STAT3 and GR cooperate to drive basal-like triple negative breast cancer gene expression and proliferation.

Authors

Megan E. Conway^{*1}, Joy M. McDaniel^{*2}, James M. Graham¹, Katrin P. Guillen¹, Patsy G. Oliver³, Stephanie L. Parker², Peibin Yue⁴, James Turkson⁴, Donald J. Buchsbaum³, Bryan E. Welm⁵, Richard M. Myers², Katherine E. Varley^{1†}

Affiliations

1. Department of Oncological Sciences, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, USA.
2. HudsonAlpha Institute for Biotechnology, Huntsville, AL, USA.
3. Department of Radiation Oncology, University of Alabama at Birmingham, Birmingham, AL, USA.
4. Department of Medicine and Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, 8700 Beverly Blvd, Los Angeles, CA, USA 90048
5. Department of Surgery, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, USA

*These authors contributed equally

† Corresponding Author: Katherine Varley, Department of Oncological Sciences, Huntsman Cancer Institute, University of Utah, HCI Research South Room 3719, 2000 Circle of Hope, Salt Lake City, UT, USA, 84112.

Email: kt.varley@hci.utah.edu

Keywords

Breast cancer

Triple negative breast cancer

Genomics

Gene expression

Transcription regulation

Transcription factor

Chromatin accessibility

Abstract

Breast cancers can be divided into subtypes with different prognoses and treatment responses based on global gene expression differences. Luminal breast cancer gene expression and proliferation are driven by the transcription factors Estrogen Receptor α (ER), FOXA1 and GATA3. Targeting ER is the most effective therapy for treating luminal breast cancer because ER is the master regulator of the luminal gene expression program. In contrast, it is unclear which transcription factors are responsible for driving the gene expression signature that defines basal-like triple negative breast cancer, and there are no targeted therapies approved to treat this aggressive subtype of the disease. This study utilized integrated analysis of DNA methylation, chromatin accessibility, transcription factor binding, and gene expression in large collections of breast cancer cell lines and patient tumors to identify transcription factors responsible for the basal-like gene expression program. The results of this study indicate that glucocorticoid receptor (GR) and signal transducer and activator of transcription 3 (STAT3) bind to the same genomic regulatory regions that are specifically open and unmethylated in basal-like breast cancer. These transcription factors cooperate to regulate expression of hundreds of genes in the basal-like gene expression signature and these downstream genes are associated with poor prognosis in patients. Furthermore, combination treatment with small molecule drugs that inhibit both transcription factors leads to synergistic decreases in cell proliferation in cell lines and patient-derived organoid models. This study demonstrates that GR and STAT3 cooperate to regulate the basal-like breast cancer gene expression program and provides the basis for improved therapy for basal-like triple negative breast cancer through rational combination of STAT3 and GR inhibitors.

Introduction

Breast cancers can be divided into subtypes with different prognoses and treatment responses based on global gene expression differences(1,2). These expression differences reflect the transcription factors (TFs), epigenetic states and gene regulatory networks in the normal cells from which the tumors arise, as well as changes that accumulate during tumorigenesis.

Normal luminal breast cells proliferate in response to estrogen through activation of the TF estrogen receptor α (ER) and ER's downstream target genes. Breast cancers that express ER are most often classified into the luminal subtype and likely arise from luminal breast cells(1). These luminal ER+ tumors are treated with hormone therapies that prevent ER from activating its pro-proliferative transcriptional program. Even though these tumors usually have wildtype ER, this is an extremely effective targeted therapy because it inhibits the master regulator of gene expression and proliferation in this cell type.

In contrast, 15-20% of breast cancer patients are diagnosed with triple negative breast cancer (TNBC), which is defined by a lack of expression of ER and progesterone receptor, and normal expression of human epidermal growth factor receptor 2. There are no FDA approved targeted therapies to treat TNBC. Even after aggressive treatment with surgery, radiation and cytotoxic chemotherapy, TNBC patients have a high rate of recurrence within 3 years (3), and a high rate of death from disease (38%) (4). There is a crucial need for alternative therapeutic strategies for TNBC. Approximately 70-80% of TNBC tumors are classified into the basal-like breast cancer gene expression subtype, and likely arise from myoepithelial breast cells(5). It is unclear which TFs are responsible for regulating the pro-proliferative transcriptional program in the basal-like TNBC subtype. The goal of this study is to identify TFs that regulate clinically relevant gene regulatory programs in basal-like breast cancer, and determine if inhibiting these TFs could provide novel therapeutic strategies to improve clinical outcomes for basal-like TNBC patients.

The role of many TFs have been studied in basal-like breast cancer, including oncogenic drivers such as MYC (6), and other TFs that regulate pathways associated with basal-like phenotypes such as EN1 (7), GLI1 (8), NF- κ B (9), HIF1 α (10), TWIST1 (11), FOXC1 (12,13), AP-1 (14), GR(15-18), STAT3(19,20), and MAFK (21). Many of these TFs regulate important pathways in basal-like breast cancer and these TFs often regulate each other's expression (22,23). Several of these TFs can be inhibited with small molecule drugs and represent promising

therapeutic targets(18,24,25). What remains unclear is which TFs are responsible for driving the gene expression signature that defines the basal-like TNBC subtype and which are downstream of the master regulator. This study utilized integrated analysis of DNA methylation, chromatin accessibility, TF binding, gene expression, and proliferation in large collections of breast cancer cell lines and patient tumors to identify TFs that drive the basal-like TNBC gene expression program. We sought to identify regulatory regions that are specifically unmethylated and have accessible chromatin in basal-like TNBC compared luminal breast cancers. We identified TFs whose motifs and binding sites are most highly enriched at these basal-specific regulatory regions. We then investigated whether these TFs were more highly expressed in basal-like TNBC compared to luminal breast cancer, whether the TFs regulated genes in the basal-like TNBC gene expression program, and whether the genes regulated by these TFs were associated with patient outcomes. This genome-wide approach led to the discovery that cooperation between two TFs (STAT3 and GR), rather than a single ‘master regulator’, drives expression of hundreds of genes in the basal-like TNBC gene expression program. This TF cooperation provides basal-like TNBC cells with their resilient proliferative capacity. Furthermore, inhibiting both TFs simultaneously led to synergistic decreases in cell proliferation in basal-like TNBC cell lines and patient-derived organoid models, which suggests a new treatment strategy for basal-like TNBC.

Results

Identification of regulatory regions and candidate TFs specific to basal-like breast cancer

We hypothesized that genomic regulatory regions that are specifically unmethylated in basal-like breast cancer compared to luminal breast cancer will be enriched for TF binding sites involved in controlling basal-specific gene expression. To identify regulatory regions specific to basal-like breast cancer, reduced representation bisulfite sequencing (RRBS) was performed on 28 breast cancer cell lines (18 basal-like and 10 luminal) in order to measure DNA methylation across the genome. Of the 479,746 CG positions in the genome with at least 10x coverage in each cell line, 3,748 CGs were significantly differentially methylated (Linear Regression Benjamini-Hochberg adjusted $p < 0.05$) between basal-like and luminal cell lines (**Figure 1A**).

There were 1,300 CGs in intergenic or promoter regions of the genome that were significantly unmethylated in luminal breast cancer. These CG positions were intersected with the locations of TF binding sites from ChIP-seq experiments performed on 149 TFs by the ENCODE Project in a variety of cell lines (26). To determine which

TFs were enriched at regions that are specifically unmethylated in luminal breast cancer cells, the fraction of binding sites for each TF that intersect these CGs was divided by the fraction of binding sites for each TF in the ENCODE dataset. The most highly enriched sequence-specific TF binding sites were ER (6.9 fold), FOXA1 (8.1 fold), and GATA3 (10.3 fold). This confirms that this approach is sound because ER, FOXA1, and GATA3 are known to be the master regulators of the luminal gene expression program (27,28) (**Figure 1B**).

There were 1,793 CGs that were in intergenic or promoter regions of the genome and were specifically unmethylated in basal-like breast cancer. These CG positions were intersected with the TF binding sites locations from ChIP-seq experiments performed on 149 TFs by the ENCODE Project (26). The most highly enriched sequence-specific TFs in these regions were JUN (6.2 fold), STAT3 (4.8 fold), and GR (4.2 fold) (**Figure 1C**). This result led us to hypothesize that JUN, STAT3, and GR are involved in regulating basal-like breast cancer gene expression.

To further investigate this hypothesis, we sought to use primary human tumors rather than cell lines, to directly investigate chromatin accessibility with ATAC-seq rather than inferring it from a lack of DNA methylation, and use TF motif enrichment rather than the ENCODE Project ChIP-seq experiments, which did not include any basal-like breast cancer cell lines. We analyzed ATAC-seq data from 59 primary breast tumors (15 basal-like and 44 luminal)(29). There were 110,156 regions of the genome that exhibit open chromatin specifically in basal-like tumors compared to luminal tumors (DESEQ2 Adjusted $p < 0.05$) (**Figure 1D**). Motif analysis found significant enrichment for the canonical motifs for JUN/AP1 ($p = 1.91 \times 10^{-1101}$), STAT3 ($p = 1.15 \times 10^{-359}$), and GR ($p = 1.15 \times 10^{-4}$) in the chromatin regions that were specifically open in basal-like tumors and closed in luminal tumors (**Figure 1E**). While the enrichment for the GR motif is significant, it is lower than that of STAT3 and JUN/AP1, which is consistent with previous studies that suggest GR can be tethered to some enhancers through protein:protein interactions, rather than direct DNA binding(30).

Figure 1. Identification of candidate TFs enriched at regions that are unmethylated and have open chromatin in basal-like compared to luminal breast cancer. **A.** Heatmap of CG positions that are significantly differentially methylated between basal-like and luminal breast cancer cell lines (Linear Regression Benjamini-Hochberg adjusted $p < 0.05$). Enrichment of TF binding sites from ChIP-seq experiments in a variety of cell lines that overlap CG positions that are specifically unmethylated in luminal (**B**) and basal-like (**C**) cell lines. **D.** Heatmap of ATAC-seq peaks that have significantly more open chromatin in basal-like patient tumors compared to luminal tumors (DESEQ2 Adjusted $p < 0.05$). **E.** Enrichment for JUN, STAT3, and GR motifs in the basal-specific ATAC-seq peaks.

Expression of GR and STAT3 is higher in basal-like breast cancer and associated with a worse prognosis

If JUN, GR and STAT3 regulate the basal-like breast cancer gene expression program, then one would expect these TFs to have higher expression in basal-like breast cancer compared to luminal breast cancer. RNA-seq was performed on the same 28 breast cancer cell lines used for DNA methylation analysis. Both GR and STAT3 have significantly higher expression in basal-like cell lines compared to luminal cell lines, while JUN does not (**Figure 2A**). An independent gene expression dataset containing 19 basal-like breast cancer cell lines and 21 luminal breast cancer cell lines (31) confirmed that both GR and STAT3 have higher expression in basal-like breast cancer (**Figure 2B**). These results confirm that GR and STAT3 have higher expression in basal-like breast cancer than luminal breast cancer.

Some luminal breast cancers do express low levels of GR and STAT3, so we investigated the association between prognosis and GR or STAT3 expression. Analysis of a dataset containing breast tumor gene expression and long-term survival data (32) revealed that higher expression of both GR and STAT3 was associated with shorter relapse-free survival in basal-like breast cancer (**Figure 2C**). In contrast, luminal breast cancer patients with higher GR and STAT3 expression were more likely to have longer relapse-free survival (**Figure 2D**). These results confirm that GR and STAT3 expression has a unique role in basal-like breast cancer and suggest that these TFs are associated with aggressive features of the disease.

Figure 2. Expression of GR and STAT3 is higher in basal-like breast cancer and associated with shorter relapse-free survival. **A.** RNA-seq data shows GR and STAT3 have higher expression in basal-like (N=18) compared to luminal (N=10) cell lines, while JUN does not. Mann Whitney p-values displayed. **B.** An independent microarray gene expression dataset shows GR and STAT3 have higher expression in basal-like (N=19) compared to luminal (N=21) cell lines. **C.** In basal-like breast cancer patients, higher expression of GR and STAT3 is associated with shorter relapse-free survival. **D.** In contrast, luminal breast cancer patients have better prognosis when GR and STAT3 are highly expressed.

GR and STAT3 bind to the same basal-specific regulatory regions across the genome

To directly test the hypothesis that GR and STAT3 bind to the genome and regulate the basal-like gene expression program, ChIP-seq was performed for GR and STAT3 in 8 breast cancer cell lines (4 basal-like, 4 luminal). GR is a steroid hormone nuclear receptor that is chaperoned in the cytoplasm until it binds glucocorticoid ligands which allow it to homo-dimerize and translocate to the nucleus and regulate gene expression. To assess the binding of GR, cell lines were treated with 100nM dexamethasone for 1 hour prior to performing ChIP-seq. STAT3 is a TF whose dimerization and translocation to the nucleus is controlled through phosphorylation by Janus kinases (JAK). Normally, JAK phosphorylates STAT3 when cytokine receptors bind ligands such as interferons, epidermal growth factor, and interleukin 5 and 6. STAT3 is constitutively phosphorylated and active in the majority of basal-like cell lines and patient tumors (19,33). Thus, STAT3 ChIP-seq was performed on cells in standard media.

The genome-wide GR and STAT3 binding sites from ChIP-seq were used to perform unsupervised hierarchical clustering of the cell lines based on Spearman correlations. This analysis revealed that the basal-like and luminal subtypes have distinct genome-wide GR and STAT3 binding profiles (**Figure 3A**). The basal-like cell lines form two distinct sub-clusters within the basal-like group, which corresponds to previous classifications as Basal A (HCC70 and HCC1937) and Basal B (SUM159 and MDA-MB-231) subtypes (34). Despite these differences, the four basal-like cell lines are still more similar to each other than they are to luminal cell lines when considering genome-wide GR and STAT3 binding. DESEQ2 was used to identify regions of the genome that exhibited subtype-specific binding of GR or STAT3. There were 2,002 regions of the genome that were bound by GR specifically in basal-like breast cancer (DESEQ2 adjusted $p < 0.05$) (**Figure 3B**), and 3,667 regions of the genome that were bound by STAT3 specifically in basal-like breast cancer (DESEQ2 adjusted $p < 0.05$) (**Figure 3C**). This result demonstrates that both GR and STAT3 bind thousands of regions of the genome in basal-like breast cancer that they do not bind in luminal breast cancer, which supports the hypothesis that these TFs regulate the basal-like gene expression program.

The correlation heatmap revealed that genome-wide GR binding sites and STAT3 binding sites were highly correlated within each of the four basal-like cell lines. To investigate this further, DESEQ2 was used to perform a multivariate analysis that combined both GR and STAT3 ChIP-seq datasets to identify sites where these TFs bound alone or together in a basal-like specific manner. This analysis revealed 1,773 sites were bound by GR

alone, 188 sites were bound by STAT3 alone, and 12,712 ChIP-seq peaks with both GR and STAT3 binding specifically in basal-like cell lines (**Figure 3D**). This result indicates that GR and STAT3 bind the same basal-specific regulatory regions more often than they bind alone.

The majority of shared GR and STAT3 basal-like specific binding sites are 50-500 kb from the nearest gene transcription start site, indicating that they are likely binding distal regulatory elements such as enhancers (**Figure 4A**). Motif analysis was performed on the shared GR and STAT3 basal-like specific binding sites. The canonical STAT3 motif was found in 5,044 of these peaks, resulting in an E-value of 4.97×10^{-425} compared to shuffled control sequences. The canonical GR motif was found in 1,309 of the peaks, resulting in an E-value of 8.00×10^{-37} compared to shuffled control sequences. This result indicates that GR and STAT3 shared binding sites are significantly enriched for motifs recognized by both TFs.

To test the hypothesis that shared GR and STAT3 basal-specific binding sites are near genes that are specifically expressed in basal-like breast cancer, RNA-seq data from 59 primary breast tumors (15 basal-like and 42 luminal) were analyzed(35). There were 7,121 genes that had significantly higher expression in basal-like breast tumors, and 3,464 genes that had significantly higher expression in luminal breast tumors (DESEQ adjusted p-value 0.05). The distance between the transcription start site of each gene and the nearest shared GR and STAT3 binding site was calculated. Genes that had significantly higher expression in basal-like breast tumors were closer to shared GR and STAT3 basal-specific binding sites than genes that were higher in luminal tumors or genes that were not differentially expressed (**Figure 4B**).

To further investigate whether the shared GR and STAT3 basal-specific binding sites discovered in cell lines are open and accessible in patient tumors, analysis of ATAC-seq data from 57 primary breast tumors (15 basal-like and 44 luminal) was performed (29). The shared GR and STAT3 basal-specific binding sites are significantly more open in basal-like patient tumors compared to luminal patient tumors (Mann Whitney $p=7.375 \times 10^{-231}$) (**Figure 4C & 4D**).

Together these results indicate that GR and STAT3 bind together at thousands of regulatory regions that are specifically open in basal-like tumors and closed in luminal tumors, and these binding sites are near genes that have significantly higher expression in basal-like tumors. These data further support the hypothesis that GR and STAT3 regulate the basal-like gene expression program.

Figure 4. **A.** Basal-specific shared GR and STAT3 binding sites are most often 50-500 kb from transcription start sites (TSS) of genes. **B.** Basal-specific shared GR and STAT3 sites are closer to genes that are expressed specifically in basal-like breast cancer compared to genes expressed in luminal breast cancer and genes that are not differentially expressed between subtypes. **C.** Basal-specific shared GR and STAT3 sites have more accessible chromatin in ATAC-seq data from basal-like patient tumors (N=15) compared to luminal patient tumors (N=44) when averaged across sites. Mean (dark points) and standard deviation (lighter band) of averaged normalized read counts across patients are depicted. **D.** Violin plot of ATAC-seq data for 12,712 Basal-specific shared GR and STAT3 shows sites have more accessible chromatin in basal-like patient tumors (N=15) compared to luminal patient tumors (N=44). Mann Whitney p-value, median (horizontal line), interquartile range (black box), and range (vertical line) are depicted.

GR and STAT3 cooperatively regulate gene expression

To determine which genes are regulated by GR and STAT3, RNA-seq was performed on two basal-like cell lines (HCC70 and MDA-MB-231) treated with 100nM dexamethasone or an equal volume of 100% ethanol vehicle control for 4 hours, and STAT3 siRNA SMARTpool or Non-targeting siRNA for 96 hours. Dexamethasone induction of GR resulted in significant changes in the expression of 546 genes in both cell lines (DESEQ2 Adjusted $p < 0.05$)(**Figure 5A**). STAT3 siRNA knockdown resulted in significant changes in the expression of 1,967 genes in both cell lines (DESEQ2 Adjusted $p < 0.05$)(**Figure 5B**).

To test the hypothesis that genes that change expression are near the shared GR and STAT3 basal-like specific binding sites, the distance between the transcription start site of each gene and the nearest shared GR and STAT3 binding site was calculated. Previous studies have reported these TFs are more likely to activate gene expression directly, and that associated repression is more likely a downstream indirect effect(36). Therefore, the genes that are activated and repressed by each TF were analyzed separately. This analysis revealed that a larger fraction of genes that are activated by GR or STAT3 are close to a shared site when compared to repressed genes, or genes that did not change expression when these TFs were modulated (**Figure 5C & 5D**). Notably, 80% of genes that were activated by dexamethasone induction of GR and 60% of genes that are activated by STAT3 are within 100,000 bp of a shared GR and STAT3 binding site.

To further elucidate how GR and STAT3 cooperate to regulate gene expression, the MDA-MB-231 cell line was treated with dexamethasone (1 μ M for 24 hours) and the STAT3 inhibitor SH4-54 (8 μ M for 24 hours) (24) alone and in combination. A multivariate linear model with a GR:STAT3 interaction term was used to identify genes that are differentially expressed when both TFs are fully active in the nucleus. There were 970 genes that had significantly higher expression when both GR and STAT3 were active in the nucleus (BH adjusted p-value<0.05) (**Figure 6A**). Strikingly, the majority of these genes (769) fit a synergistic model of activation, where the presence of both GR and STAT3 in the nucleus resulted in higher expression than the sum of the expression from GR and STAT3 alone. Additionally, 201 genes fit an additive model, where GR and STAT3 each contribute additively to the expression. There were also 628 genes that had significantly lower expression when both GR and STAT3 were active in the nucleus (BH adjusted p-value<0.05) (**Figure 6A**). Again, the majority (515 genes) fit a synergistic model, where the presence of both GR and STAT3 in the nucleus resulted in greater repression than the sum of the repression from GR and STAT3 alone.

To determine if the genes that are cooperatively regulated by GR and STAT3 are near shared GR and STAT3 basal-like specific binding sites, the distance between the transcription start site of each gene and the nearest shared GR and STAT3 binding site was calculated. Consistent with previous observations, a larger fraction of genes that are cooperatively activated by GR and STAT3 are close to a shared site, when compared to repressed genes, or genes that did not change expression when these TFs were modulated (**Figure 6B**).

Together these results indicate that GR and STAT3 cooperate to synergistically regulate the expression of hundreds of genes.

Figure 6. GR and STAT3 cooperate to regulate the expression of 1,598 genes. MDA-MB-231 cells were treated with dexamethasone to induce GR and SH4-54 to inhibit STAT3. **A.** A heatmap of genes that are regulated by both GR and STAT3 in a manner that fits additive or synergistic models for activation or repression (BH adjusted $p < 0.05$). **B.** Genes regulated by both GR and STAT3 are closer to shared GR and STAT3 binding sites than genes that are not regulated by these factors. **C.** Higher mean expression of genes that are upregulated by GR and STAT3 is associated with shorter relapse-free survival in basal-like breast cancer patients.

GR and STAT3 regulate the basal-like gene expression program

To test the hypothesis that GR and STAT3 are key regulators of the basal-like gene expression program, RNA-seq data from 28 breast cancer cell lines (18 basal-like and 10 luminal) were analyzed to identify genes that exhibit subtype-specific expression. Of the 2,485 genes that had significantly higher expression in basal-like breast cancer, 439 were significantly upregulated in HCC70 and MDA-MB-231 when comparing dexamethasone induction to baseline media, and standard media to STAT3 siRNA or SH4-54 inhibition of STAT3. These 439 basal-like genes that are up-regulated by GR and STAT3 have a significant overlap with genes that were previously characterized as more highly expressed in basal-like breast cancer compared to luminal (Hypergeometric FDR adjusted q-value = 2.6×10^{-95})(37). This result confirms that the GR and STAT3 regulated genes identified in this analysis have been associated with basal-like breast cancer in previous studies.

Further characterization of the 439 basal-like genes upregulated by GR and STAT3 revealed enrichment for cellular processes and pathways associated with the aggressive nature of basal-like breast cancer. These genes were significantly enriched for the Gene Ontology Biological Process “Positive Regulation of Cell Proliferation” (Hypergeometric FDR adjusted q-value 2.04×10^{-17}) (38-40), “hallmarks that define the epithelial-mesenchymal transition” (Hypergeometric FDR adjusted q-value 7.22×10^{-15}) (38-40), “mammary stem cells” (Hypergeometric FDR adjusted q-value = 1.04×10^{-22}) (41), and “adult tissue stem cells” (Hypergeometric FDR adjusted q-value = 2.13×10^{-20})(42).

Notably, of the 439 basal-like genes upregulated by GR and STAT3, 34 are TFs themselves, including those that are known to be upregulated by major signaling pathways involved in basal-like breast cancer growth including EGFR signaling (ELK3, ETS2, CEBPD, NFIL3, MBNL2, HIVEP2) (43), and TGFB signaling (ELK3, ETS2, SNAI1, MAFB)(44). Additionally, the genes regulated by GR and/or STAT3 include 5 TFs (ETV6, NFIL3, HIF1A, DR1, and TFCEP2L1) that were previously identified as “embryonic stem cell” TFs that are preferentially and coordinately overexpressed in the high-grade, ER-negative breast cancer tumors(45). These results suggest GR and STAT3 are upstream of the key TFs that control growth and stemness features in basal-like TNBC.

These 439 basal-like genes are upregulated by GR and STAT3 in both HCC70 and MDA-MB-231 cell lines. To determine if they are associated with prognosis in patients with basal-like breast cancer, Kaplan Meier survival plots were generated using public gene expression microarray data from 360 basal-like patient tumors for whom

five-year relapse-free survival data were available. Patients who had higher mean expression of these genes in their tumors had significantly shorter relapse-free survival than patients with lower expression of these genes (Hazard Ratio =1.76, logrank $p=0.0068$) (**Figure 6C**). This result indicates that even amongst basal-like patients, higher expression of genes that are upregulated by GR and STAT3 is associated with a worse prognosis.

Inhibiting both GR and STAT3 reduces cell proliferation synergistically

The observation that GR and STAT3 cooperate to regulate a large number of genes in the basal-like gene expression signature that are associated with cell proliferation and poor prognosis led us to investigate whether drugs that modulate GR and STAT3 would affect cell proliferation. To first evaluate the independent roles of GR and STAT3 in cell proliferation, the MDA-MB-231 cell line was treated with a range of doses of dexamethasone or the STAT3 inhibitor SH4-54 separately and was monitored for 72 hours on the Incucyte Zoom live cell imaging instrument. Reducing STAT3 by increasing doses of the STAT3 inhibitor SH4-54 resulted in decreased growth in a dose-dependent manner (**Figure 7A**). Reducing GR activity by decreasing amounts of dexamethasone resulted in decreased cell growth in a dose-dependent manner (**Figure 7B**). These results are consistent with previous publications that report that inhibiting GR or STAT3 separately decreases basal-like breast cancer proliferation (18,46-48).

To evaluate whether modulating both TFs simultaneously leads to differences in proliferation, the MDA-MB-231 and MDA-MB-468 cell lines were treated with different combinations of dexamethasone and SH4-54. The combination of inhibiting GR by withholding dexamethasone, and inhibiting STAT3 with 8 μ M SH4-54 resulted in the slowest cell proliferation (**Figure 7C & 7D**). Conversely, the activation of GR with 1 μ M dexamethasone combined with the endogenous activity of STAT3 resulted in the fastest cell proliferation (**Figure 7C & 7D**). Notably, when either GR or STAT3 were active alone, this led to intermediate proliferation rates. This result supports the hypothesis that GR and STAT3 cooperate to increase cell proliferation, and dual inhibition of both GR and STAT3 results in slower cell proliferation than inhibiting either factor alone.

To expand this analysis beyond cell lines, proliferation of patient-derived organoid cultures was analyzed in the context of GR and STAT3 inhibition. Two patient-derived organoid cultures previously derived from primary basal-like TNBC tumors (HCI-001 and HCI-002) were utilized for this experiment(49,50). The standard media for the patient-derived organoid cultures includes 1 μ g/ml hydrocortisone, which is a glucocorticoid that induces GR activity. Therefore, escalating doses of mifepristone were used to inhibit GR activity. Mifepristone also inhibits

progesterone receptor (PR) activity, but basal-like cells do not express PR, and thus this is not a confounding factor in these experiments. The STAT3 inhibitor SH4-54 was used in escalating doses to inhibit STAT3 activity. The cell line MDA-MB-231 was included in this experiment as a positive control and was grown in the same media as the organoids. The proliferation rate of the organoids is much slower than the cell line, which resulted in a narrower dynamic range for observing changes in proliferation. Despite this technical limitation, inhibition of GR and STAT3 for 72 hours resulted in significant decreases in cell proliferation in a dose-dependent manner (**Figure 7E**). An analysis of the relative decrease in cell proliferation across drug doses indicates that inhibiting both GR and STAT3 provides synergistic, rather than additive, decreases in cell proliferation in the MDA-MB-231 cell line, as well as both patient-derived organoid cultures (**Figure 7E**). This result confirms that inhibiting the activity of both GR and STAT3 reduces cell proliferation in basal-like breast cancer patient-derived organoids.

Figure 7. Inhibiting GR and STAT3 leads to decreased growth and proliferation in basal-like cell lines and patient-derived organoids. **A.** MDA-MB-231 cells treated with increasing doses of the STAT3 inhibitor SH4-54 have decreased growth rate in a dose-dependent manner. **B.** MDA-MB-231 cells treated with decreasing doses of dexamethasone have decreased growth rate in a dose-dependent manner. Inhibiting both GR and STAT3 in MDA-MB-231 (**C**) and MDA-MB-468 (**D**) cell lines leads to decreased growth compared to inhibiting either factor alone. **E.** Synergistic decreases in cell proliferation are observed after 72 hours when patient-derived organoid cultures (HCI-001 and HCI-002) and MDA-MB-231 cell lines are treated with increasing doses of both the STAT3 inhibitor SH4-54 and the GR inhibitor mifepristone.

Discussion

The goal of this study was to determine which TFs drive the basal-like TNBC gene expression program. Genome-wide analysis of DNA methylation, chromatin accessibility, TF binding, and gene expression indicate that GR and STAT3 binding is enriched at genomic regulatory regions that are specifically open in basal-like TNBC. These TFs bind to the same regulatory regions and cooperate to regulate hundreds of genes in the basal-like expression signature, including many other TFs. The genes that are cooperatively up-regulated by GR and STAT3 are transducers of major signaling pathways implicated in basal-like breast cancer growth (EGFR and TGF β), are associated with aggressive features of basal-like breast cancer (proliferation, stemness and epithelial to mesenchymal transition), and are associated with shorter relapse-free survival in patients. Furthermore, inhibiting both GR and STAT3 leads to synergistic reductions in cell proliferation in multiple cell lines and patient-derived organoid cultures. Together these data suggest that inhibiting both GR and STAT3 simultaneously can reduce expression of the basal-like gene expression signature and create a synthetic lethality in TNBC that could be exploited therapeutically.

While GR and STAT3 have been separately implicated in regulating gene expression in basal-like breast cancer previously(18,48,51,52), their cooperative role has not been previously described. This study makes it apparent that understanding this cooperation is crucial to investigating their roles in driving basal-like TNBC gene expression and phenotypes, as well as arriving at more effective therapeutic strategies. Inhibiting GR alone is not effective; basal-like cell lines proliferate and express some components of the basal-like gene signature without glucocorticoids in their media. Inhibiting STAT3 alone is not completely effective; previous studies have shown that SH4-54 killed basal-like cell lines in culture, but was insufficient to completely stop tumor growth in vivo where glucocorticoids are naturally present (24). These results indicate that inhibiting either of these TFs alone is not sufficient to abrogate the pro-proliferative gene expression program of basal-like breast cancer in vivo, but this study suggests that inhibiting both GR and STAT3 could be a novel and effective therapeutic strategy.

The study of GR's role in basal-like breast cancer has important clinical implications because potent synthetic glucocorticoids are often prescribed to breast cancer patients undergoing chemotherapy to prevent side-effects such as nausea, loss of appetite, and rare severe immune reactions. Our preliminary data supports a handful of other studies that report that induction of GR with glucocorticoids could promote tumor growth and therapy

resistance (15,17,18,52). In our study specifically, it is clear that glucocorticoid induction of GR leads to increased cell proliferation in basal-like breast cancer cell lines and patient-derived organoids. It also leads to increased expression of genes that are cooperatively regulated by STAT3, a known promoter of cancer growth and invasion (53-58). Recently, clinical trials have been initiated to test whether inhibiting GR directly with mifepristone or by inhibiting its chaperone, Hsp90, will enhance the efficacy of chemotherapy in basal-like breast cancer (16,59). Our results support this line of investigation and suggest that combination therapy that inhibits both GR and STAT3 could be even more effective in treating this disease.

Acknowledgements

This work was supported by American Cancer Society Research Scholar Grant Award # 132596-RSG-18-197-01-DMC awarded to KEV. We acknowledge support of funds in conjunction with grant P30 CA042014 awarded to Huntsman Cancer Institute and to the Nuclear Control Program at Huntsman Cancer Institute. Research reported in this publication utilized the High-Throughput Genomics and Bioinformatic Analysis Shared Resource at Huntsman Cancer Institute at the University of Utah and was supported by the National Cancer Institute of the National Institutes of Health under Award Number P30 CA042014. This research was also supported by funding from the HudsonAlpha Institute, a Cancer Research Fund from the State of Alabama to HudsonAlpha Institute, Susan G. Komen, and the NIH-National Cancer Institute Comprehensive Cancer Center Core Support Grant (5P30CA013148). JMM was supported by a fellowship from the National Science Foundation Graduate Research Fellowship Program (NSF-GRFP).

Author Contributions

MEC, JMM, JMG, KPG, PGO, and SLP performed experiments. MEC and KEV performed analysis and wrote the manuscript. PY and JT provided reagents. DJB, BEW, and RMM provided expertise and feedback.

Declaration of Interests

The authors declare no competing interests related to this study.

Materials and Methods

Cell Lines

The following basal-like breast cancer cell lines were used for this study: HCC1569, HCC1954, BT-20, HCC1143, HCC1187, HCC1599, HCC1937, HCC70, MDA-MB-468, 2-LMP, BT-549, HCC38, MDA-MB-157, MDA-MB-231, MDA-MB-436, SUM102, SUM149, SUM159.

The following luminal breast cancer cell lines were used for this study: DY36T2, ZR-75-30, MDA-MB-453, SK-BR-3, BT-474, MDA-MB-361, MCF-7, MDA-MB-134, T-47D, ZR-75-1.

Cells were obtained and cultured as described previously(60). Cell lines passaged for more than 3 months were authenticated every 6 months using Short Tandem Repeat Testing service offered through the University of Utah DNA Sequencing Core Facility.

Reduced Representation Bisulfite Sequencing (RRBS)

RRBS and primary analysis of CG methylation was performed on the panel of cell lines using the method described previously(61). CG positions with at least 10x coverage in at least 80% of the cell lines were used for subsequent analysis. To identify CGs that were differentially methylated between basal-like and luminal cell lines, Linear Regression was performed in the R software package version 3.5.0 using the lm function. P-values were adjusted for multiple hypothesis testing using the p.adjust function in R and the Benjamini-Hochberg method. CGs with adjusted p-values less than 0.05 were visualized in R using the pheatmap package version 1.0.10.

Enrichment of ENCODE Transcription Factor Binding Sites

A file named wgEncodeRegTfbsClusteredV2.bed.gz containing TF binding sites from ChIP-seq experiments performed on 149 TFs by the ENCODE Project(26) in a variety of cell lines were downloaded from the UCSC genome browser(62). Bedtools intersect (63) was used to determine which CG positions in the RRBS dataset overlapped binding site positions for each TF in the file. The number of overlaps with basal-like and luminal specific CGs was normalized to the number of overlaps to all CGs in RRBS to calculate a fold-enrichment.

Analysis of TCGA RNA-seq and ATAC-seq

RNA-seq counts file for TCGA Breast Cancer samples were downloaded from the UCSC Xena browser(64,65). These files contained log counts, so they were transformed to counts in R version 3.5.0. DESEQ2 version 1.20.0 (66) was used to identify genes that were differentially expressed between basal-like and luminal tumors using

the PAM50 Subtype classification provided by TCGA and available for download through the UCSC Xena browser(64,65) . Basal specific genes were identified using an adjusted p-value < 0.05 and coefficients less than 0. Luminal specific genes were identified using a p-value < 0.05 and coefficients greater than 0.

ATAC-seq data from TCGA breast tumors were downloaded as counts files and bigwig files from the supplementary material provided in the recent publication by Corces et al. (29). DESEQ2 version 1.20.0 (66) was used to identify peaks that were differentially accessible between basal-like and luminal tumors using the PAM50 Subtype classification provided by TCGA and available for download through the UCSC Xena browser(64,65). Pheatmap package version 1.0.10 was used to visualize basal-specific peaks that were identified using an adjusted p-value < 0.05 and coefficients less than 0. The ATAC-seq peaks specifically open in basal-like tumors were analyzed for enrichment of motifs for JUN (MA0491.1), STAT3 (MA0144.1), and GR(MA0113.2) downloaded from JASPAR2018(67). AME(68) was used to perform the analysis of sequences in ATAC-seq peaks compared to a primary sequence shuffled conserving 2-mer frequencies, and p-values were calculated using Fisher's exact test.

To analyze ATAC-seq signal from TCGA breast tumors over the basal-specific GR and STAT3 shared binding sites, the UCSC hgLiftOver was used to convert the GR and STAT3 shared sites bed file to the hg38 genome build to be compatible with TCGA ATAC-seq alignments. Deeptools version 3.1.0(69) computeMatrix was used to generate a matrix of counts over the GR and STAT3 shared sites for each sample. The mean accessibility was computed across the regions for each sample using R version 3.5.0. The Mann Whitney test was used to determine if the average normalized ATAC-seq read depth across regions was significantly different between basal-like and luminal tumors. The mean and standard deviation of the average normalized depth across basal-like and luminal samples was graphed using GraphPad Prism 7.0c.

Violin plots for averaged bigwig read counts from basal-like and luminal tumors at GR and STAT3 shared sites were created with vioplot_0.3.2, using default parameters and areaEqual=T, h=0.04, wex=0.9.

ChIP-seq

ChIP-seq for GR and STAT3 was performed on 4 basal-like cell lines (SUM159, MDA-MB-231, HCC1937, and HCC70) and 4 luminal cell lines (MDA-MB-361, BT-474, MDA-MB-453, and MCF-7).

For ChIP experiments, protein-DNA complexes were covalently cross-linked by incubating cells in 1% formaldehyde for 10 min at room temperature. Cells were incubated with 0.125 M glycine for 5 min, to quench

cross-linking reaction. Cells were washed and scraped with PBS (pH 7.4) (Lonza). Cells were lysed with Farnham Lysis Buffer (5mM PIPES at pH 8.0, 85 mM KCl, 0.5% NP-40) containing protease inhibitor (Roche). Cell lysate was centrifuged at 2,000 rpm for 5 min at 4 °C. The crude nuclear extract contained in the supernatant was stored at -80 °C. ChIP-seq was performed as previously described(70) using antibodies for GR (sc-1003, Santa Cruz Biotechnology) and STAT3 (sc-482, Santa Cruz Biotechnology). A thorough version of the ChIP-seq protocol used in this study is available on the ENCODE Project website:

https://www.encodeproject.org/documents/df9dd0ec-c1cf-4391-a745-a933ab1af7a7/@@download/attachment/Myers_Lab_ChIP-seq_Protocol_v042211.pdf

STAT3 ChIP-seq datasets have been deposited in the NCBI Gene Expression Omnibus (GEO) accession number GSE85579.

Fastq files from ChIP-seq were aligned to the hg19 build of the human genome using Bowtie with the following parameters: -m 1 -t -best -q -S -l 32 -e 80 -n 2. ChIP-seq peaks were identified by comparing GR ChIP-seq in cells induced with dexamethasone to GR ChIP-seq in cells treated with ethanol, and STAT3 ChIP-seq to input control libraries. Peaks were called using Model-Based Analysis of ChIP-seq-2 (MACS2) (71) with a p value cutoff of 1e -10 and the mfold parameter constrained between 15 and 100. Bedtools merge (63) was used to merge bed files of MACS2 narrow peak calls from each of the ChIP-seq experiments. Bedtools coverageBed (63) was used to extract read counts under each merged peak in each ChIP-seq experiment in each cell line. DESEQ2 version 1.20.0 (66) was used to identify peaks with significantly different read depth between basal-like and luminal cell lines (adjusted p-value < 0.05), and significantly different read depth between GR and STAT3 ChIP-seq experiments (adjusted p-value < 0.05). A multivariate model was used to identify shared GR and STAT3 peaks whose read depth was significantly different between subtypes in both GR and STAT3 ChIP-seq experiments (adjusted p-value < 0.05), but not significantly different between GR and STAT3 in the same subtype (adjusted p-value > 0.05). Pheatmap package version 1.0.10 and Deeptools version 3.1.0(69) computeMatrix and plotHeatmap functions were used to create the heatmaps of ChIP-seq data.

RNA-seq

RNA-seq of the panel of cell lines was described previously (72) and the data are publicly available through NCBI Gene Expression Omnibus (GEO) accession number GSE58135.

RNA-seq was also performed on MDA-MB-231 and HCC70 cell lines treated alone or in combination with dexamethasone, STAT3 siRNA, and SH4-54 as follows:

STAT3 siRNA knockdown utilized the ON-TARGETplus Human STAT3 siRNA kit from GEHealthcare (L-003544-00-0005, GE Healthcare). This SMARTpool siRNA contains four pooled siRNAs, each targeting a separate region of the STAT3 RNA sequence. ON-TARGETplus Non-targeting siRNA #1 (D-001810-01-05) as a non-targeting control. The siRNA SMARTpool and Non-targeting siRNA target sequences are below:

ON-TARGETplus SMARTpool siRNA J-003544-07, STAT3: GAGAUUGACCAGCAGUAUA

ON-TARGETplus SMARTpool siRNA J-003544-08, STAT3: CAACAUGUCAUUUGCUGAA

ON-TARGETplus SMARTpool siRNA J-003544-09, STAT3: CCAACAAUCCCAAGAAUGU

ON-TARGETplus SMARTpool siRNA J-003544-10, STAT3: CAACAGAUUGCCUGCAUUG

ON-TARGETplus Non-targeting siRNA #1: UGGUUUACAUGUCGACUAA

The siRNA transfection experiments were performed in 6-well plates in triplicate. Lipofectamine RNAiMAX Transfection Reagent (ThermoFisher Scientific) was used per manufacturer's instructions. To each well containing cells, 250 μ l siRNA-transfection reagent mix was added to each well, for a final concentration of 25 pmol siRNA in 7.5 μ l Lipofectamine RNAiMAX reagent per well.

The dexamethasone treatments were performed in 6-well plates in triplicate. Cell were treated with 100 nM dexamethasone (D4902, Sigma Aldrich) or an equal volume of 100% molecular biology grade ethanol vehicle control (E7023, Sigma Aldrich) for 4 hours. Cells were lysed with Buffer RL (Norgen Biotek) containing 10% beta-mercaptoethanol. Total RNA was extracted using the Animal Tissue RNA Purification Kit (Norgen Biotek).

RNA-seq libraries for the STAT3 siRNA and Dexamethasone treatment were prepared from 250 ng total RNA via polyA-selection (Dynabead mRNA Purification Kit, Invitrogen) followed by transposase-mediated non-stranded library construction(73). Libraries were pooled and sequenced on an Illumina HiSeq 2000 or HiSeq 2500 sequencer using paired-end 50bp reads with a 6 bp index read. TopHat v1.4.1 was used to align RNA-seq paired reads to GENCODE version 9. Cufflinks v1.3.0 and BEDTools were used to calculate raw counts for each GENCODE transcript.

DESEQ2 version 1.20.0 (66) was used to identify genes with significantly different expression between treatment and control conditions in each cell line separately. To visualize differential expression of genes across cell lines with different base-line expression, gene expression was divided by the mean in the control sample, and normalized values were displayed in a heatmap using Pheatmap package version 1.0.10 in R.

For the GR and STAT3 interaction studies, MDA-MB-231 cells were plated in 6-well plates at ~300,000 cells in 4ml of media per well and grown for 24 hours. Three replicates wells were treated for 24 hours with 1 μ M of Dexamethasone (DEX), 8 μ M SH4-54, no treatment, or combination of DEX and SH4-54. The cells were lysed using 350 μ L of Qiagen RLT buffer and 1% BME. RNA was extracted using Norgen Animal Tissue RNA Purification Kit according to manufacturer instructions. RNA-seq libraries were prepared using Illumina TruSeq Stranded RNA Kit with Ribo-Zero Gold and sequenced on the Illumina HiSeq 2500 as 50 cycle single-end reads. Reads were mapped using HISAT2(74) version 2.1.0 with parameters --dta -p 28 -t -q -U and sorted with samtools(75). A counts file was generated from the bam files using featureCounts version 1.5.1(76).

To identify genes that were additively and synergistically regulated by both GR and STAT3 the linear model (lm) function in R was used. DESEQ2 normalized counts were fit to a multivariate model that included a term for GR activity, a term for STAT3 activity, and an interaction term (GR:STAT3). Coefficients and p-values for each gene were used to identify those with additive and synergistic positive and negative regulation by GR and STAT3. Synergy was defined as a significant p-value (<0.05) for the interaction term, and non-significant p-values (>0.05) for the individual GR and STAT3 terms. Additivity was defined as a non-significant p-value for the interaction term, and significant p-values for both of the individual GR and STAT3 terms with coefficients in the same direction. Coefficients from the linear model were used to determine whether the interactions between GR and STAT3 led to increased or decreased gene expression.

RNA-seq data from STAT3 siRNA treatments are publicly available through NCBI Gene Expression Omnibus (GEO) accession number: GSE85579.

Kaplan Meier analysis

The Kaplan-Meier Plotter tool (<http://kmplot.com>)(32) was used to analyze the association between gene expression and relapse-free survival. A Kaplan Meier survival plot was generated using public gene expression microarray data from patient tumors for whom relapse-free survival data were available. The intrinsic subtype classification provided by the Kaplan-Meier Plotter tool was used to select basal-like and luminal cases for analysis (77). The analysis was performed for GR, STAT3 and 439 basal-like genes that are upregulated by GR and STAT3. The following selections were applied to all analyses: only one JetSet best probe (78) for each gene was used, multigene signatures used the mean expression of the selected probes, relapse-free survival was selected for the analysis, patients were censored at the follow-up threshold (60 or 120 months), biased arrays

were excluded, and redundant samples were removed. The most significant cutpoint was used to split patients into two groups (“autoselect best cutoff” option).

Live cell imaging of proliferation

Proliferation assays were performed using the IncuCyte ZOOM live cell imaging system (Essen BioSciences). MDA-MB-231 and MDA-MB-436 cells were plated at 20,000 cells per well in 96 well plates. For titration experiments, cells were treated a range of doses of SH4-54 between 0uM and 11uM of SH4-54, or a range of doses of dexamethasone between 0uM and 2uM. For combination treatments cells were treated with 8uM SH4-54 and 1uM dexamethasone, alone and in combination. Plates were imaged every 2 hours for 72 hours, and cell confluence was calculated by the IncuCyte ZOOM software. Relative percent confluence was calculated by subtracting the percent confluence at each timepoint by the starting percent confluence at the first timepoint.

Patient derived organoid proliferation assay

Two patient derived organoid cultures previously derived from primary basal-like TNBC tumors (HCI-001 and HCI-002) were utilized for this experiment(49,50). The organoids were cultured in media composed of Advanced DMEM/F12, (Thermo Fisher 12634028), 5% FBS ThermoFisher 26140079), 1x Hepes (ThermoFisher 15630080), 1x GlutaMAX Supplement (Thermo Fisher 35050061), 1 ug/ml Hydrocortisone (Sigma Aldrich H0888), 10 ng/ml hEGF (Sigma Aldrich E9644), 50 ug/ml Gentamicin (Genesee 25-533), 10 uM Y-27632 (Selleckchem S1049). Organoids were seeded as 50-100 organoids in 5% Matrigel (Corning 354230) per well of a 384 well plate, as described previously(49). MDA-MB-231 cell lines were included as a positive control in this experiment and plated at 500 cells per well in a 384 well plate. Cells were treated with all possible combinations of titrated doses of SH4-54 (0, 1, 2.1, 3.2, 4.3, 5.4, 6.5, 7.6, 8.8, 9.9, and 11uM) and mifepristone (0, 0.1, 1.42, 2.73, 4.05, 5.37, 6.68, and 8uM) in combination. Cell viability was measured relative to untreated cells after 72 hours using the CellTiter Glo 3D (Promega) that measures ATP. Lowe additivity analysis was used to compute synergy scores(79).

Literature Cited:

1. Prat A, Perou CM. Deconstructing the molecular portraits of breast cancer. *Mol Oncol* **2011**;5(1):5-23 doi 10.1016/j.molonc.2010.11.003.
2. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, *et al.* Molecular portraits of human breast tumours. *Nature* **2000**;406(6797):747-52 doi 10.1038/35021093.
3. 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 Res* **2007**;13(15 Pt 1):4429-34 doi 10.1158/1078-0432.CCR-06-3045.
4. Pal S, Luchtenborg M, Davies EA, Jack RH. The treatment and survival of patients with triple negative breast cancer in a London population. *Springerplus* **2014**;3:553 doi 10.1186/2193-1801-3-553.
5. Prat A, Adamo B, Cheang MC, Anders CK, Carey LA, Perou CM. Molecular characterization of basal-like and non-basal-like triple-negative breast cancer. *Oncologist* **2013**;18(2):123-33 doi 10.1634/theoncologist.2012-0397.
6. Poli V, Fagnocchi L, Fasciani A, Cherubini A, Mazzoleni S, Ferrillo S, *et al.* MYC-driven epigenetic reprogramming favors the onset of tumorigenesis by inducing a stem cell-like state. *Nat Commun* **2018**;9(1):1024 doi 10.1038/s41467-018-03264-2.
7. Peluffo G, Subedee A, Harper NW, Kingston N, Jovanovic B, Flores F, *et al.* EN1 is a transcriptional dependency in triple-negative breast cancer associated with brain metastasis. *Cancer Res* **2019** doi 10.1158/0008-5472.CAN-18-3264.
8. Sirkisoon SR, Carpenter RL, Rimkus T, Anderson A, Harrison A, Lange AM, *et al.* Interaction between STAT3 and GLI1/tGLI1 oncogenic transcription factors promotes the aggressiveness of triple-negative breast cancers and HER2-enriched breast cancer. *Oncogene* **2018**;37(19):2502-14 doi 10.1038/s41388-018-0132-4.
9. Poma P, Labbozzetta M, D'Alessandro N, Notarbartolo M. NF-kappaB Is a Potential Molecular Drug Target in Triple-Negative Breast Cancers. *OMICS* **2017**;21(4):225-31 doi 10.1089/omi.2017.0020.
10. Samanta D, Gilkes DM, Chaturvedi P, Xiang L, Semenza GL. Hypoxia-inducible factors are required for chemotherapy resistance of breast cancer stem cells. *Proc Natl Acad Sci U S A* **2014**;111(50):E5429-38 doi 10.1073/pnas.1421438111.
11. Finlay J, Roberts CM, Lowe G, Loeza J, Rossi JJ, Glackin CA. RNA-based TWIST1 inhibition via dendrimer complex to reduce breast cancer cell metastasis. *Biomed Res Int* **2015**;2015:382745 doi 10.1155/2015/382745.
12. Mott L, Su K, Pack DW. Evaluation of FOXC1 as a therapeutic target for basal-like breast cancer. *Cancer Gene Ther* **2018** doi 10.1038/s41417-018-0010-9.
13. Elian FA, Yan E, Walter MA. FOXC1, the new player in the cancer sandbox. *Oncotarget* **2018**;9(8):8165-78 doi 10.18632/oncotarget.22742.
14. Qiao Y, He H, Jonsson P, Sinha I, Zhao C, Dahlman-Wright K. AP-1 Is a Key Regulator of Proinflammatory Cytokine TNFalpha-mediated Triple-negative Breast Cancer Progression. *J Biol Chem* **2016**;291(10):5068-79 doi 10.1074/jbc.M115.702571.
15. Li Z, Dong J, Zou T, Du C, Li S, Chen C, *et al.* Dexamethasone induces docetaxel and cisplatin resistance partially through up-regulating Kruppel-like factor 5 in triple-negative breast cancer. *Oncotarget* **2017**;8(7):11555-65 doi 10.18632/oncotarget.14135.
16. Nanda R, Stringer-Reasor EM, Saha P, Kocherginsky M, Gibson J, Libao B, *et al.* A randomized phase I trial of nanoparticle albumin-bound paclitaxel with or without mifepristone for advanced breast cancer. *Springerplus* **2016**;5(1):947 doi 10.1186/s40064-016-2457-1.
17. Regan Anderson TM, Ma SH, Raj GV, Cidlowski JA, Helle TM, Knutson TP, *et al.* Breast Tumor Kinase (Brk/PTK6) Is Induced by HIF, Glucocorticoid Receptor, and PELP1-Mediated Stress Signaling in Triple-Negative Breast Cancer. *Cancer Res* **2016**;76(6):1653-63 doi 10.1158/0008-5472.CAN-15-2510.
18. Skor MN, Wonder EL, Kocherginsky M, Goyal A, Hall BA, Cai Y, *et al.* Glucocorticoid receptor antagonism as a novel therapy for triple-negative breast cancer. *Clin Cancer Res* **2013**;19(22):6163-72 doi 10.1158/1078-0432.CCR-12-3826.
19. Banerjee K, Resat H. Constitutive activation of STAT3 in breast cancer cells: A review. *Int J Cancer* **2016**;138(11):2570-8 doi 10.1002/ijc.29923.
20. McDaniel JM, Varley KE, Gertz J, Savic DS, Roberts BS, Bailey SK, *et al.* Genomic regulation of invasion by STAT3 in triple negative breast cancer. *Oncotarget* **2017**;8(5):8226-38 doi 10.18632/oncotarget.14153.

21. Okita Y, Kimura M, Xie R, Chen C, Shen LT, Kojima Y, *et al.* The transcription factor MAFK induces EMT and malignant progression of triple-negative breast cancer cells through its target GPNMB. *Sci Signal* **2017**;10(474) doi 10.1126/scisignal.aak9397.
22. Egusquiguire SP, Yeh JE, Walker SR, Liu S, Frank DA. The STAT3 Target Gene TNFRSF1A Modulates the NF-kappaB Pathway in Breast Cancer Cells. *Neoplasia* **2018**;20(5):489-98 doi 10.1016/j.neo.2018.03.004.
23. Gonzalez-Gonzalez A, Munoz-Muela E, Marchal JA, Cara FE, Molina MP, Cruz-Lozano M, *et al.* Activating Transcription Factor 4 Modulates TGFbeta-Induced Aggressiveness in Triple-Negative Breast Cancer via SMAD2/3/4 and mTORC2 Signaling. *Clin Cancer Res* **2018**;24(22):5697-709 doi 10.1158/1078-0432.CCR-17-3125.
24. Yue P, Lopez-Tapia F, Paladino D, Li Y, Chen CH, Namanja AT, *et al.* Hydroxamic Acid and Benzoic Acid-Based STAT3 Inhibitors Suppress Human Glioma and Breast Cancer Phenotypes In Vitro and In Vivo. *Cancer Res* **2016**;76(3):652-63 doi 10.1158/0008-5472.CAN-14-3558.
25. Lambert M, Jambon S, Depauw S, David-Cordonnier MH. Targeting Transcription Factors for Cancer Treatment. *Molecules* **2018**;23(6) doi 10.3390/molecules23061479.
26. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. *Nature* **2012**;489(7414):57-74 doi 10.1038/nature11247.
27. Kong SL, Li G, Loh SL, Sung WK, Liu ET. Cellular reprogramming by the conjoint action of ERalpha, FOXA1, and GATA3 to a ligand-inducible growth state. *Mol Syst Biol* **2011**;7:526 doi 10.1038/msb.2011.59.
28. Nakshatri H, Badve S. FOXA1 in breast cancer. *Expert Rev Mol Med* **2009**;11:e8 doi 10.1017/S1462399409001008.
29. Corces MR, Granja JM, Shams S, Louie BH, Seoane JA, Zhou W, *et al.* The chromatin accessibility landscape of primary human cancers. *Science* **2018**;362(6413) doi 10.1126/science.aav1898.
30. Lerner L, Henriksen MA, Zhang X, Darnell JE, Jr. STAT3-dependent enhanceosome assembly and disassembly: synergy with GR for full transcriptional increase of the alpha 2-macroglobulin gene. *Genes Dev* **2003**;17(20):2564-77 doi 10.1101/gad.1135003.
31. Heiser LM, Sadanandam A, Kuo WL, Benz SC, Goldstein TC, Ng S, *et al.* Subtype and pathway specific responses to anticancer compounds in breast cancer. *Proc Natl Acad Sci U S A* **2012**;109(8):2724-9 doi 10.1073/pnas.1018854108.
32. Györfy B, Lanczky A, Eklund AC, Denkert C, Budczies J, Li Q, *et al.* An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast Cancer Res Treat* **2010**;123(3):725-31 doi 10.1007/s10549-009-0674-9.
33. Turkson J, Jove R. STAT proteins: novel molecular targets for cancer drug discovery. *Oncogene* **2000**;19(56):6613-26 doi 10.1038/sj.onc.1204086.
34. Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T, *et al.* A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* **2006**;10(6):515-27 doi 10.1016/j.ccr.2006.10.008.
35. Cancer Genome Atlas N. Comprehensive molecular portraits of human breast tumours. *Nature* **2012**;490(7418):61-70 doi 10.1038/nature11412.
36. Latchman DS. Transcription factors: bound to activate or repress. *Trends Biochem Sci* **2001**;26(4):211-3.
37. Charafe-Jauffret E, Ginestier C, Monville F, Finetti P, Adelaide J, Cervera N, *et al.* Gene expression profiling of breast cell lines identifies potential new basal markers. *Oncogene* **2006**;25(15):2273-84 doi 10.1038/sj.onc.1209254.
38. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **2000**;25(1):25-9 doi 10.1038/75556.
39. The Gene Ontology C. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res* **2019**;47(D1):D330-D8 doi 10.1093/nar/gky1055.
40. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **2005**;102(43):15545-50 doi 10.1073/pnas.0506580102.
41. Lim E, Wu D, Pal B, Bouras T, Asselin-Labat ML, Vaillant F, *et al.* Transcriptome analyses of mouse and human mammary cell subpopulations reveal multiple conserved genes and pathways. *Breast Cancer Res* **2010**;12(2):R21 doi 10.1186/bcr2560.

42. Wong DJ, Liu H, Ridky TW, Cassarino D, Segal E, Chang HY. Module map of stem cell genes guides creation of epithelial cancer stem cells. *Cell Stem Cell* **2008**;2(4):333-44 doi 10.1016/j.stem.2008.02.009.
43. Creighton CJ, Hilger AM, Murthy S, Rae JM, Chinnaiyan AM, El-Ashry D. Activation of mitogen-activated protein kinase in estrogen receptor alpha-positive breast cancer cells in vitro induces an in vivo molecular phenotype of estrogen receptor alpha-negative human breast tumors. *Cancer Res* **2006**;66(7):3903-11 doi 10.1158/0008-5472.CAN-05-4363.
44. Padua D, Zhang XH, Wang Q, Nadal C, Gerald WL, Gomis RR, *et al.* TGFbeta primes breast tumors for lung metastasis seeding through angiopoietin-like 4. *Cell* **2008**;133(1):66-77 doi 10.1016/j.cell.2008.01.046.
45. Ben-Porath I, Thomson MW, Carey VJ, Ge R, Bell GW, Regev A, *et al.* An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat Genet* **2008**;40(5):499-507 doi 10.1038/ng.127.
46. Deng XS, Wang S, Deng A, Liu B, Edgerton SM, Lind SE, *et al.* Metformin targets Stat3 to inhibit cell growth and induce apoptosis in triple-negative breast cancers. *Cell Cycle* **2012**;11(2):367-76 doi 10.4161/cc.11.2.18813.
47. Stover DG, Gil Del Alcazar CR, Brock J, Guo H, Overmoyer B, Balko J, *et al.* Phase II study of ruxolitinib, a selective JAK1/2 inhibitor, in patients with metastatic triple-negative breast cancer. *NPJ Breast Cancer* **2018**;4:10 doi 10.1038/s41523-018-0060-z.
48. Qin JJ, Yan L, Zhang J, Zhang WD. STAT3 as a potential therapeutic target in triple negative breast cancer: a systematic review. *J Exp Clin Cancer Res* **2019**;38(1):195 doi 10.1186/s13046-019-1206-z.
49. DeRose YS, Gligorich KM, Wang G, Georgelas A, Bowman P, Courdy SJ, *et al.* Patient-derived models of human breast cancer: protocols for in vitro and in vivo applications in tumor biology and translational medicine. *Curr Protoc Pharmacol* **2013**;Chapter 14:Unit14 23 doi 10.1002/0471141755.ph1423s60.
50. DeRose YS, Wang G, Lin YC, Bernard PS, Buys SS, Ebbert MT, *et al.* Tumor grafts derived from women with breast cancer authentically reflect tumor pathology, growth, metastasis and disease outcomes. *Nat Med* **2011**;17(11):1514-20 doi 10.1038/nm.2454.
51. Yuan J, Zhang F, Niu R. Multiple regulation pathways and pivotal biological functions of STAT3 in cancer. *Sci Rep* **2015**;5:17663 doi 10.1038/srep17663.
52. Chen Z, Lan X, Wu D, Sunkel B, Ye Z, Huang J, *et al.* Ligand-dependent genomic function of glucocorticoid receptor in triple-negative breast cancer. *Nat Commun* **2015**;6:8323 doi 10.1038/ncomms9323.
53. Muellner MK, Mair B, Ibrahim Y, Kerzendorfer C, Lechtermann H, Trefzer C, *et al.* Targeting a cell state common to triple-negative breast cancers. *Mol Syst Biol* **2015**;11(1):789 doi 10.15252/msb.20145664.
54. Yang B, Shen JW, Zhou DH, Zhao YP, Wang WQ, Zhu Y, *et al.* Precise discovery of a STAT3 inhibitor from *Eupatorium lindleyanum* and evaluation of its activity of anti-triple-negative breast cancer. *Nat Prod Res* **2017**;1-9 doi 10.1080/14786419.2017.1396596.
55. Yamashita N, Kondo M, Zhao S, Li W, Koike K, Nemoto K, *et al.* Picrasidine G decreases viability of MDA-MB 468 EGFR-overexpressing triple-negative breast cancer cells through inhibition of EGFR/STAT3 signaling pathway. *Bioorg Med Chem Lett* **2017**;27(11):2608-12 doi 10.1016/j.bmcl.2017.03.061.
56. Liu CY, Su JC, Huang TT, Chu PY, Huang CT, Wang WL, *et al.* Sorafenib analogue SC-60 induces apoptosis through the SHP-1/STAT3 pathway and enhances docetaxel cytotoxicity in triple-negative breast cancer cells. *Mol Oncol* **2017**;11(3):266-79 doi 10.1002/1878-0261.12033.
57. Feng T, Cao W, Shen W, Zhang L, Gu X, Guo Y, *et al.* Arctigenin inhibits STAT3 and exhibits anticancer potential in human triple-negative breast cancer therapy. *Oncotarget* **2017**;8(1):329-44 doi 10.18632/oncotarget.13393.
58. Huynh J, Chand A, Gough D, Ernst M. Therapeutically exploiting STAT3 activity in cancer - using tissue repair as a road map. *Nat Rev Cancer* **2019**;19(2):82-96 doi 10.1038/s41568-018-0090-8.
59. Agyeman AS, Jun WJ, Proia DA, Kim CR, Skor MN, Kocherginsky M, *et al.* Hsp90 Inhibition Results in Glucocorticoid Receptor Degradation in Association with Increased Sensitivity to Paclitaxel in Triple-Negative Breast Cancer. *Horm Cancer* **2016**;7(2):114-26 doi 10.1007/s12672-016-0251-8.
60. Oliver PG, LoBuglio AF, Zhou T, Forero A, Kim H, Zinn KR, *et al.* Effect of anti-DR5 and chemotherapy on basal-like breast cancer. *Breast Cancer Res Treat* **2012**;133(2):417-26 doi 10.1007/s10549-011-1755-0.

61. Varley KE, Gertz J, Bowling KM, Parker SL, Reddy TE, Pauli-Behn F, *et al.* Dynamic DNA methylation across diverse human cell lines and tissues. *Genome Res* **2013**;23(3):555-67 doi 10.1101/gr.147942.112.
62. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, *et al.* The human genome browser at UCSC. *Genome Res* **2002**;12(6):996-1006 doi 10.1101/gr.229102.
63. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **2010**;26(6):841-2 doi 10.1093/bioinformatics/btq033.
64. Goldman M, Craft B, Kamath A, Brooks A, Zhu J, Haussler D. The UCSC Xena Platform for cancer genomics data visualization and interpretation. *bioRxiv* **2018**:326470 doi 10.1101/326470.
65. Jiang G, Zhang S, Yazdanparast A, Li M, Pawar AV, Liu Y, *et al.* Comprehensive comparison of molecular portraits between cell lines and tumors in breast cancer. *BMC Genomics* **2016**;17 Suppl 7:525 doi 10.1186/s12864-016-2911-z.
66. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **2014**;15(12):550 doi 10.1186/s13059-014-0550-8.
67. Khan A, Fornes O, Stigliani A, Gheorghe M, Castro-Mondragon JA, van der Lee R, *et al.* JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res* **2018**;46(D1):D1284 doi 10.1093/nar/gkx1188.
68. McLeay RC, Bailey TL. Motif Enrichment Analysis: a unified framework and an evaluation on ChIP data. *BMC Bioinformatics* **2010**;11:165 doi 10.1186/1471-2105-11-165.
69. Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **2016**;44(W1):W160-5 doi 10.1093/nar/gkw257.
70. Reddy TE, Pauli F, Sprouse RO, Neff NF, Newberry KM, Garabedian MJ, *et al.* Genomic determination of the glucocorticoid response reveals unexpected mechanisms of gene regulation. *Genome Res* **2009**;19(12):2163-71 doi 10.1101/gr.097022.109.
71. Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **2008**;9(9):R137 doi 10.1186/gb-2008-9-9-r137.
72. Varley KE, Gertz J, Roberts BS, Davis NS, Bowling KM, Kirby MK, *et al.* Recurrent read-through fusion transcripts in breast cancer. *Breast Cancer Res Treat* **2014**;146(2):287-97 doi 10.1007/s10549-014-3019-2.
73. Gertz J, Varley KE, Davis NS, Baas BJ, Goryshin IY, Vaidyanathan R, *et al.* Transposase mediated construction of RNA-seq libraries. *Genome Res* **2012**;22(1):134-41 doi 10.1101/gr.127373.111.
74. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* **2015**;12(4):357-60 doi 10.1038/nmeth.3317.
75. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009**;25(16):2078-9 doi 10.1093/bioinformatics/btp352.
76. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **2014**;30(7):923-30 doi 10.1093/bioinformatics/btt656.
77. Mihaly Z, Gyorffy B. Improving Pathological Assessment of Breast Cancer by Employing Array-Based Transcriptome Analysis. *Microarrays (Basel)* **2013**;2(3):228-42 doi 10.3390/microarrays2030228.
78. Li Q, Birkbak NJ, Gyorffy B, Szallasi Z, Eklund AC. Jetset: selecting the optimal microarray probe set to represent a gene. *BMC Bioinformatics* **2011**;12:474 doi 10.1186/1471-2105-12-474.
79. Di Veroli GY, Fornari C, Wang D, Mollard S, Bramhall JL, Richards FM, *et al.* CombeneFit: an interactive platform for the analysis and visualization of drug combinations. *Bioinformatics* **2016**;32(18):2866-8 doi 10.1093/bioinformatics/btw230.