Gene: X 5 (2020) 100030

Contents lists available at ScienceDirect

Gene: X

journal homepage: www.journals.elsevier.com/gene-x

Research paper

Enhanced immortalization, *HUWE1* mutations and other biological drivers of breast invasive carcinoma in Black/African American patients

Terrick Andey[a,1], Michael M. Attah[b], Nana Adwoa Akwaaba-Reynolds[a], Sana Cheema[a], Sara Parvin-Nejad[a], George K. Acquaah-Mensah[a,*,1]

[a] Department of Pharmaceutical Sciences, School of Pharmacy, MCPHS University, 19 Foster St, Worcester, MA 01608, USA
[b] University of Massachusetts, Amherst, MA 01003, USA

ARTICLE INFO

Keywords:
Race
Breast invasive carcinoma
Immortalization
African
Triple-negative breast cancer
Molecular subtype

ABSTRACT

Black/African-American (B/AA) breast cancer patients tend to have more aggressive tumor biology compared to White/Caucasians. In this study, a variety of breast tumor molecular expression profiles of patients derived from the two racial groupings were investigated. Breast invasive carcinoma sample data (RNASeq version 2, Reverse Phase Protein Array, mutation, and miRSeq data) from the Cancer Genome Atlas were examined. The results affirm that B/AA patients are more likely than Caucasian patients to harbor the aggressive *basal-like* or the poor prognosis-associated *HER2-enriched* molecular subtypes of breast cancer. There is also a higher incidence of the triple-negative breast cancer (TNBC) among B/AA patients than the general population, a fact reflected in the mutation patterns of genes such as *PIK3CA* and *TP53*. Furthermore, an immortalization signature gene set, is enriched in samples from B/AA patients. Among stage III patients, TERT, DRAP1, and PQBP1, all members of the immortalization gene signature set, are among master-regulators with increased activity in B/AA patients. Master-regulators driving differences in expression profiles between the two groups include immortalization markers, senescence markers, and immune response and redox gene products. Differences in expression, between B/AA and Caucasian patients, of *RB1*, *hsa-let-7a*, *E2F1*, *c-MYC*, *TERT*, and other biomolecules appear to cooperate to enhance entry into the S-phase of the cell cycle in B/AA patients. Higher expression of *miR-221*, an oncomiR that facilitates entry into the cell cycle S-phase, is regulated by *c-MYC*, which is expressed more in breast cancer samples from B/AA patients. Furthermore, the cell migration- and invasion-promoting miRNA, *miR-135b*, has increased relative expression in B/AA patients. Knock down of the immortalization marker *TERT* inhibited triple-negative breast cancer cell lines (MDA-MB-231 and MDA-MB-468) cell viability and decreased expression of TERT, MYC and WNT11. For those patients with available survival data, prognosis of stage II patients 50 years of age or younger at diagnosis, was distinctly poorer in B/AA patients. Also associated with this subset of B/AA patients are missense mutations in *HUWE1* and *PTEN* expression loss. Relative to Caucasian non-responders to endocrine therapy, B/AA non-responders show suppressed expression of a signature gene set on which biological processes including *signaling by interleukins*, *circadian clock*, *regulation of lipid metabolism by PPARα*, *FOXO-mediated transcription*, and *regulation of TP53 degradation* are over-represented. Thus, we identify molecular expression patterns suggesting diminished response to oxidative stress, changes in regulation of tumor suppressors/facilitators, and enhanced immortalization in B/AA patients are likely important in defining the more aggressive molecular tumor phenotype reported in B/AA patients.

Abbreviations: ARACNe, Algorithm for the Reconstruction of Accurate Cellular Networks; B/AA, Black/African-American breast cancer patients; B/AA50, Black/African-American stage II breast invasive carcinoma patients diagnosed at age 50 years or younger; BrCA, breast invasive carcinoma; DE, differential expression; DM, differential mutation; EMT, Epithelial-Mesenchymal Transition; GSEA, Gene Set Enrichment Analysis; RMA, robust multi-array average; RPPA, Reverse Phase Protein Array; TCGA, the Cancer Genome Atlas; TNBC, triple-negative breast cancer; TRN, Transcriptional Regulatory Network; VIPER, Virtual Inference of Protein activity by Enriched Regulon Analysis; W50, White stage II breast invasive carcinoma patients diagnosed at age 50 years or younger

* Corresponding author at: Department of Pharmaceutical Sciences, MCPHS University, 19 Foster St, Worcester, MA 01608, USA.
E-mail address: George.Acquaah-Mensah@mcphs.edu (G.K. Acquaah-Mensah).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.gene.2020.100030
Received 26 August 2019; Accepted 8 March 2020
2590-1583/ © 2020 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/BY-NC-ND/4.0/).

1. Introduction

African-Americans with breast cancer have more aggressive tumor biology (Ihemelandu et al., 2007). Hispanic and Black patients tend to have higher tumor sizes and higher mortality rates than Caucasians (Patel et al., 2010). Compared to Caucasian women from Norfolk (United Kingdom), black African patients from Jos (Nigeria) were found to present with breast cancer at an earlier age (64 compared to 43.1 years, respectively) (Gukas et al., 2006). In a study comparing breast cancer in black African women from Kinshasa (Democratic Republic of Congo) and Caucasian women from Leuven (Belgium), black patients tended to have the more aggressive cancers (i.e. estrogen receptor-negative and HER2/neu–positive) (Mvila et al., 2015).

At the molecular level, examples of differences previously noted include more *TP53* mutations and fewer *PIK3CA* mutations in blacks (Keenan et al., 2015). A previous study using microarrays and reverse phase protein array analysis (RPPA) found no differences in expression in breast cancer among the racial groups examined (Chavez-Macgregor et al., 2014). However, next-generation sequencing technologies have distinct advantages over microarrays in detecting transcript abundance; for instance, they are superior in measuring low-abundance transcripts and detecting the presence of variants and isoforms (Zhao et al., 2014). Furthermore, a study of breast tissues among healthy European Americans and African Americans, found differences in both gene expression and epigenetics (DNA methylation) between the two groups (Song et al., 2015).

This work further examines and characterizes the genomic differences associated with the trait differences already noted, with a focus on driver genes regulating those differences. The results from breast invasive carcinoma (BrCA) samples presented here are based on next-generation sequencing technologies (RNA-seq version 2 and miRSeq) and protein arrays.

2. Materials and methods

2.1. RNAseq

RNAseq version 2 data derived from BrCA samples were obtained from the Cancer Genome Atlas (TGCA) (McLendon et al., 2008). The TCGA-Assembler (Zhu et al., 2014) was used to download and process the data as described. Gene expression data (RNA-Seq version 2) were obtained from TCGA, July 2016 (Fig. 1). Processing was via the TCGA-Assembler. For each sample, six files were generated, including normalized gene expression values (computed via the RNA-Seq by Expectation Maximization [RSEM] algorithm) (Li and Dewey, 2011). The normalized count values were subsequently extracted. Ultimately, the resulting gene expression compendium consisted of 20,531 genes and 1213 samples. Using the Gene Ontology (GO), subsets of the table consisting of relevant samples and 5999 human genes associated with “cancer”, “apoptosis”, “senescence”, “cell cycle”, “oxidative stress”, or “signal transduction”, along with transcriptional regulators (genes associated with “transcription factor activity” and “DNA-binding”) was selected (Ashburner et al., 2000). Based on available clinical data, the samples were mostly obtained from 935 White (not Hispanic or Latino), and 179 Black or African-American (not Hispanic or Latino) patients at various stages and with various phenotypes of the disease.

2.2. miRSeq

A transcriptional regulatory network was excerpted from *TransmiR*, a database for transcription factor-microRNA regulations (Wang et al., 2009). From TCGA (July 2016), microRNA expressed at different levels between B/AA and Caucasian patients were identified using the siggenes package (Schwender, 2012). The differential microRNA expression data was then superposed on the excerpted regulatory network.

2.3. Mutation

Downloaded BrCA *Mutation Annotation Format* text files from TCGA were processed using functions from the Maftools package (Mayakonda and Koeffler, 2016). Differentially mutated genes i.e. genes with mutations occurring at different rates between the race- or receptor expression-based phenotypes of interest, were identified.

2.4. Microarrays

A breast tumor gene expression dataset deposited in the Gene Expression Omnibus, GSE47994, was also analyzed. The arrays were on the Affymetrix Human Genome U219 Array platform. The robust multi-array average (RMA) procedure was used for background correction, data normalization, and log-transformation (Irizarry et al., 2003a; Irizarry et al., 2003b). Based on associated annotation data, samples with the triple-negative and non-triple-negative phenotypes were identified.

2.5. Protein Array Data

As shown in Fig. 1, Reverse Phase Protein Array (RPPA) data derived from 937 samples from the TCGA, were processed using the TCGA Assembler (McLendon et al., 2008). The normalized data obtained

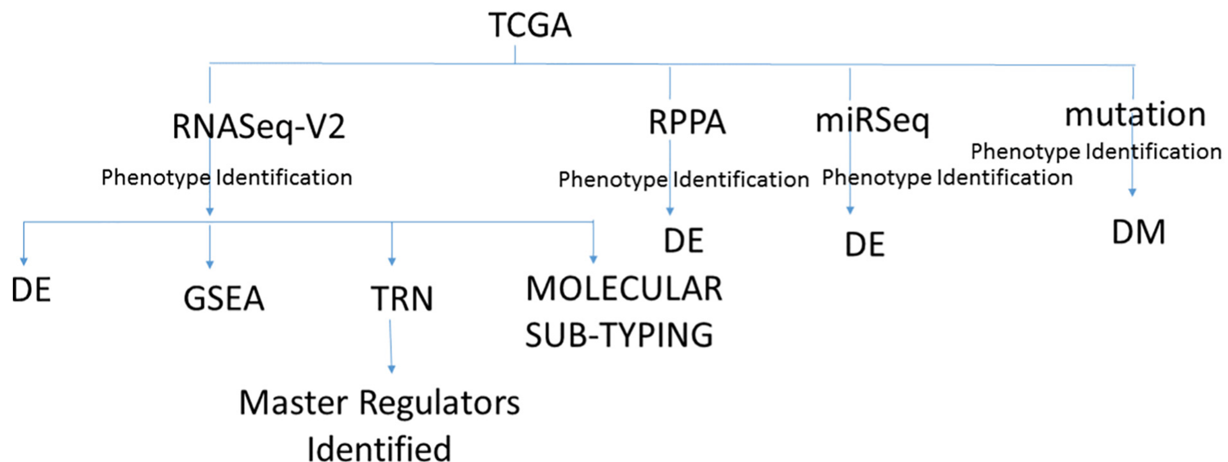

Fig. 1. A schematic laying out the steps taken in the studies in this report. RNAseq, Reverse Phase Protein Array (RPPA), miRNA-seq, and mutation data examined were from the Cancer Genome Atlas, July 2016. Specifically, the samples were from subjects diagnosed with breast invasive carcinoma. [DE → Differential Expression; GSEA → Gene Set Enrichment Analysis; TRN → Transcriptional Regulatory Network; DM → Differential Mutation].

consisted of proteins associated with 227 unique antibodies applied to 937 samples, including 165 white patients in stage 3 and 26 black patients in stage 3. The normalization process involves first subtracting the median expression value of each protein (across all samples) from each expression value for the given protein, and then subtracting the median expression value for each sample (across all proteins) from each expression value for the given sample.

2.6. Differential expression

The package, *siggenes*, was used to identify genes differentially expressed between the various phenotypes examined. These include White patients (not Hispanic or Latino) and Black or African-American (not Hispanic or Latino) (Schwender et al.). The False discovery rate used for the RNAseq data was 0.1%; for the RPPA data, it was 5% (Benjamini and Hochberg, 1995). In the case of the RPPA data, 11 proteins (linked with unique antibodies) with at least one missing expression value from across the 937 samples were not included in the analysis.

2.7. GSEA

Subsequently (as shown in Fig. 1), a Gene Set Enrichment Analysis (GSEA) was performed using RNAseq version 2 data and curated human gene sets associated with the *AHR* and *NFE2L2* in the molecular signatures database, MSigDb (Subramanian et al., 2005). Both *AHR* and *NFE2L2* have suppressed expression in Black/African-American (B/AA) patients, relative to White patients. Among the phenotypes of focus were “White” and “B/AA” patients, as well as triple-negative and non-triple BrCA samples. The signal to noise metric was used for ranking genes, and default settings were used for this analysis. The gene expression compendium consisting of 5999 genes described above was used.

2.8. PAM50

The proportion of breast cancer molecular subtypes (i.e. Basal-like, HER2-enriched, Luminal A, Luminal B, or Normal-like) was determined via the PAM50 method (Fig. 1), using the RNAseq data and the Bioconductor package *genefu* (Gendoo et al., 2016; Parker et al., 2009). The Luminal types are estrogen receptor- (ER-) positive; the HER2-enriched type which has an expanded expression of the human epidermal growth factor receptor 2; the Basal-like types which are largely triple-negative (i.e. lacking estrogen receptors, progesterone receptors and HER2). The proportion of each subtype identified in each phenotype was subsequently computed.

2.9. Virtual Inference of Protein activity by Enriched Regulon analysis (VIPER)

A transcriptional regulatory network (TRN) was reverse-engineered from the gene expression compendium using the Algorithm for the Reconstruction of Accurate Cellular Networks (ARACNe) (Lachmann et al., 2016). Regulons were identified using the TRN and a relevant function within the VIPER package (Alvarez et al., 2016). Using the regulons, and a signature generated based on the gene expression profile, VIPER was further applied to identify master regulators of regulons driving the differences in breast cancer between Caucasian and B/AA patients as well as to infer concomitant aberrant protein activity (Fig. 1).

2.9.1. Cell lines, reagents and plasmids

Triple-negative breast cancer (MDA-MB-231 and MDA-MB-468) cell monolayers (American Type Cell Culture, Manassas, VA) were established in vitro using DMEM-F12 media supplemented with 10% fetal bovine serum (FBS) and 2% penicillin-streptomycin-neomycin cocktail (PSN) (Life Technologies, Grand Island, NY). Cells were maintained under standard cell culture conditions of 95% air and 5% carbon dioxide at 37 °C. Control CRISPR/Cas9 plasmid (sense: CCAGGACCGCG CTTCCCACG), TERT knockout CRISPR/Cas9 plasmids consisting of a pool of three guide RNA (gRNA) sequences (sense: AACATGCGTCGCA AACTCTT; sense: AGGTGCAGAGCGACTACTCC; and sense: GCCAGTC TCACCTTCAACCG) and TERT Activation CRISPR/Cas9 plasmids consisting of a pool of two gRNA sequences (sense: CCAGGACCGCGCTTC CCACG; and sense: CCGCGCGGAGGAGGCGGAGC) were procured from Santa Cruz Biotechnology (Dallas, TX). UltraCruz® transfection reagent and plasmid transfection medium were obtained from Santa Cruz Biotechnology (Dallas, TX). Primary antibodies for TERT, c-MYC, Wnt11, and β-actin and anti-rabbit secondary antibodies were purchased from Cell Signaling Technology (Danvers, MA).

2.9.2. Effect of TERT knockout on human breast adenocarcinoma cell viability

MDA-MB-231 cells were seeded in a 96-well format (1.0×10^4 cells/well) in a 100-μL volume of complete growth media comprising DMEM-F12 supplemented with 10% FBS and 2% PSN for 24 h. Control and TERT knockout CRISPR/Cas9 plasmids (0.2 μg/μL) were diluted in transfection reagent to a final volume of 150 μL (0.0133 μg/μL). Transfection reagent (10 μL) was diluted up to 150 μL with plasmid transfection medium and incubated for 5 min at room temperature (RT). The plasmid solution was added drop-wise to the transfection reagent-medium mixture under vortex and incubated for 30 min at RT. Cells were refreshed with antibiotic-free growth media and 18.75 μL of plasmid mixture added drop-wise followed by incubation at 37 °C for 24 h. Transfection was verified by detection of green fluorescent protein. Cells were refreshed with media and incubated for 48 h. MDA-MB-231 cell viability 72 h post-transfection was determined using AlamarBlue® according to manufacturer's protocol (Thermo Fisher Scientific, Waltham, MA). Results were expressed as percent mean cell viability with standard deviation vs treatment.

2.9.3. Effect of TERT knockout on apoptosis induction

MDA-MB-231 and MDA-MB-468 cells (4×10^5 cells/well) were seeded in a 6-well plate format and incubated overnight followed by transfection with control, TERT activation and TERT knockout CRISPR/Cas9 plasmids as described by manufacturer's protocol (Santa Cruz Biotechnology, Dallas, TX). Cells were stained for apoptosis induction with acridine orange and ethidium bromide as previously described (Andey et al., 2013). Results were presented as qualitative fluorescence micrographs showing apoptotic cells (red to purple staining) compared to live cells (green staining).

2.9.4. Effect of TERT knockout on oncoprotein expression

MDA-MB-231 cells and MDA-MB-468 (4×10^5 cells/well) were seeded in a 6-well plate format and incubated overnight followed by transfection with control, TERT activation and TERT knockout CRISPR/Cas9 plasmids as described by manufacturer's protocol (Santa Cruz Biotechnology, Dallas, TX). Cells were expanded following transfection and lysates were prepared and subjected to SDS-PAGE as previously described (Andey et al., 2014). Proteins were immunoblotted onto a nitrocellulose membrane and probed with primary antibodies for TERT, c-MYC, Wnt11, and β-actin (internal control) and anti-rabbit secondary antibodies (Cell Signaling Technology, Danvers, MA). Band detection was performed using ImageQuant LAS4000 (GE Healthcare Bio-Sciences, Pittsburgh, PA). Measurement of protein bands was done by densitometry, using the ImageJ software (National Institute of Health, Bethesda, MD). The density of target protein bands were normalized with β-actin and the results expressed as percent mean with standard deviation.

Table 1
Incidence of BrCA molecular subtypes based on PAM50 predictions-TCGA 2016.

	Basal-like	HER2-enriched	Luminal B	Luminal A	Normal-like
All subjects (n = 1213)	16.4% (n = 199)	6.3% (n = 77)	49.8% (n = 604)	18.1% (n = 220)	9.3% (n = 113)
White subjects (n = 864)	13.8% (n = 119)	5.1% (n = 44)	48.1% (n = 416)	20.9% (n = 181)	12% (n = 104)
Black/African-American subjects (n = 189)	34.9% (n = 66)	9.5% (n = 18)	42.9% (n = 81)	10% (n = 19)	2.6% (n = 5)

2.10. Statistics

Results were computed with GraphPad Prism (version 7) and differences were analyzed using the student unpaired *t*-test with 95% confidence interval where P value < 0.05 is considered significant.

2.10.1. Results

2.10.1.1. Higher incidence of basal-like molecular subtype among B/AA. The PAM50 algorithm was used to predict the incidence of the various molecular subtypes of BrCA, based on TCGA July 2016 rnaseq version 2 data. The incidence of the aggressive Basal-like molecular subtype in B/AA subjects was 34.9%, compared to 16.4% in the general population, and 13.8% among White patients (Table 1). Further the HER2-enriched subtype, which has poor prognosis, occurred at a rate of 9.5% among B/AA patients, compared to 6.3% in the general population and 5.1% among white patients (Table 1).

A similar pattern was observed, based on incidence of the triple-negative phenotype (TNBC), as determined by immunohistochemistry. TNBC is aggressive, and can be Basal-like or not (Prat et al., 2013). Whereas TNBC occurs at a rate of 35% in the general population, it was found at a rate of 56% in the B/AA population (and 44% in the Asian population), compared to 29% in the Caucasian/White patient population (Table 2).

2.10.1.1.1. Immortalization signature

A set of analyses indicates that genes participating in pathways associated with proliferation have increased expression in B/AA patients (Fig. 2). Notable among these is the immortalization gene, *TERT*, whose expression is increased (Fig. 2A), and tumor suppressors RB1 and *hsa-let-7a*, whose expressions are suppressed in B/AA patients (Fig. 2B, C). Along with increased expression of *E2F1*, and proto-oncogene *c-MYC* in B/AA patients, the differences in gene, protein, and miR expressions suggests a concerted facilitation of entry into the cell cycle S-phase in B/AA patients, compared to Caucasian patients (Fig. 2D). Also, the RPPA data shows processes such as "Mitotic G1-G1/S phases", "S Phase", "CyclinD-associated events in G1", "G1 Phase", "Cyclin A:Cdk2-associated events at S phase entry", and "Cell Cycle, Mitotic" are over-represented among proteins with increased expression in B/AA patients (Supplementary Tables 1 and 2).

Moreover, GSEA results include an enrichment of an immortalization gene signature, which includes TERT (Fig. 3). Using the RNA-Seq data from 925 white patients, and 189 black patients, at a normalized enrichment score of −1.77 (as described by (Subramanian et al., 2005)), and a nominal p-value of 0.009, an immortalization gene set (Dairkee et al., 2007) was enriched in B/AA patients in contrast to white patients. Members of this gene set were up-regulated in breast cancer tumor cultures in which telomerase reverse transcriptase (*TERT*) expression was induced via a retroviral vector (Dairkee et al., 2007). The leading edge genes for this gene set as enriched in B/AA patients are *PPP2R2D*, *ARHGEF9*, *DNMT1*, *PITPNM1*, *NEDD8*, *DAP3*, *HDAC6*, *RAB40C*, *CDKN2C*, *RIN1*, *FAM83D*, *TELO2*, *ARPC3*, *SPNS2*, *NPRL3*, *GSTP1*, *PEBP1*, ***MYC***, *RAB4B*, *RHOT2*, *VRK1*, *ADRBK1*, *ATP6V0D1*, *ARPC1A*, *GLRX2*, *ARHGEF4*, *CHMP4B*, *CKS2*, *CNTROB*, *APOPT1*, ***DRAP1***, *POLR2E*, *CDC25C*, *CCS*, *TAF6*, ***TERT***, *AIP*, *SH2B2*, *PSMD13*, *NDUFA13*, *SLC26A6*, *MAP2K7*, *PIN1*, *GNB2*, *CDC25B*, *PPP4C*, *RFXANK*, *EMD*, *SIVA1*, *SSNA1*, *KIF2C*, *EBP*, *ARPC4*, *IKBKG*, *RNF126*, *STOML2*, ***PQBP1***, *LY6E*, *CNPY3*, *NFKBIL1*, *OGFR*, *PTTG1*, *NUP93*, *SHARPIN*, *BIRC5*, *CDC34*, *BAX*, *PFN1*, and *FAM58A*. The overall gene set has several members associated with redox and proliferative processes. Moreover of note, increased expression of these genes was inversely correlated with patient survival (Dairkee et al., 2007).

Furthermore, an examination of stage III patients for the purpose of identifying master regulators accounting for the differential expression between B/AA and white patients, identified TERT, DRAP1, PQBP1 (all among the immortalization gene set described above) among the main drivers with increased expression and activity (Fig. 4). Indeed, TERT knockout in the triple-negative human breast cancer cell line, MDA-MB-231, induces apoptosis and diminishes cell viability (Fig. 5). It also diminishes onco-protein expression i.e. c-MYC and beta-catenin protein expression (Fig. 5). c-MYC, which happens to be part of the immortalization gene set described above, targets a set of microRNAs including *miR-15a*,*-miR-17*, *miR-18a*, *miR -19a*, *miR-20a*, *miR-93*, *miR-106b*, *miR-141*, ***miR-221***, and *miR-429*, all with increased expression in B/AA patients relative to White patients (Fig. 6).

2.10.1.1.2. TERT interference inhibits cell viability and downregulates oncoprotein expression in TNBC cells. RNA interference of TERT using knockout CRISPR/Cas9 plasmid resulted in a significant decrease in MDA-MB-231 cell viability (77.39 ± 3.87%) and induction of apoptosis (Fig. 5A and B). CRISPR/Cas9-mediated activation and knockout of TERT resulted in a significant upregulation and a marginal downregulation of TERT proteins in MDA-MB-231 (148.43 ± 0.54% and 96.98 ± 0.79% respectively) compared to control (100 ± 3.92%) (Fig. 5C and D). TERT activation was concomitant with upregulation of c-Myc (139.37 ± 1.21%) and Wnt11 (152.49 ± 4.19%), compared to control (100 ± 1.96% and 100 ± 0.67% respectively). Similarly, downregulation of TERT was accompanied by decreased expression of c-Myc (53.11 ± 0184%), and a slight increase in Wnt11 expression (106.60 ± 0.62%) (Fig. 5C, E and F). In MDA-MB-468 cells, a marginal activation of TERT was achieved (101.56 ± 2.06%), resulting in a corresponding expression of Wnt11 (100.71 ± 1.86%), versus a significant upregulation of c-Myc (149.22 ± 0.54%) (Fig. 5G–J). However, significant downregulation of TERT (36.47 ± 1.72%) was associated with corresponding significant decreased expression of c-Myc (17.34 ± 0.66%) and Wnt11 (46.02 ± 1.21%).

2.10.2. Stage II B/AA BrCA patients of age 50 and younger

Out of 39 B/AA and 121 White patients 50 years old or younger at diagnosis with stage II BrCA, there were three B/AA and eleven white

Table 2
Incidence of the triple-negative phenotype, based on immunohistochemistry–TCGA 2016.

	All (n = 358)	B/AA (n = 75)	White (n = 235)	Asian (n = 18)	Other (n = 30)
TNBC	126 (35%)	42 (56%)	69 (29%)	8 (n = 44%)	7 (23%)
non-TNBC	232 (64%)	33 (44%)	166 (n = 71%)	10 (n = 56%)	23 (n = 77%)

TNBC → Triple-Negative Breast Cancer; B/AA → Black or African-American.

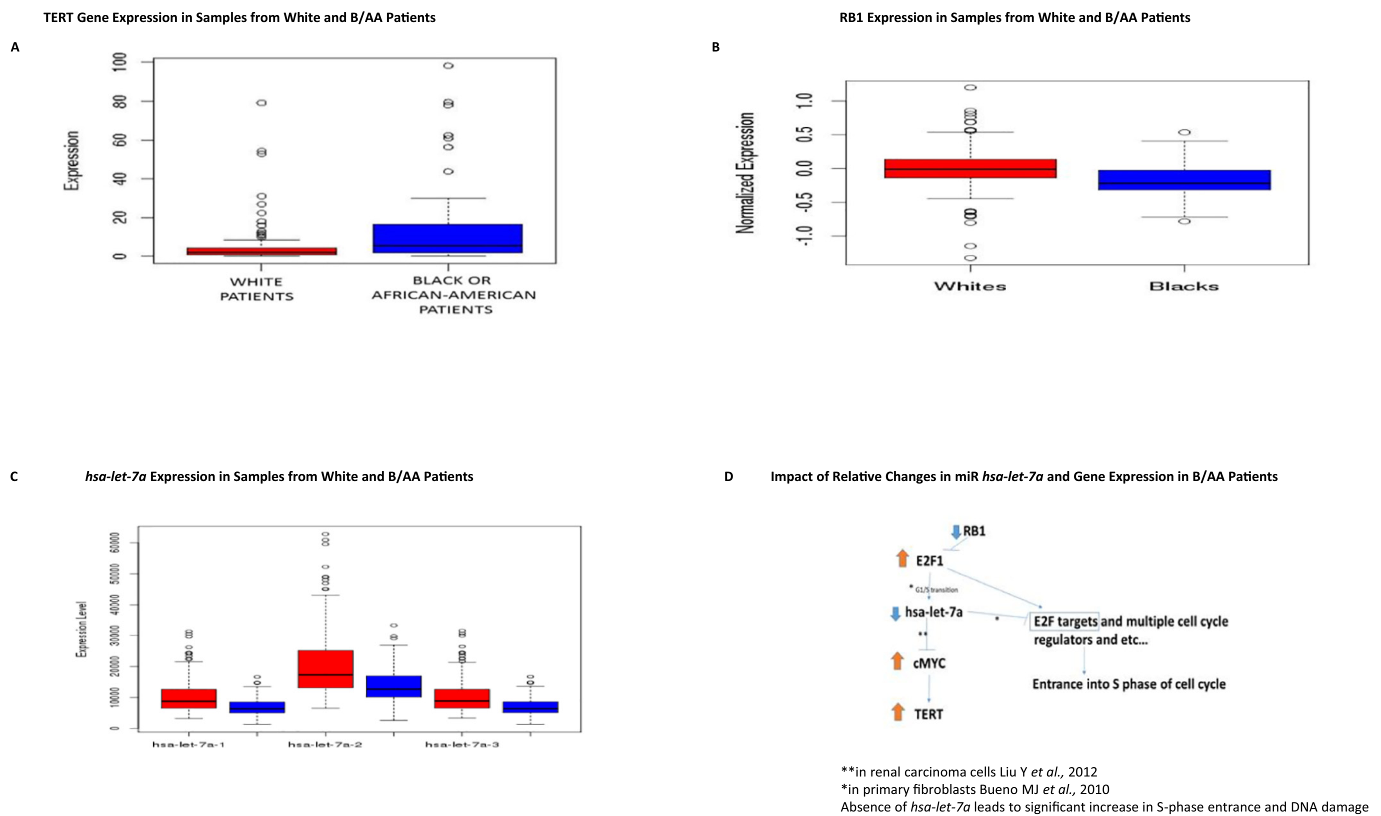

Fig. 2. Differential expression of oncogenic and tumor suppressor gene markers in Breast Invasive Carcinoma Samples from White patients and Black or African-American patients (B/AA), TCGA 2016. (A) Immortalization gene, telomerase reverse transcriptase gene (*TERT*), has increased expression in BRCA samples derived from B/AA (n = 179) relative to White (Caucasian) patients (n = 935). (B) RB1 protein expression is reduced among B/AA patients. The normalized Reverse Phase Protein Array data was derived from 677 white patients and 139 B/AA patients. RB1 was detected using the Rb-M-E antibody. The red boxplot represents normalized expression levels from Caucasian patients; the blue boxplot represents normalized expression levels from B/AA patients. (C) MicroRNA *hsa-let-7a* expression is decreased in Breast Invasive Carcinoma Samples from B/AA relative to Caucasian patients. The miRNA-seq data was derived from 849 white patients and 188 B/AA patients. Three human let-7-a precursors (*hsa-let-7a-1*, *hsa-let-7a-2*, and *hsa-let-7a-3*) all have suppressed expression in B/AA patients. The red boxplot represents normalized expression levels from white patients; the blue boxplot represents normalized expression levels from black patients. (D) Impact of Differences in MicroRNA and Gene Expression between Caucasian and B/AA patients on Entrance into S Phase of Cell Cycle. RB1 binds to E2F proteins to prevent entry into the S phase of the cell cycle. As there is decreased RB1 expression in B/AA patients, there is less constraint on E2F proteins and entry into the S phase is facilitated. Expression of tumor suppressor, microRNA *hsa-let-7a*, a negative regulator of *c-MYC* is suppressed in B/AA patients. Expression of proto-oncogene *c-MYC*, which positively regulates *TERT*, is up in B/AA patients as is *E2F1* whose targets are needed for the cell cycle S phase. Thus, the differences in expression of these appear to confederate to facilitate entrance into the S phase of the cell cycle in B/AA patients. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

patients whose survival data was available in the July 2016 BrCA data from TCGA (Tables 3a and 3b). A majority of the samples from these patients were in stage II.

Prognosis was decidedly poorer for Black/African-American stage II breast invasive carcinoma patients diagnosed at age 50 years or younger (B/AA50), compared to White patients of similar age (W50). Survival among the three B/AA50 patients for whom survival data was available was highly significantly ($p < 0.001$) of shorter duration than those in the eleven W50 patients for whom survival data was available (Fig. 7A). In patients 60 or older at diagnosis, the difference loses its statistical significance.

In the subset of patients 50 years old or younger at diagnosis with stage II, the set of master regulators driving the gene expression differences, as determined using the VIPER algorithm (Alvarez et al., 2016), overlapped with that determined across all ages, but also had its unique elements. In both analyses, PQBP1, HES4, and DRAP1 were among the top 20 master regulators with increased expression in B/AA patients, and CREB1, RB1, SP1, ARID4A and BDP1 among those with suppressed expression. Unique to those patients with stage II BrCA diagnosed at 50 years old or younger, TAF6L, DDIT3, EEF1D, E4F1, and ZNF628 were master regulators with increased expression among B/AA patients, while DDX3X, CCNT1, RBPJ, GABPA, and GTF2A1 had suppressed expression.

Regarding missense mutations, the gene encoding E3 ubiquitin ligase, *HUWE1*, which binds to the BRCA1 tumor suppressor and promotes its degradation via the ubiquitin proteasome pathway (Wang et al., 2014), is significantly more frequently mutated in the B/AA50 compared to the W50 group (Fig. 7B), and BRCA1 gene expression is lower in stage II B/AA patients. On the other hand, mutations of the *PIK3CA* gene which are known to counter the poor-prognosis associated with loss of the tumor suppressor gene *PTEN* (Lazaridis et al., 2019), occur significantly more frequently in the W50 group compared to the B/AA50 group. In those B/AA50 patients with *HUWE1* missense mutations, expression of the *PTEN* gene is severely suppressed compared to those W50 patients with *PIK3CA* missense mutations (Supplementary Table 3) i.e. B/AA50 patients with these *HUWE1* mutations lack the *PIK3CA* mutations which would otherwise have countered this lack of *PTEN* expression. Further, among B/AA50 patients with *HUWE1* missense mutations, master regulators PQBP1, DRAP1, and TAF6L have higher gene expression, while SP1, ARID4A, and CCNT1 have lower gene expression compared to those W50 patients with *PIK3CA* missense mutations (Supplementary Table 3). In addition, genes encoding upstream mitogen-activated kinases, the oncogene ERAS, Ras-association domain family of tumor suppressor proteins *RASSF1* and *RASSF8*, as

Fig. 3. An immortalization signature gene set and differences between breast invasive carcinoma samples from Black or African-American (B/AA) patients and White patients. Using the gene expression data from 925 White patients and 189 Black patients described, a GSEA was performed using MSigDb database gene sets associated with the oxidative stress response regulator, NFE2L2. NFE2L2 has suppressed expression in B/AA patients. (A) An enrichment plot showing a notable cross-section (leading edge) of an immortalization gene set (Dairkee et al., 2007) aggregates at the top (right side of figure has aggregation of bars) of a ranked list of genes in B/AA patients (but not White patients). (B) A heatmap showing the expression of the immortalization gene set. A notable cross-section of this gene set captured in the brown "box" on the heatmap has increased expression in Black patient samples. Blue boxes represent genes with suppressed expression; red boxes represent genes with increased expression. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

well as *RASA1*, and *RASEF* are all dysregulated in B/AA50 patients with *HUWE1* missense mutations compared to those W50 patients with *PIK3CA* missense mutations (Supplementary Table 3). These findings suggest participatory roles for *HUWE1* mutations and the PI3 kinase/Akt pathways in events associated with poorer prognosis in younger black breast cancer patients.

2.10.3. Response to endocrine therapy

A subset of the expression data was selected for patients who were on endocrine therapy, identified as recipients of one or a combination of the following drugs: Tamoxifen, Fulvestrant, Anastrozole, Letrozole, Exemestane, Goserelin, or Leuprolide. Represented on this list are selective estrogen receptor modulators (Tamoxifen), selective estrogen receptor down-regulators (Fulvestrant), gonadotropin releasing hormone agonists (Goserelin, Leuprolide), and aromatase inhibitors (Anastrozole, Letrozole, Exemestane). Patient samples were grouped according to response status with "responders" associated with *complete response*, and "non-responders" associated with *clinical progressive disease* or *stable disease*. Of note, carcinogens such as 2, 3, 7, 8-tetrachlorodibenzodioxin (TCDD) and benzo(a)pyrene are known to act as ligands to the aryl hydrocarbon receptor (AHR) and its dimerization partner, the AHR nuclear receptor (ARNT). These two genes were found in our studies to be significantly suppressed among B/AA patients relative to White/Caucasians. Thus, this distinction was probed further.

Along with the expression data, gene sets associated with AHR in the Molecular Signatures Database (MSigDB) were used for a Gene Set Enrichment Analysis (GSEA). A number of gene sets were enriched in B/AA and Caucasian responders and non-responders. Of particular interest are two gene sets from a study by Creighton et al. (2008) in which two mouse xenograft models of estrogen receptor–positive breast cancer, one *with* HER2 expansion ("Creighton 3") and another *without* HER2 expansion ("Creighton 5"), acquired resistance to endocrine therapies. Among non-responders, leading edge *Creighton 3* and *Creighton 5* gene set members had suppressed expression in B/AA patients relative to Caucasian patients (Fig. 8). As a collection these genes (captured in Fig. 8), which are associated with *Signaling by Interleukins*, *Circadian Clock*, *Regulation of Lipid Metabolism by PPARα*, *SUMOylation of intracellular receptors*, *FOXO-mediated Transcription*, *Regulation of TP53 Degradation*, etc., are expressed less in non-responding B/AA patients than in non-responding Caucasian patients.

Also of note, several of these genes (such as *SASH1*, *FZD4*, *KIF138*, *TCF4*, *IL 13RA1*, *SPRED1iCPEB2*, *EPAS1*, and *THES1*) have suppressed expression in Caucasian non-responders compared to Caucasian responders (Figure not shown). As an example, SASH1 is a tumor suppressor whose expression is linked to breast cancer prognosis (Burgess et al., 2016; Zeller et al., 2003), and it is note-worthy that it is suppressed in non-responding B/AA patients, compared to non-responding White patients (Fig. 8). Further MUC1, which increases resistance to Tamoxifen therapy in estrogen-positive breast cancer (Merikhian et al., 2017), has lower expression in non-responding B/AA patients, compared to non-responding White patients (Fig. 8).

An examination of the somatic mutation data of all patients associated with the two *response to endocrine drug* phenotypes (i.e. responders and non-responders) identifies no differentially mutated genes between them. However, the topmost mutated genes among responders include *PIK3CA*, *CDH1*, *CDH4*, *NEB*, *ZFHX4* and *DHAH17*, while those among non-responders include *TTN*, *PIK3CA*, *MUC16*, and *F8*.

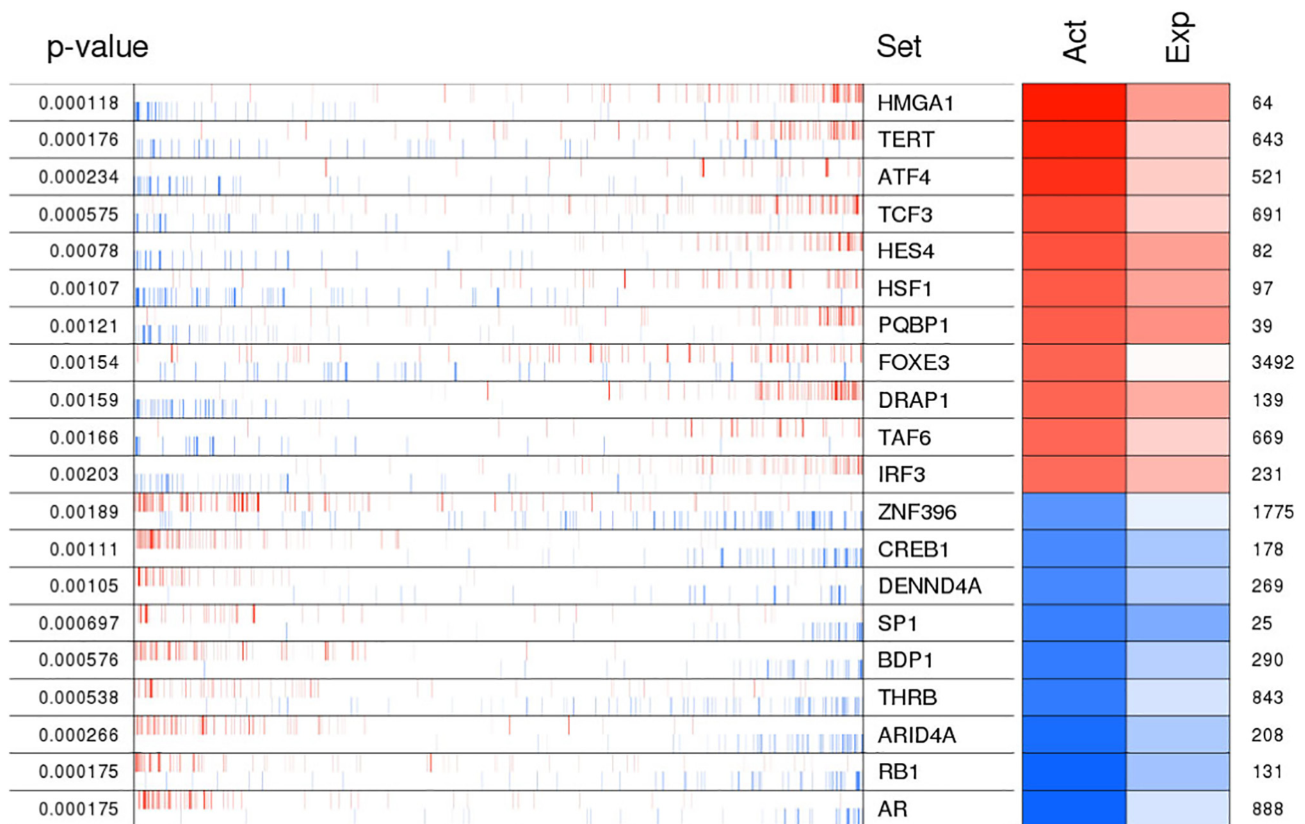

Fig. 4. The top twenty differentially active transcription regulators along with their corresponding regulons enriched (aggregated at the ends of the list of genes ranked by their levels of expression) in Black/African-American (B/AA) patients and Caucasian patients. HMGA1, TERT, ATF4, TCF3, HESF4, HSF1, PQBP1, FOXE3, DRAP1, TAF6 and IRF3 have increased expression and inferred activity in B/AA patients; ZNF396, CREB1, DENND4A, SP1, BDP1, THRB, ARID4A, RB1, and AR have decreased expression and activity in B/AA patients. In the second column, the genes are rank-sorted (left to right) from the most downregulated in B/AA patients (relative to Caucasians) to the most upregulated, and vertical bars represent targets of the transcription regulators (named in the third column) in the transcriptional regulatory network. Blue bars and boxes represent suppressed expression; red bars and boxes represent increased expression. The fourth and fifth columns represent inferred differential protein activity and gene expression respectively. The numbers in the rightmost column are the respective positions of the master regulators on the ranked gene lists. [Act. → activity; Exp. → expression]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.10.4. B/AA and TNBC

The July 2016 TCGA data indicates there is a 56% incidence rate of TNBC in B/AA patients, compared to 44% in Asian patients, and 29% in Caucasian patients (Table 2). Conversely, there is a high incidence rate of non-TNBC (71%) in White patients, compared to 44% in B/AA patients, and 10% in Asians. These observations i.e. that B/AA patients are twice as likely to present with TNBC than Caucasian patients, are reflected in mutation and gene expression patterns within those groups.

Regarding mutation patterns, there are similarities of differential mutations that occur between B/AA and White patients on the one hand, and TNBC and non-TNBC patients on the other. *PIK3CA* mutations (mostly missense mutations) occur more frequently in White stage III patients (32%) than in B/AA stage III patients (20%); they also occur more frequently in non-TNBC patients (37%), than in TNBC patients (8%) (Fig. 9). At least 75% of all patients with *PIK3CA* mutations were White and non-TNBC; 9% were B/AA and non-TNBC. Conversely, a variety of *TP53* mutations occur more frequently in stage III B/AA patients (48%), than in stage III White patients (23%); they occur more frequently in TNBC (73%), than in non-TNBC patients (20%). Further among all patients with *TP53* mutations, B/AA patients or TNBC patients are over-represented. Thus, the occurrence of these frequently mutated genes is a reflection of the high incidence rates of TNBC among B/AA patients, and high incidence rates of non-TNBC among White patients (Table 2).

Regarding over-represented pathways, samples from B/AA patients and patients with TNBC have differentially expressed genes participating in some of the same pathways and biological processes. This is evident in the triple-negative BrCA phenotypes from the Cancer Genome Atlas (TCGA), and the Gene Expression Omnibus GSE47994 dataset. For example in both datasets, and both phenotypes (B/AA and TNBC), genes associated with the *Cell Cycle* and proliferation have increased expression, while genes associated with *PI3K signaling in Cancer* and *Negative regulation of the PI3K Network* have suppressed expression (Tables 4a and 4b).

2.11. Discussion

The results suggest that a major factor in understanding why breast cancer among B/AA patients tends to be aggressive is the intrinsic breast cancer subtypes represented. Breast cancer, in reality, has a set of diverse molecular and cellular presentations and outcomes. The Basal-like intrinsic subtype, in particular, is a more aggressive presentation (Alluri and Newman, 2014). HER2-enriched breast cancers are also associated with poor prognoses (Brewster et al., 2014). The intrinsic breast cancer subtype predictor algorithm used in this study, PAM50, has been demonstrated to be a particularly robust and predictive algorithm (Parker et al., 2009). These results indicate that B/AA patients are more than two times as likely to harbor the Basal-like, and just about twice as likely as Caucasian patients to harbor the HER2-enriched molecular subtype (Table 1). The results are consistent with prior findings showing that the incidence of triple-negative breast cancers, most of which are basal-like, is high in individuals with African

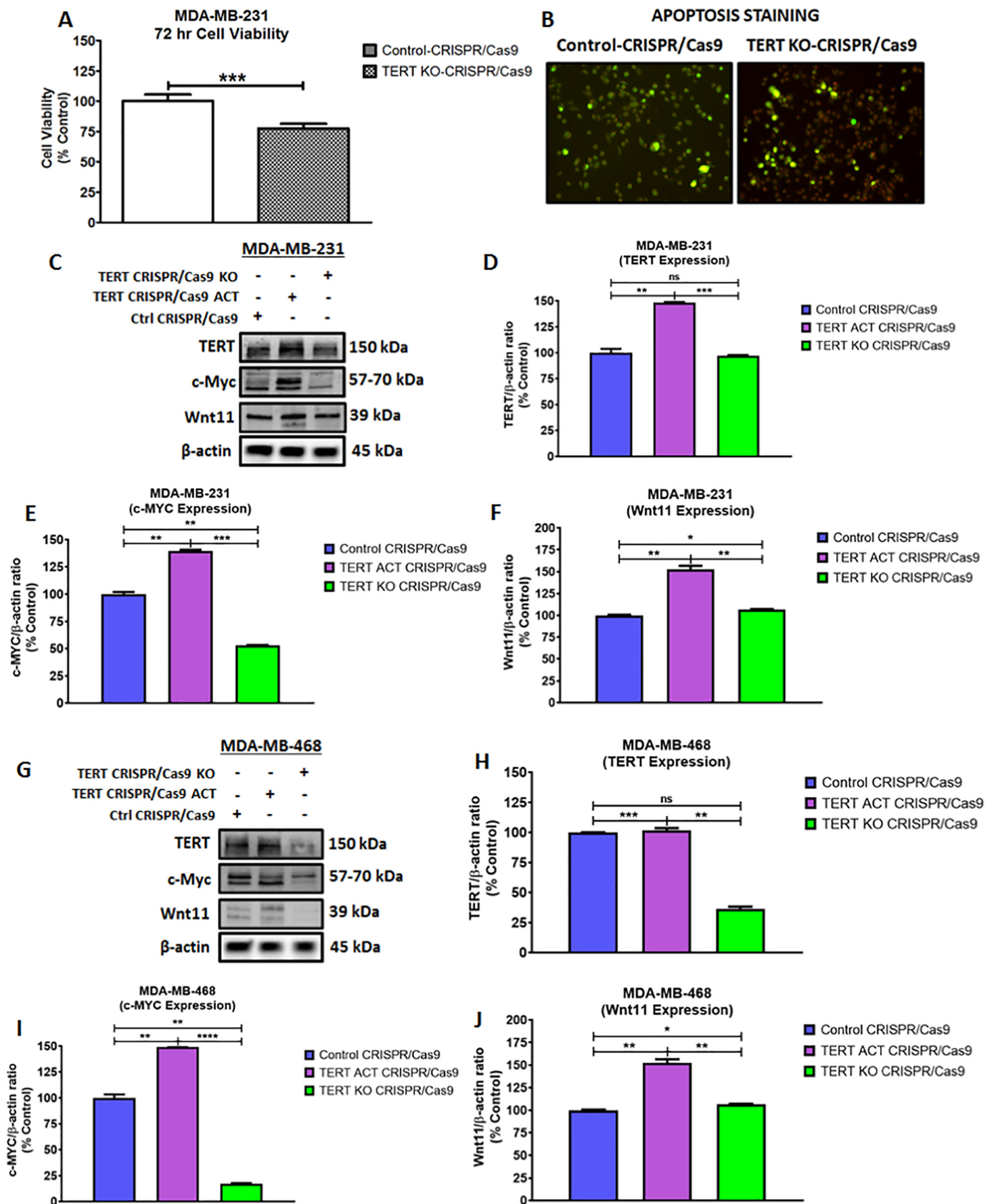

Fig. 5. TERT knockout inhibits cell viability and oncoprotein expression, and induces apoptosis TNBC cells. MDA-MB-231 cells were treated with control and TERT knockout CRISPR/Cas9 plasmids for 72 h and assayed for (A) cell viability and (B) apoptosis induction (red-purple cell nuclei). (C) MDA-MB-231 and (G) MDA-MB-468 cells were treated with control, TERT activation CRISPR/Cas9 plasmids, and TERT knockout CRISPR/Cas9 for 72 h. Cell lysates were immunoblotted for expression of (D, H) TERT, (E, I) c-MYC, and (F, J) Wnt11 proteins. Statistical analysis: Unpaired *t*-test (*$P < 0.05$, **$P < 0.01$; ***$P < 0.001$; and ****$P < 0.0001$). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

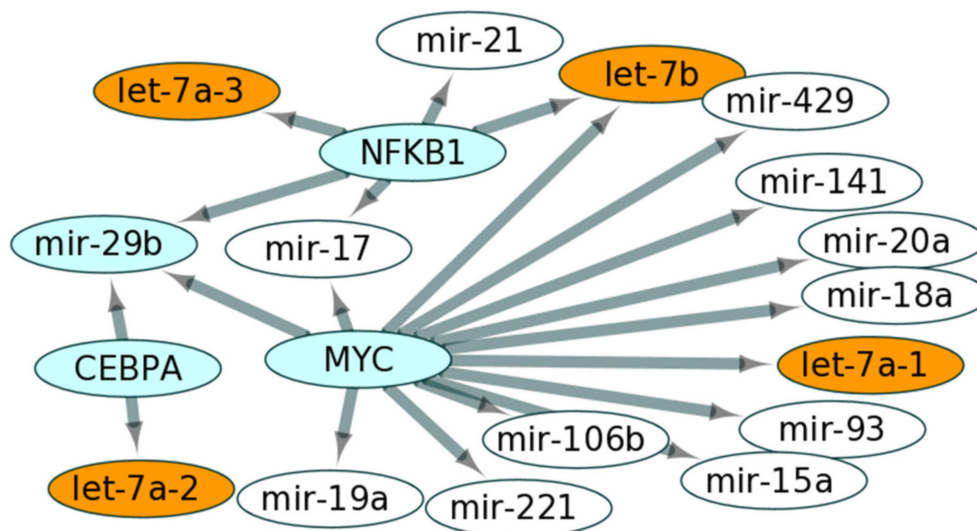

Fig. 6. The impact of increased c-MYC expression on microRNA expression. Differential (between Black/African-American and White breast cancer patients) microRNA expression data was superposed on a transcriptional regulatory network excerpted from *TransmiR*, a database for transcription factor-microRNA regulations (Wang et al., 2009). c-MYC targets microRNAs, such as miR-221, which have regulatory influence on the cancer phenotype. White nodes represent microRNAs with increased expression in Blacks, brown nodes represent microRNAs with suppressed expression in blacks. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3a
Number of patients *up to age 50* at diagnosis in dataset (TCGA, July 2016) at different stages of BrCA.

BrCA stage	Black or African-American (%)	White (%)
I/IA/IB	11 (16.7%)	33 (15.1%)
II/IIA/IIB	39 (59.1%)	121 (55.5%)
III/IIIA/IIIB/IIIC	14 (21.2%)	62 (28.4%)
IV	2 (3.0%)	2 (0.9%)

Table 3b
Number of patients of *all ages* at diagnosis in dataset at different stages of BrCA.

BrCA	Black or African-American (%)	White (%)
I/IA/IB	32 (18.0%)	142 (19.1%)
II/IIA/IIB	106 (59.6%)	415 (55.8%)
III/IIIA/IIIB/IIIC	35 (19.7%)	176 (23.7%)
IV	5 (2.8%)	11 (1.5%)

ancestry (Afghahi et al., 2016; Bauer et al., 2007).

The immortalization gene signature identified by Dairkee et al. (2007) is enriched in B/AA patients (Fig. 3) (Dairkee et al., 2007), i.e. they are among the most highly expressed (highly ranked by expression level) genes in samples derived from B/AA patients. These genes, among which are oxidoreductase and cell-proliferation pathway genes, have increased expression upon the transduction of the catalytic subunit of telomerase (*hTERT*). Of note, their expression is inversely correlated with patient survival (Dairkee et al., 2007). In line with this, this key immortalization marker gene, *TERT*, has increased expression in B/AA patients relative to White patients (Fig. 2A), and is among the leading drivers of the differences in gene expression between the two phenotypes (Fig. 4). Among B/AA patients, the immortalization gene signature is enriched in tumor samples relative to normal samples (data not shown).

The transition from the typical epithelial phenotype to the mesenchymal phenotype is associated with an increased propensity towards metastasis, invasiveness and aggressiveness (Fedele et al., 2017). Sarrio et al. (2008) generated a gene signature associated with the Epithelial-Mesenchymal Transition (EMT) (Sarrio et al., 2008). Specifically they showed that the basal-like phenotype, which has relatively high incidence among B/AA patients in the TCGA data (Table 1), was associated with the EMT.

As described above, the relative expressions, in B/AA patients, of *RB1*, *hsa-let-7a*, *E2F1*, *c-MYC*, *TERT,* and others cooperate to facilitate entry into the S-phase of the cell cycle (Fig. 2D) and, in part, explains the aggressiveness of cancer in Blacks. Also of note, TERT and RB1 are differentially expressed transcriptional factors and master regulators/drivers of regulons (including differentially expressed targets) within the transcriptional regulatory network generated (Fig. 4). These help explain the relative aggressiveness of BRCA in Blacks and are briefly described:

2.11.1. Master regulators

A number of top master regulators (i.e. differentially active (between B/AA and White patients) transcriptional regulators) were identified in this study (Fig. 4). Among them is HSF1, which has increased expression in BRCA samples from B/AA patients relative to White patients. Increased expression of HSF1 worsens breast cancer by promoting cancer cell survival and metastasis. HSF1 has a role in regulating cell stress and heat shock proteins (HSPs). Of note, increased HSF1 expression is associated with resistance to chemotherapeutic drugs, and with the emergence of the cancer stem cell phenotype (Wang et al., 2015). Several HSPs like Hsp70 and Hsp90 have been associated with breast cancer. Hsp70 has been associated with poor tumor differentiation, increased tumor growth, lack of programmed cell death (Jones et al., 2004), lymph node metastasis as well as p53 mutations in breast cancer cells (Ciocca and Calderwood, 2005). Hsp70 has also been linked to increased ERα transcriptional activity and growth in MCF-7 breast cancer cells (Spears and Barnes, 2003). Hsp90 serves as a shield to HER-2/neu and the mutated p53 by preventing the proteasome from degrading them, which leads to increased tumor growth in breast cancer (Ciocca and Calderwood, 2005).

HMGA1 was also identified as being a master regulator in B/AA patient BRCA samples (Fig. 4). This affirms the high relative incidence of TNBC among B/AA patients as HMGA1 is a known master regulator in TNBC cells (Shah et al., 2013) (Table 2). HMGA1 directly controls the urokinase plasminogen activator system which is one of the hallmark pathways in cancer metastasis (Resmini et al., 2017). The urokinase pathway is involved in biological processes such as inflammation, angiogenesis, remodeling and metastasis, hence its role in cancer cell migration. There is increasing evidence that HMGA1 plays a role in

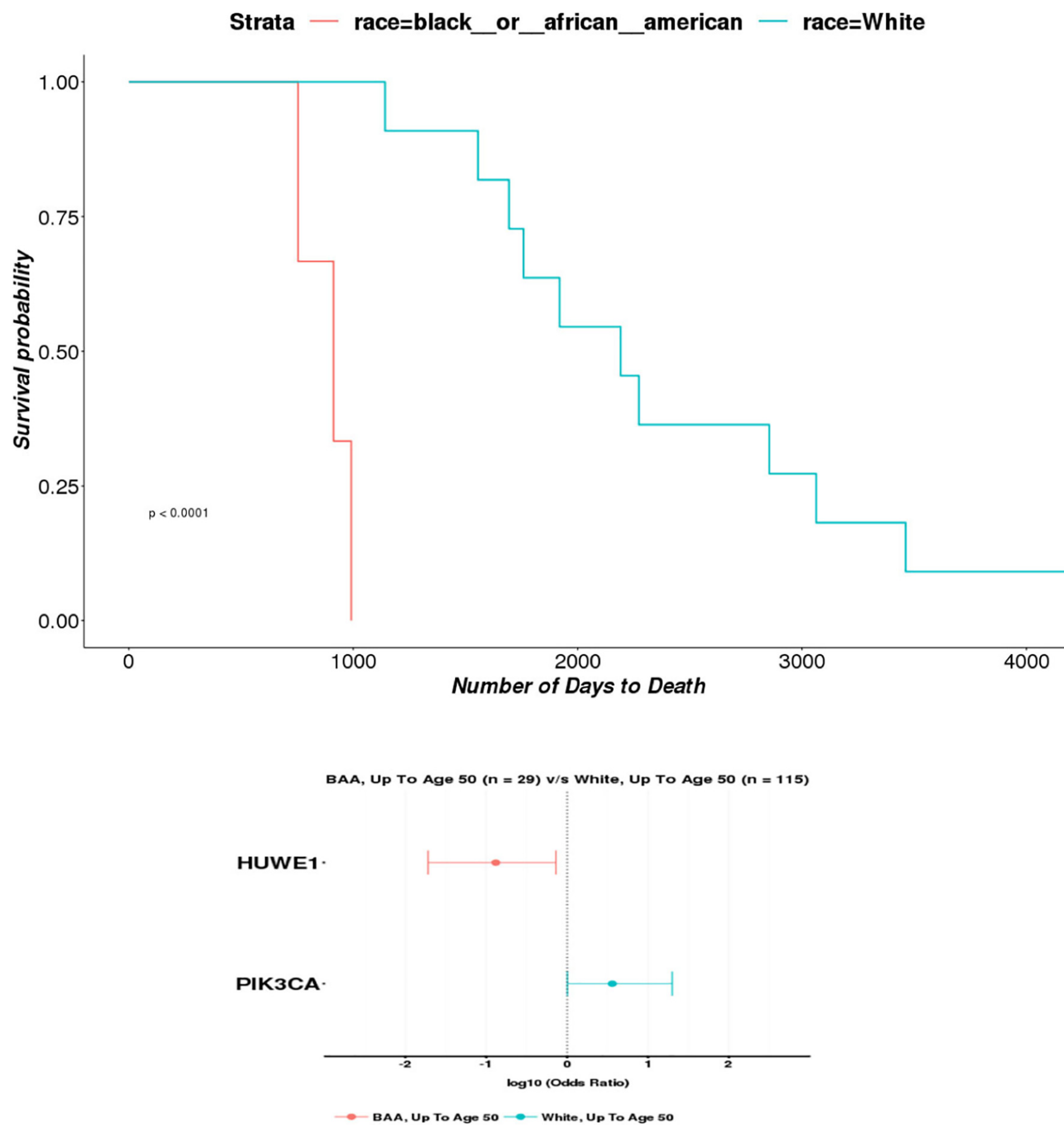

Fig. 7. (A) Kaplan-Meier (survival) plot capturing the dichotomy in survival between three Black/African-American (B/AA50) and eleven White (W50) patients at stage II breast invasive carcinoma and 50 years old or younger at diagnosis (TCGA, July 2016). The time on the horizontal axis represents the number of days to death. Survival among B/AA patients is significantly short-lived ($p < 0.0001$). (B) *HUWE1* and *PIK3CA* are differentially mutated between B/AA50 and W50. *HUWE1* is frequently mutated in B/AA50; *PIK3CA* is frequently mutated in W50.

breast cancer by inducing EMT and metastasis (Shah et al., 2013; Zhong et al., 2017). EMT is considered to be an essential contributor of the invasive behavior of tumors during cancer progression (Zhong et al., 2017).

Also among the master regulators is *TERT*, which has increased expression in BRCA samples from B/AA patients relative to White patients (Fig. 4). One way TERT contributes to cancer cell proliferation is by directly regulating protein synthesis (Khattar et al., 2016). An increase in TERT leads to increased tRNA expression and this is a mechanism by which TERT contributes to cancer development (Khattar et al., 2016). TERT promoter mutations also contribute to tumor development and have been used as a biomarker for detection and monitoring of certain cancers (Pellatt et al., 2013). Multiple TERT single nucleotide polymorphisms (SNPs) have been found to be linked directly or inversely to breast cancer risk because of TERT's involvement in the activation of telomerase (Pellatt et al., 2013). The importance of TERT in promoting the proliferative state in breast cancer was demonstrated by TERT CRISPR/Cas9 knockout in MDA-MB-231 cells with resultant inhibition of cell viability and induction of apoptosis (Fig. 5A, B). TERT has a positive reciprocal signaling effect on Wnt, which is facilitated by interaction of TERT with the β-catenin-interacting protein BRG1 (also referred to as SMARCA4). Notably, induction of TERT expression in the absence of BRG1 failed to stimulate β-catenin activity which was restored by upregulating BRG1 (Park et al., 2009). Moreover, Myc and its transcriptional target, TERT, are involved in a feed-forward oncogenic process facilitated via the regulation of degradation and chromatin stabilization of MYC by TERT (Koh et al., 2015). Of significance is the observation that a marginal decrease in TERT expression resulted in analogous and significant reduction in the expression of oncoproteins c-MYC and WNT11 in MDA-MB-231 and MDA-MB-468 cells (Fig. 5C–J). This observation was consistent with studies that highlighted a co-operative transcriptional relationship between TERT and MYC in promoting and sustaining the proliferative state in tumors (Koh et al., 2015; Tang et al., 2016).

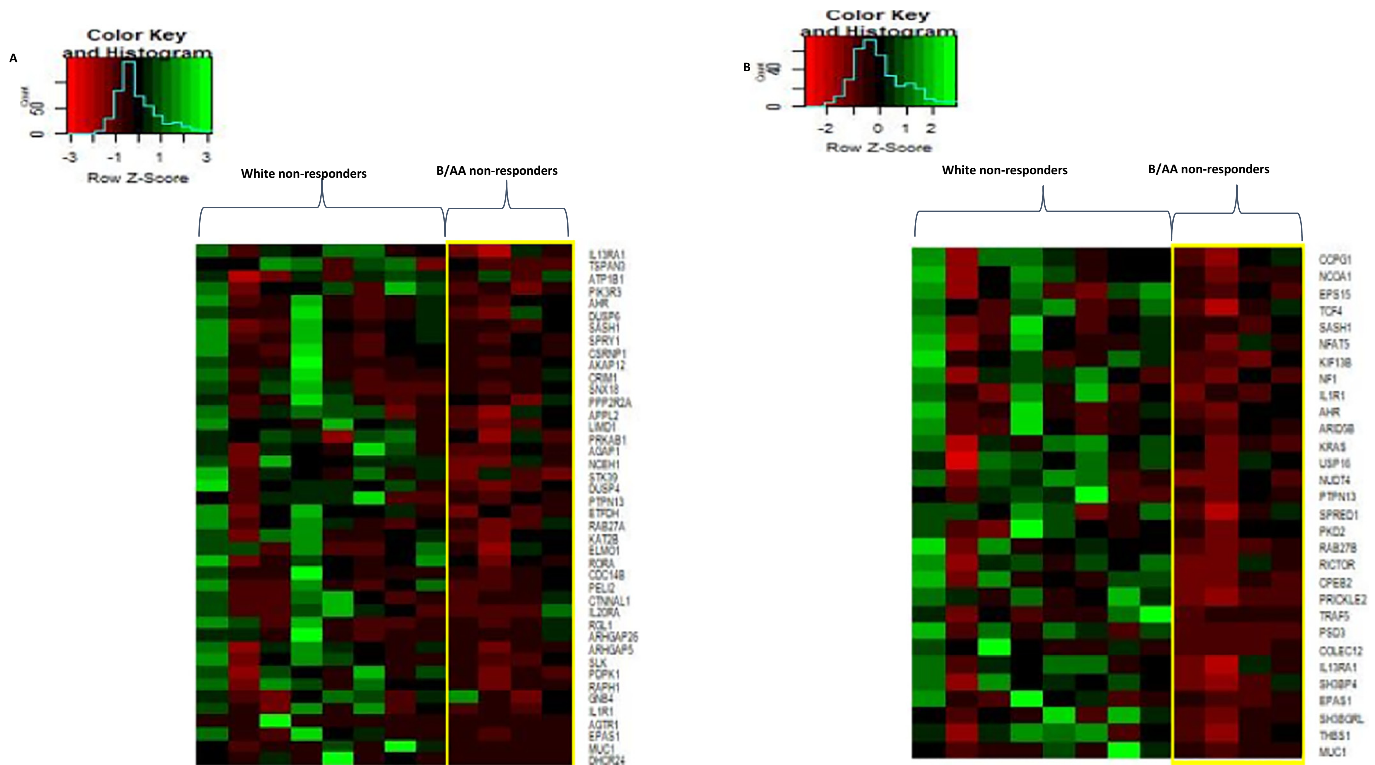

Fig. 8. Gene sets associated with breast cancer resistance to endocrine therapy are expressed differently between Black/African-American (B/AA) and White patients. Pooled gene expression profiles of responders and non-responders to therapy with endocrine drugs (Tamoxifen, Fulvestrant, Anastrozole, Letrozole, Exemestane, Goserelin, or Leuprolide). Responders to these therapies were associated with *complete response*; non-responders were associated with *stable/progressive disease*. A) The two gene sets from a study by Creighton et al. (2008) in which two mouse xenograft models of estrogen receptor–positive breast cancer, one with HER2 expansion ("Creighton 3") and another without HER2 expansion ("Creighton 5"). Leading edge "Creighton 3" (Creighton et al., 2008) genes are suppressed in B/AA non-responders relative to White non-responders. B) Leading edge "Creighton 5" (Creighton et al., 2008) genes are suppressed in B/AA non-responders relative to White non-responders.

2.11.2. microRNAs

TransmiR is a database of experimentally-validated transcription factor-microRNA regulations (Wang et al., 2009). As microRNAs regulate the expression of their target genes post-transcriptionally, changes in activity of transcription factors regulating microRNAs can have consequences on the expressions or activities of the genes regulated by those microRNAs. *miR-15a*, *miR-17*, *miR-18a*, *miR-19a*, *miR-20a*, *miR-93*, *miR-106b*, *miR-141*, ***miR-221***, and *miR-429* are among those with increased expression in Blacks relative to Whites (Fig. 5B). As an example, increased expression of *miR-221,* an oncomiR that facilitates entry into the S-phase and promotes the EMT (Li et al., 2014), is regulated by MYC (which, as noted above, has increased expression in Blacks) (Fig. 5B). Also, the cell migration and invasion promoting miRNA, *miR-135b* (Zhang et al., 2015), has increased expression in B/AA compared to Caucasian cancer patients and is regulated by the transcription factors *FOXM1* and *STAT3* in our transcriptional regulatory network (figure not shown); *FOXM1* has increased expression among B/AA patients, while *STAT3* has decreased expression.

In conclusion, the more aggressive breast cancer biology reported in B/AA patients likely involves enhanced entry into the S-phase, and immortalization. A number of likely molecular participants have been identified here and include HSF1, TERT, HMGA1, PQBP1, FOXE3, DRAP1, TAF6, IRF3, ZNF396, CREB1, DENND4A, SP1, BDP1, THRB, ARID4A, RB1, and AR. Furthermore, the *basal-like/TNBC* and *HER2-enriched* molecular subtypes are more common in B/AA patient samples than in white patient samples. Moreover, a number of genes associated with diminished response to endocrine therapy (such as *SASH1*, *FZD4*, *KIF138*, *TCF4*, *IL13RA1*, *SPRED1*, *CPEB2*, *EPAS1*, and *THES1*) have lower expression levels even among non-responding B/AA patients compared to non-responding White patients. We observe that prognosis is poorer in stage II B/AA patients diagnosed at age 50 years or younger. However, enthusiasm is assuaged by the fact that this observation is based on only three stage II B/AA and eleven White patients 50 years or younger at diagnosis, for whom survival data is available. Missense mutations in the E3 ubiquitin ligase gene *HUWE1* occur more frequently in stage II B/AA patients 50 years old or younger at diagnosis, and are associated with *PTEN* expression loss (which is, itself, associated with poor prognosis). Indeed, the gene expression profile in this group is distinctly different from that in White stage II patients 50 years or younger at diagnosis in whom missense *PIK3CA* mutations occurred more frequently. Several pathways identified here will be examined in future studies.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gene.2020.100030.

Financial disclosure

Study was carried out with support from MCPHS University.

CRediT authorship contribution statement

Terrick Andey: Conceptualization, Formal analysis, Investigation, Methodology, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing. **Michael M. Attah:** Formal analysis, Investigation, Writing - original draft, Writing - review & editing. **Nana Adwoa Akwaaba-Reynolds:** Investigation, Writing - original draft, Writing - review & editing. **Sana Cheema:** Formal analysis, Investigation, Visualization, Writing - original draft, Writing - review & editing. **Sara Parvin-Nejad:** Formal analysis, Investigation, Visualization, Writing - original draft, Writing - review & editing.

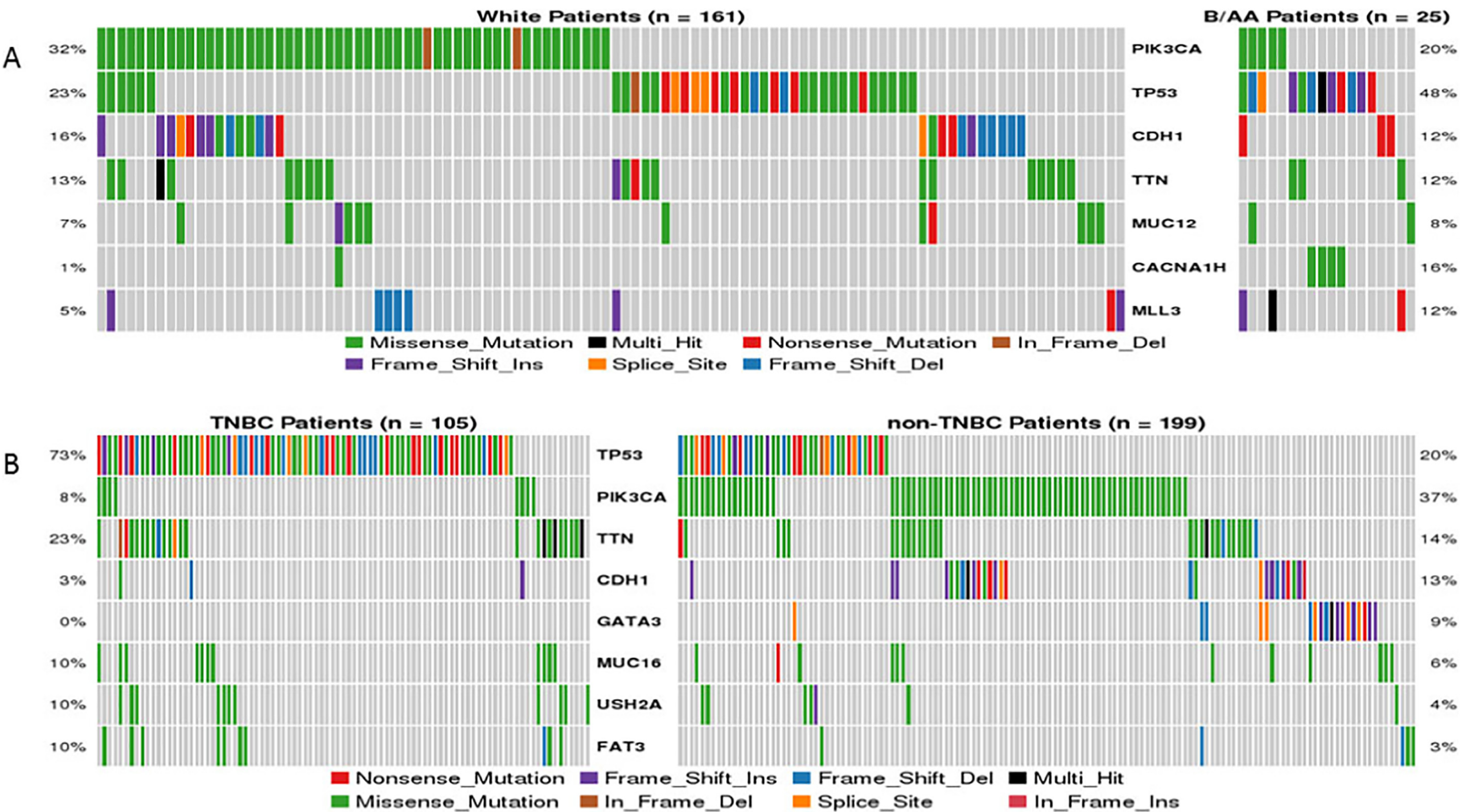

Fig. 9. Patterns of frequently mutated genes reflect dominance of triple-negative breast cancer (TNBC) among Black/African-American (B/AA) patients. Most mutated genes among A) stage III B/AA and White Caucasian BrCA patients, and B) TNBC and non-TNBC patients (TCGA, July 2016). *PIK3CA* mutations (mostly missense mutations), as an example, occur more frequently in White stage III patients (32%) than in B/AA stage III patients (20%); they also occur more frequently in non-TNBC patients (37%), than in TNBC patients (8%). TNBC patients were identified by way of immunohistochemistry (TCGA, July 2016).

Table 4a
Some pathways over-represented among genes differentially expressed between White and Black/African-American (B/AA) patients.

	Stage I (all)	Stage II (up to age 50)	Stage III (all)
Up-regulated in B/AA	Cell cycle, mitotic; G1/S transition; mitotic G2-G2/M phases; IRE1alpha activates chaperones; interleukin-1 family signaling; TLR3 cascade; TRIF-mediated TLR4 signaling; MyD88-independent TLR4 cascade; XBP1(S) activates chaperone genes; Activation of E2F1 target genes at G1/S; TNFR2 non-canonical NF-κB pathway; Unfolded protein response; signaling by Rho GTPases; TAK1 activates NFκB by phosphorylation and activation of IKKs complex; FGFR2 alternative splicing	Cell cycle, mitotic; S phase; DNA Replication; DNA replication pre-initiation; separation of sister chromatids; switching of origins to a post-replicative state; M phase; M/G1 transition; G1/S transition; degradation of beta-catenin by the destruction complex; PCP/CE pathway; APC/C:Cdh1 mediated degradation of Cdc20 and APC/C:Cdh1 targeted proteins in late mitosis/early G1; assembly of pre-replication complex; autoregulation of Cdh1 by Cdh1:APC/C; APC/C:Cdc20 mediated degradation of Securin	Cell cycle, mitotic; synthesis of DNA; G2/M checkpoints; G2/M transition; G1/S transition; p53-independent G1/S DNA damage checkpoint; p53-Independent DNA Damage Response; G2/M checkpoints; ubiquitin mediated degradation of phosphorylated Cdc25A; degradation DVL; regulation of APC/C activators between G1/S and early anaphase; APC/C:Cdh1 mediated degradation of Cdc20 and other APC/C:Cdh1 targeted proteins in late mitosis/early G1; autodegradation of Cdh1 by Cdh1:APC/C; regulation of APC/C activators between G1/S and early anaphase; APC/C:Cdc20 mediated degradation of Securin
Down-regulated in B/AA	Signaling by NOTCH1 in cancer; signaling by WNT in cancer; circadian clock; TLR9 cascade; downregulation of ERBB2 signaling; PIP3 activates AKT signaling; inhibition of replication initiation of damaged DNA by RB1/E2F1; regulation of TP53 degradation; antigen processing: ubiquitination and proteasome degradation; Rho GTPase cycle;	SUMOylation; SUMO E3 ligases SUMOylate target proteins; MyD88-independent TLR4 cascade; MyD88 dependent cascade initiated on endosome; MyD88 cascade initiated on plasma membrane; signaling by TGF-beta family members; signaling by IGFR1; nuclear signaling by ERBB4; signaling by VEGF; TLR 5 cascade; TLR 10 cascade; Rho GTPase cycle; TLR 7/8 cascade; TLR 9 cascade; TRIF-mediated TLR4 signaling; TRAF6 mediated induction of NFκB and MAP kinases upon TLR7/8 or 9 activation	Negative regulation of the PI3K/AKT network; constitutive signaling by aberrant PI3K in cancer; signaling by ERBB4; SUMOylation of intracellular receptors; SUMO E3 ligases SUMOylate target proteins; circadian clock; signaling by TGF-beta family members; Rho GTPase cycle; TLR1:TLR2 cascade;TLR6:TLR2 cascade; TLR9 cascade; PI5P, PP2A and IER3 regulate PI3K/AKT signaling; MyD88: Mal cascade initiated on plasma membrane; TRAF6 mediated induction of NFκB and MAP kinases upon TLR7/8 or 9 activation

Table 4b
Some pathways over-represented among genes differentially expressed between non-triple-negative breast cancer (non-TNBC) and TNBC patients using gene expression data from the Cancer Genome Atlas (TCGA) and the Gene Expression Omnibus (GSE47994).

	TCGA	GSE47994
Up-regulated in TNBC	Cell cycle, mitotic; G1/S transition; separation of sister chromatids; mitotic anaphase; mitotic metaphase and anaphase; mitotic prometaphase; M phase; mitotic G1-G1/S phases; interferon gamma signaling; adaptive immune system; TCR signaling; signaling by Rho GTPases; ER-phagosome pathway; antigen processing-cross presentation; M phase; antigen presentation: folding, assembly and peptide loading of class I MHC; endosomal/vacuolar pathway	Cell Cycle, Mitotic; Cell Cycle; Mitotic G1-G1/S phases; M Phase; G1/S Transition; Cell Cycle Checkpoints; Processing of Capped Intron-Containing Pre-mRNA; Cellular responses to stress; Mitotic G2-G2/M phases; Metabolism of RNA; S Phase; Cellular responses to external stimuli; DNA Replication; G2/M Transition; G2/M Checkpoints; mRNA Splicing - Major Pathway; mitotic prometaphase; Transcriptional regulation by small RNAs; Mitotic Prophase; Mitotic Anaphase; Apoptosis; Mitotic Metaphase and Anaphase; Programmed Cell Death; mRNA Splicing; DNA Replication Pre-Initiation
Down-regulated in TNBC	Negative regulation of the PI3K/AKT network; signaling by ERBB2; nuclear signaling by ERBB4; signaling by receptor tyrosine kinases; PI5P, PPA2 and IER3 Regulate PI3K/AKT signaling; RAB geranylgeranylation; PIP3 activates AKT signaling; PI3K/AKT signaling in cancer; mTOR signaling; signaling by insulin receptor; ESR-mediated signaling; estrogen-dependent gene expression; regulation of TP53 activity; signaling by PTK6; RAF/MAP kinase cascade; MAPK1/MAPK3 signaling; intrinsic pathway for apoptosis; BH3-only proteins associate with and inactivate anti-apoptotic BCL-2 members	Negative regulation of the PI3K Network; Estrogen-dependent gene expression; ESR-mediated signaling; PI3K/AKT signaling in cancer; Post-translational protein phosphorylation; TFAP2 (AP-2) family regulates transcription of growth factors and their receptors; SUMOylation of intracellular receptors; PI5P, PP2A and IER3 Regulate PI3K; Regulation of Insulin-like Growth Factor (IGF) transport and uptake by Insulin-like Growth Factor Binding Proteins (IGFBPs); Sema4D induced cell migration and growth-cone collapse; Signaling by Nuclear Receptors; The activation of arylsulfatases; ERBB2 Activates PTK6 Signaling; Molecules associated with elastic fibres; Integrin cell surface interactions; Downregulation of ERBB4 signaling; RAF-independent MAPK1 activation; ERBB2 Regulates Cell Motility; Sema4D in semaphorin signaling; GRB2 events in ERBB2 signaling; Peroxisomal protein import; Activation of gene expression by SREBF (SREBP): Elastic fibre formation

George K. Acquaah-Mensah: Conceptualization, Formal analysis, Investigation, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

As indicated under "Methods" above, the findings of this study are based in whole or in part upon data generated by the TCGA Research Network: http://cancergenome.nih.gov/. This work has been supported by the resources of Massachusetts College of Pharmacy & Health Sciences (MCPHS University).

References

Afghahi, A., Telli, M.L., Kurian, A.W., 2016. Genetics of triple-negative breast cancer: implications for patient care. Curr. Probl. Cancer 40 (2), 130–140.

Alluri, P., Newman, L.A., 2014. Basal-like and triple-negative breast cancers: searching for positives among many negatives. Surg. Oncol. Clin. N. Am. 23 (3), 567–577. https://doi.org/10.1016/j.soc.2014.03.003.

Alvarez, M.J., Shen, Y., Giorgi, F.M., Lachmann, A., Ding, B.B., Ye, B.H., Califano, A., 2016. Functional characterization of somatic mutations in cancer using network-based inference of protein activity. Nat. Genet. 48, 838–847.

Andey, T., Patel, A., Jackson, T., Safe, S., Singh, M., 2013. 1, 1-Bis (3′-indolyl)-1-(p-substitutedphenyl) methane compounds inhibit lung cancer cell and tumor growth in a metastasis model. Eur. J. Pharm. Sci. 50 (2), 227–241.

Andey, T., Marepally, S., Patel, A., Jackson, T., Sarkar, S., O'Connell, M., ... Singh, M., 2014. Cationic lipid guided short-hairpin RNA interference of annexin A2 attenuates tumor growth and metastasis in a mouse lung cancer stem cell model. Journal of Controlled Release 184, 67–78.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., ... Sherlock, G., 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nature genetics 25 (1), 25–29. https://doi.org/10.1038/75556.

Bauer, K.R., Brown, M., Cress, R.D., Parise, C.A., Caggiano, V., 2007. Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype. Cancer 109 (9), 1721–1728.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B Methodol. 289–300.

Brewster, A.M., Chavez-MacGregor, M., Brown, P., 2014. Epidemiology, biology, and treatment of triple-negative breast cancer in women of African ancestry. Lancet Oncol. 15 (13), e625–e634.

Burgess, J.T., Bolderson, E., Saunus, J.M., Zhang, S.-D., Reid, L.E., McNicol, A.M., ... Richard, D.J., 2016. SASH1 mediates sensitivity of breast cancer cells to chloropyramine and is associated with prognosis in breast cancer. Oncotarget 7 (45), 72807.

Chavez-Macgregor, M., Liu, S., De Melo-Gagliato, D., Chen, H., Do, K.A., Pusztai, L., ... Gonzalez-Angulo, A.M., 2014. Differences in gene and protein expression and the effects of race/ethnicity on breast cancer subtypes. Cancer Epidemiol. Biomarkers Prev. 23 (2), 316–323. https://doi.org/10.1158/1055-9965.EPI-13-0929.

Ciocca, D.R., Calderwood, S.K., 2005. Heat shock proteins in cancer: diagnostic, prognostic, predictive, and treatment implications. Cell Stress Chaperones 10 (2), 86–103.

Creighton, C.J., Massarweh, S., Huang, S., Tsimelzon, A., Hilsenbeck, S.G., Osborne, C.K., ... Schiff, R., 2008. Development of resistance to targeted therapies transforms the clinically associated molecular profile subtype of breast tumor xenografts. Cancer research 68 (18), 7493–7501.

Dairkee, S.H., Nicolau, M., Sayeed, A., Champion, S., Ji, Y., Moore, D.H., ... Jeffrey, S.S., 2007. Oxidative stress pathways highlighted in tumor cell immortalization: association with breast cancer outcome. Oncogene 26 (43), 6269–6279.

Fedele, M., Cerchia, L., Chiappetta, G., 2017. The epithelial-to-mesenchymal transition in breast cancer: focus on basal-like carcinomas. Cancers 9 (10), 134.

Gendoo, D.M., Ratanasirigulchai, N., Schroder, M.S., Pare, L., Parker, J.S., Prat, A., Haibe-Kains, B., 2016. Genefu: an R/bioconductor package for computation of gene expression-based signatures in breast cancer. Bioinformatics 32 (7), 1097–1099. https://doi.org/10.1093/bioinformatics/btv693.

Gukas, I.D., Jennings, B.A., Mandong, B.M., Manasseh, A.N., Harvey, I., Leinster, S.J., 2006. A comparison of the pattern of occurrence of breast cancer in Nigerian and British women. Breast 15 (1), 90–95.

Ihemelandu, C.U., Leffall, L.D., Dewitty, R.L., Naab, T.J., Mezghebe, H.M., Makambi, K.H., ... Frederick, W.A., 2007. Molecular breast cancer subtypes in premenopausal and postmenopausal African-American women: age-specific prevalence and survival. Journal of Surgical Research 143 (1), 109–118.

Irizarry, R., Gautier, L., Cope, L., 2003a. An R package for analyses of Affymetrix oligonucleotide arrays. In: The Analysis of Gene Expression Data, pp. 102–119.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., Speed, T.P., 2003b. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics (Oxford, England) 4 (2), 249–264. https://doi.org/10.1093/biostatistics/4.2.249.

Jones, E.L., Zhao, M.J., Stevenson, M.A., Calderwood, S.K., 2004. The 70kilodalton heat shock protein is an inhibitor of apoptosis in prostate cancer. Int. J. Hyperth. 20 (8), 835–849.

Keenan, T., Moy, B., Mroz, E.A., Ross, K., Niemierko, A., Rocco, J.W., ... Bardia, A., 2015. Comparison of the Genomic Landscape Between Primary Breast Cancer in African American Versus White Women and the Association of Racial Differences With Tumor Recurrence. Journal of clinical oncology: official journal of the American Society of Clinical Oncology 33 (31), 3621–3627. https://doi.org/10.1200/JCO.2015.62.2126.

Khattar, E., Kumar, P., Liu, C.Y., Akıncılar, S.C., Raju, A., Lakshmanan, M., ... Tan, E.Y., 2016. Telomerase reverse transcriptase promotes cancer cell proliferation by augmenting tRNA expression. The Journal of clinical investigation 126 (10), 4045–4060.

Koh, C.M., Khattar, E., Leow, S.C., Liu, C.Y., Muller, J., Ang, W.X., ... Guccione, E., 2015. Telomerase regulates MYC-driven oncogenesis independent of its reverse transcriptase activity. The Journal of clinical investigation 125 (5), 2109–2122.

Lachmann, A., Giorgi, F.M., Lopez, G., Califano, A., 2016. ARACNe-AP: gene network reverse engineering through adaptive partitioning inference of mutual information. Bioinformatics 32 (14), 2233–2235. https://doi.org/10.1093/bioinformatics/btw216.

Lazaridis, G., Kotoula, V., Vrettou, E., Kostopoulos, I., Manousou, K., Papadopoulou, K., ... Pentheroudakis, G., 2019. Opposite Prognostic Impact of Single PTEN-loss and PIK3CA Mutations in Early High-risk Breast Cancer. Cancer Genom. Proteom. 16 (3), 195–206.

Li, B., Dewey, C.N., 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12 (1), 1.

Li, Y., Liang, C., Ma, H., Zhao, Q., Lu, Y., Xiang, Z., ... Cho, W.C., 2014. miR-221/222 promotes S-phase entry and cellular migration in control of basal-like breast cancer. Molecules 19 (6), 7122–7137.

Mayakonda, A., Koeffler, H.P., 2016. Maftools: efficient Analysis, Visualization and Summarization of MAF Files from Large-scale Cohort Based Cancer Studies. BioRxiv, pp. 52662.

McLendon, R., Friedman, A., Bigner, D., Van Meir, E.G., Brat, D.J., Mastrogianakis, G.M., ... Aldape, K., 2008. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature 455 (7216), 1061–1068.

Merikhian, P., Ghadirian, R., Farahmand, L., Mansouri, S., Majidzadeh-A, K., 2017. MUC1 induces tamoxifen resistance in estrogen receptor-positive breast cancer. Expert. Rev. Anticancer. Ther. 17 (7), 607–613.

Mvila, G.L., Batalansi, D., Praet, M., Marchal, G., Laenen, A., Christiaens, M.-R., ... Van Ongeval, C., 2015. Prognostic features of breast cancer differ between women in the Democratic Republic of Congo and Belgium. The breast 24 (5), 642–648.

Park, J.-I., Venteicher, A.S., Hong, J.Y., Choi, J., Jun, S., Shkreli, M., ... Ji, H., 2009. Telomerase modulates Wnt signalling by association with target gene chromatin. Nature 460 (7251), 66.

Parker, J.S., Mullins, M., Cheang, M.C., Leung, S., Voduc, D., Vickery, T., ... Bernard, P.S., 2009. Supervised risk predictor of breast cancer based on intrinsic subtypes. Journal of clinical oncology: official journal of the American Society of Clinical Oncology 27 (8), 1160–1167. https://doi.org/10.1200/JCO.2008.18.1370.

Patel, T.A., Colon-Otero, G., Bueno Hume, C., Copland 3rd, J.A., Perez, E.A., 2010. Breast cancer in Latinas: gene expression, differential response to treatments, and differential toxicities in Latinas compared with other population groups. Oncologist 15 (5), 466–475. https://doi.org/10.1634/theoncologist.2010-0004.

Pellatt, A.J., Wolff, R.K., Torres-Mejia, G., John, E.M., Herrick, J.S., Lundgreen, A., ... Fejerman, L., 2013. Telomere length, telomere-related genes, and breast cancer risk: the breast cancer health disparities study. Genes, Chromosomes and Cancer 52 (7), 595–609.

Prat, A., Adamo, B., Cheang, M.C., Anders, C.K., Carey, L.A., Perou, C.M., 2013. Molecular characterization of basal-like and non-basal-like triple-negative breast cancer. Oncologist 18 (2), 123–133.

Resmini, G., Rizzo, S., Franchin, C., Zanin, R., Penzo, C., Pegoraro, S., ... Sgarra, R., 2017. HMGA1 regulates the Plasminogen activation system in the secretome of breast cancer cells. Scientific reports 7 (1), 11768.

Sarrio, D., Rodriguez-Pinilla, S.M., Hardisson, D., Cano, A., Moreno-Bueno, G., Palacios, J., 2008. Epithelial-mesenchymal transition in breast cancer relates to the basal-like phenotype. Cancer Res. 68 (4), 989–997. https://doi.org/10.1158/0008-5472.CAN-07-2017.

Schwender, H., 2012. Siggenes: Multiple Testing Using SAM and Efron's Empirical Bayes Approaches. R Package Version, 1(0). pp. 57.

Schwender, H., Schwender, M. H., Sam, M., MultipleComparisons, M., GeneExpression, S. N. P., & ExonArray, D. Package 'siggenes'.

Shah, S.N., Cope, L., Poh, W., Belton, A., Roy, S., Talbot Jr, C.C., ... Resar, L.M.S., 2013. HMGA1: a master regulator of tumor progression in triple-negative breast cancer cells. PloS one 8 (5), e63419.

Song, M.-A., Brasky, T.M., Marian, C., Weng, D.Y., Taslim, C., Dumitrescu, R.G., ... Shields, P.G., 2015. Racial differences in genome-wide methylation profiling and gene expression in breast tissues from healthy women. Epigenetics 10 (12), 1177–1187.

Spears, P.A., Barnes, J.A., 2003. HSP70 enhances MCF-7 cell growth and estrogen receptor activity. In: Paper Presented at the Proc 1st Int Cong Stress Responses Biol Med, pp. 2003.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., ... Mesirov, J.P., 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America 102 (43), 15545–15550. https://doi.org/10.1073/pnas.0506580102.

Tang, B., Xie, R., Qin, Y., Xiao, Y.F., Yong, X., Zheng, L., ... Yang, S.M., 2016. Human telomerase reverse transcriptase (hTERT) promotes gastric cancer invasion through cooperating with c-Myc to upregulate heparanase expression. Oncotarget 7 (10), 11364–11379. https://doi.org/10.18632/oncotarget.6575.

Wang, J., Lu, M., Qiu, C., Cui, Q., 2009. TransmiR: a transcription factor–microRNA regulation database. Nucleic Acids Res. 38 (suppl_1), D119–D122.

Wang, X., Lu, G., Li, L., Yi, J., Yan, K., Wang, Y., ... Zhang, S., 2014. HUWE1 interacts with BRCA1 and promotes its degradation in the ubiquitin–proteasome pathway. Biochemical and biophysical research communications 444 (3), 290–295.

Wang, B., Lee, C.-W., Witt, A., Thakkar, A., Ince, T.A., 2015. Heat shock factor 1 induces cancer stem cell phenotype in breast cancer cell lines. Breast Cancer Res. Treat. 153 (1), 57–66.

Zeller, C., Hinzmann, B., Seitz, S., Prokoph, H., Burkhard-Goettges, E., Fischer, J., ... Scherneck, S., 2003. SASH1: a candidate tumor suppressor gene on chromosome 6q24. 3 is downregulated in breast cancer. Oncogene 22 (19), 2972.

Zhang, Y., Wang, Y., Guo, Y., Liao, Z., Xu, R., Ruan, Z., 2015. miR-135b promotes the invasion and metastasis of hepatocellular carcinoma cells. Xi bao yu fen zi mian yi xue za zhi = Chinese Journal of Cellular and Molecular Immunology 31 (10), 1316–1321.

Zhao, S., Fung-Leung, W.-P., Bittner, A., Ngo, K., Liu, X., 2014. Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. PLoS One 9 (1), e78644.

Zhong, J., Liu, C., Zhang, Q.-H., Chen, L., Shen, Y.-Y., Chen, Y.-J., ... Cao, R.-X., 2017. TGF-β1 induces HMGA1 expression: The role of HMGA1 in thyroid cancer proliferation and invasion. International journal of oncology 50 (5), 1567–1578.

Zhu, Y., Qiu, P., Ji, Y., 2014. TCGA-assembler: open-source software for retrieving and processing TCGA data. Nat. Methods 11 (6), 599–600.