apoptosis when breast cancer cells were treated by combining JQ1 with paclitaxel. We speculate that the increase in apoptosis observed upon JQ1 and doxorubicin combination depends mainly on the mechanism of action of the chemotherapeutic agent. For instance, while doxorubicin induces DNA damage, JQ1, through SE inhibition, may be downregulating the expression of antiapoptotic genes or genes involved in DNA damage repair (including NSMCE2), thus leaving the cells with no other choice but to activate the apoptotic pathway upon doxorubicin treatment. In line with this, we found that NSMCE2 RNA expression increases upon doxorubicin treatment in breast cancer cells, suggesting that NSMCE2 upregulation could be required to overcome doxorubicin induced DNA damage. This upregulation is inhibited by BET inhibitors and, in consequence, mediated by SEs. Importantly, by generating NSMCE2 knockdown breast cancer cell lines, we showed that NSMCE2 plays a key role in resistance to doxorubicin treatment, probably due to its DNA repairing function, which reduces doxorubicin's DNA damaging effect and prevents the activation of the apoptotic program. In line with our results, depletion of the NSMCE2 homologous in yeast, MMS21, was reported to hypersensitize cells to doxorubicin-induced apoptosis [65]. Altogether, our experimental findings imply that combining SE blockade with doxorubicin can be an effective therapy to treat breast tumors that express high levels of NSMCE2. Importantly, our computational analyses show that high NSMCE2 gene expression levels correlate to patients' poor response to chemotherapy, especially in grade III TN and HER2 + breast cancers. Although we ignore the epigenetic modifications as well as the gene expression changes occurring in response to treatment

in these breast tumors, this correlation also points to the role of NSMCE2 on resistance to chemotherapy. Taken together, our results suggest that reducing NSMCE2 levels could help to overcome resistance to DNA damage-inducing therapies in specific cohorts of patients.

Conclusions

In conclusion, our work demonstrates for the first time that in breast tumors NSMCE2 and MAL2 can be dysregulated by SEs, are frequently amplified, and thus, overexpressed. Moreover, we showed that high expression levels of these genes are associated to patients' poor survival. We propose that targeting these genes either by reducing their expression through SE blockade or by directly inhibiting their function through pharmacological inhibition could be effective strategies to prevent their tumor progression activities and thus these approaches need to be further explored.

Abbreviations

AUC: Area Under the Curve; BET: Bromodomain and Extraterminal Proteins; BRD: Bromodomain; ChIP-Seq: Chromatin immuno precipitation followed by high throughput sequencing; CI: Combination index; ER: Estrogen receptor; GEO: Gene Expression Omnibus; GTEX: Genotype-Tissue Expression study; H3K27Ac: Acetylation at histone 3 lysine 27; HER2: Human epidermal growth factor receptor 2; METABRIC: Molecular Taxonomy of Breast Cancer International Consortium; PDX: Patient Derived Xeno transplant; PR: Progesterone receptor; RNA-Seq: RNA sequencing; ROC: Receiver Operating Characteristic; ROSE: Rank Ordering of Super-Enhancers; SE: Super-enhancer; shRNA: Short hairpin RNA; SMC 5/6: Structural Maintenance of Chromosomes 5/6 complex; SUMO: Small ubiquitin-related modifier; TCGA: The Cancer Genome Atlas; TN: Triple negative.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12885-022-10157-7.

> **Additional file 1.**
>
> **Additional file 2.**
>
> **Additional file 3.**
>
> **Additional file 4.**
>
> **Additional file 5.**
>
> **Additional file 6.**
>
> **Additional file 7.**
>
> **Additional file 8.**

Author's information

Corresponding author: Paola Betancur; email address: paola.betancur@ucsf.edu.

Availability of data and materials

The PDX breast tumor H3K27ac ChIP-Seq dataset analyzed during the current study is available in the Genotypes and Phenotypes repository, under accession number phs001264.v1.p1, https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001264.v1.p1 [18]. The TCGA cancer patient datasets analyzed in this study are available in the GDC Data Portal repository, https://portal.gdc.cancer.gov/repository [66]. For this study, publicly available TCGA data were downloaded from UCSC Xena-Browser Platform, https://xenabrowser.net/ [22]. The GTEX datasets analyzed during the current study are available in the Genotypes and Phenotypes repository, under accession number phs000424.v3.p1, https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000424.v3.p1 [25]. For this study, publicly available GTEX and TCGA data were downloaded from UCSC XenaBrowser Platform, https://xenabrowser.net/ [22]. The METABRIC dataset analyzed during the current study is available in the European Genome-phenome Archive (EGA) repository, https://ega-archive.org/studies/EGAS00000000083 [23]. For this study, publicly available META-BRIC data were downloaded from cBioPortal, https://www.cbioportal.org/ (24).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics. CA Cancer J Clin. 2021;71(1):7–33.
2. Bauer K, Parise C, Caggiano V. Use of ER/PR/HER2 subtypes in conjunction with the 2007 St Gallen Consensus Statement for early breast cancer. BMC Cancer. 2010;10:228.
3. Perou CM, Sørlie T, Eisen MB, Van De Rijn M, Jeffrey SS, Ress CA, et al. Molecular portraits of human breast tumours. Nature. 2000;406(6797):747–52.
4. Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A. 2001;98(19):10869–74.
5. Koboldt DC, Fulton RS, McLellan MD, Schmidt H, Kalicki-Veizer J, McMichael JF, et al. Comprehensive molecular portraits of human breast tumours. Nature. 2012;490(7418):61–70.
6. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thu B. Strategies for subtypes — dealing with the diversity of breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. Ann Oncol. 2011;(June):1736–47.
7. Waks AG, Winer EP. Breast Cancer Treatment: A Review. JAMA. 2019;321(3):288–300.
8. Dent R, Trudeau M, Pritchard KI, Hanna WM, Kahn HK, Sawka CA, et al. Triple-negative breast cancer: clinical features and patterns of recurrence. Clin Cancer. 2007;13(15 Pt 1):4429–34.

Di Benedetto *et al. BMC Cancer*   (2022) 22:1056

Page 15 of 16

9. Ensenyat-Mendez M, Llinàs-Arias P, Orozco JIJ, Íñiguez-Muñoz S, Salomon MP, Sesé B, et al. Current Triple-Negative Breast Cancer Subtypes: Dissecting the Most Aggressive Form of Breast Cancer. Front Oncol. 2021;11:681476.

10. Nedeljković M, Damjanović A. Mechanisms of chemotherapy resistance in triple-negative breast cancer-how we can rise to the challenge. Cells. 2019;8(9):957.

11. Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-André V, Sigova AA, et al. Super-enhancers in the control of cell identity and disease. Cell. 2013;155(4):934.

12. Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell. 2013;153(2):307–19.

13. Martinez MF, Medrano S, Brown EA, Tufan T, Shang S, Bertoncello N, et al. Super-enhancers maintain renin-expressing cell identity and memory to preserve multi-system homeostasis. J Clin Invest. 2018;128(11):4787–803.

14. Adam RC, Yang H, Ge Y, Infarinato NR, Gur-Cohen S, Miao Y, et al. NFI transcription factors provide chromatin access to maintain stem cell identity while preventing unintended lineage fate choices. Nat Cell Biol. 2020;22(6):640–50.

15. Ge Y, Gomez NC, Adam RC, Nikolova M, Yang H, Verma A, et al. Stem Cell Lineage Infidelity Drives Wound Repair and Cancer. Cell. 2017;169(4):636-650.e14.

16. Lovén J, Hoke HA, Lin CY, Lau A, Orlando DA, Vakoc CR, et al. Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell. 2013;153(2):320–34.

17. Mansour MR, Abraham BJ, Anders L, Berezovskaya A, Gutierrez A, Durbin AD, et al. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. Science. 2014;346(6215):1373–7.

18. Betancur PA, Abraham BJ, Yiu YY, Willingham SB, Khameneh F, Zarnegar M, et al. A CD47-associated super-enhancer links pro-inflammatory signalling to CD47 upregulation in breast cancer. Nat Commun. 2017;8(May 2016):14802.

19. Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci. 2010;107:21931–6.

20. Nakamura Y, Umehara T, Nakano K, Jang MK, Shirouzu M, Morita S, et al. Crystal structure of the human BRD2 bromodomain: insights into dimerization and recognition of acetylated histone H4. J Biol Chem. 2007;282(6):4193–201.

21. Chapuy B, McKeown MR, Lin CY, Monti S, Roemer MGM, Qi J, et al. Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. Cancer Cell. 2013;24(6):777–90.

22. Goldman MJ, Craft B, Hastie M, Repečka K, McDade F, Kamath A, et al. Visualizing and interpreting cancer genomics data via the Xena platform. Nat Biotechnol. 2020;38(6):675–8.

23. Curtis C, Shah SP, Chin S-F, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature. 2012;486(7403):346–52.

24. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012;2(5):401–4.

25. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The Genotype-Tissue Expression (GTEx) project. Nat Genet. 2013;45(6):580–5.

26. Ghandi M, Huang FW, Jané-Valbuena J, Kryukov GV, Lo CC, McDonald ER 3rd, et al. Next-generation characterization of the Cancer Cell Line Encyclopedia. Nature. 2019;569(7757):503–8.

27 Fekete JT, Győrffy B. ROCplot.org: Validating predictive biomarkers of chemotherapy/hormonal therapy/anti-HER2 therapy using transcriptomic data of 3,104 breast cancer patients. Int J Cancer. 2019;145(11):3140–51.

28. Slinker BK. The statistics of synergism. J Mol Cell Cardiol. 1998;30(4):723–31.

29. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative CT method. Nat Protoc. 2008;3(6):1101–8.

30. Rank Ordering of Super-Enhancers (ROSE) [Internet]. [cited 2022 Apr 4]. Available from: https://bitbucket.org/young_computation/rose/src/master/.

31. Zhang CQ, Williams H, Prince TL, Ho ES. Overexpressed HSF1 cancer signature genes cluster in human chromosome 8q. Hum Genomics. 2017;11(1):35.

32. Zhao K, Zhao Y, Zhu J-Y, Dong H, Cong W-M, Yu Y, et al. A Panel of Genes Identified as Targets for 8q24.13-24.3 Gain Contributing to Unfavorable Overall Survival in Patients with Hepatocellular Carcinoma. Curr Med Sci. 2018;38(4):590–6.

33. Reon BJ, Takao Real Karia B, Kiran M, Dutta A. LINC00152 Promotes Invasion through a 3'-Hairpin Structure and Associates with Prognosis in Glioblastoma. Mol Cancer Res. 2018;16(10):1470–82.

34. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun. 2019;10(1):1523.

35. Madanikia SA, Bergner A, Ye X, Blakeley JO. Increased risk of breast cancer in women with NF1. Am J Med Genet A. 2012;158A(12):3056–60.

36. Głodzik D, Purdie C, Rye IH, Simpson PT, Staaf J, Span PN, et al. Mutational mechanisms of amplifications revealed by analysis of clustered rearrangements in breast cancers. Ann Oncol Off J Eur Soc Med Oncol. 2018;29(11):2223–31.

37. Beroukhim R, Mermel CH, Porter D, Wei G, Raychaudhuri S, Donovan J, et al. The landscape of somatic copy-number alteration across human cancers. Nature. 2010;463(7283):899–905.

38. Glodzik D, Morganella S, Davies H, Simpson PT, Li Y, Zou X, et al. A somatic-mutational process recurrently duplicates germline susceptibility loci and tissue-specific super-enhancers in breast cancers. Nat Genet. 2017;49(11):1661.

39. Tang H, Sebti S, Titone R, Zhou Y, Isidoro C, Ross TS, et al. Decreased BECN1 mRNA Expression in Human Breast Cancer is Associated With Estrogen Receptor-Negative Subtypes and Poor Prognosis. EBioMedicine. 2015;2(3):255–63.

40. Potts PR, Yu H. Human MMS21/NSE2 is a SUMO ligase required for DNA repair. Mol Cell Biol. 2005;25(16):7021–32.

41. Potts PR, Porteus MH, Yu H. Human SMC5/6 complex promotes sister chromatid homologous recombination by recruiting the SMC1/3 cohesin complex to double-strand breaks. EMBO J. 2006;25(14):3377–88.

42. Behlke-Steinert S, Touat-Todeschini L, Skoufias DA, Margolis RL. SMC5 and MMS21 are required for chromosome cohesion and mitotic progression. Cell Cycle. 2009;8(14):2211–8.

43. de Marco MC, Martín-Belmonte F, Kremer L, Albar JP, Correas I, Vaerman JP, et al. MAL2, a novel raft protein of the MAL family, is an essential component of the machinery for transcytosis in hepatoma HepG2 cells. J Cell Biol. 2002;159(1):37–44.

44. Filippakopoulos P, Qi J, Picaud S, Shen Y, Smith WB, Fedorov O, et al. Selective inhibition of BET bromodomains. Nature. 2010;468(7327):1067–73.

45. Seal J, Lamotte Y, Donche F, Bouillot A, Mirguet O, Gellibert F, et al. Identification of a novel series of BET family bromodomain inhibitors: binding mode and profile of I-BET151 (GSK1210151A). Bioorg Med Chem Lett. 2012;22(8):2968–72.

46. ROC Plotter [Internet]. [cited 2022 Apr 4]. Available from: http://www.rocplot.org/.

47. Fekete JT, Ősz Á, Pete I, Nagy GR, Vereczkey I, Győrffy B. Predictive biomarkers of platinum and taxane resistance using the transcriptomic data of 1816 ovarian cancer patients. Gynecol Oncol. 2020;156(3):654–61.

48. Menyhárt O, Fekete JT, Győrffy B. Gene expression-based biomarkers designating glioblastomas resistant to multiple treatment strategies. Carcinogenesis. 2021;42(6):804–13.

49. Pond KW, de Renty C, Yagle MK, Ellis NA. Rescue of collapsed replication forks is dependent on NSMCE2 to prevent mitotic DNA damage. PLoS Genet. 2019;15(2):e1007942–e1007942.

50. Deiss K, Lockwood N, Howell M, Segeren HA, Saunders RE, Chakravarty P, et al. A genome-wide RNAi screen identifies the SMC5/6 complex as a non-redundant regulator of a Topo2a-dependent G2 arrest. Nucleic Acids Res. 2019;47(6):2906–21.

51. Zhao X, Blobel G. A SUMO ligase is part of a nuclear multiprotein complex that affects DNA repair and chromosomal organization. Proc Natl Acad Sci U S A. 2005;102(13):4777–82.

52. Abu Samaan TM, Samec M, Liskova A, Kubatka P, Büsselberg D. Paclitaxel's mechanistic and clinical effects on breast cancer. Biomolecules. 2019;9(12):789.

53. Potts PR, Yu H. The SMC5/6 complex maintains telomere length in ALT cancer cells through SUMOylation of telomere-binding proteins. Nat Struct Mol Biol. 2007;14(7):581–90.

Di Benedetto *et al. BMC Cancer*   (2022) 22:1056

Page 16 of 16

54. Ni H-J, Chang Y-N, Kao P-H, Chai S-P, Hsieh Y-H, Wang D-H, et al. Depletion of SUMO ligase hMMS21 impairs G1 to S transition in MCF-7 breast cancer cells. Biochim Biophys Acta. 2012;1820(12):1893–900.

55. Fang Y, Wang L, Wan C, Sun Y, Van der Jeught K, Zhou Z, et al. MAL2 drives immune evasion in breast cancer by suppressing tumor antigen presentation. J Clin Invest. 2021;131(1):e140837.

56. Jeong J, Shin JH, Li W, Hong JY, Lim J, Hwang JY, et al. MAL2 mediates the formation of stable HER2 signaling complexes within lipid raft-rich membrane protrusions in breast cancer cells. Cell Rep. 2021;37(13):110160.

57. Zack TI, Schumacher SE, Carter SL, Cherniack AD, Saksena G, Tabak B, et al. Pan-cancer patterns of somatic copy number alteration. Nat Genet. 2013;45(10):1134–40.

58. Shorstova T, Foulkes WD, Witcher M. Achieving clinical success with BET inhibitors as anti-cancer agents. Br J Cancer. 2021;124(9):1478–90.

59. Ameratunga M, Braña I, Bono P, Postel-Vinay S, Plummer R, Aspegren J, et al. First-in-human Phase 1 open label study of the BET inhibitor ODM-207 in patients with selected solid tumours. Br J Cancer. 2020;123(12):1730–6.

60. Berthon C, Raffoux E, Thomas X, Vey N, Gomez-Roca C, Yee K, et al. Bromodomain inhibitor OTX015 in patients with acute leukaemia: a dose-escalation, phase 1 study. Lancet Haematol. 2016;3(4):e186–95.

61. Amorim S, Stathis A, Gleeson M, Iyengar S, Magarotto V, Leleu X, et al. Bromodomain inhibitor OTX015 in patients with lymphoma or multiple myeloma: a dose-escalation, open-label, pharmacokinetic, phase 1 study. Lancet Haematol. 2016;3(4):e196-204.

62. Stewart HJS, Horne GA, Bastow S, Chevassut TJT. BRD4 associates with p53 in DNMT3A-mutated leukemia cells and is implicated in apoptosis by the bromodomain inhibitor JQ1. Cancer Med. 2013;2(6):826–35.

63. Ge JY, Shu S, Kwon M, Jovanović B, Murphy K, Gulvady A, et al. Acquired resistance to combined BET and CDK4/6 inhibition in triple-negative breast cancer. Nat Commun. 2020;11(1):2350.

64. Baker EK, Taylor S, Gupte A, Sharp PP, Walia M, Walsh NC, et al. BET inhibitors induce apoptosis through a MYC independent mechanism and synergise with CDK inhibitors to kill osteosarcoma cells. Sci Rep. 2015;5:10120.

65. Huang R-Y, Kowalski D, Minderman H, Gandhi N, Johnson ES. Small ubiquitin-related modifier pathway is a major determinant of doxorubicin cytotoxicity in Saccharomyces cerevisiae. Cancer Res. 2007;67(2):765–72.

66. Hoadley KA, Yau C, Wolf DM, Cherniack AD, Tamborero D, Ng S, et al. Multiplatform Analysis of 12 Cancer Types Reveals Molecular Classification within and across Tissues of Origin. Cell. 2014;158(4):929–44.

Publisher's Note