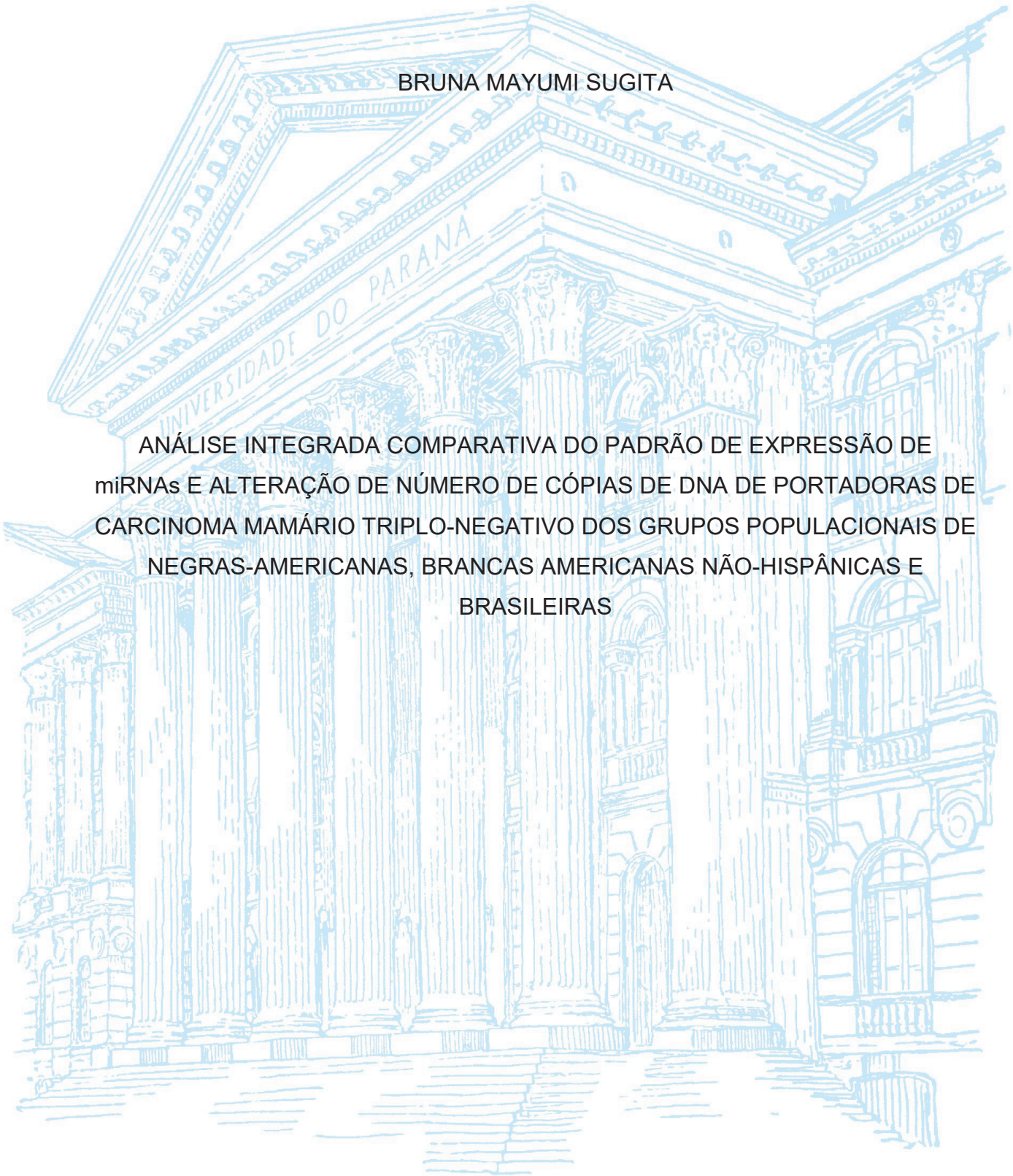
UNIVERSIDADE FEDERAL DO PARANÁ

BRUNA MAYUMI SUGITA



ANÁLISE INTEGRADA COMPARATIVA DO PADRÃO DE EXPRESSÃO DE miRNAs E ALTERAÇÃO DE NÚMERO DE CÓPIAS DE DNA DE PORTADORAS DE CARCINOMA MAMÁRIO TRIPLO-NEGATIVO DOS GRUPOS POPULACIONAIS DE NEGRAS-AMERICANAS, BRANCAS AMERICANAS NÃO-HISPÂNICAS E BRASILEIRAS

CURITIBA

2018

BRUNA MAYUMI SUGITA

ANÁLISE INTEGRADA COMPARATIVA DO PADRÃO DE EXPRESSÃO DE miRNAs E ALTERAÇÃO DE NÚMERO DE CÓPIAS DE DNA DE PORTADORAS DE CARCINOMA MAMÁRIO TRIPLO-NEGATIVO DOS GRUPOS POPULACIONAIS DE NEGRAS-AMERICANAS, BRANCAS AMERICANAS NÃO-HISPÂNICAS E BRASILEIRAS

Tese apresentada como requisito parcial à obtenção do grau de Doutor em Genética, no Curso de Pós-Graduação em Genética, Setor de Ciências Biológicas, da Universidade Federal do Paraná.

Orientadora: Prof[a]. Dr[a]. Enilze M. S. F. Ribeiro
Coorientadores: Prof[a]. Dr[a]. Luciane R. Cavalli
Prof. Dr. Iglenir J. Cavalli

CURITIBA

2018

Universidade Federal do Paraná. Sistema de Bibliotecas.
Biblioteca de Ciências Biológicas.
(Telma Terezinha Stresser de Assis –CRB/9-944)

Sugita, Bruna Mayumi
Análise integrada comparativa do padrão de expressão de miRNAS e alteração de número de cópias de DNA de portadores de carcinoma mamário triplo-negativo dos grupos populacionais de negras-americanas, brancas americanas não hispânicas e brasileiras. / Bruna Mayumi Sugita. – Curitiba, 2018.
145 p.: il. ; 30cm.

Orientadora: Enilze M. S. F. Ribeiro
Coorientadora: Luciane R. Cavalli
Coorientador: Iglenir J. Cavalli
Tese (Doutorado) - Universidade Federal do Paraná, Setor de Ciências Biológicas. Programa de Pós-Graduação em Genética.

1. Mamas - Cancer. 2. MicroRNAs. I. Título. II. Ribeiro, Enilze M. S. F. III. Cavalli, Luciane R. IV. Cavalli, Iglenir J. V. Universidade Federal do Paraná. Setor de Ciências Biológicas. Programa de Pós-Graduação em Genética.

CDD (20. ed.) 574.873224

MINISTÉRIO DA EDUCAÇÃO
SETOR CIÊNCIAS BIOLÓGICAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO GENÉTICA

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em GENÉTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **BRUNA MAYUMI SUGITA** intitulada: **Análise integrada comparativa do padrão de expressão de miRNAs e alteração de número de cópias de DNA de portadoras de carcinoma mamário triplo-negativo dos grupos populacionais de negras-americanas, brancas americanas não-hispânicas e brasileiras.**, após terem inquirido a aluna e realizado a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de doutor está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 23 de Março de 2018.

ENILZE MARIA DE SOUZA FONSECA RIBEIRO
Presidente da Banca Examinadora

DANIELA FIORI GRADIA
Avaliador Interno

IRIS RABINOVICH
Avaliador Externo

SARAH FRANCO VIEIRA DE OLIVEIRA MACIEL
Avaliador Externo

SELENE ELIFIO ESPOSITO
Avaliador Externo

RESUMO

As disparidades da saúde consistem em diferenças no status de saúde de indivíduos estratificados em grupos específicos devido a alguma característica ou fator epidemiológico específico. Os tumores mamários triplo-negativos (TNBC) constituem um subtipo de carcinoma mamário clinicamente agressivo, frequentemente diagnosticado em mulheres jovens e de grupos populacionais específicos como negras americanas, hispânicas e latinas. Essas diferenças na incidência e prognósticos do TNBC em determinadas populações indicam a existência de disparidades da saúde do câncer de mama. Os TNBC são um grupo bastante heterogêneo uma vez que são caracterizados apenas pela ausência de expressão dos receptores hormonais de estrogênio e progesterona, e a não superexpressão do HER2. Neste estudo, foram analisados os padrões de alterações de números de cópias de DNA (CNA) e de expressão de microRNA de amostras TNBC e não-TNBC (NTNBC) de pacientes negras americanas (AA), brancas americanas não-hispânicas (NHW) e brasileiras (LA). Um total de 164 amostras foi obtido do Hospital Nossa Senhora das Graças (Curitiba, PR, Brasil) e do *Lombardi Comprehensive Cancer Center Georgetown University* (Washington, DC, EUA). A análise comparativa da expressão de microRNAs foi realizada entre os subtipos TNBC e NTNBC de cada grupo populacional, resultando em 194 miRNAs diferencialmente expressos (DEmiRNAs) entre as pacientes AA, 336 DEmiRNAs entre as pacientes NHW, e 163 DEmiRNAs entre as pacientes LA.

A determinação dos DEmiRNAs entre as amostras AA TNBC e NHW TNBC resultou em 256 miRNAs diferencialmente expressos entre os dois grupos; a análise integrada com dados de CNAs das amostras AA TNBC resultou em um painel de 26 miRNAs que apresentaram dados concordantes de expressão de miRNAs e CNAs. Análises de curva de ROC indicaram um bom poder discriminatório desses miRNAs individuais e do painel em diferenciar as amostras AA TNBC e NHW TNBC. Análises de bioinformática indicaram que os 26 miRNAs podem estar envolvidos em vias de sinalização importantes na carcinogênese como via do PI3K-Akt, MAPK e Insulina.

Entre os 26 miRNAs, o miR-150-5p foi selecionado para estudos funcionais que permitiram a determinação de seu potencial oncogênico no câncer de mama, principalmente em AA TNBC. A modulação da expressão do miR-150-5p através da utilização de inibidores e mimetizadores foi realizada seguida de análises da capacidade proliferativa (MTS e Ensaio Clonogênico), migratória, expressão de marcadores de transição epitelial-mesenquimal (EMT) e resistência a tratamento.

A análise integrada dos dados de expressão de miRNAs e CNAs das pacientes LA resultou em um painel de 17 miRNAs com dados concordantes das duas análises. Este painel apresentou alto poder de discriminar amostras TNBC e NTNBC de pacientes brasileiras, com valor de AUC de 0.953.

Em resumo, foram observadas diferenças biológicas no nível molecular (DNA e RNA) entre os tumores mamários das três populações estudadas que podem auxiliar na elucidação das disparidades de saúde observadas. Estes resultados indicam que os miRNAs são importantes na tumorigênese mamária e a integração com dados de CNAs permitem uma seleção de miRNAs para estudos como biomarcadores de diagnóstico, prognóstico e desenvolvimento de tratamentos.

Palavras-chave: Câncer de mama. microRNAs. Disparidades de saúde.

ABSTRACT

Health disparities are differences in health status of individuals separated in specific groups because of specific characteristic or epidemiological factor. Triple negative breast cancer (TNBC) is a clinically aggressive subtype of breast cancer frequently presented in younger women from specific populations as African American, Hispanic and Latina. The differences in TNBC incidence and prognosis of specific populations show that breast cancer presents health disparities. TNBC is a highly heterogeneous group of diseases characterized by lack of expression of estrogen and progesterone receptors and lack of superexpression of HER2. In this study we analyzed DNA copy number alterations (CNA) and microRNAS expression patterns of TNBC and non-TNBC samples from African American (AA), non-Hispanic white (NHW) and Brazilian (LA) women. One hundred sixty four samples were obtained from *Hospital Nossa Senhora das Graças* (Curitiba, PR, Brazil) and Lombardi Comprehensive Cancer Center Georgetown University (Washington, DC, EUA). Comparative analysis of miRNA expression were performed between TNBC and NTNBC samples of each population, resulting in 194 differentially expressed miRNAs (DEmiRNAs) between AA subgroups, 336 DEmiRNAs between NHW subgroups, and 163 DEmiRNAs between LA subgroups.

Two hundred and fifty six DEmiRNAs were obtained when AA TNBC and NHW TNBC samples were compared; integrative analysis of miRNA expression and CNA data of TNBC samples resulted in 26 miRNAs with concordant data of both analysis. ROC analysis indicates a good power of both individual miRNAs and combined panel of 26 miRNAs in discriminating AA TNBC and NHW TMBC samples. Bioinformatic analysis indicate that these 26 miRNAs may be involved in important signaling pathways for carcinogenesis, as PI3K-Akt, MAPK and Insulin pathways.

Among the 26 miRNAs, miR-150-5p was selected for further functional analysis uncovering its oncogenic potential in breast cancer, especially in AA TNBC. Modulation of miR-150-5p expression by using inhibitors and mimics was performed followed by proliferation (MTS and Clonogenic Assay), migration, epithelial-to-mesenchymal transition (EMT) markers expression and cytotoxic assays.

Integrative analysis of miRNA expression and CNA data of LA patients resulted in 17 miRNAs with concordant data. This panel presented high power in discriminating TNBC and NTNBC subtypes of the Brazilian patients, with AUC of 0.953.

In conclusion, there are biological differences at molecular levels (DNA and RNA) among breast tumors of the three populations of this study that might help elucidate health disparities observed. These results indicate that miRNAs are important for breast tumorigenesis and integration with CNA data allow a fine selection of miRNAs as diagnostic, prognostic biomarkers and for treatment development.

Key-words**:** Breast cancer. microRNAs. Health disparities

LISTA DE FIGURAS

CAPÍTULO I

CAPÍTULO II

CAPÍTULO III

LISTA DE TABELAS

CAPÍTULO I

CAPÍTULO III

LISTA DE ABREVIATURAS E SIGLAS

AA	*African American*
aCGH	Hibridização Genômica Comparativa por microarranjo
AIMs	Marcadores informativos de ancestralidade
BL1	*Basal Like 1*
BL2	*Basal Like 2*
BSA	*Bovine serum albumine*
CNA	*Copy number alterations*
EMT	*Epithelial-Mesenchymal Transition*
ER	Receptor de estrogênio
HER2	Receptor de crescimento epidérmico 2
IM	*Immunomodulatory*
LA	Latinas / Brasileiras
LAR	Luminal Receptor de Andrógeno
LNA	*locked nucleic acid*
miRNAs	microRNAs
ML	*Mesenchymal like*
MSL	*Mesenchymal stem cell like*
NHW	*Non-Hispanic white*
NTNBC	Carcinoma Mamário não-triplo negativo
pCR	Resposta completa ao tratamento
PR	Receptor de progesterona
RT-qPCR	Reação em Cadeia da Polimerase Quantitativa em tempo real
SNP	Polimorfismo de nucleotídeo único
TNBC	Carcinoma Mamário triplo negativo

SUMÁRIO

1 INTRODUÇÃO

1.1 DISPARIDADES DA SAÚDE - “HEALTH DISPARITIES”

O termo disparidades da saúde (do inglês, *health disparities*) tem sido amplamente utilizado por agências de saúde principalmente dos Estados Unidos da América (EUA) e Europa para explicar as diferenças nos indicadores de saúde entre os diferentes grupos de indivíduos que compõem os seus países. O Instituto Nacional da Saúde dos EUA (NIH, do inglês *National Institute of* Health) considera disparidades da saúde como diferenças no *status* de saúde de indivíduos estratificados devido ao grupo populacional, gênero, idade, *status* socioeconômico, localização geográfica ou qualquer outra característica/fator epidemiológico que permite a distinção de grupos específicos (HEALTH, 2017).

Apesar da popularidade do termo, ainda não há um consenso quanto ao conceito de disparidades da saúde, uma vez que ele pode ser utilizado em um contexto não apenas biológico, mas também social em que se consideram o acesso diferencial a sistemas de saúde, a comorbidades e renda que podem afetar tanto o diagnóstico como a progressão, tratamento e potencial cura de diversas doenças, como doenças cardiovasculares, HIV/AIDS, hepatite viral, doenças sexualmente transmissíveis, tuberculose e câncer (CDC, 2017). Dessa forma, é importante considerar a diferença no conceito entre diferenças e inequidades da saúde, que têm sido utilizadas como sinônimos de Disparidades da Saúde, mas que apresentam conotações distintas. Em 1992, Margaret Whitehead publicou o manuscrito “*The Concepts and Principles of Equity and Health”* no qual a definição de equidade em saíde (do inglês, *equity in health*) e o seu significado no contexto da Organização Mundial de Saúde (WHO, do inglês *World Health Organization*) foram definidas, resultantes de diversas reuniões de membros do programa *Equity in Health in WHO’s Regional Office for Europe* (EURO) (WHITEHEAD, 1992).

As diferenças da saúde consistem em condições de saúde distintas devido a variações naturais biológicas, dano à saúde devido a comportamento escolhido pelo indivíduo (esportistas, por exemplo) e vantagem devido a comportamento adotado pelo indivíduo ou população que resulta em promoção à saúde. Em geral, esses fatores podem resultar em alterações do *status* de saúde do grupo estudado que são “desnecessárias e evitáveis”, mas não conferem a conotação de injustiça como as inequidades de saúde. As inequidades de saúde incluem comportamentos danosos

à saúde nos quais a escolha do estilo de vida do indivíduo é restrita (devido principalmente à classe sócioeconômica), a exposição a fatores que resultam em situações de estresse anormal ou condições insalubres de vida e de trabalho, acesso inadequado a serviços de saúde essenciais e públicos, e a tendência de redução da condição sócio-econômica de indivíduos não-sadios (WHITEHEAD, 1992). Assim, as inequidades de saúde resultam de fatores que surgem devido a uma injustiça social que ao ser reparada pode resultar na igualdade das condições de saúde dos diferentes grupos.

1.1.1 População Americana

Os Estados Unidos da América apresentam-se como uma população multicultural formada por diversos grupos populacionais distintos devido a ondas de imigração de diversos países que ocorreram durante os últimos séculos. O *United States Census Bureau* estimou a população estado-unidense em 2016 em 323.127.513 habitantes, sendo a sua grande maioria formada por indivíduos brancos (76,91%), seguida de negros americanos (13,31%), asiáticos (5,67%), ameríndios e nativos do Alasca (1,26%), e nativos do Havaí e outras ilhas do Pacífico (0,24%); indivíduos que se auto-declararam como pertencentes a dois ou mais grupos populacionais representam os 2.62% restantes (United States Census Bureau, 2018). O grupo denominado Hispânico e Latino Americano é considerado um grupo étnico formado por indivíduos de todos os grupos populacionais: 57.470.287 dos 323.127.513 habitantes dos EUA apresentam origem “hispânica”, resultando em 17,79% da população.

Por ser uma população composta por indivíduos de diversas origens com histórias de migração distintas, a disparidade da saúde pode ser facilmente observada, sendo em sua maioria inequidades, principalmente pelo acesso diferenciado a sistemas de saúde, mas também diferenças de saúde. Assim, o Departamento de Saúde e Serviços Humanos dos Estados Unidos da América (HHS, *U.S. Department of Health and Human Services*) criou em 2 de dezembro de 2010 o programa denominado *Healthy People 2020*, apresentando objetivos de 10 anos para promoção de saúde e prevenção de doenças de todos os americanos, o que inclui os grupos populacionais denominados *minorities* ou minorias, entre eles latinos, hispânicos, asiáticos e negros americanos (HHS, 2010). O câncer é uma das doenças apresentadas no programa Healthy People, tendo como objetivo a redução

de novos casos, assim como dos efeitos causados por ela, como os sintomas e deficiências físicas, assim como a morte.

1.1.2 População Brasileira

A população brasileira é composta por indivíduos provenientes de diversos continentes, como Europa, Ásia, África e Américas e apresenta alto índice de miscigenação devido a casamentos entre esses grupos populacionais, assim como outros países da América Latina (SALZANO e SANS, 2014). Essa miscigenação resulta em uma população com uma ampla variabilidade genética, conferindo à população brasileira uma alta complexidade biológica (GIOLO *et al.*, 2012)

Do ponto de vista de cuidados da saúde, essa miscigenação apresenta-se ambiguamente como uma vantagem e um desafio: testes clínicos com novas terapias que apresentam sucesso na população brasileira muito provavelmente serão bem sucedidos em outras populações mundiais; ao mesmo tempo tratamentos testados e aprovados em populações específicas (europeias geralmente) podem ser aproveitadas por apenas uma pequena fração da população brasileira.

1.2 DISPARIDADES DA SAÚDE NO CÂNCER DE MAMA

As taxas de incidência e mortalidade do câncer de mama variam consideravelmente entre grupos populacionais, com valores geralmente altos em indivíduos com ascendência africana e baixos em asiáticos ou das ilhas do Pacífico (SIEGEL *et al.*, 2017).

Mulheres brancas americanas não-Hispânicas (NHW, do inglês *non-Hipanic white women*) apresentam maior chance de serem diagnosticadas com câncer de mama, entretanto mulheres negras americanas (AA, do inglês *adrican american*) apresentam maior probabilidade de ir a óbito devido à doença (DALY e OLOPADE, 2015). Isso ocorre principalmente devido a condições sócioeconômicas distintas, resultando na disponibilidade e acessibilidade a métodos de diagnóstico precisos diferentes, o que confere certa vantagem a determinado grupo populacional (NHW) em relação a outro (AA): o diagnóstico preciso precoce facilita o tratamento e potencial cura do câncer, enquanto o diagnóstico tardio pode resultar no contrário, frequentemente levando a paciente à óbito. Entretanto, o componente biológico das

diferenças no *status* de saúde dessas populações permanece desconhecido e camuflado pelo alto poder das diferenças sócioeconômicas. Dessa forma, estudos com populações NHW e AA com condições sócioeconômicas similares podem auxiliar na descoberta desses fatores biológicos, o que pode resultar no desenvolvimento de métodos de diagnóstico, prognóstico e também de novos tratamentos mais específicos.

Diversos estudos epidemiológicos e moleculares comprovaram a influência da ancestralidade/grupo populacional dos indivíduos na distribuição dos diferentes subtipos de câncer de mama (O'BRIEN *et al.*, 2010), incluindo estudos que analisam casos da doença em homens (CHAVEZ-MACGREGOR *et al.*, 2013; KROENKE *et al.*, 2014; SCHINKEL *et al.*, 2014; KOHLER *et al.*, 2015). Mulheres diagnosticadas com câncer de mama do subtipo triplo negativo (TNBC, do inglês *triple negative breast cancer*) apresentam diferentes taxas de incidencia de acordo com a origem etnica, sendo que as negras americanas (24-28%) e latinas/hispânicas (12%) apresentam uma maior incidência e frequência quando comparadas com mulheres brancas não-hispânicas (12%) (SCHINKEL *et al.*, 2014; PLASILOVA *et al.*, 2016). Ainda, apresentam maior risco de ir a óbito, uma vez que o tumor apresenta-se em idade mais precoce, em estágios avançados, com maiores chances de metástase e baixa resposta a tratamento (PATEL *et al.*, 2010; LARA-MEDINA *et al.*, 2011; KILLELEA *et al.*, 2015; SHEPPARD *et al.*, 2016; WARNER *et al.*, 2016)

1.3 CÂNCER DE MAMA

O câncer de mama é o tipo mais frequente (com exceção dos tumores de pele não melanoma) e uma das principais causa de morte de mulheres por doenças malignas no mundo (SIEGEL *et al.*, 2017). Para o biênio de 2018/2019 estima-se a ocorrência de 59.700 novos casos no Brasil (risco de 56,33 a cada 100 mil mulheres) e 266.120 nos Estados Unidos (INCA/MS, 2018; SIEGEL *et al.*, 2018). Nos Estados Unidos, as estimativas indicam que o câncer de mama corresponderá a 30% dos novos casos e 14% das causas de morte por câncer no ano de 2018 (SIEGEL *et al.*, 2018).

Apesar de ser uma doença tratável e de prognóstico favorável quando detectado precocemente, foram observadas ainda 15.593 mortes por câncer de mama no Brasil no ano de 2015 (MS/SVS/CGIAE, 2015) e estimadas 40.920 nos

Estados Unidos no ano de 2018 (SIEGEL *et al.*, 2018). Um estudo realizado em 2012 analisou as taxas de mortalidade do câncer de mama no Brasil como um todo e por regiões, observando uma tendência à redução do número de mortes nas regiões Sudeste e Sul e a um aumento na região Nordeste, principalmente. Esse padrão foi explicado pelo número de mamógrafos presentes por região, sendo que as regiões com maior número de mortes apresentaram menor número de mamógrafos (FREITAS-JUNIOR *et al.*, 2012). Isso indica a importância e necessidade de uma conscientização maior da população no que diz respeito à doença e de políticas públicas de saúde do governo que auxiliem no diagnóstico precoce do câncer de mama, bem como métodos alternativos a exames radiológicos que frequentemente dependem de máquinas caras, como os métodos laboratoriais utilizando-se de marcadores biológicos específicos.

O câncer de mama compreende um grupo heterogêneo de doenças, que apresentam características histológicas, patológicas e moleculares distintas, resultando na progressão tumoral, potencial metastático e respostas diferentes aos tratamentos (NORUM *et al.*, 2014).

1.3.1 Classificação Histológica do Câncer de Mama

Histologicamente, os carcinomas mamários são amplamente distribuídos em pelo menos 21 subtipos, sendo nove tipos especiais e suas variantes, onze tipos muito raros e o carcinoma invasivo do tipo não especial (NST, do inglês *no special type*). Entre os tipos especiais, incluem-se carcinomas mamários lobulares invasivos, carcinomas tubulares, carcinoma cribiforme invasivo, carcinoma com elementos medulares, carcinoma mucinoso, carcinomas com diferenciação em celulas em anel de sinete, carcinoma micropapilar invasivo, carcinoma com diferenciação apócrina e carcinoma metaplásico (GOBBI, 2012; SINN e KREIPE, 2013).

Os carcinomas mamários invasivos NST, previamente conhecidos como carcinomas ductais invasivos sem outra especificação (CDI-SOE), constituem um grupo formado por carcinomas mamários que não apresentam características patológicas específicas que os diferenciem (SINN e KREIPE, 2013).

Ainda, os carcinomas mamários são classificados de acordo com a o Sistema de Graduação de Nottingham (do inglês *Nottingham Grading* System), de acordo com os indicadores de diferenciação celular do tecido tumoral: grau nuclear, formação tubular e atividade mitótica. Cada uma dessas características histológicas

é analisada por médicos patologistas e designada por uma pontuação de 1 a 3 (ELSTON e ELLIS, 1991).

1.3.2 Classificação Molecular do Câncer de Mama

Devido à heterogeneidade molecular do câncer de mama, diversos pesquisadores têm buscado meios de classificar os tumores em grupos moleculares distintos. Perou *et al*. foram os primeiros a utilizar análises genômicas para classificar os carcinomas mamários de acordo com seus perfis de expressão gênica (PEROU *et al.*, 2000). Em 2000, esses autores analisaram 65 tumores mamários que foram então classificados em quatro subtipos tumorais, os chamados subtipos intrínsecos: luminal, basal, HER2 enriquecido (do inglês, *HER2 enriched*) e semelhante ao normal (do inglês, *normal like*). Nos anos de 2001 e 2003, o mesmo grupo validou esses dados em um novo grupo de amostras, e dividiram o subtipo luminal em A e B (SORLIE *et al.*, 2001; SORLIE *et al*., 2003).

Após a publicação dos dados de Perou e Sorlie e com a evolução dos estudos genômicos e das plataformas de expressão gênica, outros subtipos moleculares de carcinomas mamários foram identificados. O primeiro novo subtipo molecular identificado foi denominado Molecular Apócrino (do inglês, *apocrine molecular*), caracterizado pela alta atividade da via de sinalização andrógena e características histológicas apócrinas (FARMER, P. *et al.*, 2005).

Em 2007, a análise de amostras de tumores mamários humanos em conjunto com modelos de tumores mamários murinos resultou na identiicação de um grupo específico de tumores que apresentou expressão reduzida de genes envolvidos em junções oclusivas (*tight junctions*) e adesão célula a célula, como as claudinas 3, 4,7, ocludina e E-caderina, além da baixa expressão de genes luminais, expressão inconsistente de genes basais, alta expressão de genes envolvidos com o sistema imune como *CD4* e *CD79a*. Esse grupo de tumores com padrão de expressão gênica singular foi denominado Claudina baixa (do inglês, *claudin low*) (HERSCHKOWITZ *et al.*, 2007).

1.3.3 Classificação Imunohistoquímica

A classificação dos carcinomas mamários em grupos moleculares tem sido amplamente utilizada na pesquisa básica, no entanto não tem se mostrado de muita

utilidade na prática clínica uma vez que análises de expressão gênica apresentam um custo elevado, sendo pouco acessível para a população. Assim, clinicamente os *status* dos receptores de estrogênio e progesterona (PR, do inglês *progesterone receptor)* e a amplificação do receptor de crescimento epidérmico 2 (*HER2,* do inglês *human epidermal growth factor receptor 2*) se apresentam com maior utilidade, principalmente devido a maior acessibilidade e menor custo das técnicas imunoistoquímicas quando comparadas às análises de expressão gênica. Ainda, não há uma relação perfeita entre os subtipos moleculares e imunoistoquímicos, uma vez que o *status* dos receptores hormonais é variável dentro dos subtipos moleculares, como pode ser observado na TABELA 1. Um exemplo é o fato de a maioria dos carcinomas mamários clinicamente/imunoistoquimicamente classificados como HER2 positivos e potencialmente tratados com trastuzumab apresenta-se na classe molecular dos tumores luminais, devido à expressão concomitante de ER (PRAT e BASELGA, 2008; MONTEMURRO *et al.*, 2013), e não ao subtipo HER2 enriquecido. Além disso, 29 a 34% dos tumores classificados como HER2 enriquecido não apresentam a superexpressão do HER2, mas apenas o aumento da expressão de genes relacionados à via do HER2 (REIS-FILHO e PUSZTAI, 2011; ALIZART *et al.*, 2012). Interessantemente, em um estudo de 2012, 17% das amostras classificadas como HER2-enriquecido apresentaram ausência de ER, PR e da superexpressão do HER2 (padrão imunoistoquímico triplo-negativo) (BASTIEN *et al.*, 2012).

Após uma conferência internacional realizada em 2011 (*12th St Gallen International Breast Cancer Conference*) foi adotada uma nova abordagem para a classificação de pacientes com o objetivo de padronização do tratamento do câncer de mama. Nessa conferência, foram utilizadas as características imunoistoquimicas das amostras para classifica-las em subtipos tumorais: luminal A (ER e PR positivo, HER2 negativo, e Ki67 baixo), luminal B/HER2 negativo (ER e PR positivo, HER2 negativo e Ki67 alto), luminal B/HER2 positivo (ER e PR positivo, HER2 positivo e Ki67 indiferente), HER2 positivo (ER e PR negativos, HER2 positivo) e triplo-negativos. Essa classificação permite a escolha e a melhor predição do resultado dos tratamentos padrões indicados para cada subtipo (GOLDHIRSCH *et al.*, 2011).

TABELA 1 – SUBTIPOS MOLECULARES DE CARCINOMAS MAMÁRIOS

Subtipo Intrínseco	**Características Gerais**	**Imunoistoquímica**	**Prognóstico**
Luminal A	Apresentam um perfil de expressão gênico semelhante ao das células luminais do epitélio normal da mama, geralmente associado à alta expressão de ER e genes associados (*SLC39A6*) com baixo grau histológico, apresentando prognóstico favorável. As células cancerosas geralmente são diploides, apresentam mutação em *PI3KCA, GATA3, MAP3K1*, e superexpressão de ciclina D1.	ER+ (91 – 100%) PR+ (70 – 74%) HER2+ (8 – 11%) Ki67 ≤20%	Bom
Luminal B	Apresentam um perfil de expressão gênico semelhante ao das células luminais do epitélio normal da mama, mas com alta taxa de proliferação de células cancerosas e prognóstico pior do que o luminal A.	ER+ (91 – 100%) PR+ (41 – 53%) HER2+ (15 – 24%) Ki67 >20%	Ruim
HER2 enriquecido	Apresentam alta expressão de genes associados com a ativação da via de sinalização do receptor HER2. Geralmente apresentam alta expressão da proteína HER2 e GRB7, e baixa expressão de genes das células luminais. Apresentam prognóstico ruim antes da utilização de medicamentos como herceptina.	ER- (41 – 71%) PR- (70 - 75%) HER2+ (66 – 71%)	Ruim
Semelhante ao Normal	Apresentam expressão semelhante às células normais do epitélio mamário, como genes basais/mioepiteliais e genes específicos do tecido adiposo.	ER+ (44 – 100%) PR+ (22 – 63%) HER2+ (6 – 22%)	Intermediário
Basal	Apresenta expressão gênica similar a células basais ou mioepiteliais, como citoqueratinas 5/6 e 17, mas não expressam genes relacionados com células luminais ou com o gene HER2. Esse subtipo tumoral está associado com alta taxa de proliferação de células cancerosas e prognóstico desfavorável. Representa cerca de 15% de todos os tipos de câncer de mama	ER- (81 – 100%) PR- (87 – 94%) HER2- (87 – 91%)	Ruim
Apócrino Molecular	Geralmente apresentam alta expressão de receptor de andrógeno (AR, do inglês *androgen receptor*) e genes associados.	ER- e PR- HER2-/+ AR: +	Ruim
Claudina Baixa	Ausência da expressão de proteínas claudinas que auxiliam na adesão célula-célula, baixa expressão de marcadores luminais, mas alta expressão de marcadores da transição epitelial-mesenquimal (EMT, do inglês *epithelial-to-mesenchymal transition*) e de células tronco putativas.	ER- (67 – 88%) PR- (77 – 78%) HER2- (78 – 94%)	Intermediário

FONTE: Adaptado de REIS FILHO & PUSZTAI (2011); NORUM, ANDERSEN & SORLIE (2014)
Legenda: ER: receptor de estrogênio, PR: receptor de progesterona, HER2: receptor de crescimento epidérmico 2, AR: receptor de andrógeno

1.3.4 Câncer de mama triplo-negativo

O Câncer de Mama Triplo-Negativo (TNBC, do inglês *Triple Negative Breast Cancer*) consiste em um grupo heterogêneo que compreende os carcinomas mamários que não apresentam expressão de ER, PR e nem a superexpressão do HER2. Apesar de apresentarem características similares aos carcinomas do tipo basal e serem considerados como sinônimos por alguns pesquisadores, os TNBC podem apresentar padrões de expressão gênica distintos entre si e nem sempre se sobrepõem aos carcinomas basais. BASTIEN *et al*. (2012) mostraram que 81% dos tumores tipo basal em seu estudo (57 das 70 amostras) apresentaram padrão imunoistoquímico TNBC, entretanto apenas 56% (57 das 110 amostras) e 67% (60 das 90 amostras) dos TNBC clasificados por IHC e qPCR respectivamente apresentaram padrão de expressão gênica basal. Do restante das 110 amostras classificadas como TNBC por IHC, 30 (30%) foram classificadas como HER2-enriquecidas, 10 (10%) como Luminal B e 4 (4%) como Luminal A (BASTIEN *et al.*, 2012).

LEHMANN *et al*. (2011) realizou uma análise de agrupamento (do inglês *clustering*) de dados de perfil de expressão gênica de 587 tumores triplo negativos de 21 bancos de dados. Essa análise resultou na classificação dos TNBC em seis grupos estáveis com características moleculares e ontológicas distintas: basal 1 (BL-1, do inglês *basal-like 1*), basal 2 (BL-2, do inglês *basal-like 2*), imunomodulatório (IM), mesenquimal (M), mesenquimal similar às células tronco (MSL, do inglês *mesenchymal stem-like*), luminal com receptor de andrógeno (LAR, do inglês *luminal androgen receptor*) (LEHMANN *et al.*, 2011). Estudos posteriores do mesmo grupo resultaram no refinamento da classificação previamente proposta com seis subtipos, resultando em uma nova classificação em quatro subtipos de TNBC: BL1, BL2, M e LAR, uma vez que os subtipos MSL e IM eram produtos de células mesenquimais associadas a tumores e, de tumores com grande infiltração de linfócitos respectivamente (LEHMANN *et al.*, 2016).

Tumores mamários do subtipo BL1 apresentam alta expressão de genes envolvidos no ciclo celular, divisão celular e resposta a dano no DNA, como *CCNA2, MYC, RAS, CHEK1* e *RAD51.* O subtipo BL2 apresenta alterações em genes de vias de sinalização de fatores de crescimento (EGF, NGF, MET, Wnt/β-catenina, IGFR1) e na glicólise e glucogenólise. Ainda, ambos os subtipos basais apresentam alta

expressão de genes de citoqueratinas basais (*KRT5, KRT6A, KRT6B, KRT4, KRT16, KRT17, KRT23 e KRT81*) (LEHMANN *et al.*, 2011).

O padrão de expressão do sutipo M é enriquecido em genes envolvidos em motilidade celular, interação de receptores de matriz extracelular (ECM, do inglês *extracellular matrix*) e vias de difereciação celular (LEHMANN *et al.*, 2011). Pacientes diagnosticadas com TNBC subtipo M apresentaram maior frequência de metástases de pulmão (46%) quando comparadas aos outros subtipos (25%, p=0,0388) (LEHMANN *et al.*, 2016).

O subtipo LAR apresenta alta expressão de genes envolvidos em vias reguladas por hormônios, além das citoqueratinas luminais (*KRT7, KRT8, KRT18 e KRT19*) e outros marcadores luminais como *FOXA1* e *XBP1* (LEHMANN *et al.*, 2011). Assim como outros cânceres regulados por hormônios, tumores do subtipo LAR apresentam o tecido ósseo como sítio preferencial metastático (46%), sendo essa preferência significativa quando comparada aos outros subtipos (16%, p= 0,0456) (LEHMANN *et al.*, 2016).

1.3.4.1 Tratamento do TNBC

Devido às características moleculares dos TNBC, as tratamentos específicos como hormonioterpia com Tamoxifeno e a terapia-alvo com Trastuzumab não são eficazes no tratamento desse subtipo tumoral, dessa forma pacientes acometidas por TNBC são tratadas com combinações de cirurgia, radioterapia e quimioterapia.

Estudos indicam que o TNBC apresenta boa resposta à quimioterapia neoadjuvante, alcançando 22 (LIEDTKE *et al*., 2008) e 19% (KHOKER *et al*., 2013) de resposta clínica completa em dois estudos realizados separadamente (LIEDTKE *et al.*, 2008; KHOKHER *et al.*, 2013). Apesar do sucesso, os efeitos adversos da quimioterapia ainda são graves por ser um tratamento sistêmico que age tanto em células cancerosas quanto em células saudáveis, dessa forma faz-se necessário o desenvolvimento de novos medicamentos que possam ser utilizados no tratamento do TNBC e que atinjam apenas as células do câncer.

O tratamento de escolha para TNBC frequentemente é a quimioterapia neoadjuvante com taxanos e antraciclina. Devido a heterogeneidade já demonstrada dos TNBCs, os diversos subtipos apresentam respostas diferentes ao tratamento tradicional. Em seu estudo, MASUDA *et al.* (2013) obtiveram Resposta Completa ao tratamento (pCR) em 37 dos 130 (28%) pacientes diagnosticados com TNBC.

Quando classificados em subtipos, os tumores apresentaram valores de pCR significativamente distintos BL1= 52%, BL2= 0%, M= 31% e LAR= 10% (Teste exato de Fisher, p= 0,04379) e o subtipo de TNBC foi considerado um preditor independente do *status* de pCR (Teste de razão de verossimilhança, p= 0,022) (MASUDA *et al.*, 2013).

Algumas terapias alvos foram desenvolvidas e tem sido testadas para o tratamento do TNBC, como os inibidores da PARP, EGFR, VEGF e SRC. Apesar dos esforços em estudos clínicos, apenas o tratamento com inibidores da PARP tem apresentado resultados promissores.

As PARPs são uma família de enzimas envolvidas em diversos processos celulares, como a transcrição, o reparo do DNA, manutenção da estabilidade genômica, regulação do ciclo celular e morte celular não-apoptótica (MICHELS *et al.*, 2014). Entre as 18 enzimas da família, a PARP1 é a mais abundante, sendo relacionada com reparo do DNA, regulação da estrutura da cromatina, transcrição, sobrevivência e morte celular, e transdução de sinais inflamatórios (WEIL e CHEN, 2011). A PARP1 age como um sensor de quebras do DNA, resultando na ativação da resposta de dano ao DNA reparando as quebras simples e dupla-fita, auxiliando na manutenção da estabilidade do genoma. Eventos que levem a quebras do DNA resultam no aumento dos níveis de PARP1 que se liga ao DNA quebrado e catalisa o processo de *PARylation* de diversas proteínas próximas a quebra do DNA como a PARP1, histonas, proteínas de reparo do DNA, fatores de transcrição e reguladores da cromatina. A ativação dessas proteínas desencadeia então o reparo do DNA (WEIL e CHEN, 2011).

Uma vez que grande parte das terapias antineoplásicas age provocando danos ao DNA, acredita-se que a utilização de Inibidores da PARP pode reduzir a ativação da resposta ao dano ao DNA, o que impediria o reparo das quebras de DNA e a morte de células cancerosas. Um estudo em 2001 demonstrou a maior atividade dos inibidores da PARP em células cancerosas, quando comparadas às células saudáveis, o que reduziria efeitos adversos do uso dessa nova terapia (SHIOBARA *et al.*, 2001). A seletividade da ação dos inibidores da PARP pode ser explicada pela maior instabilidade genômica em células tumorais quando comparadas com células do tecido normal (WEIL e CHEN, 2011).

O tratamento com inibidores da PARP parece ser mais efetivo em carcinomas mamários que apresentam alguma deficiência nos genes *BRCA1* e/ou *BRCA2*. Com

a inibição da PARP, danos do DNA se acumulam e poderiam ser corrigidos por Recombinação Homóloga, processo de reparo de DNA que necessita de BRCA1 e BRCA2 funcional. Dessa forma, células deficientes em BRCA1 ou BRCA2 apresentariam maior suscetibilidade ao uso desse tratamento, resultando em morte celular (FARMER, H. *et al.*, 2005). Um estudo clínico de fase II (NCT00494234) demonstrou a atividade anti-tumorigênica do inibidor da PARP *olaparib* em pacientes com câncer de mama que apresentam mutação germinativa BRCA1/2 (KAUFMAN *et al.*, 2015). A Fase III do estudo, denominado OlympiAD (NCT02000622) determinou que o uso do olaparib como terapia adjuvante de pacientes HER2 negativas com mutação germinativa BRCA1/2 apresentou benefício significativo quanto a sobrevida livre da doença (PFS, do inglês *progression-free survival*) quando comparado ao tratamento de escolha do oncologista (quimioterapia, *capecitabine, vinotelbine* ou *eribuline*) (ROBSON *et al.*, 2017).

Cerca de 10 a 30,8% de mulheres diagnosticadas com TNBC apresentam mutações em *BRCA1* ou *BRCA2* (GREENUP *et al.*, 2013; SHARMA *et al.*, 2014), assim acredita-se que este grupo apresentaria maiores benefícios com o tratamento individual com inibidores da PARP ou combinado com outras terapias. O efeito da administração de um inibidor da PARP (*iniparib*) associada a quimioterapia (gencitabina + carboplatina) contra apenas a quimioterapia foi avaliado em pacientes diagnosticados com TNBC metastáticos. Como resultado, observaram benefícios clínicos e na sobrevida de pacientes tratadas com *iniparib*, sem aumento significativo de efeitos tóxicos adversos (O'SHAUGHNESSY *et al.*, 2011).

Apesar dos esforços na avaliação da utilização de terapias existentes para o tratamento de TNBC, até o momento os estudos clínicos não têm demonstrado eficiência satisfatória. Dessa forma, ainda fazem-se necessários estudos que foquem nas características moleculares desse subgrupo tumoral. Os miRNAs surgem como biomarcadores moleculares, para diagnostico, prognostico e também como alvos no desenvolvimento de terapias especificas.

1.4 microRNAS

Os microRNAs (miRNAs) são uma classe de pequenos RNAs não codificadores (sncRNA, do inglês *small non coding RNA*), com aproximadamente 22pb de comprimento que atuam na regulação da expressão de genes responsáveis por diversos processos biológicos importantes para o bom funcionamento de célula,

tecidos e organismo, como o desenvolvimento, proliferação, diferenciação e metabolismo celular (AMBROS, 2004; LI *et al.*, 2009). Dessa forma, a alteração da expressão de um ou mais miRNAs pode resultar na desregulação desses genes, apresentando-se como protagonistas no desenvolvimento de doenças malignas como o câncer (ZHANG *et al.*, 2014; TAKAHASHI *et al.*, 2015).

Estima-se que um único miRNA é capaz de regular mais de 30% dos genes codificadores de proteínas no genoma humano, e um único gene pode ser regulado por mais de um miRNA (LEWIS *et al.*, 2003; DOENCH e SHARP, 2004). Essas características demonstram a complexidade da regulação gênica pelos miRNAs e a importância desses pequenos transcritos para o bom funcionamento do organismo.

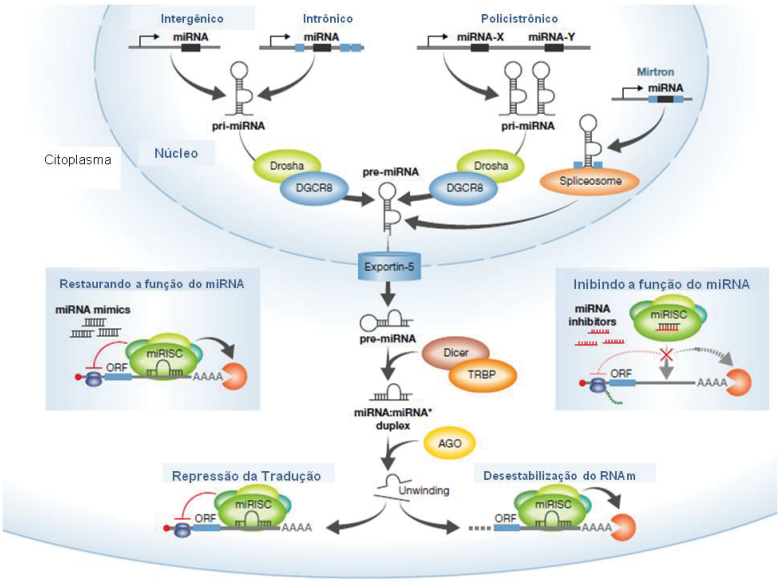
1.4.1 Biogênese dos miRNAs

A biogênese dos miRNAs é um processo de múltiplas etapas (FIGURA 1) que se inicia no núcleo da célula com a transcrição do precursor pri-miRNA (do inglês, *primary miRNAs*) com 1-3kb pela RNA polimerase II. O pri-miRNA é então clivado por um complexo multiprotéico denominado *Microprocessor* (formado pela enzima RNase II DROSHA e pela proteína DGCR8, do inglês *DiGeorge syndrome critical region gene 8*) em pre-miRNAs, que consistem em estruturas secundárias com 70 a 120 nucleotídeos apresentados em forma de grampo. Os pre-miRNAs são exportados para o citoplasma com o auxílio da EXPO5, e então processados pela DICER (RNase III) em um dúplex miRNA-miRNA com 18 a 24 nucleotídeos. O dúplex miRNA-miRNA é separado e a fita guia (do inglês, *guide strand*) forma o Complexo RISC (do inglês, *RNA-induced silencing complex*) juntamente com proteínas ligadoras de RNA (entre elas TNRC6A) e as proteínas catalíticas Argonautas (AGO). A fita passageira (do inglês, *passenger strand*) é frequentemente degradada, mas pode também se associar a proteínas AGO formando outro complexo RISC funcional.

O resultado da ligação do Complexo RISC com os RNAm alvos depende da complementariedade com o miRNA: a complementariedade completa resulta na degradação do RNAm, levando à redução da concentração do transcrito e da proteína. O pareamento incompleto miRNA-RNAm pode resultar na deadenilação do RNAm (GIRALDEZ *et al.*, 2006), regulação da tradução (CHEKULAEVA e FILIPOWICZ, 2009; FABIAN *et al.*, 2010) e no aumento da expressão do RNAm em resposta a condições celulares específicas como a quiescência (VASUDEVAN *et al.*,

2008; MORTENSEN *et al.*, 2011). A regulação positiva da transcrição pode ocorrer diretamente pela ação do complexo RISC ou indiretamente como consequência dos efeitos regulatórios da repressão mediada por miRNA, efeito denominado como "afrouxamento da repressão" (do inglês, *relief of repression*) (BHATTACHARYYA *et al.*, 2006; SARAIYA *et al.*, 2013).

FIGURA 1 - BIOGÊNESE DOS miRNAS



FONTE: Adaptado de (VAN ROOIJ e KAUPPINEN, 2014)

A maioria dos genes de miRNAs encontra-se localizada em íntrons de genes codificadores e não-codificadores de proteínas, o que pode permitir uma regulação coordenada dos miRNAs sobre esses genes (RODRIGUEZ *et al.*, 2004; BHASKARAN e MOHAN, 2014). A outra porção dos genes de miRNAs encontra-se em regiões intergênicas que possuem suas próprias unidades de transcrição independente com região promotora e sítios de ligação de fatores de transcrição (GARZON *et al.*, 2009; BHASKARAN e MOHAN, 2014).

1.4.2 miRNAs e Câncer

Estudos de expressão indicam uma redução global da expressão de miRNAs em diversos tipos de neoplasias malignas, incluindo câncer de mama (HURTEAU *et al.*, 2007), próstata (SHI *et al.*, 2007), fígado (GRAMANTIERI *et al.*, 2007), pâncreas (LEE *et al.*, 2007) e ovário (YANG *et al.*, 2008), entretanto alguns miRNAs específicos apresentam expressão aumentada em cânceres humanos, pricipalmente aqueles relacionados com a capacidade metastática das células cancerosas

(ZHANG *et al.*, 2014). Os níveis de expressão reduzidos da maioria dos miRNAs pode resultar em células em estágios pouco diferenciados, características comuns encontradas em células cancerosas (DI LEVA e CROCE, 2010). KUMAR *et al*. (2007) realizaram o *knockdown* da maquinaria de processamento de miRNAs e a depleção global desses transcritos estimulou a transformação celular e a tumorigênse *in vivo* (KUMAR *et al.*, 2007).

Os miRNAs podem alterar a expressão de oncogenes e genes supressores de tumor, assim como podem ter seus níveis de expressão alterados por eles (KAVITHA *et al.*, 2014; ZHANG *et al.*, 2014). O resultado biológico obtido através dos níveis alterados de determinado miRNA é a soma de todas as interações desse miRNA com seus alvos, o que depende não apenas do tipo celular e do tecido em que ocorre a expressão do miRNA, mas também da quantidade e variedade dos genes expressos por essas células. Dessa forma, os miRNAs podem apresentar função dupla de oncogene e gene supressor de tumor, dependendo principalmente dos genes que estão afetando.

1.5 VALIDAÇÃO EXPERIMENTAL DE miRNAS

A detecção dos níveis de expressão dos miRNAs é um passo importante nos estudos de doenças malignas humanas para elucidação da patogênese. Entretanto é importante ainda a determinação da função biológica desse miRNA no tecido específico.

Cada miRNA apresenta centenas de alvos evolutivamente conservados e um número ainda maior de alvos não-conservados (BENTWICH *et al.*, 2005), o que torna a identificação experimental desses alvos um desafio. Dessa forma, estudos que envolvem a análise de miRNAs e seus respectivos genes alvos devem ser realizados para confirmar a interação miRNA-RNAm predita, assim como o resultado dessa interação no contexto biológico.

1.5.1 Predição de Alvos

Os miRNAs reconhecem seus RNAms alvos pelo pareamento entre seus nucleotídeos 2-7 (região *seed*) e os nucleotídeos complementares na região 3'UTR dos mRNAs. As regras de pareamento da região *seed* são importantes e informativas para a predição de alvos de miRNAs e têm sido amplamente usadas

por ferramentas de bioinformáticas atuais. Entretanto, evidências sugerem que o pareamento perfeito entre as duas sequências não é um indicador confiável de interação funcional entre o miRNA e o mRNA alvo, não sendo necessária e nem suficiente para essa interação, dessa forma alguns mRNAs alvos preditos podem não ser biologicamente funcionais (DIDIANO e HOBERT, 2006).

As ferramentas de bioinformática de predição de mRNAs alvos utilizam algoritmos computacionais que analisam as sequências individuais de cada gene para prever a interação com os miRNAs em estudo. Para cada ferramenta são utilizados parâmetros específicos para auxiliar na predição dos alvos, dessa forma, ferramentas diferentes podem resultar em listas diferentes de alvos prováveis para um determinado miRNA. Assim, recomenda-se a utilização de pelo menos três algoritmos de predição de alvos de miRNA, sendo considerados os alvos preditos por pelo menos dois dos algoritmos para validação experimental posterior (KUHN *et al*., 2008).

TargetScan (http://www.targetscan.org/)

O TargetScan é uma ferramenta de bioinformática de predição de alvos de miRNA que analisa a presença de sítios 8mer e 7mer conservados que correspondem a região *seed* de determinado miRNA (LEWIS *et al*., 2005). Os sítios 7mer e 8mer são definidos como:

- 7mer-m8: complementariedade completa com os nucleotídeos 2-8 do miRNA maduro (região *seed* + posição 8)

- 7mer-1A: complementariedade completa com os nucleotídeos 2-7 do miRNA maduro (região *seed*) seguidos por uma Adenina.

- 8mer: complementariedade completa com os nucleotídeos 2-8 do miRNA maduro (região *seed* + posição 8) seguidos por uma Adenina.

miRDB (http://mirdb.org/)

O miRDB é uma ferramenta *online* de predição de alvos de miRNA que utiliza a abordagem de “aprendizagem de máquinas” (do inglês, *machine learning approaches*) pelo método de máquinas de vetores suporte (SVM, do inglês, *support vector machines*) (WANG e EL NAQA, 2008). O algoritmo SVM desenvolvido para essa ferramenta combina 131 características heterogêneas de predição de alvo, como conservação da combinação da região *seed*, combinação da base terminal e estrutura secundária formada pelo duplex miRNA-mRNA.

Diana micro-T-CDS (http://www.microrna.gr/webServer)

O Diana micro-T-CDS é uma ferramenta de bioinformática que utiliza parâmetros calculados individualmente para cada miRNA e cada elemento de reconhecimento de miRNA (MRE, do inglês *miRNA recognition element*) dos genes alvos (MARAGKAKIS *et al.*, 2009). Os MREs são os sítios UTR que possuem 7, 8 ou 9 nucleotídeos que pareiam de forma Watson-Crick (guanina-citosina e adenina-timina/uracila) com o miRNA, começando pela posição 1 ou 2 do terminal 5' do miRNA. O score obtido para cada par de interação miRNA-mRNA considera os MREs conservados e não-conservados dos genes alvos.

1.5.2 Efeito do miRNA sobre a expressão da proteína alvo

Para a avaliação da ação dos miRNAs sobre as proteínas alvos e processos biológicos específicos, é necessária a realização de análises que permitam um melhor entendimento do papel dos miRNAs nas células e o impacto da modulação dos níveis de expressão desses miRNAs sobre o comportamento da célula. Essas análises são denominadas análises de ganho de função e perda de função e utilizam-se mimetizadores e inibidores de miRNA respectivamente.

Os ensaios de ganhos de função ou de aumento de níveis de expressão são realizados utilizando vetores ou oligonucleotídeos quimicamente desenhados que imitam os miRNAs endógenos, ou seja, os miRNA-mimetizadores (do inglês, *miRNA mimics*). Uma vez que os miRNAs agem reduzindo a expressão dos genes alvos, espera-se que o aumento dos níveis do miRNA em estudo pela adição do miRNA-mimetizador resulte na redução da expressão do gene alvo respectivo (ZHAO *et al*., 2007).

Os estudos de perda de função ou inibição de expressão dos miRNAs utilizam oligonucleotídeos complementares aos miRNAs maduros (anti-miRs) com LNAs (do inglês, *locked nucleic acid*) que são inseridos nas células através da técnica de transferência gênica, com o uso de vetores virais ou não-virais. A interação LNA/DNA com o miRNA em estudo é mais estável e facilita o sequestro do miRNA, dessa forma não há a interação com o gene alvo (NAGUIBNEVA *et al*., 2006).

Os resultados dos ensaios de ganho e perda de função devem ser preferencialmente analisados de forma a avaliar a ação sobre a expressão de RNA (*Northern Blotting*) e de proteína. A alteração da expressão da proteína respectiva pode então ser observada através de análises de *Western Blotting*, ensaio de

imunoabsorção enzimática (ELISA, do inglês *enzyme-linked immunosorbent assay*) ou experimentos de imunoistoquímica (ZHAO *et al*., 2007; KUHN *et al*., 2008), além de análises proteômicas com gel 2D-DIGE e espectrometria de massa (ZHU *et al*., 2007).

1.5.3 Efeito do miRNA sobre as funções biológicas alvo

Após a confirmação da ação do miRNA em estudo sobre os genes alvos selecionados, é necessária a realização de análises que avaliem se há alguma alteração dos processos e funções biológicas pelos quais esses genes são responsáveis ou atuam. Nessa etapa, análises a serem utilizadas dependerão da função do gene, incluindo análises de proliferação celular, diferenciação celular, migração, apoptose, entre outros.

1.6 ALTERAÇÕES DE NÚMERO DE CÓPIAS DE DNA

Alterações de número de cópias de DNA (CNA, do inglês *copy number alterations*) são alterações somáticas de ganho e perda de segmentos cromossômicos que ocorrem durante a vida do indivíduo e podem desencadear diversas doenças, principalmente o câncer (CIRIELLO *et al*., 2013; ZACK *et al.*, 2013).

A determinação de padrões de CNAs em diversos tipos de câncer tem sido possível graças a projetos como *The Cancer Genome Atlas* (TCGA), no qual trinta e três tipos de tumores, incluindo o câncer de mama, tem sido exaustivamente analisados de forma a elucidar os mecanismos moleculares do desenvolvimento do câncer. A partir desses estudos, foi observado que tumores originados de mesmos tecidos ou órgãos podem apresentar alterações genômicas subtancialmente distintas (heterogeneidade intracâncer, do inglês *intracancer heterogeneity*), assim como tumores originados de tecidos ou orgãos diferentes podem apresentar padrões de alteração genômica similares (similaridade entre tipos de câncer, do inglês *cross-cancer similarity*) (CIRIELLO *et al.*, 2013). Esse conhecimento auxilia no desenvolvimento de estratégias terapêuticas mais eficazes e efetivas para subtipos tumorais específicos (como o TNBC), e também terapias que possam ser utilizadas para o tratamento de diversos tipos de neoplasias malignas.

As CNAs podem resultar na ativação de oncogenes e na inativação de genes supressores de tumor, resultando na desregulação de processos importantes na iniciação e progressão da tumorigênese. Essas alterações podem não apenas afetar diretamente oncogenes e genes supressores de tumor, mas também reguladores da expressão como co-ativadores de transcrição, RNAs longos não-codificadores e miRNAs. Os genes de miRNAs são frequentemente encontrados em regiões que apresentam alta instabilidade cromossômica em câncer (CALIN *et al.*, 2004; ZHANG *et al.*, 2006), assim perdas e amplificações de segmentos cromossômicos podem ocorrer resultando na alteração de seus níveis de expressão.

2 JUSTIFICATIVA

A abordagem biológica das disparidades de saúde é importante pois implica na compreensão das diferenças biológicas entre as populações que podem resultar em diferentes histórias naturais da doença assim como em resultados distintos. Estudos que analisam a variabilidade genética humana já indicam a importância da influência de polimorfismos no desenvolvimento de doenças como o câncer, entretanto poucos estudos analisam as diferenças somáticas entre os tumores mamários de grupos populacionais distintos. Diante disto, a caracterização de tumores mamários de diferentes grupos populacionais como as de mulheres brancas americanas não-hispânicas (NHW), negras americanas (AA) e brasileiras (LA) é importante no auxílio da compreensão do componente biológico das disparidades da saúde.

Os tumores mamários triplo-negativos (TNBC) constituem um grupo heterogêneo de carcinomas mamários, com características histopatológicas e moleculares distintas, resultando em prognósticos e resposta a tratamentos diferentes. Geralmente apresentam-se com maior frequência e mais agressivamente em grupos populacionais distintos como negras americanas, latinas e hispânicas, assim a caracterização molecular dos TNBC dessas populações é importante para o melhor entendimento dessa disparidade de saúde no contexto biológico.

Os microRNAs surgiram como importantes candidatos a biomarcadores moleculares, capazes de diferenciar tipos e subtipos tumorais de forma eficaz, uma vez que apresentam níveis de expressão tecido-específico, com alterações sensíveis a fatores endógenos (ex: hormônios) e exógenos. Por se encontrarem em regiões de instabilidade cromossômica, a análise integrada dos dados de expressão de miRNAs e de alterações de números de cópias de DNA permite uma melhor seleção de miRNAs para estudos funcionais que determinem o seu papel na tumorigênese.

3 OBJETIVOS

Neste trabalho temos como objetivo principal identificar e comparar o padrão de expressão de miRNAs e de alteração de números de cópias de DNA de amostras de carcinomas mamários de três populações (negras americanas, brancas americanas não-hispânicas e brasileiras), a fim de avaliar as possíveis diferenças biológicas entre os carcinomas mamários apresentados pelos grupos.

Os objetivos específicos são:

- Obter o perfil de expressão de miRNAs de pacientes diagnosticadas com tumores mamários triplo-negativos (TNBC) dos três grupos populacionais;
- Comparar o padrão de expressão global de miRNAs em amostras TNBC e NTNBC de pacientes negras americanas, brancas americanas não-hispânicas e brasileiras;
- Obter o padrão de alterações de números de cópias de DNA (CNAs) de pacientes com TNBC dos três grupos populacionais,
- Integrar os dados de expressão de miRNAs e de CNAs das três populações estudadas para identificar e selecionar miRNAs potencialmente importantes na tumorigênese de TNBC das três populações estudadas;
- Avaliar o potencial preditivo e prognóstico da expressão dos miRNAs selecionados (de acordo com o valor de p) em cada grupo de pacientes;
- Validar os dados de expressão dos miRNAs obtidos pelo *screening* global selecionados utilizando o método de RT-qPCR nas mesmas amostras de tumor e em linhagens celulares comerciais de câncer de mama;
- Selecionar os miRNAs validados acima e os correspondentes genes alvos para subsequentes estudos funcionais, utilizando sistemas *in vitro* de modulação de níveis de expressão de miRNAs.

4 DESCRIÇÃO DOS CAPÍTULOS

Esta tese está organizada em capítulos, contendo três artigos científicos. O primeiro capítulo consiste no artigo publicado em 2 de novembro de 2016 na revista *Oncotarget* (2016; 7(48):79274-79291), doi:10.18632/oncotarget.13024) intitulado ***Differentially expressed miRNAs in triple negative breast cancer between African American and Non-Hispanic White women***. Este artigo descreve as diferenças encontradas entre amostras de tumores triplo-negativos de pacientes de dois grupos populacionais americanos distintos, considerando a expressão global de miRNAs e o padrão de alterações de números de cópias de DNA. Entre os resultados obtidos, o painel de 26 miRNAs diferencialmente expressos selecionados apresenta-se como um material de estudo interessante para melhor compreensão das diferenças biológicas entre as duas populações estudadas.

O segundo capítulo apresenta o artigo intitulado ***The oncogenic role of miR-150-5p in triple negative breast cancer*** a ser submetido para publicação. A expressão diferencial do miR-150-5p foi observada em nossas análises como níveis mais altos em amostras de pacientes AA-TNBC quando comparadas a NHW-TNBC. Assim, o objetivo desse segundo artigo foi analisar a influência da modulação dos níveis de mR-150-5p em processos celulares encontrados desregulados na carcinogênese. Como resultado, foi observado um papel oncogênico do miR-150-5p, principalmente na linhagem celular HCC1806 (derivada de paciente negra) possivelmente pela interação com a via SRC.

O artigo que compõem o terceiro capítulo desta tese será submetido para publicação e se intitula **Integrated DNA copy number alteration and global miRNA expression analysis: a panel of 17 miRNAs for triple negative breast cancer in Brazilian patients**. Neste projeto, foram utilizadas amostras de pacientes brasileiras da cidade de Curitiba, com o objetivo de determinar as diferenças do padrão de expressão de miRNAs e de alterações de números de cópias de DNA entre os subtipos tumorais TNBC e NTNBC. Como resultado da análise integrada realizada, foi obtido um painel com 17 miRNAs que apresentaram um ótimo poder discriminatório entre os subtipos tumorais analisados.

5 CAPÍTULO I

Differentially expressed miRNAs in triple negative breast cancer between African-American and non-Hispanic White women

Bruna Sugita[1], Mandeep Gill[2], Akanksha Mahajan[2], Anju Duttargi[2], Saurabh Kirolikar[2] Rodrigo Almeida[1], Kenny Regis[2], Olusayo L. Oluwasanmi[2], Fabio Marchi[3], Catalin Marian[4], Kepher Malambi[2,5], Bhaskar Kallakury[6], Laura Sheahan[7], Iglenir J. Cavalli[1], Enilze M. Ribeiro[1], Subha Madhavan[2,7], Simina Boca[2,7], Yuriy Gusev[2,7], Luciane R. Cavalli[2*].

[1] Department of Genetics, Federal University of Parana, Curitiba, PR Brazil

[2] Department of Oncology, Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, Washington, DC, USA

[3]International Research Center-CIPE, A. C. Camargo Cancer Center, São Paulo, SP, Brazil

[4] The Ohio State University Comprehensive Cancer Center, Division of Cancer Prevention and Control, College of Medicine, The Ohio State University, Columbus, Ohio and the University of Medicine and Pharmacy Timisoara, Timisoara, Romania

[5] Department of Biostatistics, Bioinformatics and Biomathematics, Georgetown University, Washington, DC USA

[6]Department of Pathology, Georgetown University Medical Center, Washington, DC, USA

[7] Innovation Center for Biomedical Informatics, Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, Washington, DC, USA

*corresponding author:

Luciane R. Cavalli, PhD, Assistant Professor of Oncology, Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, 3800 Reservoir Rd, NW, room E512, Washington DC, 20007. Office number: 202-687-2029/email: lrc@georgetown.edu.

Artigo publicado na revista Oncotarget

ISNN: 1949-2553

Fator de Impacto: 5.168

Aceito: 25 de outubro de 2016

Publicado: 2 de novembro de 2016

ABSTRACT

Triple Negative Breast Cancer (TNBC), a clinically aggressive subtype of breast cancer, disproportionately affects African American (AA) women when compared to non-Hispanic Whites (NHW). MiRNAs(miRNAs) play a critical role in these tumors, through the regulation of cancer driver genes. In this study, our goal was to characterize and compare the patterns of miRNA expression in TNBC of AA (*n* = 27) and NHW women (*n* = 30). A total of 256 miRNAs were differentially expressed between these groups, and distinct from the ones observed in their respective non- TNBC subtypes. Fifty-five of these miRNAs were mapped in cytobands carrying copy number alterations (CNAs); 26 of them presented expression levels concordant with the observed CNAs. Receiving operating characteristic (ROC) analysis showed a good power (AUC ≥ 0.80; 95% CI) for over 65% of the individual miRNAs and a high combined power with superior sensitivity and specificity (AUC = 0.88 (0.78−0.99); 95% CI) of the 26 miRNA panel in discriminating TNBC between these populations. Subsequent miRNA target analysis revealed their involvement in the interconnected PI3K/AKT, MAPK and insulin signaling pathways. Additionally, three miRNAs of this panel were associated with early age at diagnosis. Altogether, these findings indicated that there are different patterns of miRNA expression between TNBC of AA and NHW women and that their mapping in genomic regions with high levels of CNAs is not merely physical, but biologically relevant to the TNBC phenotype. Once validated in distinct cohorts of AA women, this panel can potentially represent their intrinsic TNBC genome signature.

INTRODUCTION

Triple negative breast cancer (TNBC) is a clinically aggressive subtype of breast cancer that confers a high risk of metastasis development, usually shortly after the initial diagnosis [1]. Despite numerous advances identifying and testing potential biomarkers and their corresponding therapeutic compounds in TNBC clinical trials, an effective and approved targeted therapy for these tumors is not yet available [2, 3].

Variation among different racial and ethnic groups in the incidence of breast cancer molecular subtypes and clinical outcomes is well documented [4, 5]. The

basal-like tumors, in particular TNBC, are present at higher frequencies in African American (AA) women when compared to non-Hispanic Whites (NHW), Hispanics or Asian women [5]. AA women with TNBC disease are usually diagnosed at an earlier age and more advanced stage, and likely to develop early metastasis compared to NHW women [5]. An increased number of studies have characterized the differences in tumor biology between AA and NHW patients and have shown that breast tumors from AA patients present increased cell proliferation, elevated expression of angiogenesis markers and higher migration and invasive properties [6–8]. These findings suggest that although socio-economic and cultural factors play a role, biological factors can also be the major drivers of these disparities [8, 9].

MicroRNAs (miRNAs) are a class of non-coding endogenous RNA molecules that have been identified to play a role in breast cancer, through the regulation of cellular processes associated with aggressive tumor phenotypes, such as TNBC, that rapidly progress to metastatic disease and develop treatment resistance [10]. A number of miRNAs were observed with differential expression in TNBC in comparison to non-TNBC subtypes [11]. Interestingly, as in gene expression studies, miRNA expression profiling has been shown to discriminate the intrinsic molecular breast cancer subtypes [12]. MiRNA expression varies according to ethnicity [13]. A number of studies have shown germline miRNA polymorphisms in association with the susceptibility risk of breast cancer in specific ethnic populations [14–19]. However, there are limited reports on somatic miRNA expression levels in the breast tissue of these populations [20, 21]. Consequently, the variation of the miRNA expression levels in the tumors of AA in comparison to NHW or other groups is not well known.

In this study our primary goal was to characterize the main patterns of miRNA expression in the breast tumor tissue of AA patients with TNBC in comparison to that of NHW patients with TNBC by using genome-wide miRNA profiling. A number of significant miRNAs were observed to be differentially expressed between these groups. These miRNAs were distinct from the ones differentially expressed in the non-TNBC subtypes of both AA and NHW patients. The association of the miRNA expression with copy number data, performed by array-CGH analysis in the same TNBC specimens of the AA patients, revealed a panel of 26 miRNAs which mapped in the most frequent cytobands with copy number alterations (CNAs) and with their expression levels directly corresponded to copy number gains or losses. Receiving

operating characteristic (ROC) curve analysis of the individual miRNAs of this panel showed that over 65% of them presented a good discriminatory power value between the TNBC of AA and NHW patients. Subsequently, by applying a pipeline of comprehensive computational analysis, several critical and interconnected cancer gene networks and signaling pathways were found to be regulated by these miRNAs and their validated targets. Finally, association with clinical-pathological data from the patients revealed that three miRNAs from this 26 miRNA panel were associated with early patients' age at onset.

RESULTS

Differentially expressed miRNAs of TNBC and non-TNBC in African American and NHW women

MiRNA profiling was successfully performed in 88.9% (24/27) of the TNBC and non-TNBC (24/27) cases of the AA patients and in 93% (28/30) of the TNBC and 83.3% (25/30) of the non-TNBC cases of the NHW patients. The differentially expressed miRNAs were initially compared between each tumor subtype (TNBC and non-TNBC) for each group of patients according to the workflow of FIGURE 1.

The comparison of the miRNA expression levels of the TNBC and non-TNBC cases in the AA group revealed 194 miRNAs differentially expressed (*t*-test; $P < 0.01$; FDR < 0.05). The top 15 significant miRNAs observed up- and down-regulated, based on log2 fold change value, are presented in Supplementary TABLE S1. Unsupervised and Supervised Hierarchical Clustering analysis (Pearson correlation; $P < 0.01$, FDR < 0.05) applied to these tumors distinctly clustered the TNBC and non-TNBC tumors, with the exception of six and four cases, respectively (FIGURE 2).

To explore the function of each of the 194 identified miRNAs, we used DIANA miRPath analysis to perform KEGG pathway enrichment analysis. Among the top 15 pathways identified, based on *P* value, were the ones related to Pathways in Cancer, PI3K/AKT and MAPK signaling pathways (Supplementary TABLE S2).

The comparison of the miRNA expression levels of the TNBC and non-TNBC cases from the NHW patients revealed 336 miRNAs differentially expressed (*t*-test; $P < 0.01$; FDR < 0.05). The top 15 significant miRNAs observed up-regulated, based on log2 fold change value, are presented in Supplementary TABLE S3. For the down-regulated miRNAs, only a set of 12 miRNAs was observed in this analysis

(Supplementary TABLE S3). Unsupervised and Supervised Hierarchical Clustering analysis (Pearson correlation; $P < 0.01$, FDR < 0.05) applied to these tumors distinctly clustered the TNBC and non-TNBC cases, with the exception of one and two cases, respectively (FIGURE 2). As for the AA group, among the top 15 KEGG pathways identified by DIANA miRPath analyses in the NHW group, based on *P* value, were the ones related to Pathways in Cancer, PI3K/AKT and MAPK signaling pathways (Supplementary TABLE S4).

Finally, the comparison of the miRNA expression levels of the TNBC cases from both groups of patients, revealed 256 miRNAs differentially expressed (*t*-test; $P < 0.01$, FDR < 0.05). The top 15 significant miRNAs observed up- and down-regulated, based on log2 fold change value, are presented in TABLE 1. Unsupervised and Supervised Hierarchical Clustering analysis (Pearson correlation; $P < 0.01$, FDR < 0.05) distinctly clustered the TNBC from the AA and NHW patients, with the exception of ten and four cases, respectively (FIGURE 3).

DNA copy number analysis of TNBC in African American patients and association with miRNA expression levels

Array-CGH analysis was performed in the 27 TNBC cases of the AA group, 24 of which were also profiled for miRNA expression. A total number of 388 copy number alterations (CNAs) (as measured by the “number of calls”) were identified, with an average of 14.4 CNAs per case. The most frequent cytobands affected by CNAs were: 1q21.1-q44 and 8q11.1-q24 (in 55% of the cases), 3q11.1-q29 and 6p25.3-p12.1 (44% of the cases), 9p24.3-p13.1, 12p13.33-p11.1 and Xp22.33-p11.21 (39% of the cases) and 2p25.3-p11.2, 5p15.33-p12, 6q16.1-q25.3, 7q11.23-q36.3, 10p15.3-p11.1, 13q21.2-q34, 16p13.3-p11.1, 18p11.32-p11.21, 19p13.3-p12 and 19q12-q13.33 (26–33% of the cases) (FIGURE 4). A number of 7,362 genes were found to be located in these cytobands with CNAs, as generated by the Agilent Cytogenomics probe report.

For the association of CNAs with miRNA expression, the genomic location of the initial set of 256 miRNAs found differentially expressed between the TNBC of AA and NHW patients was verified. Fifty-five of them were located in the cytobands mostly affected by CNAs in the same AA-TNBC cases profiled by array-CGH as described above. From these 55 miRNAs, 26 presented expression levels in concordance with the observed CNAs (i.e. up-regulated miRNA expression/cytoband

with copy number gain and/or down-regulated miRNA expression /cytoband with copy number loss) at their respective genome locus (TABLE 2, FIGURE 4), including the miR- 150-5p, miR-200c-3p and miR-205-5p that were among the top 15 miRNAs with highest fold changes observed differentially expressed between the AA and NHW groups of patients (TABLE 1).

The average of the individual expression level (box plots) of the 26 miRNA panel in the TNBC of the AA and NHW's groups is shown in FIGURE 5. In the AA-TNBC group of patients, 9 to 24 of these selected 26 miRNAs were observed with alteration in their expression levels, with an average of 19.25 ± 0.85 miRNAs with expression changes per case. In the NHW-TNBC group, 8 to 24 of these miRNAs presented expression changes, with an average of 17.96 ± 0.88 miRNAs with expression changes per case. This difference was not statistically significant at $P < 0.05$.

Next, we integrated the copy number and miRNA expression data to determine gene targets that were potentially affected by both of these mechanisms. The 7,362 genes identified to be located in the cytobands affected by CNAs were "matched" with the miRNA targets of the 26 selected miRNAs. A number of 5,010 miRNA targets were identified, predicted by at least two miRNA target databases. The integration of these data revealed 1,557 common genes, reducing the number of targets to 69%. However, when the miRNA targets were integrated with the genes located in the most frequent cytobands (altered in ≥ 40% of the cases: 1q21.1-q44, 3q11.1-q29, 6p25.3-p12.1 and 8q11.1-q24.3) a larger reduction (86%: from 1,557 to 711) of the targets was observed (FIGURE 6).

Receiver operating characteristic (ROC) curve analysis

Individual ROC analysis of the 26 miRNAs of the identified panel showed that over 65% of them presented a good power in discriminating between the TNBC cases of AA and NHW group of patients [(Area Under the Curve (AUC) ≥ 0.80 with 95% Confidence Interval (CI)]. The highest discriminatory power was observed for miR-1125-3p (AUC = 0.89; 95%CI: 0.81-0.98), followed by miR-216a-5p, miR-532-5p, miR-580-3p, miR-599, miR-769-5p, miR-18a-5p, miR-28-5p, miR-182-5p, miR-183-5p, miR-1263, miR-4284, miR-4458, miR-205-5p, miR-614 and miR-940 (AUC values from 0.87 to 0.80). The remaining nine miRNAs of the 26 miRNA panel, presented AUC values ranging from 0.75 to 0.79. The combined analysis of the panel

showed a AUC value of 0.88, demonstrating the robust power of the 26 miRNA panel, with a high and combined superior level of sensitivity and specificity (0.78 and 0.99, respectively) in discriminating TNBC between these populations. The ROC plots, AUC values and the corresponding 95% confidence intervals for each of the 26 miRNAs and for the combined panel are presented in FIGURE 7 and Supplementary TABLE S5, respectively.

Functional enriched pathways

To explore the function of each of the miRNAs composing the 26 miRNA panel of differentially expressed miRNAs in the TNBC of AA and NHW groups and their corresponding targets, we performed KEGG pathway enrichment analysis. The top pathway identified (based on the lowest *P* value) was the neurotrophin signaling pathway. Twenty-three (88.5%) miRNAs of this panel were involved in this pathway. The next most significant pathways were the MAPK, insulin and PI3K/AKT signaling pathways, involving 92.3%, 88.5% and 88.5% of the 26 miRNAs, respectively. For the MAPK pathways the miRNAs not present were miR-614 and miR-4431 and for the insulin and PIK3/AKT were the miR-548ad-3p, miR-614 and miR-4431 (TABLE 3). These three pathways are interconnected, presenting several miRNA target genes in common (Supplementary FIGURE S1). IPA analysis showed that the selected 26 miRNAs are involved in the processes of cellular growth and proliferation, cell cycle, cell-to-cell signaling and interaction, and cell death and survival. This analysis also generated a gene network with 12 out of the 26 miRNAs of our panel targeting established cancer-related genes, such as *TP53*, *MYC*, *MYB*, *ZEB1*, *CCND3* and *TGFB* (FIGURE 8).

Association of the 26 miRNA panel with the clinical-pathological variables of the TNBC and non-TNBC cases of the AA and NHW groups of patients

Clinical-pathological variables at diagnosis, such as age, tumor size, stage and grade, and lymph node metastasis status and at follow up including local and/or distant metastasis recurrence were analyzed separately for each ethnic group in the TNBC and non-TNBC subtypes (TABLE 4). AA patients with the TNBC subtype presented an earlier age at diagnosis when compared to patients with the non-TNBC subtype (50.74 ± 1.55 and 57.41 ± 2.79, respectively; $P = 0.0413$). AA patients with the TNBC subtype also presented a higher frequency of stage III (88.9%) and grade 3 (88.5%) tumors when compared to patients with the non-TNBC subtype (42.3%

and 50%, *P* = 0.001 and *P* = 0.009, respectively). For the other clinical parameters, tumor size, lymph node, local recurrence and distant metastasis status, no significant difference was observed between the TNBC and non- TNBC cases in the AA group of patients.

In the NHW group of patients, no difference was observed in the mean age at diagnosis between the TNBC and non-TNBC subtypes. However, tumor size was significantly higher in the TNBC subtype of the NHW patients when compared to the non-TNBC subtype (2.96 ± 0.31 and 1.99 ± 0.25, respectively; *P* = 0.019). A higher level of significance was observed regarding tumor stage and grade in this group of patients: TNBC patients presented 89.7% of stage III and 92.8% of grade 3 tumors when compared to the non-TNBC patients (36.4% of both stage III and grade 3 tumors; *P* = 0.0002 and *P* < 0.0001 respectively). As for the AA patients, the other clinical parameters evaluated were not differently distributed in the tumor subtypes of the NHW patients.

Next, these analyses were performed comparing the tumor subtypes between the AA and NHW groups. A “borderline” significant difference was observed in relation to the presence of local recurrence in the TNBC group; AA patients presented a higher frequency of local recurrence when compared to the NHW patients (33.3% and 10.3%, respectively; *P* = 0.0519). None of the other parameters were significantly different in the TNBC or non-TNBC subtypes between these patients’ groups (TABLE 4).

The association of the expression levels of the 26 miRNA panel and clinical-pathological variables from the TNBC patients revealed that three miRNAs were significantly associated with age at the time of surgery, while adjusting for ethnicity, tumor size, and lymph node status: miR-216a-5p, miR-580-3p and miR-4458 (adjusted *P* value = 0.026 for all three). They all showed a positive association with age. Neither tumor size nor lymph node status was significantly associated with the expression levels of any of the 26 miRNAs evaluated in this analysis. This analysis excluded 7 samples which had missing clinical-pathological parameters.

DISCUSSION

The triple negative breast cancer (TNBC) subtype is observed with higher frequency in African American (AA) women in comparison to non-Hispanic White

(NHW) women, usually in association with poor prognosis and high mortality rate [4, 5, 8]. This disparity has been attributed to several causes, including tumor characteristics at diagnosis, such as age, tumor size, stage and grade [8]. In our study, however, none of these factors were significantly different between the AA and NHW' patients studied; the same occurred for the non- TNBC subtypes between the groups. These findings were in agreement with previous studies showing that age and pathology stage were not different between these populations, specially in the TNBC subtype [8]. However, within each group of patients we did observe significant differences at these parameters. In the AA group, a younger age at diagnosis, higher frequency of advanced tumor stage and poorly differentiated tumors were observed in the TNBC when compared to the non-TNBC subtypes. In the NHW group the same was observed for tumor grade and stage, but not for age at diagnosis, which is in agreement with several other reports, showing that age at diagnosis has a higher impact in the AA population than in other populations [5].

MiRNA expression has been shown to present an extraordinary power in classifying breast tumor subtypes [12]. In this study, genome-wide miRNA profiling distinctively clustered most of the TNBC and non-TNBC subtypes of both AA and NHW patients. In the NHW group only two (3.8%) cases were "misclassified" by the miRNA profiling (336 miRNAs differentially expressed) as opposed to four (8.3%) cases in the AA group (194 miRNAs differentially expressed). These results confirm the robust power of miRNA profiling in differentiating the intrinsic breast cancer molecular subtypes.

One of the most challenging aspects of analyzing tumor cells using whole genome miRNA platforms is the selection and "triage" of the most biologically relevant miRNAs and their corresponding targets among the large output data generated by these high-density platforms. One of the strategies utilized includes multi-platforms integration and functional enriched pathway analysis [21]. MiRNAs are frequently located in regions of genomic instability, markedly characterized by the presence of gains and losses of genomic regions [22–24]. Previous studies have integrated the data from miRNA expression and copy number profiling in human tumors [25–27]. Most of these studies were however, not performed in tumor tissue specimens that were concomitantly analyzed by these methods; in fact the vast majority of them were based on the miRNA and array-CGH profiling data extracted from public genomic datasets [26, 27], which invariably introduce a high level of

technical and sample heterogeneity considering that array platforms with different annotations and variable sample sources and material are “uniformly” combined.

In our study, these technical issues were eliminated, given that both analyses were performed in the same tissue specimens. This strategy allowed for the direct mapping of the total initial number of miRNAs differentially expressed in the TNBC of AA and NHW patients, in the genomic regions carrying CNAs, from which 26 miRNAs presented expression levels directly corresponding to copy number gains or losses. These findings are in agreement with the integration analysis performed by others, where tumor’ miRNAs were located in genomic regions frequently amplified and/or deleted in cancer [22–24]. It is relevant to point out, however, that because some of the regions with genomics gains and losses are defined by large cytobands, it is not unusual to find the same region with either gain or loss of copy number, which can explain some of the lack of concordance of the miRNA expression levels mapped on these genomic regions [24]. In addition, several other mechanisms can impact miRNA expression regulation other than CNAs [28].

However, among the 26 miRNAs of our study, fourteen (miRs 205-5p, 15b-5p, 1263, 28-5p, 4284, 93-5p, 182-5p, 183-5p, 599, 661, 200c-3p, 17-5p, 18a-5p and 23a-3p) were located in cytobands frequently amplified in TNBC [20, 29], including the 1q21.1-q44, 3q11.1-q29, 7q11.23-q36.3, 8q11.1-q24.3, 12p13, 13q21.2-q34, and 19p13, respectively. The miRs 200c-3p, 205-5p, 548ad-3p, 661, 17-5p, 18a-5p and 93-5p which were among the top 15 miRNAs with highest fold changes observed differentially expressed between the AA and NHW groups, are mapped in these locations (12p13, 1q32.2, 1q32.2, 8q23-24, and 13q31.3, 13q31.3 and 7q22.1, respectively). These miRNAs have been previously described with deregulated expression levels in breast cancer in association with cancer related pathways such as the ones involved in epithelial mesenchymal transition (EMT), cell migration, invasion and treatment resistance [30–33]. In addition, ROC curve analysis of the individual 26 miRNAs showed for 65.4% of these miRNAs a good discriminatory power in discriminating TNBC of AA and NHW patients. The ROC curve analysis of the combined miRNAs (AUC of 0.88) showed a superior power in discriminating these two populations, supporting the overall robustness of this panel. These results indicate that these 26 miRNAs are not randomly affected in TNBC, and may constitute significant differences in the biology of TNBC in these populations.

MiRNA expression levels have been demonstrated to vary according to ethnicity [13]. Several studies have shown the presence of genetic variants, mostly of single nucleotide polymorphisms (SNPs) in miRNA sites in association with the susceptibility risk of breast cancer in specific ethnic populations [14–19]. While in this study, we reported on “somatic” miRNA expression levels among the tumor tissue of the AA and NHW populations, which is not as frequently described, it is important to consider that the presence of these polymorphisms can modify miRNA expression, which could imply that some of the differences noted were not necessarily related to the tumor etiology, but rather to population stratification. Therefore, we queried the available miRNAs-SNPs databases (former HapMap [34], 1000 Genome Project [35]) and other reports on miRNA-SNPs in human populations [13, 15, 19], which include datasets of Lymphoblastoid cell lines (LCLs) derived from CEU (Utah residents with northern and western European ancestry) and YRI (Yoruba people from Ibadan, Nigeria) [34] for the presence of these genetic variants, that could affect the expression of the miRNAs composing our panel in the populations studied. Although SNPs were reported enriched in the AA and/or White population in few of the miRNAs of our 26 miRNA panel, e.g. miR-183-5p, miR-661, they were mostly present in the seeding sequences of the miRNAs (3'UTR sequences), which could present a higher impact in the impairment of their interaction with their corresponding targets and functional activity, rather than their expression levels. We evaluated by qRT-PCR the miRNA expression of few miRNAs composing our panel, including the miR-661, in the normal breast tissue from a subset of the AA and NHW breast cancer patients of this study (data not shown) and did not observed the altered expression levels that was shown in the tumor, ruling out this possible polymorphism “effect” in their expression levels.

The pathway and function enrichment analysis of the 26 miRNAs differentially expressed in TNBC of AA and NHW of our study, confirmed their association with breast cancer tumorigenic processes, especially the ones that confer clinically aggressive tumor phenotypes, such as the ones in TNBC.

The miR-205-5p in particular, have been described down-regulated in the TNBC subtype, which is in alignment with its tumor suppressor role in the inhibition of proliferation, migration and invasion of cancer cells [36, 37]; others, as in our study, have reported its up-regulation, compatible with its oncogenic role in tumor initiation and proliferation [38]. Due to its differential expression in serum among cancer

patients and healthy individuals, this miRNA has been considered a new biomarker for early detection of cancer [39, 40].

MiR-599 and miR-661, which were up-regulated in the TNBC cases of the AA group of this study, are mapped in cytobands with high level of amplification, the 8q23-24 region, which is often associated with basal tumors that present with focal amplification of the *C-MYC* oncogene [41, 42]. Although this is a region frequently affected in breast cancer irrespective of the molecular subtype and ethnicity [20], it has been observed preferentially amplified in TNBC, including the ones that present *BRCA1* mutations [43], which are frequently in AA women [44, 45].

Two clusters of miRNAs in our study were previously observed up-regulated in TNBC, with copy number directly influencing expression levels [25]; the cluster comprising miR-17-5p and miR-18a-5p, and the cluster of miR-93-5p, both located in 13q31.3, a commonly amplified genomic region in the TNBC of the AA patients. Targets of these miRNAs clusters are involved in gene networks associated with tumor aggressiveness, including the ones involving PIK3CA, C-MYC and PTEN genes [46–51]. In fact, the identification of the main pathways and gene networks affected by the corresponding miRNA targets of our 26 miRNA panel supported these reported findings. The neurotrophin signaling pathway was the one mostly affected, followed by the interconnected PI3K/ AKT, MAPK kinase and insulin pathways. With the exception of three miRNAs, miR-548ad-3p, miR-614 and miR-4431, 23 miRNAs in our panel were involved in at least one of these pathways. Growth factors of the neurotrophin family and their receptors have been shown to be involved in breast cancer, affecting tumor cell growth and metastasis [52–54]. PI3K/AKT, MAPK kinase and insulin pathways are frequently described in breast cancer, which downstream targets control cell proliferation, cell survival and glucose metabolism [55].

The target pathways of our study were also among the most common ones observed in the TNBC cases of the TCGA and other studies [20, 56–58], with downstream genes involved in the VEGF, C-MYC and PIK3 gene networks. In the TCGA study [20], the PIK3 network presented with the highest levels of gene deregulations, involving critical downstream genes, such as PTEN, one of the most relevant mRNA targets of our 26 miRNA panel. In fact, our previous study of TNBC in Latina women, also suffering from TNBC disparities, showed a high frequency of PTEN loss (62%) in significant association with advanced tumor grades [59].

Interestingly, the insulin pathway, which has also been shown to be regulated by miRNAs [60], including the miR-1225-5p in our study [61], was among the top pathways of the differentially expressed miRNAs in the TNBC and non-TNBC subtypes of the AA patients, but not in the NHW group of patients. These findings are supported by previous reports showing differential expression of the Insulin Growth Factor (IGF-1) gene and its receptors in association with breast cancer subtypes and ethnicity [62, 63]. In AA patients in particular, alterations in this pathway and its downstream targets have been shown to contribute to the increased risk of malignant transformation in young women and to confer more aggressive breast cancer subtypes [56–58]. Consistent with our findings, previous studies have shown that the transcriptional profile of the TNBC basal 1 subtype, which is often the most frequent TNBC subtype in AA [64], was associated with a low IGF1 signature score, a marker of high IGF1 receptor (IGF1R) expression and up regulation of the MAPK and AKT growth pathways [56]. A number of IGF1 receptor inhibitors are currently in preclinical and clinical trials [65, 66]; our and others data suggest that inhibitors of the IGF1 pathway may be a sensitive target in TNBC, particularly benefiting AA patients.

Interestingly, in our study, the insulin pathway was not specifically associated with the TNBC subtype in the AA patients, which indicates its relevance in the etiology of breast cancer in general in this population, which may be directly correlated with the high incidence of obesity and other co-morbidities, such as diabetes and other metabolic syndromes in this population [67]. Considering the critical role of the identified miRNAs in directing regulating targets associated with these co-morbidities, overall, these findings can form the basis to build a race specific genomic signature associated with co-morbidities, which can promote and augment prevention and intervention strategies, stratify and select appropriate patients for treatment and clinical trials and personalize cancer care.

In conclusion, we observed different patterns of miRNA expression in the TNBC of AA and NHW women in this study. The 26 miRNA panel observed in association with CNAs in the AA patients, presented a high power in discriminating TNBC between AA and NHW patients and affected critical cancer related gene networks and signaling pathways. These findings indicate that the mapping of the miRNAs in genomic regions with high levels of CNAs of this study was not merely physical, but biologically relevant to the TNBC genome of AA patients. The validation of this 26 miRNA panel in independent and larger samples sets from these ethnic

groups is required to ensure that their expression patterns are reflected across populations and are not unique to a particular cohort of patients. The future determination of their functional regulatory role in conferring the aggressive TNBC phenotype, including early development of metastasis and drug resistance, can open new opportunities to develop novel therapeutic targets for TNBC, holding promises towards the improvement of the overall survival rate of AA women with TNBC.

MATERIALS AND METHODS

General study design

Genome-wide miRNA profiling was performed in TNBC and non-TNBC cases from AA and NHW patients. The array-CGH analysis performed in the same TNBC specimens of the AA patients was associated with the miRNA expression profiling data as presented in the workflow of FIGURE 1. ROC analysis was performed to determine the individual discriminatory power of the identified panel of miRNAs in the TNBC of the studied populations. The identified miRNAs and corresponding targets were then biologically selected for their relevance in the breast cancer/TNBC phenotype by applying a comprehensive computational analysis using combinatorial target prediction algorithms in conjunction with Gene Ontology and pathway enrichment analysis. Finally association with clinical-pathological data from the patients was performed.

Patient accrual and sample collection

Formalin-fixed paraffin-embedded (FFPE) tumor sections were obtained from 27 and 30 cases of TNBC from AA and NHW breast cancer patients, respectively, from patients that undergone primary surgery for tumor removal, prior to any treatment, at the MedStar Georgetown Hospital, Washington DC. Tumors with a non-TNBC subtype were also obtained from both groups in the same number of patients: 27 non-TNBC from AA and 30 non-TNBC cases from NHW patients. Altogether 54 AA and 60 NHW patients were evaluated in this study. All the samples were procured from the patients from all groups of tumors (TNBC and non-TNBC) and ethnic groups (AA and NHW) under informed consent, performed by the personal from Non-Therapeutic Shared Resource (NTSR) of Lombardi Comprehensive Cancer Center according to their establish SOP. The specimens were received de-codified, with no

patient identifiers, under the Histopathology Tissue Shared Resources (HTSR)-IRB approved protocol (IRB#1992-048).

Clinical and pathological information was retrieved by the HTSR personnel and included: age at diagnosis, tumor size, stage and grade, and presence of lymph node metastasis. Breast cancer recurrence and distant metastasis status were also obtained with a follow-up period that varied from 2 to 9 years (TABLE 4). The analyzed breast cancer tissues were from primary tumor lesions, obtained prior to any cancer treatment, at the time of the surgery. The classification of the breast cancer TNBC and non-TNBC phenotype was determined by the three IHC surrogate markers ER, PR and HER2, following the American Society of Clinical Oncology (ASCO)/College of American Pathology (CAP) guidelines [68,69]. In the non- TNBC group, 65% of the cases were ER+/PR+/HER2-, followed by 19% of ER+/PR+/HER2+, when combining AA and NHW patients. Prior to the genomic analysis all of the tumor tissues were inspected for the presence of > 80% of tumor cells followed by microdissection according to a previous protocol [70].

Ancestral markers analysis

Ethnicity information was primarily obtained from self-reported patients' records. To obtain a genomic based information and assess ancestry we genotyped a subset of patients (13 AA and 14 NHW) using SNP chip Illumina Infinium QC Array (Illumina Inc., CA), which contains 15,949 markers (including ~3,000 ancestral informative markers (AIMs). The genotype calling was performed using GenomeStudio Software v. 2011.1 using the default settings. SNPs with MAF ≤ 0.01 were excluded from analysis. Further, we merged our dataset with the 1000 Genomes Project phase 1 (*n* =1,902 samples) [71], which present an overlap of 14,718 variants between ours and these data sets. Finally, we performed Principal Components Analysis (PCA) using PLINK 1.9 [72], which uses the EIGENSTRAT method [73] to calculate model ancestry differences between different samples. Based on the result of PC1 and PC2 we were able to differentiate the two main population groups (European (EUR) and African (AFR) in our samples confirming the self-report ethnicity information (Supplementary FIGURE S2).

MiRNA analysis

MiRNA expression analysis was performed using the Human v2 miRNA Expression Assay from NanoString nCounter Technology (Seattle, WA, USA), that contains 800 human probes derived from miRBase v.18. The raw data was pre-processed by NanoString's nCounter RCC collector and the miRNAs were normalized using the geometric mean. Unsupervised and supervised hierarchical cluster analysis was performed on miRNAs that were found to be significantly differentially expressed ($P < 0.01$, FDR < 0.05), using Pearson's correlation coefficient and average linkage by using the Multiexperiment Viewer software (MeV 4.9.0). Fold changes, represented on the log2 scale (log2FC), were calculated for all differentially expressed miRNAs.

Array-CGH analysis

DNA copy number analysis was performed using an oligonucleotide array-CGH platform (SurePrint G3 Human CGH Microarray 8x60K; Agilent Technologies Inc., Santa Clara, CA), according to the protocol for FFPE samples that we have established in our lab [70]. DNA was isolated from consecutive FFPE sections of the cases profiled for miRNA. DNA isolated from peripheral blood from multiple normal individuals was used as control DNA. The array data was analyzed using the Feature Extraction (FE) v.10.10 and Agilent CGH Analytics v.7.0 software (Agilent Technologies Inc., Santa Clara, CA), using the ADM-2 algorithm, threshold 6.0 and an aberration filter with a minimum number > 3 probes. Gene amplifications and deletions were defined as minimum average absolute log2 ratio (intensity of the Cy5 dye (reference DNA)/ intensity of the Cy3 dye (test DNA) value of > 0.25 and <−0.25, respectively, as per the CGH analytics analysis.

Association of miRNA expression levels and copy number alterations (CNAs)

To determine the potential association of miRNA expression levels and CNAs, we integrated the miRNA and copy number data that was performed in the same TNBC tissue specimens of the AA patients. Initially, the genomic location of each miRNA differentially expressed between the TNBC of AA and NHW patients was assessed at the miRbase.org and determined whether they were mapped at the most commonly affected cytobands (CNAs present in ≥ 33% of the cases). Next, it was determined whether the miRNAs residing in these cytobands presented the

corresponding changes in expression levels (i.e. cytoband with copy number gain/up-regulated miRNA expression and cytoband with copy number loss/down-regulated miRNA expression). A second integration approach was to determine the common gene targets that may be affected by both CNAs and miRNA expression alterations. For this approach, gene targets were queried using the available miRNA target databases and integrated with the genes located in the most commonly cytobands above. Only miRNA target genes that were present in two out of the three miRNA databases were selected. Considering that one miRNA can target several different genes irrespectively of their genomic location, all of the most frequently affected cytobands (present in at least 33% of the TNBC-AA cases) were included in this analysis.

Receiver operating characteristic (ROC) curve analysis

The ability of the identified 26 miRNA panel to discriminate between AA-TNBC and NHW-TNBC was examined by constructing receiver operating characteristic (ROC) curves and calculating the area under the curve (AUC). Sensitivity was plotted against 1-specificity for the binary classifier (AA-TNBC and NHW-TNBC). An AUC of 100% denotes perfect discrimination by the miRNA, whereas an AUC of 50% denotes complete lack of discrimination by the miRNA. AUCs and 95% corresponding confidence intervals were calculated for each miRNA.

Biological function and pathway analysis

The targeted pathways of the identified miRNAs were determined by DIANA miRPath v.2.0 software [74]. A detailed functional analysis to identify miRNA-mediated, cancer-related and statistically significant networks, biological functions and canonical signaling pathways for both differentially expressed miRNAs and target genes was performed by Ingenuity Pathway Analysis System (IPA v.8.5, Ingenuity Systems, Edwood, CA).

Analysis of clinical-pathological variables of the TNBC and non-TNBC of AA and NHW patients

The Student *t* test was used to analyze the differences of the mean age at diagnosis and tumor size in the TNBC and non-TNBC cases within and between the patient groups. The Chi-square (χ^2) test was used to evaluate tumor stage and grade

and the Fisher Exact test was used to evaluate lymph node, local recurrence, and distant metastasis status in both tumor subtypes. Significance level was considered to be $P < 0.05$. For the 26 miRNAs that were selected from the integration with the array-CGH data, linear regression models were considered having the transformed miRNA values as the outcomes and ethnicity, age, tumor size, and lymph node status as the regressors. For each of the clinical-pathological parameters (age, tumor size, lymph node status), a significance level of FDR < 0.05 using the Benjamin and Hochberg FDR control method [75] was considered. Tumor stage and grade were not considered since none of the AA and NHW patients with the TNBC subtype showed tumors with stage I and grade 1.

ACKNOWLEDGMENTS

The authors thank Dr. Deborah Berry and Krysta Chaldekas from The Histopathology and Tissue Shared Resource (HTSR) of Lombardi Comprehensive Cancer Center (LCCC) for assistance with specimen collection and Dr. Janice Rone for the careful reading of the manuscript and language revision. We are also in debt with Dr. Ira Shoulson and Erin Wilhelm, principal investigator and executive director, respectively, of the Georgetown University Center of Excellence in Regulatory Science and Innovation (CERSI) program.

CONFLICTS OF INTEREST

All the authors declare no conflicts of interest.

FINANCIAL SUPPORT

This research was supported by the Georgetown University Center of Excellence in Regulatory Science and Innovation (CERSI; U01FD004319), a collaborative effort between the University and the U.S. Food and Drug Administration to promote regulatory science through innovative research and education. This research does not necessarily reflect the views of the FDA. Scholarship to B.S. was provided by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and to K.R. by the Susan G. Komen Post Baccalaureate Training in Breast Cancer Health (PBTDR12228366). The Histopathology and Tissue

Shared Resource (HTSR) and Genomics and Epigenomics Shared Resource of the Lombardi Comprehensive Cancer Center (LCCC), partially supported by the NIH/NCI grant P30-CA051008, were utilized in this study.

REFERENCES

1. Foulkes WD, Smith IE, Reis-Filho JS. Triple-negative breast cancer. N Engl J Med. 2010; 363:1938–1948.

2. Pal SK, Childs BH, Pegram M. Triple negative breast cancer: unmet medical needs.Breast Cancer Res Treat. 2011; 125:627–636.

3. Crown J, O'Shaughnessy J, Gullo G. Emerging targeted therapies in triple-negative breast cancer. Ann. Oncol. 2012; 23:vi56–65.

4. Kohler BA, Sherman RL, Howlader N, Jemal A, Ryerson AB, Henry KA, Boscoe FP, Cronin KA, Lake A, Noone AM, Henley SJ, Eheman CR, Anderson RN, et al. Annual Report to the Nation on the Status of Cancer, 1975–2011, Featuring Incidence of Breast Cancer Subtypes by Race/Ethnicity, Poverty, and State. J Natl Cancer Inst. 2015; 107:djv048.

5. Tao L, Gomez SL, Keegan TH, Kurian AW, Clarke CA. Breast Cancer Mortality in African-American and Non- Hispanic White Women by Molecular Subtype and Stage at Diagnosis: A Population-Based Study. Cancer Epidemiol Biomarkers Prev. 2015; 24:1039–1045.

6. Loo LW, Wang Y, Flynn EM, Lund MJ, Bowles EJ, Buist DS, Liff JM, Flagg EW, Coates RJ, Eley JW, Hsu L, Porter PL. Genome-wide copy number alterations in subtypes of invasive breast cancers in young white and African American women. Breast Cancer Res Treat. 2011; 127:297–308.

7. Grunda JM, Steg AD, He Q, Steciuk MR, Byan-Parker S, Johnson MR, Grizzle WE. Differential expression of breast cancer-associated genes between stage- and age-matched tumor specimens from African- and Caucasian-American Women diagnosed with breast cancer. BMC Res Notes. 2012; 5:248.

8. Sturtz LA, Melley J, Mamula K, Shriver CD, Ellsworth RE. Outcome disparities in African American women with triple negative breast cancer: a comparison of epidemiological and molecular factors between African American and Caucasian women with triple negative breast cancer. BMC cancer. 2014; 14:62.

9. Dietze EC, Sistrunk C, Miranda-Carboni G, O'Regan R, Seewaldt VL. Triple-negative breast cancer in African- American women: disparities versus biology. Nat Rev Cancer. 2015; 15:248–254.

10. Takahashi RU, Miyazaki H, Ochiya T. The Roles of MicroRNAs in Breast Cancer. Cancers. 2015; 7:598–616.

11. Yang F, Zhang W, Shen Y, Guan X. Identification of dysregulated microRNAs in triple-negative breast cancer (review). Int J Oncol. 2015; 46:927–932.

12. Riaz M, van Jaarsveld MT, Hollestelle A, Prager-van der Smissen WJ, Heine AA, Boersma AW, Liu J, Helmijr J, Ozturk B, Smid M, Wiemer EA, Foekens JA, Martens JW. miRNA expression profiling of 51 human breast cancer cell lines reveals subtype and driver mutation-specific miRNAs. Breast Cancer Res. 2013; 15:R33.

13. Rawlings-Goss RA, Campbell MC, Tishkoff SA. Global population-specific variation in miRNA associated with cancer risk and clinical biomarkers. BMC Med Genomics. 2014; 7:53.

14. Pelletier C, Weidhaas JB. MicroRNA binding site polymorphisms as biomarkers of cancer risk. Expert Rev Mol Diagn. 2010; 10:817–829.

15. Huang RS, Gamazon ER, Ziliak D, Wen Y, Im HK, Zhang W, Wing C, Duan S, Bleibel WK, Cox NJ, Dolan ME. Population differences in microRNA expression and biological implications. RNA Biol. 2011; 8:692–701.

16. Bensen JT, Tse CK, Nyante SJ, Barnholtz-Sloan JS, Cole SR, Millikan RC. Association of germline microRNA SNPs in pre-miRNA flanking region and breast cancer risk and survival: the Carolina Breast Cancer Study. Cancer Causes Control. 2013; 24:1099–1109.

17. Yao S, Graham K, Shen J, Campbell LE, Singh P, Zirpoli G, Roberts M, Ciupak G, Davis W, Hwang H, Khoury T, Bovbjerg DH, Jandorf L, et al. Genetic variants in microRNAs and breast cancer risk in African American and European American women. Breast Cancer Res Treat. 2013; 141:447–459.

18. Toyama T, Kondo N, Endo Y, Sugiura H, Yoshimoto N, Iwasa M, Takahashi S, Fujii Y, Yamashita H. High expression of microRNA-210 is an independent factor indicating a poor prognosis in Japanese triple-negative breast cancer patients. Jpn J Clin Oncol. 2012; 42:256–263.

19. Khan S, Greco D, Michailidou K, Milne RL, Muranen TA, Heikkinen T, Aaltonen K, Dennis J, Bolla MK, Liu J, Hall P, Irwanto A, Humphreys K, et al. MicroRNA related polymorphisms and breast cancer risk. PloS one. 2014; 9:e109973.

20. Comprehensive molecular portraits of human breast tumours. Nature. 2012; 490:61–70.

21. Kristensen VN, Lingjaerde OC, Russnes HG, Vollan HK, Frigessi A, Borresen-Dale AL. Principles and methods of integrative genomic analyses in cancer. Nat Rev Cancer. 2014; 14:299–313.

22. Calin GA, Sevignani C, Dumitru CD, Hyslop T, Noch E, Yendamuri S, Shimizu M, Rattan S, Bullrich F, Negrini M, Croce CM. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc Natl Acad Sci USA. 2004; 101:2999–3004.

23. Zhang L, Huang J, Yang N, Greshock J, Megraw MS, Giannakakis A, Liang S, Naylor TL, Barchetti A, Ward MR, Yao G, Medina A, O'Brien-Jenkins A, et al. microRNAs exhibit high frequency genomic alterations in human cancer. Proc Natl Acad Sci USA. 2006; 103:9136–9141.

24. Selcuklu SD, Yakicier MC, Erson AE. An investigation of microRNAs mapping to breast cancer related genomic gain and loss regions. Cancer Genet Cytogenet. 2009; 189:15–23.

25. de Rinaldis E, Gazinska P, Mera A, Modrusan Z, Fedorowicz GM, Burford B, Gillett C, Marra P, Grigoriadis A, Dornan D, Holmberg L, Pinder S, Tutt A. Integrated genomic analysis of triple-negative breast cancers reveals novel microRNAs associated with clinical and molecular phenotypes and sheds light on the pathways they control. BMC Genomics. 2013; 14:643.

26. Cava C, Bertoli G, Ripamonti M, Mauri G, Zoppis I, Della Rosa PA, Gilardi MC, Castiglioni I. Integration of mRNA expression profile, copy number alterations, and microRNA expression levels in breast cancer to improve grade definition. PloS one. 2014; 9:e97681.

27. Li D, Xia H, Li ZY, Hua L, Li L. Identification of Novel Breast Cancer Subtype-Specific Biomarkers by Integrating Genomics Analysis of DNA Copy Number Aberrations and miRNA-mRNA Dual Expression Profiling. BioMed Res Int. 2015; 2015:746970.

28. Iorio MV, Croce CM. Causes and consequences of microRNA dysregulation. Cancer J. 2012; 18:215–222.

29. Shah SP, Roth A, Goya R, Oloumi A, Ha G, Zhao Y, Turashvili G, Ding J, Tse K, Haffari G, Bashashati A, Prentice LM, Khattra J, et al. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. Nature. 2012; 486:395–399.

30. Korpal M, Lee ES, Hu G, Kang Y. The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2. J Biol Chem. 2008; 283:14910–14914.

31. Chen J, Tian W, Cai H, He H, Deng Y. Down-regulation of microRNA-200c is associated with drug resistance in human breast cancer. Med Oncol. 2012; 29:2527–2534.

32. Humphries B, Yang C. The microRNA-200 family: small molecules with novel roles in cancer development, progression and therapy. Oncotarget. 2015; 6:6472–6498.

33. Vosgha H, Salajegheh A, Smith RA, Lam AK. The important roles of miR-205 in normal physiology, cancers and as a potential therapeutic target. Curr Cancer Drug Tar. 2014; 14:621–637.

34. The International HapMap Project. Nature. 2003; 426:789–796.

35. Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, Abecasis GR. A global reference for human genetic variation. Nature. 2015; 526:68–74.

36. Qin AY, Zhang XW, Liu L, Yu JP, Li H, Wang SZ, Ren XB, Cao S. MiR-205 in cancer: an angel or a devil? Eur J Cell Biol. 2013; 92:54–60.

37. Su CM, Lee WH, Wu AT, Lin YK, Wang LS, Wu CH, Yeh CT. Pterostilbene inhibits triple-negative breast cancer metastasis via inducing microRNA-205 expression and negatively modulates epithelial-to-mesenchymal transition. J Nutr Biochem. 2015; 26:675–685.

38. Piovan C, Palmieri D, Di Leva G, Braccioli L, Casalini P, Nuovo G, Tortoreto M, Sasso M, Plantamura I, Triulzi T, Taccioli C, Tagliabue E, Iorio MV, et al. Oncosuppressive role of p53-induced miR-205 in triple negative breast cancer. Mol Oncol. 2012; 6:458–472.

39. Zhang H, Li B, Zhao H, Chang J. The expression and clinical significance of serum miR-205 for breast cancer and its role in detection of human cancers. Int J Clin Exp Med. 2015; 8:3034–3043.

40. Halvorsen AR, Bjaanaes M, LeBlanc M, Holm AM, Bolstad N, Rubio L, Penalver JC, Cervera J, Mojarrieta JC, Lopez-Guerrero JA, Brustugun OT, Helland A. A unique set of 6 circulating microRNAs for early detection of non-small cell lung cancer. Oncotarget. 2016.

41. Vetter G, Saumet A, Moes M, Vallar L, Le Bechec A, Laurini C, Sabbah M, Arar K, Theillet C, Lecellier CH, Friederich E. miR-661 expression in SNAI1-induced epithelial to mesenchymal transition contributes to breast cancer cell invasion by targeting Nectin-1 and StarD10 messengers. Oncogene. 2010; 29:4436–4448.

42. Hoffman Y, Bublik DR, Pilpel Y, Oren M. miR-661 downregulates both Mdm2 and Mdm4 to activate p53. Cell Death Differ. 2014; 21:302–309.

43. Toffoli S, Bar I, Abdel-Sater F, Delree P, Hilbert P, Cavallin F, Moreau F, Van Criekinge W, Lacroix-Triki M, Campone M, Martin AL, Roche H, Machiels JP, et al. Identification by array comparative genomic hybridization of a new amplicon on chromosome 17q highly recurrent in BRCA1 mutated triple negative breast cancer. Breast Cancer Res. 2014; 16:466.

44. Churpek JE, Walsh T, Zheng Y, Moton Z, Thornton AM, Lee MK, Casadei S, Watts A, Neistadt B, Churpek MM, Huo D, Zvosec C, Liu F, et al. Inherited predisposition to breast cancer among African American women. Breast Cancer Res Treat. 2015; 149:31–39.

45. Zhang J, Fackenthal JD, Zheng Y, Huo D, Hou N, Niu Q, Zvosec C, Ogundiran TO, Hennis AJ, Leske MC, Nemesure B, Wu SY, Olopade OI. Recurrent BRCA1 and BRCA2 mutations in breast cancer patients of African ancestry. Breast Cancer Res Treat. 2012; 134:889–894.

46. Aguda BD, Kim Y, Piper-Hunter MG, Friedman A, Marsh CB. MicroRNA regulation of a cancer network: consequences of the feedback loops involving miR-17-92, E2F, and Myc. Proc Natl Acad Sci USA. 2008; 105:19678–19683.

47. Petrocca F, Vecchione A, Croce CM. Emerging role of miR- 106b-25/miR-17-92 clusters in the control of transforming growth factor beta signaling. Cancer Res. 2008; 68:8191–8194.

48. Fu X, Tian J, Zhang L, Chen Y, Hao Q. Involvement of microRNA-93, a new regulator of PTEN/Akt signaling pathway, in regulation of chemotherapeutic drug cisplatin chemosensitivity in ovarian cancer cells. FEBS Lett. 2012; 586:1279–1286.

49. Jiang L, Wang C, Lei F, Zhang L, Zhang X, Liu A, Wu G, Zhu J, Song L. miR-93 promotes cell proliferation in gliomas through activation of PI3K/Akt signaling pathway. Oncotarget. 2015; 6:8286–8299.

50. Ohta K, Hoshino H, Wang J, Ono S, Iida Y, Hata K, Huang SK, Colquhoun S, Hoon DS. MicroRNA-93 activates c-Met/PI3K/Akt pathway activity in hepatocellular carcinoma by directly inhibiting PTEN and CDKN1A. Oncotarget. 2015; 6:3211–3224.

51. Ren J, Jin F, Yu Z, Zhao L, Wang L, Bai X, Zhao H, Yao W, Mi X, Wang E, Olopade OI, Wei M. MYC overexpression and poor prognosis in sporadic breast cancer with BRCA1 deficiency. Tumour Biol. 2013; 34:3945–3958.

52. Hondermarck H. Neurotrophins and their receptors in breast cancer. Cytokine Growth Factor Rev. 2012; 23:357–365.

53. Momeny M, Saunus JM, Marturana F, McCart Reed AE, Black D, Sala G, Iacobelli S, Holland JD, Yu D, Da Silva L, Simpson PT, Khanna KK, Chenevix-Trench G, et al. Heregulin-HER3-HER2 signaling promotes matrix metalloproteinase-dependent blood-brain-barrier transendothelial migration of human breast cancer cell lines. Oncotarget. 2015; 6:3932–3946.

54. De la Cruz-Morcillo MA, Berger J, Sanchez-Prieto R, Saada S, Naves T, Guillaudeau A, Perraud A, Sindou P, Lacroix A, Descazeaud A, Lalloue F, Jauberteau MO. p75 neurotrophin receptor and pro-BDNF promote cell survival and migration in clear cell renal cell carcinoma. Oncotarget. 2016.

55. Cossu-Rocca P, Orru S, Muroni MR, Sanges F, Sotgiu G, Ena S, Pira G, Murgia L, Manca A, Uras MG, Sarobba MG, Urru S, De Miglio MR. Analysis of PIK3CA Mutations and Activation Pathways in Triple Negative Breast Cancer. PloS one. 2015; 10:e0141763.

56. Stewart PA, Luks J, Roycik MD, Sang QX, Zhang J. Differentially expressed transcripts and dysregulated signaling pathways and networks in African American breast cancer. PloS one. 2013; 8:e82460.

57. Keenan T, Moy B, Mroz EA, Ross K, Niemierko A, Rocco JW, Isakoff S, Ellisen LW, Bardia A. Comparison of the Genomic Landscape Between Primary Breast Cancer in African American Versus White Women and the Association of Racial Differences With Tumor Recurrence. J Clin Oncol. 2015; 33:3621–3627.

58. Haddad SA, Lunetta KL, Ruiz-Narvaez EA, Bensen JT, Hong CC, Sucheston-Campbell LE, Yao S, Bandera EV, Rosenberg L, Haiman CA, Troester MA, Ambrosone CB, Palmer JR. Hormone-related pathways and risk of breast cancer subtypes in African American women. Breast Cancer Res Treat. 2015; 154:145–154.

59. Oliveira MC, Pereira SR, Ribeiro EMF, Lima RS, Urban CA, Sebastião APM, Marian C, Haddad BR, Cavalli IJ, Cavalli LR. DNA copy number changes in the PI3K signaling pathway genes in triple negative breast cancer. Cancer Res. 2011; 71.

60. Jung HJ, Suh Y. Regulation of IGF -1 signaling by microRNAs. Frontiers in genetics. 2014; 5:472.

61. Zheng H, Zhang F, Lin X, Huang C, Zhang Y, Li Y, Lin J, Chen W. MicroRNA-1225-5p inhibits proliferation and metastasis of gastric carcinoma through repressing insulin receptor substrate-1 and activation of beta-catenin signaling. Oncotarget. 2016; 7:4647–4663.

62. Christopoulos PF, Msaouel P, Koutsilieris M. The role of the insulin-like growth factor-1 system in breast cancer. Mol Cancer. 2015; 14:43.

63. Fowke JH, Matthews CE, Yu H, Cai Q, Cohen S, Buchowski MS, Zheng W, Blot WJ. Racial differences in the association between body mass index and serum IGF1, IGF2, and IGFBP3. Endocr Relat Cancer. 2010; 17:51–60.

64. Lindner R, Sullivan C, Offor O, Lezon-Geyda K, Halligan K, Fischbach N, Shah M, Bossuyt V, Schulz V, Tuck DP, Harris LN. Molecular phenotypes in triple negative breast cancer from African American patients suggest targets for therapy. PloS one. 2013; 8:e71915.

65. Mu L, Tuck D, Katsaros D, Lu L, Schulz V, Perincheri S, Menato G, Scarampi L, Harris L, Yu H. Favorable outcome associated with an IGF-1 ligand signature in breast cancer. Breast Cancer Res Treat. 2012; 133:321–331.

66. Arcaro A. Targeting the insulin-like growth factor-1 receptor in human cancer. Front Pharmacol. 2013; 4:30.

67. Ashing K, Rosales M, Lai L, Hurria A. Occurrence of comorbidities among African-American and Latina breast cancer survivors. J Cancer Surviv. 2014; 8:312–318.

68. Hammond ME, Hayes DF, Dowsett M, Allred DC, Hagerty KL, Badve S, Fitzgibbons PL, Francis G, Goldstein NS, Hayes M, Hicks DG, Lester S, Love R, et al. American Society of Clinical Oncology/College Of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. J Clin Oncol. 2010; 28:2784–2795.

69. Wolff AC, Hammond ME, Hicks DG, Dowsett M, McShane LM, Allison KH, Allred DC, Bartlett JM, Bilous M, Fitzgibbons P, Hanna W, Jenkins RB, Mangu PB, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. Arch Pathol Lab Med. 2014; 138:241–256.

70. Torresan C, Oliveira MM, Pereira SR, Ribeiro EM, Marian C, Gusev Y, Lima RS, Urban CA, Berg PE, Haddad BR, Cavalli IJ, Cavalli LR. Increased copy number of the DLX4 homeobox gene in breast axillary lymph node metastasis. Cancer Genet. 2014; 207:177–187.

71. Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012; 491:56–65.

72. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience. 2015; 4:7.

73. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. PLoS genetics. 2006; 2:e190.

74. Vlachos IS, Kostoulas N, Vergoulis T, Georgakilas G, Reczko M, Maragkakis M, Paraskevopoulou MD, Prionidis K, Dalamagas T, Hatzigeorgiou AG. DIANA miRPath v.2.0: investigating the combinatorial effect of microRNAs in pathways. Nucleic Acids Res. 2012; 40:W498–504.

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

FIGURE 1 - Workflow of miRNA expression and copy number profiling and downstream comprehensive computational analysis performed in the TNBC and non-TNBC cases of AA and NHW group of patients

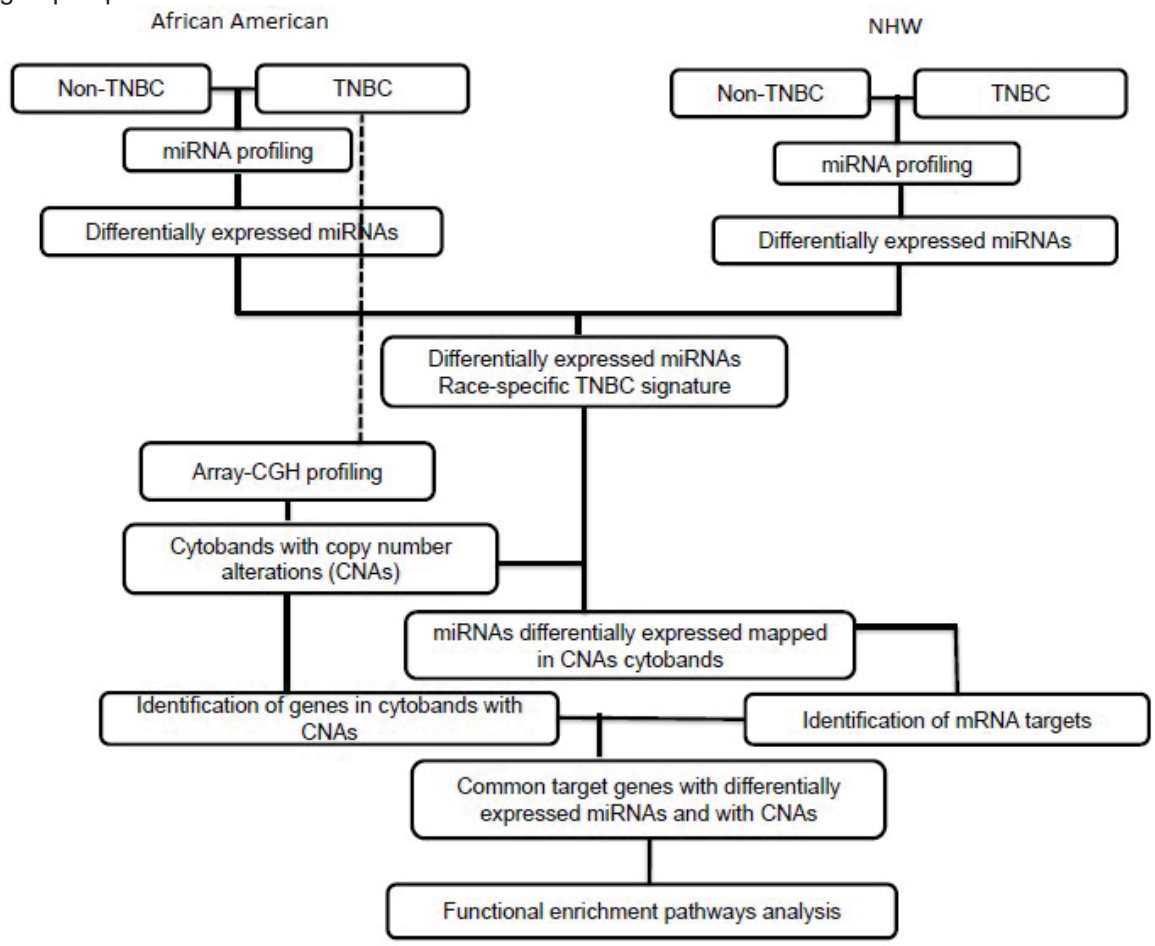


FIGURE 2 - Unsupervised (A) and Supervised (B) Hierarchical Clustering analysis to the TNBC (green bars) and non-TNBC (yellow bars) cases of the AA and NHW group of patients (left and right panel respectively). Up-regulated miRNAs (yellow) and down-regulated miRNAs (blue). (MeV 4.9, Pearson Correlation, P<0.01, FDR <0.05)

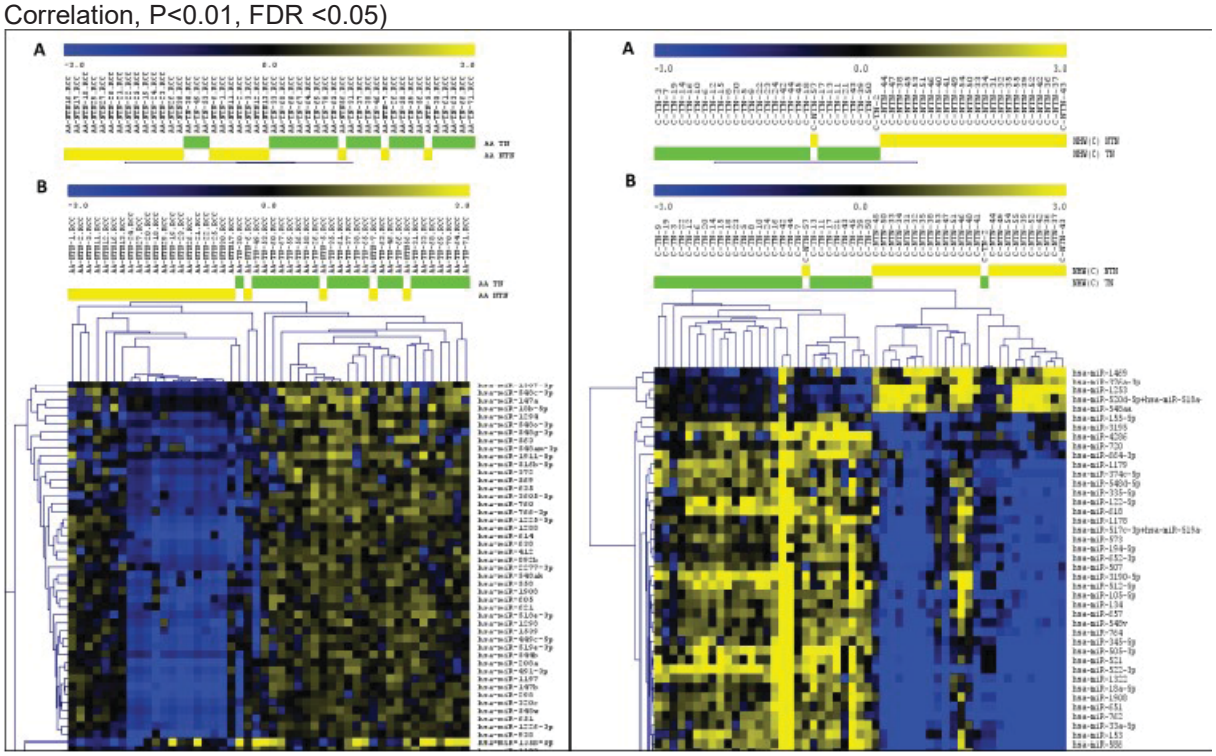


TABLE 1 - Top 15 miRNAs (based on log2FC) observed up- and down-regulated in the TNBC subtype of the AA and NHW group of patients

MiRNAs up-regulated	Log2 FC	*P*-value	FDR	MiRNAs down-regulated	Log2 FC	*P*-value	FDR
hsa-miR-9-5p	3.138	8.61E-06	6.75E-05	hsa-miR-1253	−5.063	5.99E-10	5.99E-08
hsa-miR-127-3p	3.065	1.33E-07	2.87E-06	hsa-miR-1283	−4.601	9.57E-06	7.36E-05
hsa-miR-451a	2.993	1.56E-05	1.08E-04	hsa-miR-378e	−4.241	6.90E-06	5.75E-05
hsa-miR-205-5p	2.925	7.82E-05	3.86E-04	hsa-miR-549	−4.136	3.45E-08	1.25E-06
hsa-miR-548p	2.850	4.07E-08	1.36E-06	hsa-miR-1268b	−3.082	3.41E-13	2.72E-10
hsa-miR-4508	2.838	8.25E-08	2.06E-06	hsa-miR-1265	−2.781	3.80E-11	1.01E-08
hsa-miR-4425	2.829	2.36E-08	9.45E-07	hsa-miR-433	−2.750	2.22E-12	8.88E-10
hsa-miR-150-5p	2.730	6.20E-04	0.002119733	hsa-miR-1305	−2.748	1.30E-08	5.80E-07
hsa-miR-374a-5p	2.682	2.03E-05	1.33E-04	hsa-miR-518f-3p	−2.746	4.11E-11	8.23E-09
hsa-miR-455-3p	2.507	2.51E-09	2.01E-07	hsa-miR-649	−2.731	2.93E-10	3.35E-08
hsa-miR-424-5p	2.499	3.97E-07	5.78E-06	hsa-miR-520e	−2.689	4.79E-11	7.67E-09
hsa-miR-423-5p	2.467	3.15E-07	5.25E-06	hsa-miR-206	−2.671	3.38E-08	1.29E-06
hsa-miR-200c-3p	2.465	6.10E-04	0.002093624	hsa-miR-520d-5p, 518a-5p, 527	−2.647	1.46E-06	1.60E-05
hsa-miR-606	2.453	3.54E-08	1.23E-06				
hsa-miR-26a-5p	2.449	7.54E-05	3.84E-04				

FIGURE 3 - Unsupervised (A) and Supervised (B) Hierarchical Clustering analysis to the AA (green bars) and NHW (blue bars) group of patients. Up-regulated miRNAs (yellow) and down-regulated miRNAs (blue). (MeV 4.9, Pearson Correlation, P<0.01, FDR <0.05)

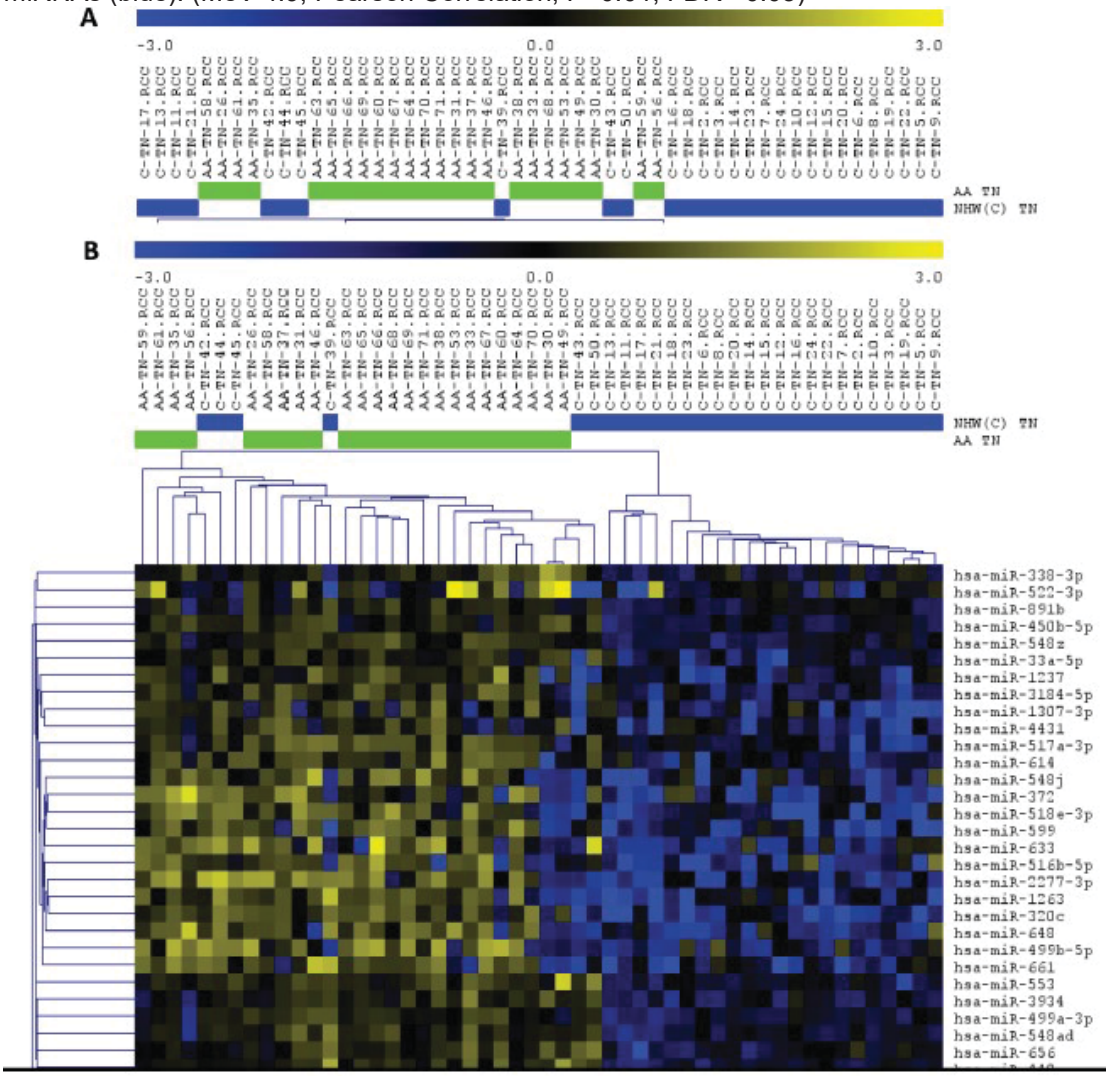


FIGURE 4 - Penetrance plot of the array-CGH profiling of the TNBC cases from the AA patients analyzed, showing the corresponding genome location (arrows) of the 26 miRNAs of the identified panel. Vertical lines represent each chromosome number. Red peaks indicate copy number gains and green peaks indicate copy number losses. MiRNAs with up- and down-regulated expression levels are annotated in red and green color boxes, respectively.

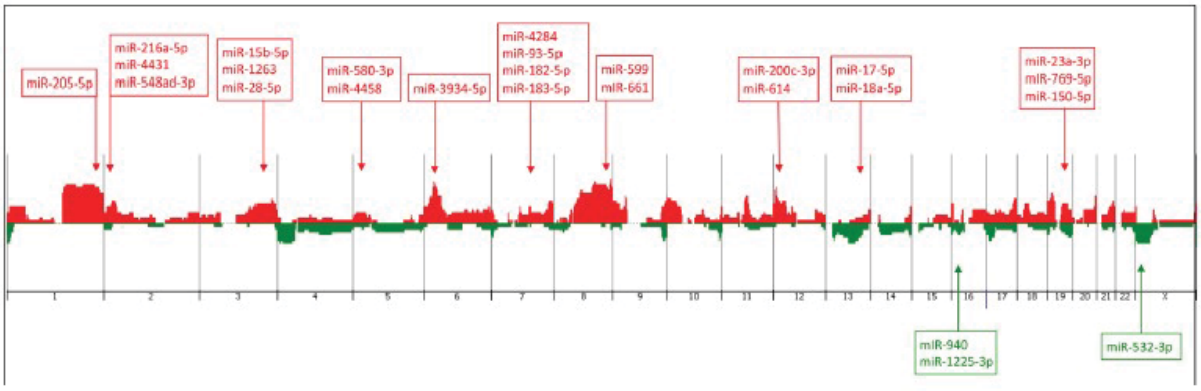


TABLE 2 - Twenty-six miRNAs differentially expressed between the TNBC of the AA and NHW patients, with expression levels in concordance with copy number alterations (CNAs) (presented by chromosome numerical order).

MiRNAs	Cytoband	Start	Stop	CNAs	miRNA expression	Log2FC	*P* value
hsa-miR-205-5p	1q32.2	209432133	209432242	gain	up-regulated	2.925	7.82E-05
hsa-miR-216a-5p	2p16.1	55988950	55989059	gain	up-regulated	1.059	7.65E-07
hsa-miR-4431	2p16.2	52702522	52702615	gain	up-regulated	1.220	2.42E-04
hsa-miR-548ad-3p	2p25.1	35471405	35471486	gain	up-regulated	0.690	8.22E-04
hsa-miR-15b-5p	3q25.33	160404588	160404685	gain	up-regulated	2.202	2.97E-04
hsa-miR-1263	3q26.1	164171471	164171556	gain	up-regulated	1.375	2.69E-04
hsa-miR-28-5p	3q28	188688781	188688866	gain	up-regulated	1.655	2.28E-05
hsa-miR-580-3p	5p13.2	36147892	36147988	gain	up-regulated	1.526	7.47E-07
hsa-miR-4458	5p15.31	8460925	8460999	gain	up-regulated	2.224	4.79E-07
hsa-miR-3934-5p	6p21.31	33698128	33698234	gain	up-regulated	0.754	2.99E-04
hsa-miR-4284	7q11.23	73711317	73711397	gain	up-regulated	1.981	4.65E-06
hsa-miR-93-5p	7q22.1	100093768	100093847	gain	up-regulated	2.230	1.88E-04
hsa-miR-182-5p	7q32.2	129770383	129770492	gain	up-regulated	2.026	7.77E-05
hsa-miR-183-5p	7q32.2	129774905	129775014	gain	up-regulated	1.717	7.54E-05
hsa-miR-599	8q22.2	99536636	99536730	gain	up-regulated	1.755	7.81E-06
hsa-miR-661	8q24.3	143945191	143945279	gain	up-regulated	1.307	7.37E-04
hsa-miR-614	12p13.1	12915829	12915918	gain	up-regulated	1.021	4.31E-04
hsa-miR-200c-3p	12p13.31	6963699	6963766	gain	up-regulated	2.465	6.10E-04
hsa-miR-17-5p	13q31.3	91350605	91350688	gain	up-regulated	1.990	3.21E-04
hsa-miR-18a-5p	13q31.3	91350751	91350821	gain	up-regulated	1.160	3.60E-05
hsa-miR-940	16p13.3	2271747	2271840	loss	down-regulated	−1.118	2.05E-04
hsa-miR-1225-3p	16p13.3	2090195	2090284	loss	down-regulated	−1.953	1.29E-07
hsa-miR-23a-3p	19p13.12	13836587	13836659	gain	up-regulated	2.334	9.01E-05
hsa-miR-769-5p	19q13.32	46018932	46019049	gain	up-regulated	1.355	2.46E-06
hsa-miR-150-5p	19q13.33	49500762	49500873	gain	up-regulated	2.730	6.20E-04
hsa-miR-532-5p	Xp11.23	50003148	50003238	loss	down-regulated	−1.875	9.57E-08

FIGURE 5 - Expression levels of the 26 differentially expressed miRNAs observed between the AA and NHW TNBC cases. In (A) and (B): miRNAs up- and down-regulated, respectively, in the AA group of [patients when compared to the NHW group.

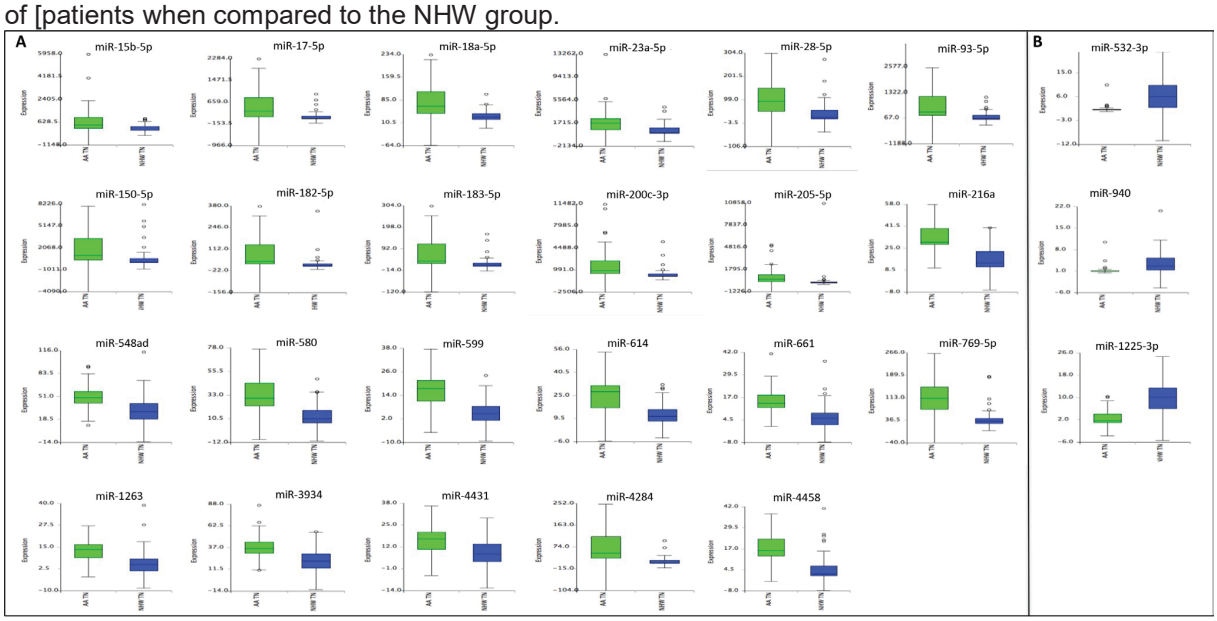


FIGURE 6 - Venn diagrams showing integration of genes located at the identified cytobands (A) and in the most frequent cytobands (present in greater than or equal to 50% f the cases) (B) with CNAs in the TNBC in the TNBC-AA cases and the corresponding miRNA target genes.

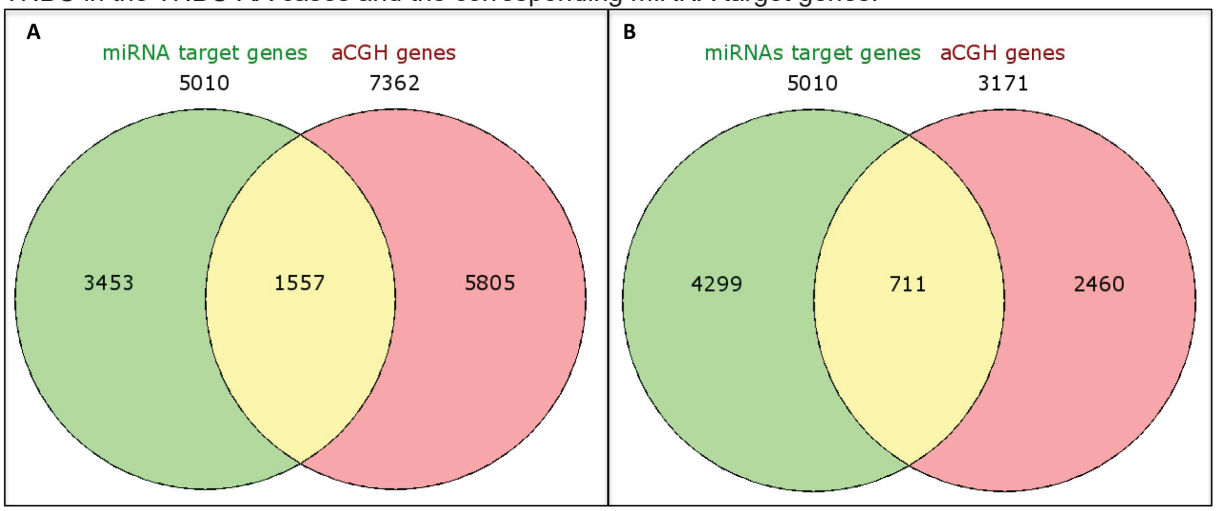


FIGURE 7 - ROC plots of the individual and combined 26 miRNAs differentially expressed between the AA and NHW-TNBC group of patients.

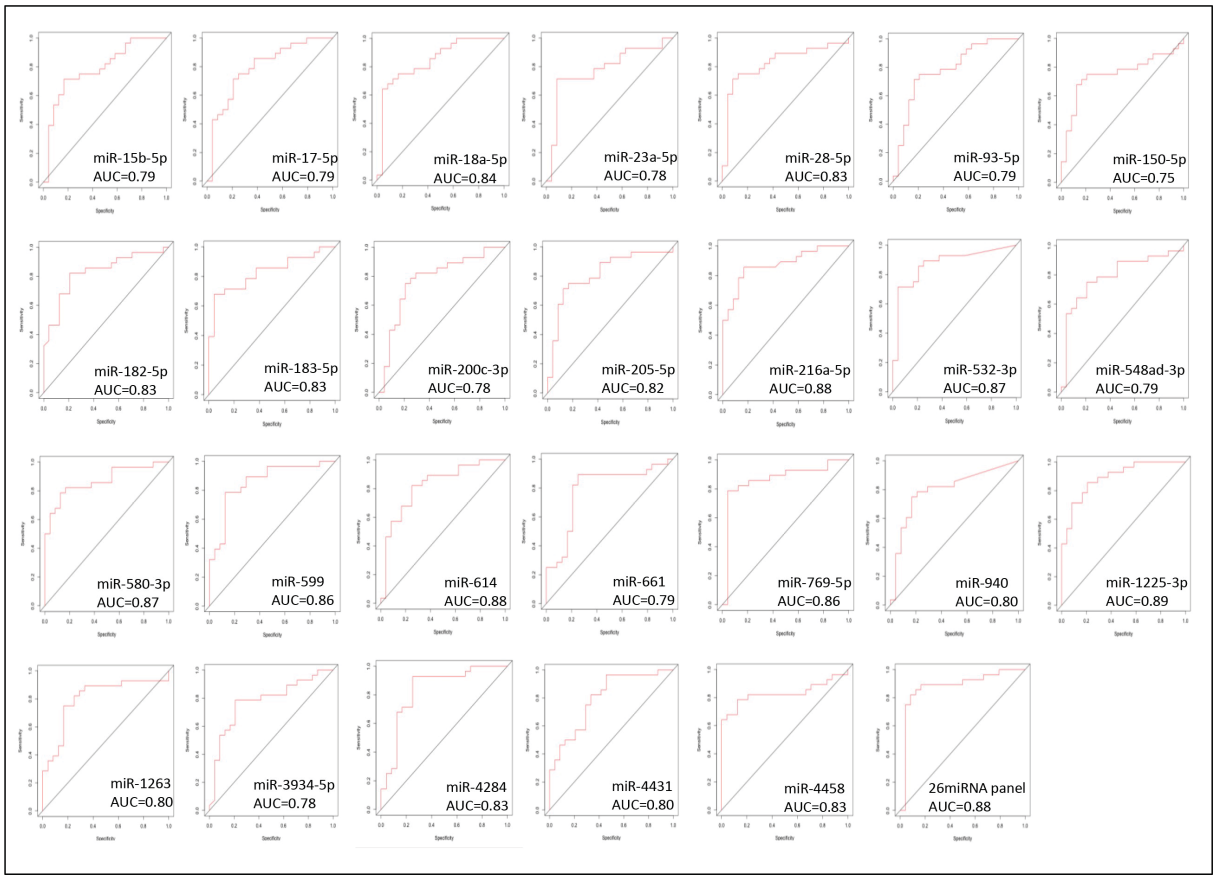


TABLE 3 - Top 15 pathways (based on P value) mostly affected by the 26 miRNAs differentially expressed in the TNBC of AA and NHW groups of patients (DIANA miRPath v.2.0)

#	KEGG pathway	*P*-value	# Genes	# MiRNAs
1	Neurotrophin signaling pathway	7.39E-32	68	23
2	MAPK signaling pathway	1.61E-27	118	24
3	Insulin signaling pathway	1.15E-25	67	23
4	PI3K-AKT signaling pathway	3.75E-25	144	23
5	Pathways in cancer	3.75E-25	150	24
6	ERBB2 signaling pathway	1.06E-24	47	20
7	TGF-beta signaling pathway	2.75E-24	45	20
8	Focal adhesion	1.22E-23	92	23
9	Prostate cancer	3.45E-23	48	21
10	GNRH signaling pathway	5.42E-23	48	22
11	Long-term depression	1.37E-22	39	17
12	Endocytosis	1.25E-21	92	23
13	Chronic myeloid leukemia	1.09E-20	41	21
14	Renal cell carcinoma	3.06E-19	40	21
15	Axon guidance	3.26E-19	65	24

KEGG = Kyoto Encyclopedia of Genes and Genomes

FIGURE 8 - Ingenuity Pathway Analysis (IPA) showing the main gene network interaction of 12 out of the 26 miRNA panel identified.

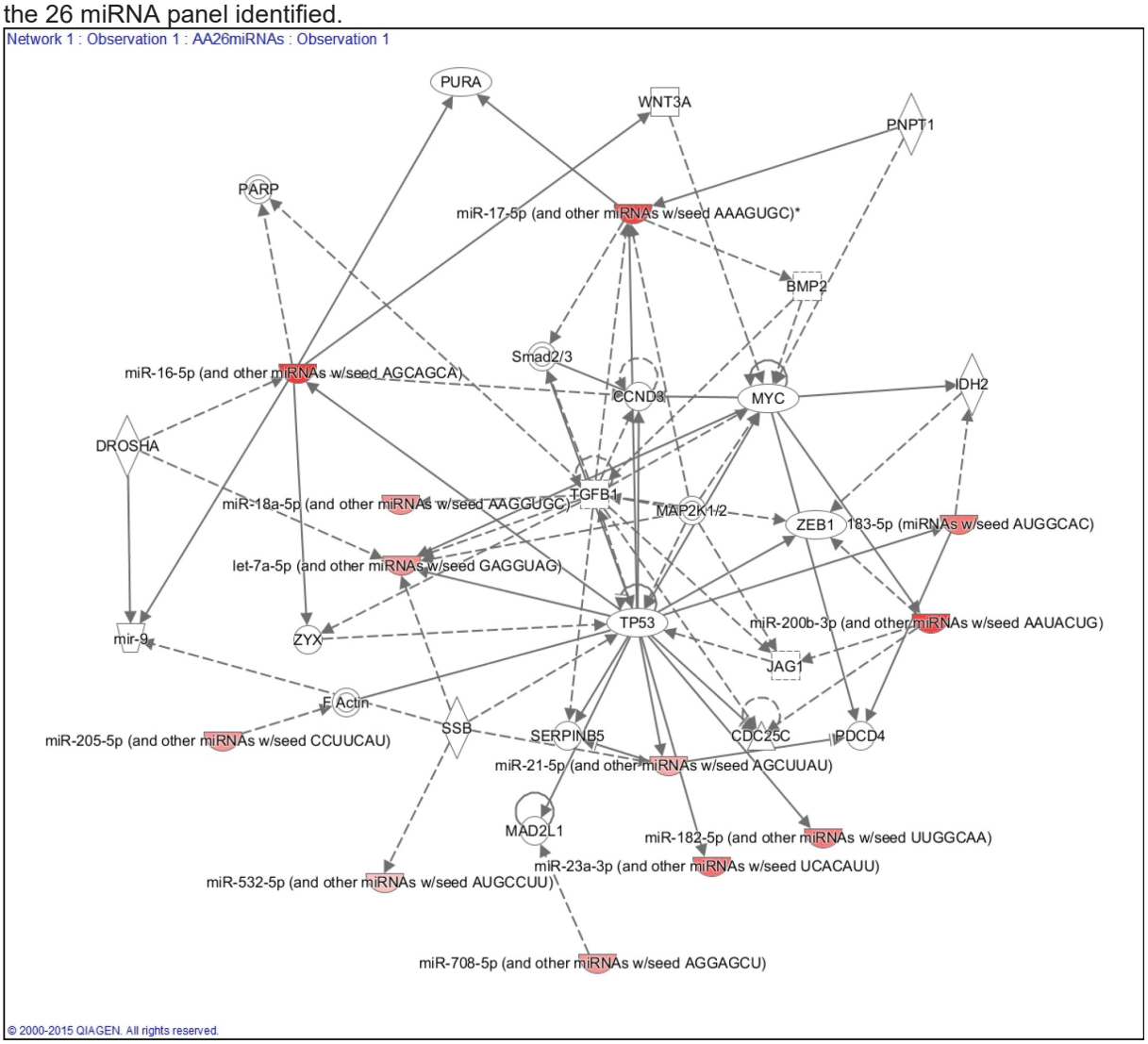
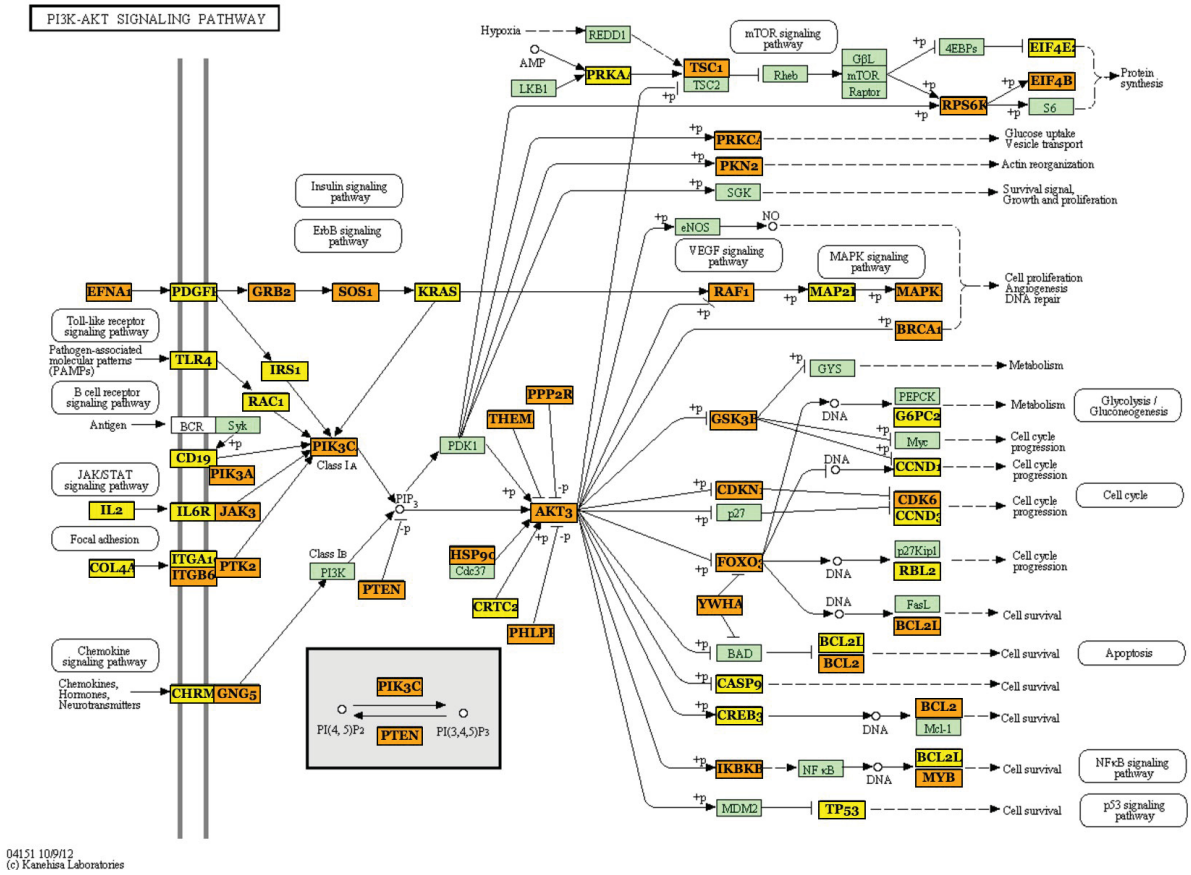


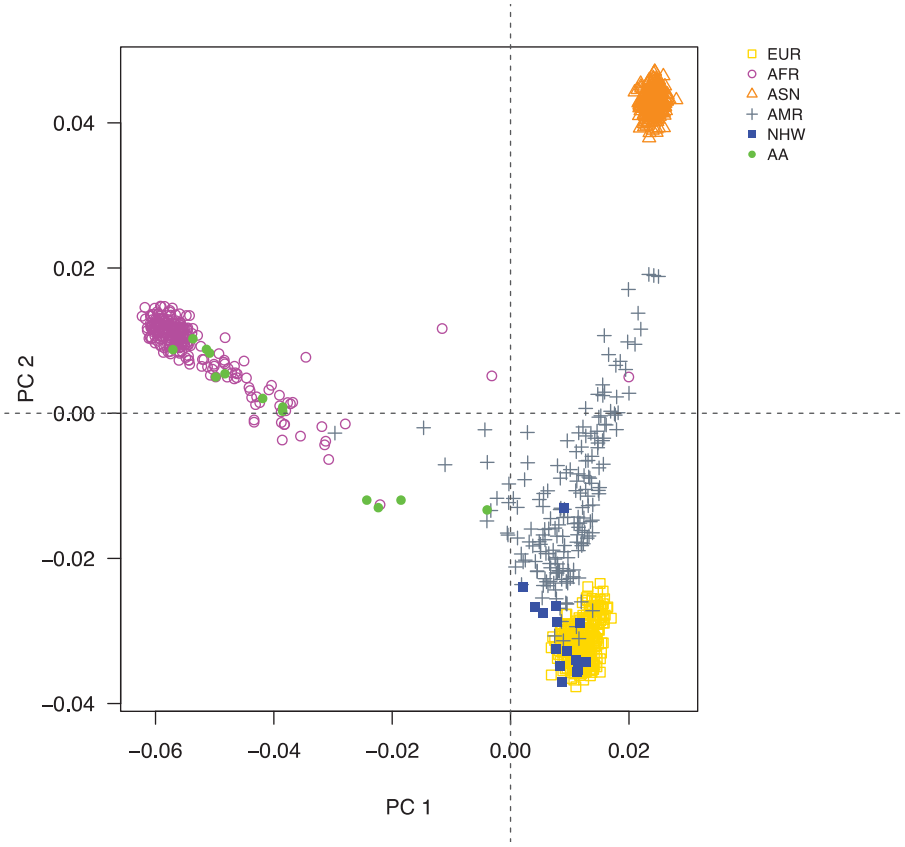
TABLE 4 - Analysis of clinical-pathological parameters of the TNBC and non-TNBC cases in the AA and NHW groups of patients

	AA			NHW			AA and NHW	
	TNBC	**Non-TNBC**	***P* value**	**TNBC**	**Non-TNBC**	***P* value**	**TNBC**	**Non-TNBC**
Mean age (yrs)	50.74 ± 1.549, *n* = 27	57.41 ± 2.785, *n* = 27	*P* = 0.0413*	53.70 ± 2.089, *n* = 30	53.93 ± 2.498, *n* = 30	*P* = 0.4358	*P* = 0.2687	*P* = 0.3557
Tumor size (cm)	2.819 ± 0.3821, *n* = 26	2.800 ± 0.4553, *n* = 27	*P* > 0.9999	2.964 ± 0.3074, *n* = 28	1.988 ± 0.2525, *n* = 26	*P* = 0.0185*	*P* = 0.7671	*P* = 0.1291
Tumor stage I	0	11.5% (3/26)	*P* = 0.0014**	0	27.3% (6/22)	*P* = 0.0002***	*P* = 1.0	*P* = 0.3765
Tumor stage II	11.1% (3/27)	46.2% (12/26)		10.3% (3/29)	36.35% (8/22)			
Tumor stage III	88.9% (24/27)	42.3% (11/26)		89.7% (26/29)	36.35% (8/22)			
Tumor grade 1	0	11.5% (3/26)	*P* = 0.0085**	0	27.3% (6/22)	*P* < 0.0001****	*P* = 0.6633	
Tumor grade 2	11.5% (3/26)	38.5% (10/26)		7.1% (2/28)	36.35% (8/22)			
Tumor grade 3	88.5% (23/26)	50% (13/26)		92.9% (26/28)	36.35% (8/22)			
LN- positive	50% (13/26)	63.2% (12/19)	*P* = 0.5446	38.5% (10/26)	56.3% (9/16)	*P* = 0.3437	*P* = 0.5771	*P* = 0.7391
LN- negative	50% (13/26)	36.8% (7/19)		61.5% (16/26)	43.7% (7/16)			
BC rec-positive	33.3% (9/27)	11.1% (3/27)	*P* = 0.0994	10.3% (3/29)	3.8% (1/26)	*P* = 0.6131	*P* = 0.0519*	*P* = 0.604
BC rec-negative	66.7% (18/27)	88.9% (24/27)		89.7% (26/29)	96.2% (25/26)			
DM-positive	25.9% (7/27)	19.2% (5/26)	*P* = 0.7445	17.2% (5/29)	19.2% (5/26)	*P* = 1.0	*P* = 0.5225	*P* = 1.0
DM-negative	74.1% (20/27)	80.8% (21/26)		82.8% (24/29)	80.8% (21/26)			

Supplementary FIGURE S1 - PIK3-AKT and other signaling related pathways affected by the 26 miRNAs of our panel (DIANA miRPath v.2.0). White boxes indicate gene entries in non-metabolic pathways or enzyme and reaction entries in metabolic pathways (Reference Pathway). Green boxes indicate presence of the gene in the genome (Organism-specific pathway). Yellow boxes indicate genes targeted by one miRNA. Orange boxes indicate genes targeted by two or more miRNAs.



Supplementary FIGURE S2 - Principal component analysis (PCA) showing that a subset genotype dataset of African American (AA) and Non-Hispanic White (NHW) patients of this study, distinctly clustered with the European (EUR) and African (AFR) main ethnic groups from the 1000 Genome Project (phase 1). **ASN: East Asian, AMR: Latin American**.



Supplementary TABLE S1 - Top 15 miRNAs (based on log2FC) observed up- and down-regulated in the TNBC subtype in comparison to the non-TNBC subtype of the AA group of patients

miRNA up-regulated	Log2 FC	P-value	FDR	miRNAs down-regulated	Log2 FC	P-value	FDR
hsa-miR-384	2.559	4.05E-05	7.20E-04	hsa-miR-497-5p	-4.079	1.89E-08	1.51E-05
hsa-miR-135b-5p	2.360	9.31E-04	0.00443	hsa-miR-145-5p	-3.251	1.28E-06	1.28E-04
hsa-miR-1277-3p	2.358	5.79E-05	8.91E-04	hsa-miR-29c-3p	-3.105	1.75E-07	7.00E-05
hsa-miR-1185-5p	2.304	7.74E-06	3.10E-04	hsa-miR-143-3p	-3.036	1.47E-06	1.31E-04
hsa-miR-626	2.245	2.08E-06	1.19E-04	hsa-miR-26a-5p	-2.976	3.20E-06	1.70E-04
hsa-miR-219-5p	2.194	7.87E-06	3.00E-04	hsa-miR-1253	-2.879	4.51E-04	0.002713
hsa-miR-576-5p	2.190	1.69E-06	1.23E-04	hsa-miR-199b-5p	-2.825	1.65E-06	1.32E-04
hsa-miR-606	2.164	2.79E-07	7.43E-05	hsa-miR-342-3p	-2.785	4.42E-07	7.07E-05
hsa-miR-499a-3p	2.132	4.86E-05	7.94E-04	hsa-miR-1290	-2.763	7.36E-06	3.10E-04
hsa-miR-548al	2.093	8.52E-06	2.96E-04	hsa-miR-29b-3p	-2.742	8.27E-05	0.00105
hsa-miR-522-3p	2.091	1.72E-06	1.15E-04	hsa-miR-30a-5p	-2.720	1.41E-04	0.00143
hsa-miR-516a-5p	2.029	7.06E-05	9.57E-04	hsa-miR-196a-5p	-2.668	4.62E-05	7.69E-04
hsa-miR-659-3p	2.020	4.01E-05	7.29E-04	hsa-miR-195-5p	-2.631	3.72E-07	7.45E-05
hsa-miR-568	2.016	1.43E-05	4.08E-04	hsa-miR-451a	-2.613	8.64E-05	0.001064
hsa-miR-518c-3p	1.993	1.91E-05	4.50E-04	hsa-miR-30b-5p	-2.555	3.15E-05	6.45E-04

FC= fold change, FDR= false discovery rate

Supplementary TABLE S2 - Top 15 KEGG pathways (based on P value) mostly affected by the differentially expressed miRNAs in the TNBC and non-TNBC subtypes of the AA group of patients (DIANA miRPath v.2.0)

#	KEGG pathway	P-value	#Genes	#miRNAs
1	Pathways in cancer	8.22E-34	190	49
2	PI3K-AKT signaling pathway	1.49E-33	206	49
3	MAPK signaling pathway	1.43E-26	155	46
4	Endocytosis	1.24E-21	118	45
5	Focal adhesion	3.00E-21	127	47
6	Regulation of actin cytoskeleton	1.10E-20	116	45
7	WNT signaling pathway	4.65E-18	98	47
8	Chemokine signaling pathway	1.87E-17	102	47
9	Hepatitis B	3.58E-17	90	45
10	Axon guidance	5.81E-16	82	44
11	Ubiquitin mediated proteolysis	1.09E-14	84	44
12	Insulin signaling pathway	1.09E-14	85	49
13	RNA transport	1.69E-14	82	43
14	Neutrophin signaling pathway	1.69E-14	82	46
15	Measles	1.92E-14	79	43

KEGG= Kyoto Encyclopedia of Genes and Genomes

Supplementary TABLE S3 - Top 15 and 12 miRNAs observed up- and down-regulated, respectively, in the TNBC subtype in comparison to the non-TNBC subtype of the NHW group of patients

miRNA	Log2FC	p-value	FDR	miRNA	Log2FC	p-value	FDR
hsa-miR-522-3p	8.41	1.46E-18	0	hsa-miR-1253	-4.45	2.99E-07	2.78E-06
hsa-miR-451a	7.12	8.63E-12	8.63E-10	hsa-miR-1469	-3.26	1.84E-07	1.82E-06
hsa-miR-150-5p	6.68	2.61E-09	6.14E-08	hsa-miR-548aa	-2.74	1.18E-08	2.01E-07
hsa-miR-126-3p	6.53	3.38E-10	1.18E-08	hsa-miR-520d-5p+hsa-miR-518a-5p+hsa-miR-527	-2.52	1.79E-07	1.81E-06
hsa-miR-223-3p	6.41	3.29E-10	1.20E-08	hsa-miR-1234	-2.25	5.18E-06	2.82E-05
hsa-miR-374a-5p	6.01	1.00E-10	6.17E-09	hsa-miR-877-5p	-1.89	2.01E-05	9.20E-05
hsa-miR-125a-5p	5.98	8.25E-12	9.42E-10	hsa-miR-376a-3p	-1.81	1.33E-06	8.87E-06
hsa-miR-651	5.95	1.79E-13	4.78E-11	hsa-miR-874	-1.50	3.11E-04	9.73E-04
hsa-miR-182-5p	5.75	4.61E-13	9.22E-11	hsa-miR-1224-5p	-1.50	8.42E-04	0.002
hsa-miR-762	5.70	2.31E-12	3.07E-10	hsa-miR-940	-1.49	0.001	0.003
hsa-miR-34a-5p	5.60	1.78E-10	7.50E-09				
hsa-miR-185-5p	5.55	2.84E-14	1.14E-11				
hsa-miR-181a-5p	5.19	8.43E-09	1.57E-07				
hsa-miR-18a-5p	5.16	5.85E-11	4.68E-09				
hsa-miR-191-5p	5.13	2.56E-09	6.20E-08				

FC= fold change, FDR= false discovery rate

Supplementary TABLE S4 - Top 15 KEGG pathways (based on P value) mostly affected by the differentially expressed miRNAs in the TNBC and non-TNBC subtypes of the NHW group of patients (DIANA miRPath v.2.0)

#	KEGG pathway	P-value	#genes	#miRNAs
1	Pathways in cancer	5.02E-43	191	43
2	PI3K-Akt signaling pathway	2.95E-38	181	44
3	MAPK signaling pathway	2.27E-34	159	45
4	HTLV-I infection	5.00E-30	137	45
5	Endocytosis	1.37E-27	119	39
6	Regulation of actin cytoskeleton	1.44E-27	124	39
7	Epstein-Barr virus infection	1.92E-24	109	37
8	Focal adhesion	1.46E-23	110	41
9	Transcriptional misregulation in cancer	2.08E-21	102	43
10	Wnt signaling pathway	2.38E-21	93	41
11	Chemokine signaling pathway	4.49E-21	97	40
12	Axon guidance	1.40E-20	82	38
13	Hepatitis B	2.68E-20	80	38
14	Neurotrophin signaling pathway	1.27E-18	83	41
15	Ubiquitin mediated proteolysis	3.19E-18	82	40

KEGG= Kyoto Encyclopedia of Genes and Genomes

Supplementary TABLE S5: Area Under the Curve (AUC) values of the individual 26 miRNAs differentially expressed between the AA and NHW-TNBC group of patients

MiRNA	AUC (95% CI)
hsa-miR-1225-3p	0.89 (0.81–0.98)
hsa-miR-216a-5p	0.88 (0.78–0.97)
hsa-miR-532-3p	0.87 (0.77–0.97)
hsa-miR-580-3p	0.87 (0.77–0.97)
hsa-miR-599	0.86 (0.75–0.96)
hsa-miR-769-5p	0.86 (0.75–0.98)
hsa-miR-18a-5p	0.84 (0.73–0.95)
hsa-miR-28-5p	0.83 (0.70–0.95)
hsa-miR-182-5p	0.83 (0.71–0.94)
hsa-miR-183-5p	0.83 (0.72–0.94)
hsa-miR-205-5p	0.82 (0.69–0.94)
hsa-miR-4284	0.83 (0.72–0.95)
hsa-miR-4458	0.83 (0.71–0.95)
hsa-miR-614	0.82 (0.70–0.94)
hsa-miR-940	0.80 (0.67–0.92)
hsa-miR-1263	0.80 (0.68–0.93)
hsa-miR-4431	0.80 (0.68–0.92)
hsa-miR-15b-5p	0.79 (0.66–0.91)
hsa-miR-17-5p	0.79 (0.67–0.92)
hsa-miR-93-5p	0.79 (0.66–0.91)
hsa-miR-548ad-3p	0.79 (0.66–0.92)
hsa-miR-661	0.79 (0.65–0.92)
hsa-miR-23a-3p	0.78 (0.64–0.91)
hsa-miR-200c-3p	0.78 (0.64–0.91)
hsa-miR-3934-5p	0.78 (0.65–0.91)
hsa-miR-150-5p	0.75 (0.61–0.89)
Score 1 (combined panel)	0.88 (0.78–0.99)

6 CAPÍTULO II

The oncogenic role of miR-150-5p in triple negative breast cancer

Bruna M. Sugita[1], Yara Zabala[2], Aline Fonseca[2], Rodrigo Almeida[1], Yuriy Gusev[2,3], Simina Boca[2], Iglenir J. Cavalli[1], Enilze M. S. F. Ribeiro[1], Luciane R. Cavalli[2]

[1] Department of Genetics, Federal University of Parana, Curitiba, PR Brazil
[2] Department of Oncology, Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, Washington, DC, USA
[3] Innovation Center for Biomedical Informatics, Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, Washington, DC, USA

*corresponding author:
Dr. Luciane Regina Cavalli
Georgetown Lombardi Comprehensive Cancer Center, Oncology Department, Washington, DC
e-mail: lrc@georgetown.edu

ABSTRACT

Altered expression of miR-150-5p has been shown in several types of cancer, including breast cancer, suggesting its important role in carcinogenesis. In this study, our aim was to investigate the expression levels of miR-150-5p and its functional role in conferring the aggressive tumor phenotype in triple negative breast cancer (TNBC). In clinical cases, we observed significant over-expressed levels of miR-150-5p in the tumor tissue of the patients when compared to the adjacent normal tissue (ANT) and in the TNBC cases when compared to the non-TNBC cases. Also, miR-150-5p was found to be over-expressed among samples of African American and non-Hispanic white women, and in cases positive for LN metastasis and breast cancer recurrence. In addition, its over-expression was shown to correlate with tumor progression, seen with significant increase in expression from the primary tumor, to LN and distant metastatic lesions. These results show the association of up-regulation of miR-150-5p with poor prognostic factors in TNBC patients. The manipulation of miR-150-5p expression levels, using miRNA-150-5p mimic and/or inhibitor assays in transfected/repressed systems in TNBC cell lines, demonstrated its oncogenic function in modulating cell proliferation, migration, expression of EMT markers and drug resistance.

INTRODUCTION

Triple negative breast cancer (TNBC) is a highly aggressive breast cancer subtype frequently associated with poor prognosis, resistance to therapy and rapid progression to metastatic disease [1]. These tumors, are associated with classical “hallmarks” of highly invasive tumors, including absence of estrogen-signaling, BRCA1 deficiency, increase genomic instability, high mitotic and angiogenesis activity [2,3].

MiRNAs are a class of non-coding endogenous RNA molecules that act as regulators of important biological processes and therefore are involved with cancer initiation and progression. These molecules can act as oncogenes or tumor suppressor genes, regulating targets involved in multiple cancer signaling pathways such as the ones that control cell proliferation, differentiation, apoptosis, metastasis, and angiogenesis [4,5].

The miR-150-5p, localized on chromosome 19q13, has been shown to regulate the expression of several driver oncogenes and/or tumor suppressor genes involved in these critical signaling pathways. Its deregulated expression has been observed in different types of human cancers; in hematologic diseases it has been mostly found to be significantly down-regulated [6], while in solid tumors its abnormal expression varies according to the affected tissue; in esophageal, hepatic and colorectal carcinomas and gliomas it has been observed down-regulated [7-11] while in lung, gastric, prostate, thyroid and breast cancer it was mostly found to be up-regulated [12-16].

In vitro assays investigating the functional role of miR-150-5p have indicated that its deregulated expression is associated with aggressive tumor phenotypes, such as high tumor proliferation, invasion, migration and metastasis [11,17,18]. The molecular mechanisms underliying its regulatory role in modulating these phenotypes are not completely known; *C-MYB, P53, SRC1N1, MUC4, ZEB1* and other EMT associated genes have been shown to be some of the targets through which miR-150-5p exerts its action [7,8,13,17-21].

In this present study, our aim was to determine the level of expression of miR-150-5p in clinical samples and commercially available cell lines of breast cancer, and elucidate its functional role in TNBC.

MATERIAL AND METHODS

Clinical cases

One-hundred and fourteen formalin-fixed paraffin-embedded (FFPE) cases diagnosed with ductal breast carcinoma from Washington-DC area were analysed for miR-150-5p expression. The cases were acquired under IRB protocol from the Histopathology Shared Resources (HSTR) at the Lombardi Comprehensive Cancer Center, Georgetown University and analyzed by the investigators in a de-codified manner. Clinical-histopathological information was obtained by HSTR personal and included age at diagnosis, self-reported race, tumor size (mm), grade and stage, presence of lymph node metastasis and disease recurrence.

All tumors were subclassified in triple negative breast cancer (TNBC) and non-TNBC after ER, PR and HER2 immunohistochemical evaluation following American Society of Clinical Oncology (ASCO)/College of American Pathology (CAP)

guidelines [22,23]. Altogether 61 African-american (29 AA-TNBC and 32 AA-NTNBC) and 53 non-Hispanic White (28 NHW-TNBC and 25 NHW-NTNBC) patients were evaluated in this study.

Tumor tissue and adjacent non-tumoral tissue were obtained from the same slide by microdissection following pathology evaluation for forty-one samples, according to previous protocol [24].

Cell lines

Nine commercially established cell lines of human breast carcinoma were acquired from the Tissue Culture Shared Resource (TCSR) at the Lombardi Comprehensive Cancer Center, Georgetown University: HCC1806, BT549, DU4475, MCF10A, MCF7, MDA-MB-157, MDA-MB-231, MDA-MB-453 and MDA-MB-468. The cell lines were genomic verified by STR allelotype analysis. All cell lines were cultivated in RPMI or DMEM with 10% FBS, 5% P/S and 1% Fungizone in T75 flasks at 37^{o}C + 5% CO_2. Before each functional assay, all cells were cultivated until a confluence of 70 to 80% .

RNA extraction

The FFPE tumor blocks were sectioned into 10 µm tissue sections and the selected areas with more than 80% of tumor cells (delineated by the pathologist Dr.Kallakury) were carefully microdissected according to our previous protocols [24] to ensure the analysis of a pure cell population. Microdissected tissue were isolated for RNA using RecoverAll Total Nucleic Acid Isolation kit (Thermo Fisher), accordingly to manufacture's protocol. RNA quality and quantity were assessed using RNA LabChip and 2100 Bioanalyzer (Agilent Technologies Inc) and NanoDrop spectrophotometer (Thermo Scientific Inc) respectively. The adjacent non-tumoral tissue (ANT) was obtained microdissected from the same 10 µm tissue sections, after pathology evaluation.

RNA from the cell lines was isolated using the same commercial kit above.

Global miRNA expression analysis

Global miRNA expression was performed using NanoString nCounter Technology Human V2/2.1 miRNA expression platform. Raw data was pre-processed by NanoString's nCounter RCC collector worksheet and normalized by NanoString

nSolver 3.0 software. MultiExperiment Viewer 4.9 (MeV 4.9) was used to perform differential expression analysis berween the groups (unpaired or paired t test) and Pvalues and Log2 Fold Changes were also calculated.

RT-qPCR

Expression of miR-150-5p was assessed in both the clinical samples and breast cell lines using Taqman MicroRNA Assay (Life Technologies, CA) specific primer hsa-miR-150-5p (Cat#4427975, Assay# 000473). RNU44 and RNU48 were used as endogenous control. MiRNA expression was calculated using $2^{-\Delta\Delta Ct}$ method [25,] using MCF-10A as reference sample.

Transfection

Modulation of miR-150-5p levels was performed in MCF-10A, MDA-MB-231 and HCC1806 cell lines using mirVanaTM miRNA mimics (miR-150m) and inhibitors (miR-150i) assays (Life Technologies, CA). Negative control (miRNC) and positive control (let-7c) were included in all the transfections, performed in triplicates. Prior to transfection different concentrations of lipofectamine were tested in relation to cell viability. Cells were incubated at 37oC in 5% CO2 in RPMI 1640 with 10% FBS and 1% P/S, and reversed-transfected accordingly to manufacture's protocol.

Transfection efficiency was also determined by Western Blot assessing the inhibition of miR-let-7c (positive control) by the expression levels of its target, the HMGA2 protein (FIGURE S1A). MiR-150-5p activity after transfection was verified by Western Blot by checking the expression of MYB protein, its direct target (FIGURE S1B)

Western blotting

Cells were lysed in RIPA buffer (Invitrogen) supplemented with protease inhibitor cocktail (Invitrogen) and collected using a cell scraper. Cell lysates were maintained 20 minutes in agitation at 4°C and centrifugated 20 minutes at 12 000 rpm. Proteins were resolved in Bolt-PAGE Novex 10% Bis-Tris gel (Invitrogen) and transferred to PDV membranes (Life Technologies) for western blot analysis. The primary antibodies used were c-Myb (1:300 dilution, Invitrogen), p-Erk(1/2) (1:1000, Cell Signaling), p-Src (Y416) (1:1000, Cell Signaling), p-Scr (Y517) (1:1000, Cell Signaling), E-cadherin (1:1000, Cell Signaling) Vimentin (1:1000, Cell Signaling,

USA), snail (1:200, Cell Signaling, USA), Slug (1:200, Cell Signaling, USA), GAPDH (1:1000, Life Technologies, USA) and AP-conjugated secondary antibody (1:10 000, Life Technologies). Proteins were visualized using Supersignal west pico ECL (Thermo Scientific, USA) and detected by Amersham Imager 600 (Amersham Biosciences).

Cell proliferation assay

The Cell Titer Aqueous One Solution Cell Proliferation Assay (MTS),(Promega, Corp.) was used to assess the effect of miR-150-5p expression levels modulation on the proliferation of the TNBC cell lines of this study. Briefly, $4x10^3$ MDA-MB-231 and HCC1806 cells were plated in 96-well culture plates and reverse transfected with miR-150-5p inhibitor, miR-negative control (NC) and, miR-150-5p mimic. Cell Titer Solution was added to each well (20 ml/well) an absorbance was measured, using a ELISA reader at 490nm, at 24, 48, 72 and 96 hours. Day 0 was set as 48h after transfection.

Clonogenic Assay.

One thousand HCC1806 cells were reverse transfected with miR-150-5p inhibitor, miR-negative control and, miR-150-5p mimic as described above. Following 48 hours transfection, cells were trypzinized, counted and transferred to 6-well plates at a density of 1000 cells/well for colony formation. After 14 days, colonies were fixed with 3:1 methanol:acetic acid and stained for 1h with 0.5% crystal violet (Sigma Aldrich, Corp). The number of colonies were counted in a light microscope. Clusters with >50 cells were considered colonies. Each experiment was performed in triplicate.

Wound Healing Assay

MDA-MB-231 and HCC1806 cells were seeded at $3x10^4$ cells/well and reverse transfected with miR-150-5p inhibitor, miR-negative control and, miR-150-5p mimic directly into 2 well culture-inserts in μ-Dish 35 mm (Ibidi Culture Inserts). After 48h of transfection the insert was removed and cells were washed 4 times with PBS to allow cell migration. Images were captured and analyzed by ImageJ software at designated times to assess wound closure.

Cytotoxic assays

MDA-MB-231 and HCC1806 cells were plated at $3x10^4$ cells/well and reverse transfected with miR-150-5p inhibitor, miR-negative control and treated with Doxorubicin in concentrations of 500nM. The MTT [3- (4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide] assay was used, per standard protocols, to measure cell viability. Absorbance was measured, using a ELISA reader at 490nm, at 24, 48, 72 and 96 hours after treatment. Experiments were performed in triplicates.

Statistical analysis

Statistical analyses were carried out using Graphpad. The differences between the means were tested by one-way analysis of variance (ANOVA) with Newman-Keuls Multiple Comparison Test. Each experiment was performed at least three times. Measurement data were presented as mean and standard error of the mean (SEM). $P<0.05$ was considered statistically significant.

RESULTS

miR-150-5p is significantly up-regulated in tumor breast tissue when compared to non-tumor breast tissue and in TNBC cases when compared to non-TNBC cases

Expression levels of mIR-150-5p were assessed by RT-qPCR in 41 and 97 samples of adjacent non-tumor (ANT) and tumor tissue respectively. Significant up-regulation of miR-150-5p was found in tumor tissue ($p=0.0001$) when conpared to ANT (**FIGURE 1A**). Higher expression levels of miR-150-5p were also found in TNBC cases ($n=42$) when compared to non-TNBC cases ($n=57$) ($p=0.0205$) (**FIGURE 1B**).

miR-150-5p is significantly up-regulated in TNBC cases of African-American (AA) patients when compared to TNBC of Non-Hispanic White (NHW) patients presenting a good power in discriminating these groups

Global miRNA expression analysis was perfomed in 24 and 28 AA-TNBC and NHW-TNBC clinical cases respectively using Nanostring platform in our previous study[26]. Differential expression analysis resulted in 256 differentially expressed mIRNAs between the two groups (MeV 4.9, tTest $p<0.01$, FDR<0.05), among them mIR-150-5p presented higher expression in AA-TNBC cases when compared to

NHW-TNBC cases (p=6.20E-04; Log2FC 2.73) (**FIGURE 1C**). Receiver Operating Characteristic Curve analysis was performed showing a good power of miR-150-5p in discriminating AA-TNBC and NHW-TNBC cases (AUC= 0.75; CI 95% 0.61-0.79) (**FIGURE 1D**).

miR-150-5p is significantly up-regulated in distant and lymph node breast cancer metastasis when compared to primary tumors

Global miRNA expression analysis was perfomed in 17 paired lesions of primary breast tumor (PBT) and correspondent sentinel lymph node (SLN) metastasis using a Nanostring miRNA platform. A total of 47 miRNAs were found differentially expressed among these lesions (MeV 4.9, paired tTest, $p<0.05$), inlcuding miR-150-5p, which was found with higher expression in the SLN metastasis when compared to the primary breast tumors (p=0.01, Log2FC=1.70) (**FIGURE 1E**). The global expression levels of miR-150-5p were also confimed by RT-qPCR in the same paired PBT and SNL metastasis (FC= 2.125, p=0.0445**;** **FIGURE 1F**).

miR-150-5p is up-regulated in TNBC cell lines when compared to non-TNBC and non-tumor breast cell lines

Expression of miR-150-5p was assessed in nine breast cancer cell lines: MCF-10A (non-tumorigenic cell line), MCF-7 (non-TNBC cell line), MDA-MB-157, MDA-MB-231, MDA-MB-453, MDA-MB-468, DU4475, BT549 and HCC1806 (TNBC cell lines). Higher expression levels were found in five out of seven TNBC cell lines (BT549=3.07, MDA-MB-231=4.46, MDA-MB-453=6.72, HCC1806=1.84 and DU4475=176) with exception of MDA-MB-157 (FC=0.90) and MDA-MB-468 (FC= 0.75). Although DU4475 presented the highest levels of miR-150-5p, under our cell culture conditions this cell line presents itself in suspension, and therefore it was not chosen for further functional analysis.

The inhibition of miR-150-5p reduces cell proliferation and clonogenic growth

To determine the role of miR-150-5p in cell growth we performed an MTS and colony formation assay in MDA-MB-231 and HCC1806 cells transfected with (miR-150i) and miR-150-5p Mimic (miR-150m) in comparison with miRNC. Proliferation of MDA-MB-231 was not affected by manipulation of miR-150-5p expression levels at

any of the time points (**FIGURE 3A**), however HCC1806 cells showed a significant decrease of 34.6% in cell proliferation inhibition of miR-150 expression after 72h when compared with the NC (**FIGURE 3B**). The ability to form colonies in these cells was also significantly decreased by 63.75% after inhibition of miR-150-5p expression (**FIGURE 3C**). However, the ectopic expression of miR-150-5p did not show any significant change in clonogenicity compared to the NC (data not shwon).

The inhibition of miR-150-5p reduces cell migration

Wound healing assays were performed to determined whether miR-150-5p plays a role in cell migration of MDA-MB-231 and HCC1806 breast cancer cell lines. Our results showed that the inhibition of miR-150-5p expression levels significantly decreased the percentage of wound clousure by 26.12% and 30.68% in MDA-MB-231 (**FIGURE 4A**) and HCC1806 (**FIGURE 4B**) cell lines, respectively. The ectopic expression of miR-150 was not able to increase the migration capacity of any of the studied cell lines.

The inhibition of miR-150-5p reduces resistance to doxorrubicin

In order to assess whether miR-150-5p expression levels affects the cytotoxic response to commonly used therapies against TNBC, we treated transfected HCC1806 cells with Doxorrubicin (50nM) and Abraxane (50nM). Inhibition of miR-150-5p significantly decreased cell viability after Doxorubicin treatment at 24h, 48h and 72h (**FIGURE 5**). No changes were observed upon ectopic expression of this miRNA.

miR-150-5p could be involved in triggering the expression of early inducers of Epithelial to Mesenchymal Transition (EMT) markers

To experimentally determine the mechanisms by which miR-150-5p could affect the tumorigenic phenotypes above, by modulating the expression of epithelial mesenchymal transition (EMT) expression markers, we conducted Western Blot analysis in the MDA-MB-231 transfected cells with miR-150-5p assays. We observed that with suppression of miR-150-5p, there was a down-regulation in the protein expression levels of the EMT induced markers, Snail and Slug. No changes were observed in the levels of E-cadherin, Vimentin and Zeb1 markers (**FIGURE 6A**). Interestingly, no changes in this cell lines in the EMT markers were observed with the

ectopic expression of miR-150-5p. In HCC 1806 cells, the EMT markers expression were not changed after modulation of miR-150-5p (data not shown).

miR-150-5p could behave as an oncogene through an indirect mechanism that leads to the activation of ERK(1/2) pathway

In order to analyze the molecular pathways mediating the observed effects in vitro, we assessed the levels of expression of two predicted targets of miR-150-5p, SRCIN1 and TP53, in the MDA-MB-231 cell line. As a result, modulation of miR-150-5p expression did not affect the levels of active Src (p-Src Y416) or inactive (p-Scr Y527) Src and TP53 (**FIGURE 6B**). Nevertheless, we observed a downregulation of p-ERK(1/2) upon inhibition of miR-150-5p. Furthermore, transfection of these cells with miR-150-5p mimic increased the protein levels of p-ERK(1/2) compared to the control.

DISCUSSION

The triple negative breast cancer (TNBC) subtype has been extensively studied in both basic science and clinical research, considering its aggressive phenotype and poor outcome. In this study, our main goal was to determine the role of miR-150-5p as a contributor to the TNBC natural history. Analyzing well-annotated clinical samples, we found that this miRNA presents higher expression in patients with TNBC subtype when compared to the ones with the non-TNBC. Interestingly, this association was significantly associated with race, as African-American (AA) patients with TNBC presented higher levels of miR-150-5p expression when compared to non-Hispanic Whites (NHW). As our previous miRNA global profiling study in these populations [26], these results support the existence of biological differences in the transcriptome between populations and their impact in patients prognosis and treatment.

We have also shown that the expression of miR-150-5p is potentially associated with tumor progression, as it was found with increased expression in lymph node (LN) metastasis when compared to primary tumors (PBT). The association of this miRNA and metastasis has been previously reported, however with distinct roles (either anti- or pro-metastatic) accordingly to the cancer type [17,27,28]. In breast cancer, however, miR-150-5p metastatic role was not clearly defined.

Consistent with its pro-metastatic role, we observed that suppression of miR-150-5p expression in transfected TNBC cells, reduced their tumorigenicity, as evidenced by a significant decrease in cell proliferation, clonogenic growth, cell migration and resistance to Doxurubicin. However, in most of these phenotypes, we did not observe an opposite effect or a change in the rate of cell proliferation, clonogenic growth, migration or cytotoxicity, after increasing levels of miR-150-5p by adding miR-150-5p mimic when compared to the transfected cells with negative control (scramble miRNA). These results may indicate that the existent levels of miR-150-5p in the TNBC cell lines evaluated are sufficient to confer their observed tumorigenic phenotype. Also, they can point out to the fact, that the complete impairment of these phenotypes, are not driven entirely by miR-150-5p expression, and may require synergy with other miRNAs to exert its full oncogenic action in TNBC. The “flexible” miRNA/mRNA target interaction can confer miRNAs the ability to cooperatively regulate a single biological process by targeting common components of that process.

Enrichment function analysis of miR-150-5p showed their involvement in common functions associated with cell adhesion and adherens junction. These functions are critical for the epithelial-mesenchymal transition (EMT) process, which is essential to confer cell migration and invasion and stem cell tumor capabilities. EMT plays pivotal roles in cancer progression and metastasis, through intricate and complex regulations of EMT associated transcription factors and proteins, such as E-cadherin, Slug, Vimentin, Snail, Zeb1 and Twist. Alterations in these EMT markers, leads to changes in the cytoskeleton, connection and polarity of the tumor cells, and directly impacts their tumorigenic behavior. Therefore, we have shown that the EMT process is also controlled post-transcriptionally by a complex and regulated network of miRNAs, we verified whether the manipulation of the expression of miR-150-5p would directly affect the expression of EMT markers. Indeed, we observed that the inhibition of the expression of miR-150-5p in MDA-MB-231 cells, resulted in lower expression of the EMT induced markers, SNAIL and SLUG. By evaluating the SRC signaling pathway, as previously reported to be affected by miR-150-5p in lung cancer [13], we observed that the inhibition of miR-150-5p expression resulted in decreased expression levels of the phosphorylated ERK protein (pERK), a SRC downstream protein. Interestingly, pERK and SNAIL relationship was been previously reported [29]. C-Src is a member of the family of membrane-associated non-receptor

protein tyrosine kinases, named Src family kinases (SFKs), and presents an important function as regulator of proliferation, differentiation, survival, motility and angiogenesis [30]. Altered expression of c-Src was already reported in several types of cancer, including elevated levels in breast cancer; an abnormally high tyrosine kinase activity can shift the balance of key biological signaling pathways and processes, which might lead to cancer [31]. SRC Kinase Signaling Inhibitor 1 (SRCIN1) is a negative regulator of SRC; it represses SRC activity by activating CSK by phosphorylation of the tyrosine 597 and has been identified as a target of miR-150-5p [13]. Therefore, transcriptional repression of SRCIN1 by miR-150-5p could lead to a relief of repression of the SRC pathway, resulting in increased proliferation and migration activities, as we observed in our study.

In any event, we showed that in our transfected TNBC cell models SNAIL and SLUG down-regulation are associated with cells with lower cell proliferation, clonogenic growth, drug resistance and migration, mediated by miR-150-5p inhibition.

Considering the previous association of miR-150-5p and the SRC pathway and our own observed results, we suggest that miR-150-5p, can exert its oncogenic action in TNBC, by targeting this oncogenic pathway. Additional functional studies in TNBC models are required to substantiate this hypothesis.

CONCLUSION

In conclusion, we have showed that miR-150-5p may act as an oncogene in TNBC by targeting important biological functions as proliferation, migration and resistance to doxorrubicin. This oncogenic role may be more pronounced in AA patients; therefore further studies may help elucidate its function in order to use it as target for drug development especially for these patients.

FIGURE 1. Expression levels of miR-150-5p in normal breast tissue, tumor tissue (TNBC and NTNBC cases), and paired cases of primary breast tumor (PBT) and sentinel lymph node (SLN). Higher levels of miR-150-5p was found in adjacent tumor breast tissue when compared to non-tumor breast tissue (**A**) in TNBC clinical cases when compared to NTNBC clinical cases (**B**). In a global expression analysis, miR-150-5p was found differentially expressed between AA-TNBC and NHW-TNBC cases, with higher expression in AA-TNBC patients (p=6.20E-04; Log2FC= 2.73) (**C**), showing a robust power in discriminating these groups (ROC analysis, AUC=0.75) (**D**). Expression levels of miR-150-5p was also increased in SLN samples when compared to respective PBT in global expression analysis (Nanostring; p= 0.01, Log2FC= 1.70) and RT-qPCR analysis (FC= 2.125, p= 0.0445).

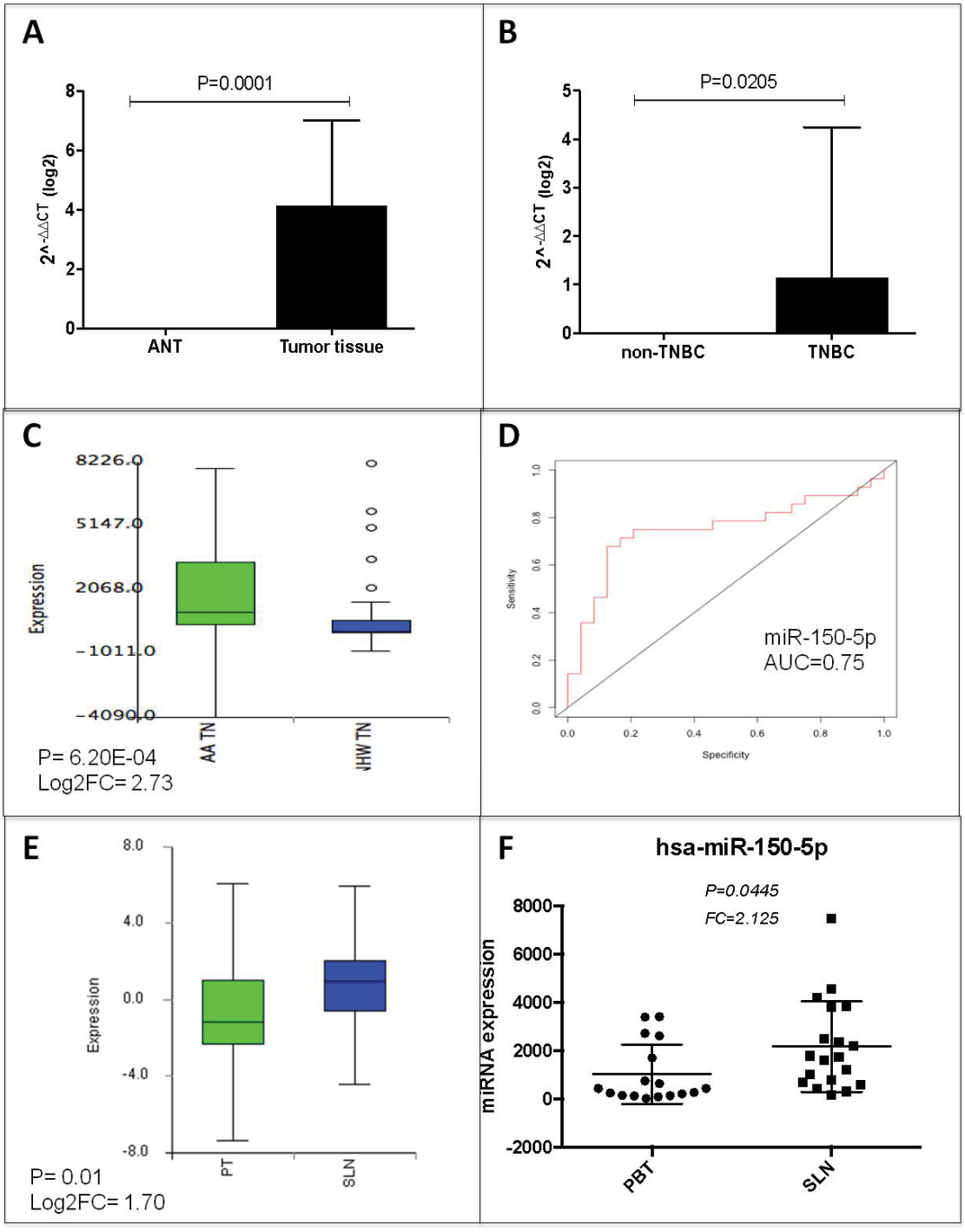


FIGURE 2. **Expression levels of miR-150-5p in 9 breast cell lines.** Higher levels of miR-150-5p were observed in almost all TNBC cell lines when compared to MCF10A, with exception of MDA-MB-157 and MDA-MB-468 cell lines. DU4475 cell line presented the highest expression levels, with 176 times the amount of MCF10A expression.

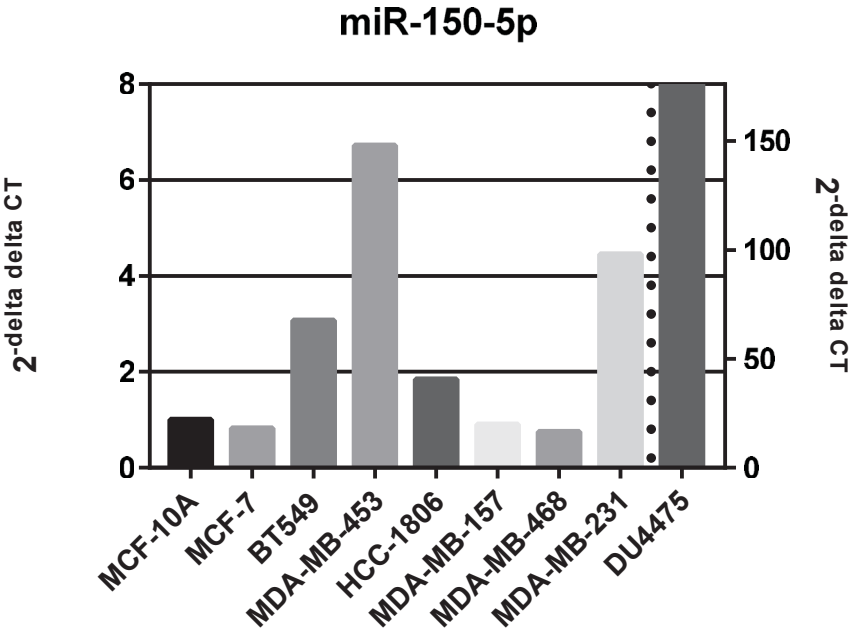


FIGURE 3. **The inhibition of miR-150-5p reduces clonogenic growth of breast cancer cell lines.** Proliferation of MDA-MB-231 cells was not affected by treatment at any of the time points (**A**), however HCC1806 cells showed significant decrease of cell proliferation at 72h when compared to negative control (34.6%, $p<0.05$) (**B**). Inhibition of miR-150-5p also decreased colony formation of HCC1806 cells; up-regulation of miR-150-5p did not produce significant change in clonogenicity compared to the control (**C**).

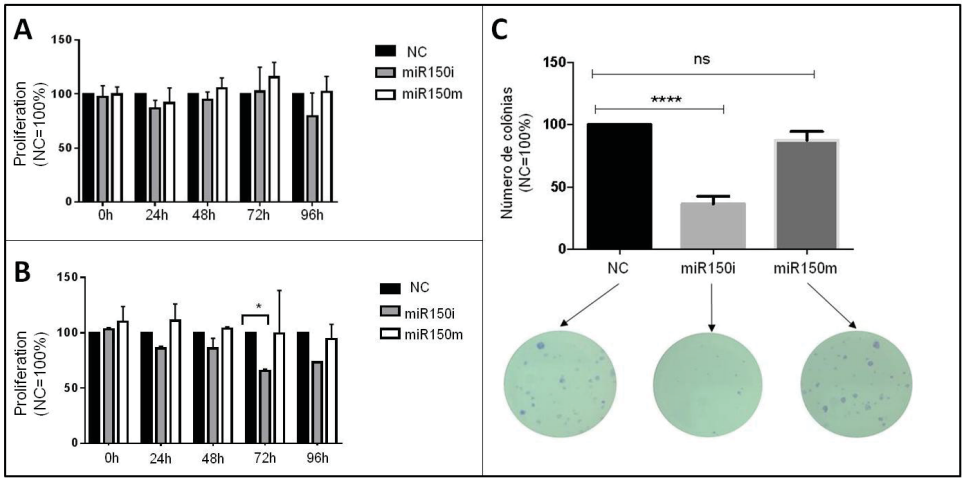


FIGURE 4. Migration of MDA-MB-231 and HCC1806 cell lines was inhibited by the inhibition of miR-150. Inhibition of miR-150-5p significantly decreased the percentage of wound closure 26.12%in MDA-MB-231 (after 72h, **A** and **B**) and 30.68% in HCC1806 (after 24h, **C** and **D**) cell lines respectively when compared to the control. The up-regulation of miR-150 was not able to increase the migration capacity of any of the studied cell lines compared to the control (**A** and **C**).

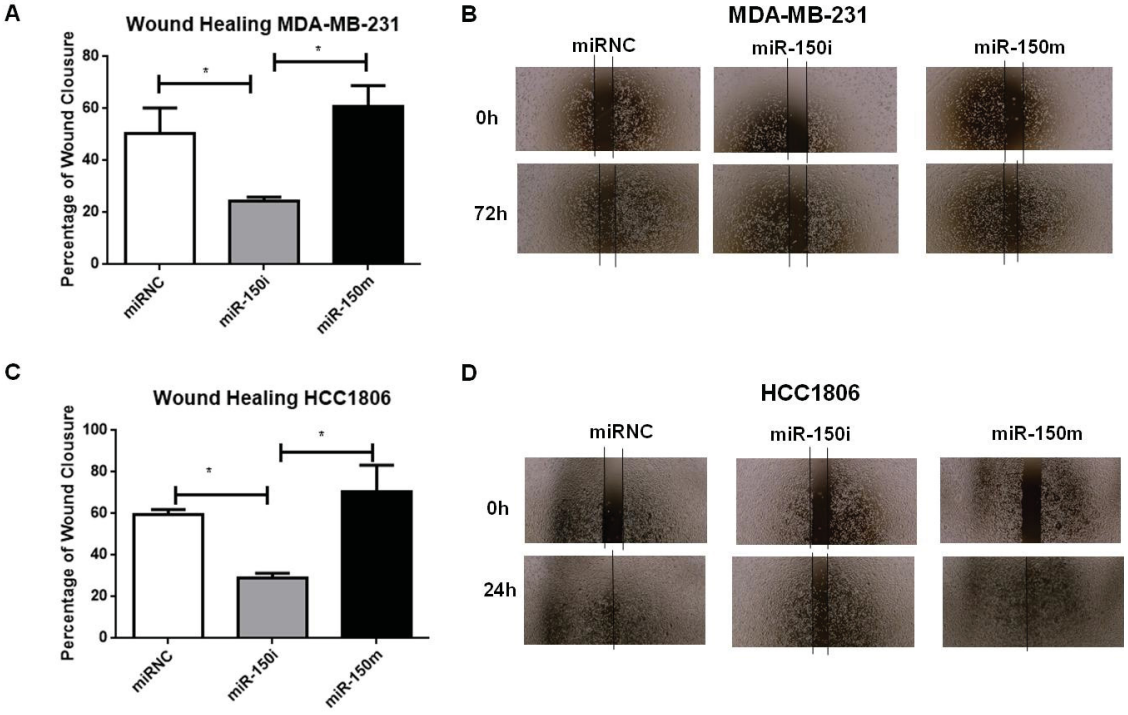


FIGURE 5. **miR-150-5p affects breast cancer cell line HCC1806 sensitivity to doxorrubicin.** Inhibition of miR-150-5p decreases HCC1806 cell resistance to doxorubicin in 24, 48 and 72h (compared to control)

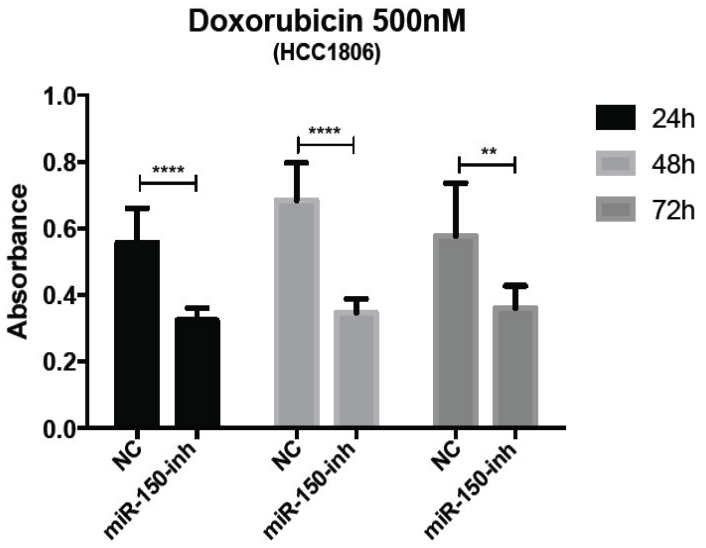


FIGURE 6. miR-150-5p could be involved in triggering the expression of early inducers of Epithelial to Mesenchymal Transition (EMT) markers and ERK(1/2) pathway proteins in MDA-MB-231. Although levels of E-cadherin, Vimentin and Zeb1 did not change, we observed an up-regulation in the protein levels of Snail and Slug, which are inducers of EMT (A). Inactive and active forms of Src

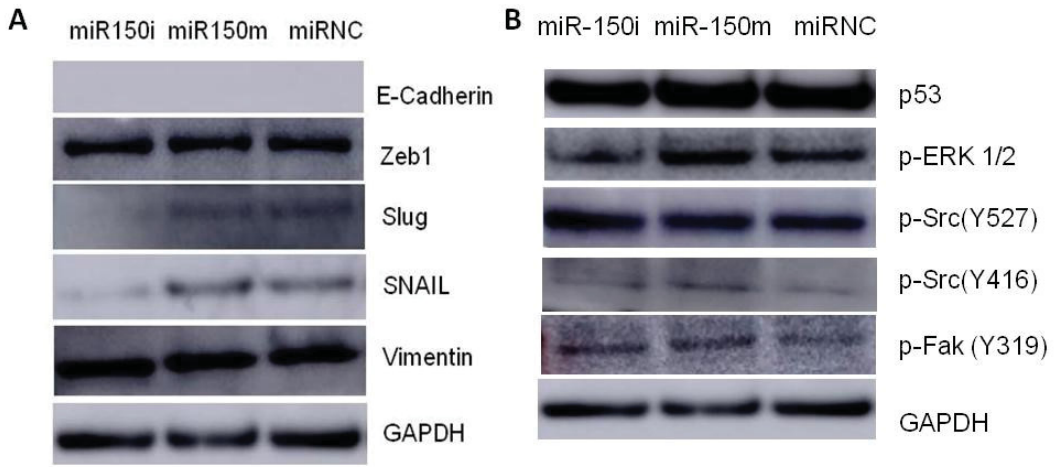
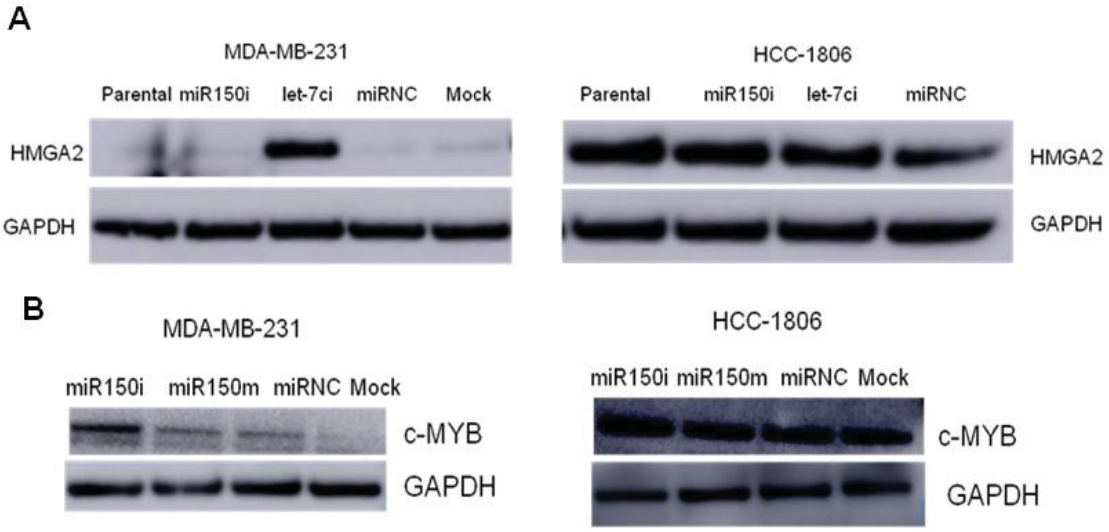


FIGURE S1. Transfection success assessment. (A) Inhibition of let-7c resulted in increased protein levels of validated target HMGA2 proving transfection success of miRvana inhibitors & mimics through reverse transfection. (B) Modulation of miR-150 resulted in altered levels of c-MYB.



REFERENCES

1 Foulkes, W. D., Smith, I. E. & Reis-Filho, J. S. Triple-negative breast cancer. N Engl J Med 363, 1938-1948 (2010).

2 Dai, X. et al. Breast cancer intrinsic subtype classification, clinical use and future trends. Am J Cancer Res 5, 2929-2943 (2015).

3 Sharma, P. Biology and management of patients with triple-negative breast cancer. Oncologist 21, 1050-1062 (2016).

4 Wang, J., Chen, J. & Sen, S. Microrna as biomarkers and diagnostics. J Cell Physiol 231, 25-30 (2016).

5 Zhou, X. & Izzoti, A. Microrna: A newcomer to biological science. Microrna 4, 2 (2015).

6 He, Y., Jiang, X. & Chen, J. The role of mir-150 in normal and malignant hematopoiesis. Oncogene 33, 3887-3893 (2014).

7 Yokobori, T. et al. Mir-150 is associated with poor prognosis in esophageal squamous cell carcinoma via targeting the emt inducer zeb1. Cancer Sci 104, 48-54 (2013).

8 Feng, J. et al. Mir-150 functions as a tumour suppressor in human colorectal cancer by targeting c-myb. J Cell Mol Med 18, 2125-2134 (2014).

9 Wang, W. H. et al. Mir-150-5p suppresses colorectal cancer cell migration and invasion through targeting muc4. Asian Pac J Cancer Prev 15, 6269-6273 (2014).

10 Sakr, M. et al. Mir-150-5p and mir-133a suppress glioma cell proliferation and migration through targeting membrane-type-1 matrix metalloproteinase. Gene 587, 155-162 (2016).

11 Sun, J. et al. Targeting of mir-150 on gli1 gene to inhibit proliferation and cell cycle of esophageal carcinoma ec9706. Cancer Biomark 21, 203-210 (2017).

12 Yanaihara, N. et al. Unique microrna molecular profiles in lung cancer diagnosis and prognosis. Cancer Cell 9, 189-198 (2006).

13 Cao, M. et al. Mir-150 promotes the proliferation and migration of lung cancer cells by targeting src kinase signalling inhibitor 1. Eur J Cancer 50, 1013-1024 (2014).

14 Wang, P. Y. et al. Regulating a549 cells growth by aso inhibiting mirna expression. Mol Cell Biochem 339, 163-171 (2010).

15 Dettmer, M. S. et al. Microrna profile of poorly differentiated thyroid carcinomas: New diagnostic and prognostic insights. J Mol Endocrinol 52, 181-189 (2014).

16 Dezhong, L. et al. Mir-150 is a factor of survival in prostate cancer patients. J BUON 20, 173-179 (2015).

17 Li, H. et al. Mir-150 promotes cellular metastasis in non-small cell lung cancer by targeting foxo4. Sci Rep 6, 39001 (2016).

18 Guo, Y. H. et al. Wnt/beta-catenin pathway transactivates microrna-150 that promotes emt of colorectal cancer cells by suppressing creb signaling. Oncotarget 7, 42513-42526 (2016).

19 Huang, S. et al. Mir-150 promotes human breast cancer growth and malignant behavior by targeting the pro-apoptotic purinergic p2x7 receptor. PLoS One 8, e80707 (2013).

20 Srivastava, S. K. et al. Microrna-150 directly targets muc4 and suppresses growth and malignant behavior of pancreatic cancer cells. Carcinogenesis 32, 1832-1839 (2011).

21 Wang, D. T. et al. Mir-150, p53 protein and relevant mirnas consist of a regulatory network in nsclc tumorigenesis. Oncol Rep 30, 492-498 (2013).

22 Wolff, A. C. et al. American society of clinical oncology/college of american pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. J Clin Oncol 25, 118-145 (2007).

23 Hammond, M. E. et al. American society of clinical oncology/college of american pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. Arch Pathol Lab Med 134, 907-922 (2010).

24 Torresan, C. et al. Increased copy number of the dlx4 homeobox gene in breast axillary lymph node metastasis. Cancer Genet 207, 177-187 (2014).

25 Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative pcr and the 2(-delta delta c(t)) method. Methods 25, 402-408 (2001).

26 Sugita, B. et al. Differentially expressed mirnas in triple negative breast cancer between african-american and non-hispanic white women. Oncotarget 7, 79274-79291 (2016).

27 Yue, P. Y. et al. Microrna profiling study reveals mir-150 in association with metastasis in nasopharyngeal carcinoma. Sci Rep 7, 12012 (2017).

28 Zhao, Y., Zhu, Y. & Song, J. Mir-150 promotes the cell invasion of prostate cancer cells by directly regulating the expression of p53. Pharmazie 71, 533-536 (2016).

29 Nagarajan, D., Melo, T., Deng, Z., Almeida, C. & Zhao, W. Erk/gsk3beta/snail signaling mediates radiation-induced alveolar epithelial-to-mesenchymal transition. Free Radic Biol Med 52, 983-992 (2012).

30 Summy, J. M. & Gallick, G. E. Src family kinases in tumor progression and metastasis. Cancer Metastasis Rev 22, 337-358 (2003).

31 Finn, R. S. Targeting src in breast cancer. Ann Oncol 19, 1379-1386 (2008).

7 CAPÍTULO III

Integrated DNA copy number alteration and global miRNA expression analysis: a panel of 17 miRNAs for triple negative breast cancer in Brazilian patients

Bruna M. Sugita[1], Silma R. Pereira[2], Rodrigo Almeida[1], Mandeep Gill[3], Akanksha Mahajan[3], Anju Duttargi[3], Saurabh Kirolikar[3], Rubens S. de Lima[4], Cicero A. Urban[4], Kepher Makambi[5], Subha Madhavan[3,6], Simina Boca[3,6], Yuriy Gusev[3,6], Iglenir J. Cavalli[1], Enilze M. S. F. Ribeiro[1], Luciane R. Cavalli[3*].

[1] Department of Genetics, Federal University of Parana, Curitiba, PR Brazil
[2] Department of Biology, Federal University of Maranhão, São Luis, MA, Brazil.
[3] Department of Oncology, Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, Washington, DC, USA
[4] Breast Unit, Hospital Nossa Senhora das Graças, Curitiba, Brazil.
[5] Innovation Center for Biomedical Informatics, Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, Washington, DC, USA

*corresponding author:
Dr. Luciane Regina Cavalli
Georgetown Lombardi Comprehensive Cancer Center, Oncology Department, Washington, DC
e-mail: lrc@georgetown.edu

ABSTRACT

Triple Negative Breast Cancer (TNBC) is a highly aggressive disease with poor overall survival, more frequently observed in younger patients from African American and Latin/Hispanic populations, when compared to non-Hispanic whites. Understanding the biological mechanism underlying this disparity is important in developing more efficient treatments to this disease. Therefore, in this study our aim was to determine the patterns of copy number and miRNA expression of the TNBC and non-TNBC cases of patients from Latin (or South) America, Brazil. Upon a direct integration of the copy number and global miRNA profiling data, performed in the same TNBC specimens, we identified a specific panel of 17 miRNAs associated with TNBC of Brazilian patients. This panel presented a high power in discriminating the TNBC subtype from the non-TNBC subtype of the population studied, and can represent the intrinsic biology of their TNBC transcriptomes. Alterations in the expression levels of these miRNAs and the corresponding impact in their gene targets associated with cancer signaling pathways can provide biological insights into the observed differences in the TNBC clinical outcome among racial/ethnic groups, taking into consideration their genetic ancestry.

INTRODUCTION

Breast cancer is the leading cause of death in women from Latin America living in US or in Latin American countries (referred here as Latinas). The overall incidence and mortality rates of breast cancer in this population in US were 91.9 and 14.0 per 100,000 in 2015 [1, 2]. In Brazil 59,700 new cases of breast cancer are expected in 2018, with an estimated risk of 56.3 per 100,000 [3]. Higher mortality rates were observed among Cubans (17.89), Mexicans (18.78), and Puerto Ricans (19.04), and a lower rate was observed among Central and South Americans (10.15) [4]. The breast cancer survival rate in Latin America hardly exceeds 70%, which is usually correlated with late diagnosis; approximately 30%–40% of breast cancer patients in the Latin American countries, including Brazil, are diagnosed in stages III and IV of the disease [5].

There is a well-documented disparity of breast cancer in Latinas when compared to non-Hispanic Whites (NHW); Latinas are more likely to present with

non-localized disease, receive less aggressive therapy, and have a disproportionately low survival rate when compared to NHW [6-8]. Several factors can contribute to the increased mortality rate in this population group, including socio-economic barriers, which can limit their overall access to early detection and cancer prevention services [9, 10] and presence of co-morbidities, such as obesity and diabetes [11-16]. Many deaths in this population can also be attributed to after-diagnosis factors, such as inadequate access to appropriate treatment or early treatment interruption or discontinuation [17-19].

Epidemiological and molecular studies have shown that breast cancer subtypes/phenotypes are distributed unevenly among various racial/ethnic groups [6, 7, 20-22]. The incidence of the triple negative breast cancer (TNBC) subtype in particular, one of the most clinically aggressive breast cancer subtypes, largely varies according to ethnicity, being present at higher frequencies in African-American (AA) (24-42%) and Hispanic/Latina (15-33%) women when compared to NHWs (11-28%) or Asians (11-33%) [6, 23]. As in AA women, Latina patients with TNBC are more often diagnosed at an earlier age, with advanced stage, likely to experience metastasis and be refractory to treatment [24-28]. In addition to the general attributed socio-economic status mentioned above, tumor biology and genetic background play a significant role in such disparity [11, 14], [29-32], however the individual contribution of each of these factors to their observed poor clinical outcome remains unclear. There is also lack of basic science studies and clinical trials that are conducted in Latinas [33, 34] which limits the knowledge of the biological causes that contributes to these tumors' clinical aggressiveness.

MicroRNAs (miRNAs) are a class of non-coding endogenous RNA molecules that have been identified to play a role in breast tumorigenesis [35, 36]. MiRNA expression has been shown to present different expression patterns according to the intrinsic molecular breast cancer subtypes [37, 38]. In TNBC, for instance, distinct miRNAs were reported mediating cellular process associated with aggressive tumors phenotypes, such as the ones that promote metastatic development [39-42] and treatment resistance [43, 44]. Interestingly, MiRNAs were shown to present variable expression according to race and/or ethnic groups [45-48]; a number of studies have shown miRNA polymorphisms in association with the susceptibility risk of breast cancer in specific ethnic populations [48-52]. However, there are limited reports on somatic miRNA expression levels in the breast tissue of these populations [53, 54].

Although the characterization of the genomic profiles in TNBC has been extensively performed, few studies have characterized them in specific ethnic groups, such as Latinas. This translates to a deficiency in the understanding of the intrinsic characteristics of their tumors' genome, which can differentially impact their tumor phenotypes and aggressive clinical behavior. Therefore, in this present study our aim was to obtain DNA copy number alterations (CNAs) and miRNA expression profiles of TNBC and NTNBC cases from Brazilian patients. In addition, we aimed to determine whether these CNAs could impact miRNA expression levels and whether there were common miRNA target genes affected by both CNAs and miRNA deregulation.

MATERIAL AND METHODS

General study design

Genome-wide array-CGH and miRNA profiling were performed in TNBC and non-TNBC cases of patients from Latin America, Brazil, to detect the patterns of copy number alterations (CNAs) and changes in global miRNA expression, respectively. The differentially expressed miRNAs among these subtypes were integrated with copy number profiling data performed in the same TNBC tissue specimens: the miRNAs were first mapped on the main cytobands harboring CNAs and further selected based on the concordance with CNAs (i.e. cytoband with gains/amplifications of copy number/ up-regulated miRNA expression and/or losses/deletions of copy number/ down-regulated miRNA expression); subsequently the mRNA targets of these miRNAs were identified and correlated with genes mapped in the cytobands affected by CNAs. Combinatorial target predicted algorithms in conjunction with functional and pathway annotation enrichment systems were applied to the selected miRNAs to identify the most relevant miRNAs and their corresponding targets associated with TNBC. Receiver Operating Characteristic (ROC) curve analysis was performed to determine the individual and combined power of the differentially expressed miRNAs in discriminating the TNBC and non-TNBC cases of the patients. Finally, the molecular data was associated with clinical-pathological information from the patients to determine their potential prognostic relevance.

Patient accrual and sample collection

Fifty-four formalin-fixed paraffin-embedded (FFPE) tissue samples of non-treated primary breast tumors were collected from the pathology tumor bank at the Hospital Nossa Senhora das Graças (HNSG), Paraná, Brazil. All samples were transferred to Georgetown University with no patient's identifiers under patient informed consent and through the IRB approval of Georgetown University, HNSG and the National Review Board of Ethics in Research (CONEP-Brazil). The TNBC and non-TNBC cases were determined by ER, PR and HER2 receptors status by immunohistochemistry (IHC) analysis performed at the time of diagnosis, following international guidelines [55, 56]. Briefly, the Monoclonal Mouse Anti-Human Estrogen Receptor α and Polyclonal Rabbit Anti-Human Progesterone Receptor were used for ER and PR analysis, respectively. ER and PR positivity were considered using a cut-off of 1%. The HercepTest (Dako North America Inc, Carpinteria, CA) was used for HER2/Neu+ status. Using these criteria, 27 patients presented tumors of the TNBC phenotype and 27 of the non-TNBC subtype (10 ER+/PR+/HER2-, 3 ER+/PR+/HER2+, 3 ER-/PR+/HER2-, 2 ER+/PR-/HER2- and 1 ER+/PR-/HER2+). In the remaining eight non-TNBC cases, 5 were ER+/HER2- and 1 ER+/HER2+, one presented no information of HER2 status (ER+/PR+) and one presented only HER2 status (HER2+).

The clinical and histopathological information from the patients was retrieved in a de-codified manner from the pathology reports, and included age, tumor size and grade and lymph node status (TABLE 1).

Ancestral analysis

The Brazilian population is highly heterogeneous and comprise individuals of several different genetic ancestry [57-60]. To obtain a genomic based information of the population studied in relation to ethnicity, a subset of the patients (15 patients) was genotyped using the SNP chip Illumina Infinium QC Array (Illumina Inc., CA), which contains 15,949 markers (including ~3,000 ancestral informative markers (AIMs). The genotype calling was performed as we previously described [54], using the GenomeStudio Software v. 2011.1. SNPs with MAF ≤ 0.01 were excluded from analysis. The data obtained was subsequently merged with the 1000 Genomes Project phase 1 (n =1,902 samples) dataset [61], which present an overlap of 14,718 variants with the one from our study. Finally, Principal Components Analysis (PCA)

was performed using PLINK 1.9 [62], which uses the EIGENSTRAT method [63] to calculate model ancestry differences between different samples. Based on the result of PC1 and PC2 the patients that were genotyped for this analysis clustered with the European (EUR) defined group from the 1000 Genome Project as well as with the Admixed Americans (AMR) main group, mainly composed of Colombians and Mexicans. This data was not surprising considering the South of Brazil presents high concentration of European individuals. It is of note, that the patients with TNBC and non-TNBC cases were equally distributed within these genotyped-based clusters (FIGURE S1).

DNA and RNA isolation

Prior to DNA and RNA isolation, the FFPE specimens were evaluated by the pathologist for the presence of at least 80% of tumor cells. Specific tumor areas were microdissected from unstained 10μm FFPE tissue sections and used for the subsequent molecular analysis. Consecutive tissue sections from the same tissue blocks were used to isolate DNA and RNA ensuring a direct correlation of miRNA and DNA copy number data, as we previously performed [54].

DNA isolation was performed using phenol-clorophorm protocol optimized for FFPE material [64] and RNA isolation was performed using MasterPureTM Complete DNA and RNA Purification kit (Epicentre Biotechnologies) following manufacture protocol. DNA and RNA quantity and quality were assessed using NanoDropTM Spectrophotometer (Thermo Scientific Inc.) and the Bioanalyzer (Agilent Technologies Inc.), respectively.

Array-CGH analysis

Genome-wide copy number profiling was performed by array-CGH using the SurePrint G3 Human CGH Microarray (Agilent, Santa Clara, CA) according to our previous protocol for FFPE samples[64]. DNA isolated from peripheral blood from multiple normal individuals was used as control (reference) DNA. Control and case samples were directly labeled using the Bioprimer a-CGH Genomic Labeling kit and hybridized to the arrays for 40 hours. The arrays were scanned using Scanner Agilent G2565CA, and the data extracted using Feature Extraction (FE) software v10.10 (Agilent Tech. Inc.). The Agilent Cytogenomics v.3.0 software (Agilent Technologies Inc., Santa Clara, CA), was used to analyze the data, using the

algorithm ADM-2, threshold of 6.0 and an aberration filter with a minimum of 3 probes. Gene amplifications and deletions were defined as minimum average absolute log2 ratio (intensity of the Cy5 dye (reference DNA)/intensity of the Cy3 dye (test DNA) value of >0.25 and <-0.25, respectively. The number of calls (total significant number of CNAs) and the specifically affected cytobands were obtained from the generated aberration interval base reports (Agilent Cytogenomics v.3.0).

MiRNA expression analysis

Global miRNA expression profiling was performed using NanoString nCounter technology Human v2 miRNA Expression Assay, as we previously performed [54]. This specific assay contains 800 endogenous miRNAs, six positive miRNA assay controls, six negative miRNA assay controls, and five housekeeping transcripts (ACTB, B2M, GAPDH, RPL19, RPLP0). Raw miRNA expression data was pre-processed and normalized using NanoString's nCounter RCC collector and nSolver v2 software respectively. Fold changes, represented on the log2 scale (log2FC), were calculated for all differentially expressed miRNAs. Adjusted p-values were used to rank miRNAs of interest, and correction for multiple tests was done by the Benjamini-Hochberg method to estimate False Discovery Rate (FDR). Unsupervised and supervised hierarchical cluster analysis was performed on significantly differentially expressed miRNAs ($P<0.05$) among the patients' subtypes, using Pearson's correlation coefficient and average linkage on the Multiexperiment Viewer software (MeV 4.9.0).

Integrated analysis of array-CGH and miRNA data

Direct integration of the most differentially expressed miRNAs associated with the TNBC subtype with CNAs from the same tissue samples was performed using two distinct approaches, as previously described [54]: 1. Mapping of the miRNAs at the cytobands with high levels of CNAs and further selection based on their concordance level (i.e. cytoband with gains/amplifications of copy number/ up-regulated miRNA expression and/or losses/deletions of copy number/ down-regulated miRNA expression). Only the significant DNA segments affected by CNAs that were present in more than 33% of the cases, as identified in the aberration interval base reports (Agilent Cytogenomics v.3.0) were considered in this analysis. The location of each miRNA was determined using miRBase (http://www.mirbase.org); 2. Identification of

common gene targets of the selected miRNAs above, that may be affected by both CNAs and miRNA expression alterations. For this second approach, for the previously selected miRNAs, gene targets were queried using the available miRNA target databases (Diana micro-T-CDS v.5.0 (diana.imis.athena-innovation/gr/DianaTools/index.phpr=microT_CDS/index) 65, miRDB (http://www.mirdb.org/miRDB/)66 and TargetScan Release 7.1 (http://www.targetscan.org/vert_71/)67; only miRNA target genes that were present in two out of the three miRNA databases were selected.

Biological function and pathway analysis

In order to assess the potential impact of the deregulated miRNAs identified above in cancer associated biological processes and pathways, Diana miRPath v.3.0 was used (http://snf-515788.vm.okeanos.grnet.gr/) [68]. Enrichment analysis of multiple miRNA target genes comparing each set of miRNA targets to all known KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways was obtained and selected by significant p-value ($p < 0.05$) and cancer-associated biological functions. CyKEGG Parser [69], GENEMANIA [70] and CyTargetLinker [71] applications were used to create KEGG pathways including miRNAs and respective target genes on Cytoscape 3.5.1 software [72].

Receiver operating characteristic (ROC) curve analysis

Receiver operating characteristic (ROC) curves, by calculating the area under the curve (AUC), was used to identify the power of the selected miRNAs in discriminating the TNBC and non-TNBC cases. Sensitivity was plotted against 1-specificity for the binary classifier (TNBC and non-TNBC). An AUC of 100% denotes perfect discrimination by the miRNA, whereas an AUC of 50% denotes complete lack of discrimination by the miRNA. AUCs and 95% corresponding confidence intervals were calculated for each miRNA and for the combined panel of 17 miRNAs.

Analysis of clinical-pathological variables

The Student t test was used to analyze the differences of the mean age at diagnosis and tumor size in the TNBC and non-TNBC group of cases. The Chi-square ($\chi 2$) test was used to evaluate tumor stage and grade and the Fisher Exact

test was used to evaluate lymph node, local recurrence, and distant metastasis status in both tumor subtypes. Significance level was considered to be $P<0.05$.

Kaplan Meier Plot analysis

The KM Plotter Tool (http://kmplot.com/analysis/) was used to calculate hazard ratios, confidence intervals, and log-rank P values for each of the selected 17 miRNAs in relation to survival in the aggregated breast cancer clinical studies from The Cancer Genome Atlas (TCGA) database and Molecular Taxomony of Breast Cancer International Consortium (METABRIC) database (selected specifically for the TNBC subtype).

RESULTS

Copy number profiling

Copy number analysis was performed by array-CGH in 25 cases of TNBC and 16 cases of non-TNBC cases. A total number of 292 and 204 CNAs (as measured by the “number of calls” per the aberration interval base reports) were identified in the TNBC and non-TNBC cases, with an average of 18.25±5.98 and 14.57±3.09 CNAs per case, respectively. These differences were not significant at $p<0.05$ (Unpaired t test; $p=0.5902$).

In the TNBC cases, 64.0% of the cases (16/25) showed CNAs; the most frequent cytobands affected were: 8q21.3-q24.3 (43.8 % of the cases), 3q24-29 and 6p25.3-p12.2 (37.5% of the cases) and 1q21.1-q44, 5q11.1-q22.1, 11p13-p11.2, 13q12.11-q14.3, 17q24.2-q25.3 and Xp22.33-p11.21 (31.3% of the cases) (TABLE 2). Sixty-percent of these cytobands were affected by gains of copy number and 40% by losses. The identification of the genes mapped in these cytobands, per the aberration interval base reports, revealed a number of 4,585 genes (ranging from 51 to 989 genes per cytoband). In the non-TNBC cases, 87.5% of the cases (14/16) showed CNAs, and the most frequent cytobands affected were 1q21.11-q44 (64.3% of the cases), 8q21.13-q24.3 (50.0% of the cases), 7q11.21-q36.3 (42.9% of the cases), followed by 6p22.3-p21.2, 17q21.32-q25.3 and 20q13.31-q13.33 (35.7% of the cases) (TABLE 2). In these cases, only gains of copy number were observed. A total of 3,057 genes (ranging from 324 to 1029 genes per cytoband) were found

located in these affected cytobands. The combined array-CGH profiling of the TNBC and non-TNBC cases are presented in FIGURE 1.

Global miRNA expression profiling

Global miRNA expression profiling was successfully performed using the Nanostring technology in 19 and 24 cases of the TNBC and non-TNBC cases, respectively. A number of 163 miRNAs were identified with significantly differentially expression between these two groups (p-value ≤0.05) (TABLE S1). Unsupervised (UHC) and Supervised hierarchical clustering (SHC) analysis of these miRNAs showed a more concise cluster of the TNBC cases, while most of the non-TNBC did not clustered uniformly (FIGURE 2A and 2B respectively). Of the 163 miRNAs, 87 showed increased miRNA expression and 76 showed decreased miRNA expression in TNBC when compared to the non-TNBC cases. The miRNAs that presented with the highest changes (log2FC>2) in expression between these groups were: miR-187-3p, miR-601, miR-663a, miR-421, miR-378b, miR-1305 (TABLE 3). To explore the function of each of the 163 identified diferentially expressed miRNAs, we used DIANA miRPath analysis to perform pathway enrichment analysis with the KEGG database. Among the top 15 pathways identified, based on p value, were the ones related to ECM-receptor interaction, adherens junction, biosynthesis O-glycan mucin, morphin addiction, proteoglycans in cancer (TABLE S2).

Integration of miRNA and CNAs analysis

1. Mapping of miRNAs in cytobands affected by CNAs

In order to identify miRNAs that could be affected by CNAs in TNBC, the genomic location of the initial set of 163 miRNAs found differentially expressed between the TNBC and non-TNBC cases was verified. Forty-five of them (27.6%) were located in the cytobands mostly affected by CNAs in the same TNBC cases profiled by array-CGH as described above. From these 45 miRNAs, 17 (37.8%) presented expression levels in concordance with the observed CNAs at their respective genome locus (TABLE 4, FIGURE 3). The analysis of each individual case per subtype, showed that in the TNBC group, 3 to 12 of the selected 17 miRNAs were observed with altered expression levels, with an average of 6.32±0.52 miRNAs with alterations per case. In the non-TNBC group, 0 to 10 of these miRNAs presented expression alterations, with an average of 3.58±0.51 miRNAs with alterations per case. This difference was statistically significant (p=0.0006), showing

that the TNBC cases presented individually, a higher number of miRNA expression alteration levels of the selected 17 miRNA panel. Unsupervised Hierarchical Analysis using expression levels of these miRNAs was able to distintincly cluster all but four non-TNBC samples (FIGURE 4).

2. Identification of miRNA target genes with concomitant miRNA expression and CNAs

In addition to the mapping of miRNAs in the cytobands with CNAs as described above, we identified miRNA gene targets that could be potentially affected by the miRNAs mapped in the cytobands affected by CNAs, as identified above. Using miRNA target prediction databases, we identified a total of 10,675 targets of the 17 miRNAs that were mapped in the most affected cytobands, as predicted from more than two independent databases. MiR-608, mapped at 10q24.31, was the miRNA with the highest number of target genes (5,268 genes) and miR-129-2-3p, mapped at 11p11.2, was the miRNA with the lowest (131 genes), Next, these miRNA targets were matched with the genes that were mapped in the cytobands affected by CNAs, as generated by the aberration interval base report (array-CGH gene list) (n= 4,585). The integration of these data revealed 2,098 common genes (FIGURE 5), i.e. genes that could be potentially affected by both CNAs and miRNA expression regulation. This integration reduced the initial number of the total miRNA targets in 80.3% (from 10,675 to 2,098).

3. Functional analysis of the miRNAs/mRNA pairings affected by miRNA expression and CNAs.

In order to identify the main function of the 17 selected miRNAs and correspondent miRNA targets identified to be affected by both miRNA expression and CNAs, we performed pathway enrichment analysis (KEGG pathways) using Diana miRPath v.3.0 analysis. We found two major miRNAs clusters among the 17 miRNAs regulating these pathways, one of the clusters was formed by four miRNAs (miR-539, miR-548, miR-607 and miR-944) and the other by thirteen miRNAs (miR-129-2-3p, miR-135b-5p, miR-188-5p, miR-323a-5p, miR-323b-3p, miR-323b-5p, miR-342-3p, miR-378c, miR-608, miR-634, miR-668-3p, miR-1260a and miR-1275) (FIGURE 5). The top 10 pathways identified were the Axon guidance (hsa04360), Glycosaminoglycan biosynthesis-chondroitin sulfate/dermatan sulfate (hsa00532), Mucin type-O-Glycan biosynthesis (hsa00512), Thyroid hormone signaling pathway

(hsa04919), Signaling pathways regulating pluripotency of stem cells (hsa04550), Proteoglycans in cancer (hsa05205), Wnt signaling pathway (hsa04310), Hippo signaling pathway (hsa04390), Ras signaling pathway (hsa0414) and Pathways in cancer (hsa05200). Remarkably, all the 17 selected miRNAs are presented in four of the top 10 pathways identified. The less representative pathways by these miRNAs were the Mucin type-O-Glycan biosynthesis and the Glycosaminoglycan biosynthesis-chondroitin sulfate/dermatan sulfate, with 29.4% and 41.2% of the miRNAs, respectively (**TABLE S3**).

Identification of KEGG pathways potentially affected by the 17 miRNAs were also conducted considering only the 2,098 putative target genes presented in CNAs previously selected in the integrated analysis. As a result, three KEGG pathways were found to be affected: Glycosaminoglycan biosynthesis - chondroitin sulfate / dermatan sulfate (hsa00532, p= 0.01967), Biosynthesis of unsaturated fatty acids (hsa01040, p=0.01967) and Hippo signaling pathway (hsa04390, p= 0,01967)(**TABLE 5**) generated a KEGG pathway with the 7 miRNAs (upregulated miRs: miR-944, miR-135b-5p and miR-1275 and downregulated miRs: miR-548p, miR-607, miR-323b-3p and miR-342-3p out of the 17) and their corresponding target genes found to specifically affect this pathway. Using Cytoscape we observed that these miRNAs can act targeting key regulators of Hippo Signaling Pathway, potentially damaging cell control over several biological functions (**FIGURE 7**). MiR-944 was involved in Tight junction and Adherens junction by regulating CRB1, MPP5 and WWTR1 genes. Down-regulated miRs were shown to target known proto-oncogenes as CCND2, YWHAZ, SOX2 and also genes related to Wnt signaling pathway (WNT8B, FZD6 and CSNK1D).

Individual and combined discriminatory power of the 17 miRNA panel among the TNBC and non-TNBC cases

Each miRNA composing the identified 17 miRNA panel was evaluated in relation to its power in discriminating the TNBC and non-TNBC cases of the Brazilian patients (**FIGURE 8**). ROC analysis showed that 70% of the miRNAs presented an Area Under the Curve (AUC) value superior than 0.7. The miR-539-5p, miR-634, miR-323a-5p, miR944 and miR-1260a presented the highest discriminatory power, with AUC values ranging from 0.742 (miR-1260a) to 0.780 (miR-539-5p). The combined analysis of the panel showed a AUC value of 0.953, demonstrating a

robust power of the 17 miRNA panel in discriminating the TNBC and non-TNBC cases of the cases studied.

Analysis of the clinical-pathological variables of the patients studied

This analysis was conducted independently of the molecular data to determine whether there were differences in the clinical-pathological variables at diagnosis from both group of the patients. Age, tumor size were not found to have significant differences between the groups analyzed ($p>0.05$): average age of 54.25+3.59 and 58.15+2.34 and tumor size of 3.1+0.47 and 3.48+0.39 for the TNBC and non-TNBC patients, respectively. TNBC patients presented higher frequency of grade III tumors when compared to non-TNBC patients (11/23 and 5/27 respectively, $p=0.0268$), however non-TNBC patients presented more lymph nodes metastasis than the non-TNBC patient (21/24 and 10/20, respectively $p=0.0066$).

Association with survival (Kaplan Meier Plot (KM Plot) database

The miRNA expression of 13 out of the 17 miRNAs observed differentially expressed in the TNBC and non-TNBC cases of our study, were previously associated with TNBC survival in the analysis of the KMPlot dataset (**TABLE 6**). This dataset contains a large set of patients with basal-like/TNBC cases from the TCGA and METABRIC cohorts of patients. In these TNBC cohort, higher expression levels of miR-129-3p, miR-135b-5p, and miR-634, as observed in our study, were significantly associated with reduced overall survival (OS) (months) (**FIGURE 9A**). Lower expression levels of miR-548p, miR-607 and miR-668, as we also observed in our study, were also associated with lower survival (**FIGURE 9B**). For the remaining 7 miRNAs of our panel included in the TCGA and metabric TNBC cohort of patients, no significant differences in OS was found to be associated with their expression levels ($p<0.05$) (**FIGURE 9C**).

DISCUSSION

MiRNA expression patterns have been shown, as gene expression profile, to discriminate the major breast cancer subytpes in the analysis of both cell lines and clinical cases [73, 74]. Race and ethnicity has also been shown to play a factor in

miRNA expression patterns; most of the data is however in germline miRNA alterations and their associated risk with breast cancer [75].

In this study, using an integrated analysis of genome-wide copy number and global miRNA profiling we identified a panel of 17 miRNAs differentially expressed among the TNBC and non-TNBC groups of patients from Latin America, Brazil. ROC analysis showed a remarkable high power of this panel in discriminating the TNBC and non-TNBC cases.

In TNBC, 5 of these miRNAs were observed up-regulated (miR-135b-5p, miR-944, miR-1275, miR-129-2-3p and miR-534) and 12 down-regulated (miR-548p, miR-607, miR-608, miR-378c, miR-1260a, miR-342-3p, miR-323a-5p, miR-539-5p, miR-668, miR-323b-3p, miR-323b-5p and mIR-188-5p) when compared to the non-TNBC patients. These miRNAs were selected among the significant differentiated expressed miRNAs among these subtypes and based on their localization in genomic regions affected by copy number alterations (amplifications and/or deletions), as we detected in the same TNBC specimens by array-CGH analysis. These regions involved mainly the 1q, 3q, 6p, 10q, 11p, 14q, 17q and Xp cytobands, some of which recurrently described in TNBC [76].

Four out of five (except miR-524) of our up-regulated miRNAs were previously reported with altered expression in breast cancer studies: miR-135b was found upregulated both in breast cancer cell lines and clinical cases, and shown to confer higher proliferation, migration and invasion activity in MCF-7 and MDA-MB-231 cells [77, 78]. Higher expression of miR-135b was also associated to ER-negative breast cancer [79], and found to regulate ERα protein levels by interacting with its 3'UTR regions [78]. Similar oncogene activity was also shown for miR-944 in the MCF7 and MDA-MB-231 cells, in addition to promote resistance to cisplatin, by regulating the BNIP3 gene [80]. Upregulation of miR-1275 expression, was particularly found in breast cancer of young women (<35 years) when compared to older women [81] and also found hypomethylated in samples of healthy individuals who developed breast cancer when compared to individuals who remained healthy [82]. Although these evidences warrants further validations, it is suggested that this miRNA might present a role in the initiation of breast cancer and could be used as a preditive biomarker. Finally, miR-129-2-5p, that targets the Progesterone Receptor (PR) has been described upregulated in patients with low PR levels (PR-) [83]. This target

suppression, is compatible with our finds, showing its up-regulation in the TNBC cases when compared to the non-TNBC cases.

Among the twelve down-regulated miRNAs, only miR-548 and miR-539 were previously reported with altered expression in breast cancer, with possible function as tumor supressors. The miR-548 ‘s anti-oncogenic activity was observed in two different studies in breast cancer, as its up-regulation was found to inhibit cell proliferation and induce apoptosis by targeting Enoyl Coenzyme A Hydratase short chain 1 (ECHS1) and Nuclear Paraspeckle Assembly Transcript (NEAT) genes, respectively [84, 85]. MiR-539 was also found to interfere with proliferation, in addition to migration activity, by regulating EGFR expression in MCF-7 and MDA-MB-231 cell lines [86].

Identification of the main pathways and networks potentially affected by the 17 miRNAs and putative target genes resulted in forty-six KEGG pathways, among them Axon guidance, Glycosamin biosynthesis – chondroitin sufate/ dermatan sulfate and Mucin type O-GLycan biosynthesis. Interestingly, all of the 17 miRNAs of the identified panel, were found to regulate gene targets associated with cancer signaling pathways, such Wnt, Ras, ErbB and Rap1. It is of note, that 15 out of the 17 of the miRNAs (were found to affect the Hippo signaling pathway.

A second integration analysis of the genes that were potentially affected by the CNAs (identified by the copy number profiling) and the corresponding mRNA targets of the 17 miRNAs located at the same affected cytobands, revealed a total of 2,098 genes that may be commonly affected by both of these mechanisms. Among them, genes critical for TNBC tumorigenesis and pathogenesis (and miRNA biogenesis), such as CDKN1A, DICER1, ETV6, IGF1R, MYC, PIK3CA. Using the selected genes of this analysis, three signaling pathways were identified to be involved, including the Hippo signaling pathway, observed in the first data integration approach above. Seven out of our 17 miRNAs (miR-1275, miR-135b-5p, miR-323b-3p, miR-342-3p, miR-5395p, miR-548p, miR-607 and mIR-944) were found related to this pathway, potentially targeting 18 genes: DVL3, PPP2RD2, LATS2, TCF7L2, CCND2, FZD6, WNT8B, BTRC, CSNK1D, FRMD6, STK3, YWHAZ, MPP5, SAV1, SOX2, CRB1, GDF6 and WWTR1. A search on miRTarBase database showed that few of these interactions were validated: TCF7L2 and miR-539, and YWHAZ and miR-548p, miR-607 and miR944 by PAR-CLIP (photoactivaTABLE-ribonucleoside-enhanced crosslinking and imminoprecipitation) method in the same study [87]. Considering the

critical role of the Hippo signaling pathway in regulating cell proliferation and apoptosis, and other tumorigenic process, it is relevant to pursue downstream functional studies to confirm these interactions and determine their biological significance to the TNBC phenotype.

CONCLUSION

In conclusion, our study aimed to understand the biological differences among TNBC and NTNBC of Brazilian patients. Through our integrated analysis of DNA copy number alterations and miRNAs expression levels we suggest the use of this 17 miRNA panel as biomarker for diagnostic and prognostic purposes, and also for further new target therapies development.

TABLE 1 - Clinical information regarding TNBC and NTNBC samples

	TNBC (n=27)	NTNBC (n=27)
Age	54.25 ± 3.59	58.15 ± 2.34
	16-83	34-88
	(n=20)	(n=27)
Tumor size (cm)	3.10 ± 0.47	3.48 ± 0.39
	0.9-10.5	0.7-8.0
	(n=23)	(n=26)
Grade		
I/II	12	22
III	11	5
Lymph node		
Positive	10	21
Negative	10	3

TABLE 2 - Most common cytobands and corresponding number of genes affected by CNAs in the TNBC and non-TNBC cases analyzed

TNBC cases						
Chr	**Cytoband**	**Start**	**Stop**	**CNA**	**Cases (%)**	**# Genes***
chr1	q21.1 - q44	144988715	247737874	gain	5 (31.25%)	989
chr3	q24 - q29	148802495	194903188	gain	6 (37.50%)	242
chr5	q11.1 - q22.1	49690172	111370979	loss	5 (31.25%)	245
chr6	p25.3 - p12.2	2117686	52103799	gain	6 (37.50%)	609
chr8	q21.3 - q24.3	88884192	146066584	gain	7 (43.75%)	289
chr10	q23.31 - q26.3	89507004	135372492	loss	4 (25.00%)	360
chr11	p13 - p11.2	34322106	46565735	gain	5 (31.25%)	51
chr12	p13.33 - p13.1	309062	14132896	gain	4 (25.00%)	205
chr13	q12.11 - q14.3	19703703	53876286	loss	5 (31.25%)	216
chr14	q21.1 - q32.32	38723471	103447263	loss	4 (25.00%)	485
chr15	q26.1 - q26.3	90276459	102241406	gain	4 (25.00%)	63
chr16	q11.2 - q22.1	46693731	67933130	loss	4 (25.00%)	169
chr17	q24.2 - q25.3	65989022	80993001	gain	5 (31.25%)	232
chr19	p13.11 - p12	18266482	21108358	gain	4 (25.00%)	66
chrX	p22.33 - p11.21	1314894	58051765	loss	5 (31.25%)	364
Non-TNBC cases						
Chr	**Cytoband**	**Start**	**Stop**	**CNA**	**Samples (%)**	**# Genes**
chr1	q21.1 - q44	145103876	249118400	gain	9 (64.28%)	1029
chr6	p22.3 - p21.1	18093033	43409896	gain	5 (35.71%)	458
chr7	q11.21 - q36.3	62516153	158909738	gain	6 (42.86%)	687
chr8	q21.13 - q24.3	82193925	146280020	gain	7 (50.00%)	324
chr17	q21.32 - q25.3	46048958	81029941	gain	5 (35.71%)	445
chr20	q13.31 - q13.33	55212094	62893189	gain	5 (35.71%)	114

** cytobands locations, positions, size and # genes and miRNAs affected by CNAs based on the aberration interval base reports (Agilent Cytogenomics v. 3.0)

FIGURE 1 - Genomic view/penetrance plot of the array-CGH profiling of the 25 TNBC (A) and 16 non-TNBC (B) cases analyzed by array-CGH (Agilent Genomic Workbench 7.0) showing the copy number alterations (CNAs) (vertical blue lines) observed in each chromosome. Red peaks indicate copy number gains and green peaks indicate copy number losses

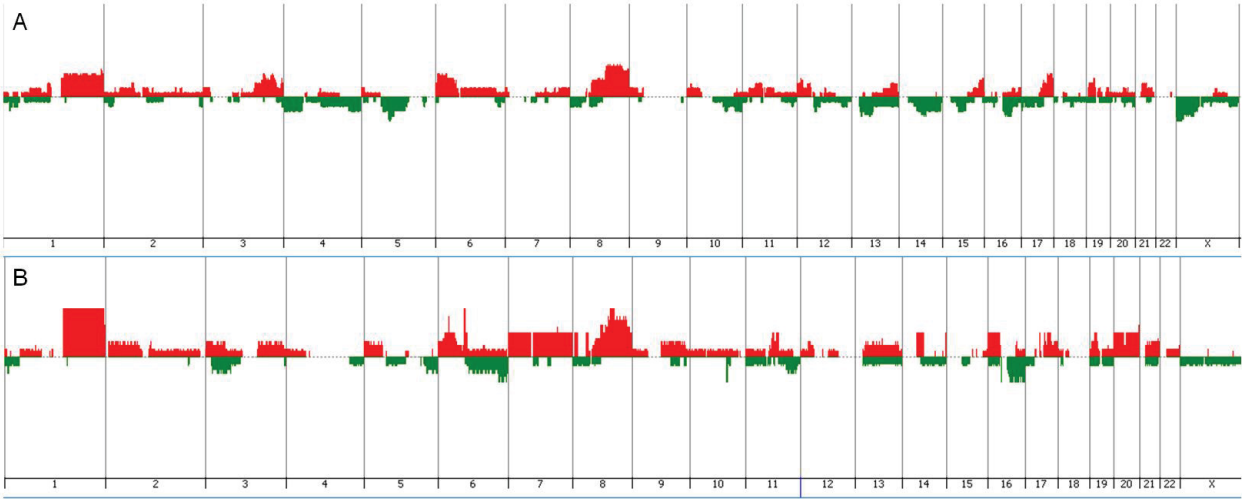


FIGURE 2 - Unsupervised (A) and Supervised (B) Hierarchical Clustering Analysis applied to the TNBC (blue bars) and non-TNBC (green bars) cases analyzed. Up-regulated miRNAs (yellow) and down-regulated miRNAs (blue). (MeV 4.9, Pearson Correlation, p<0.05)

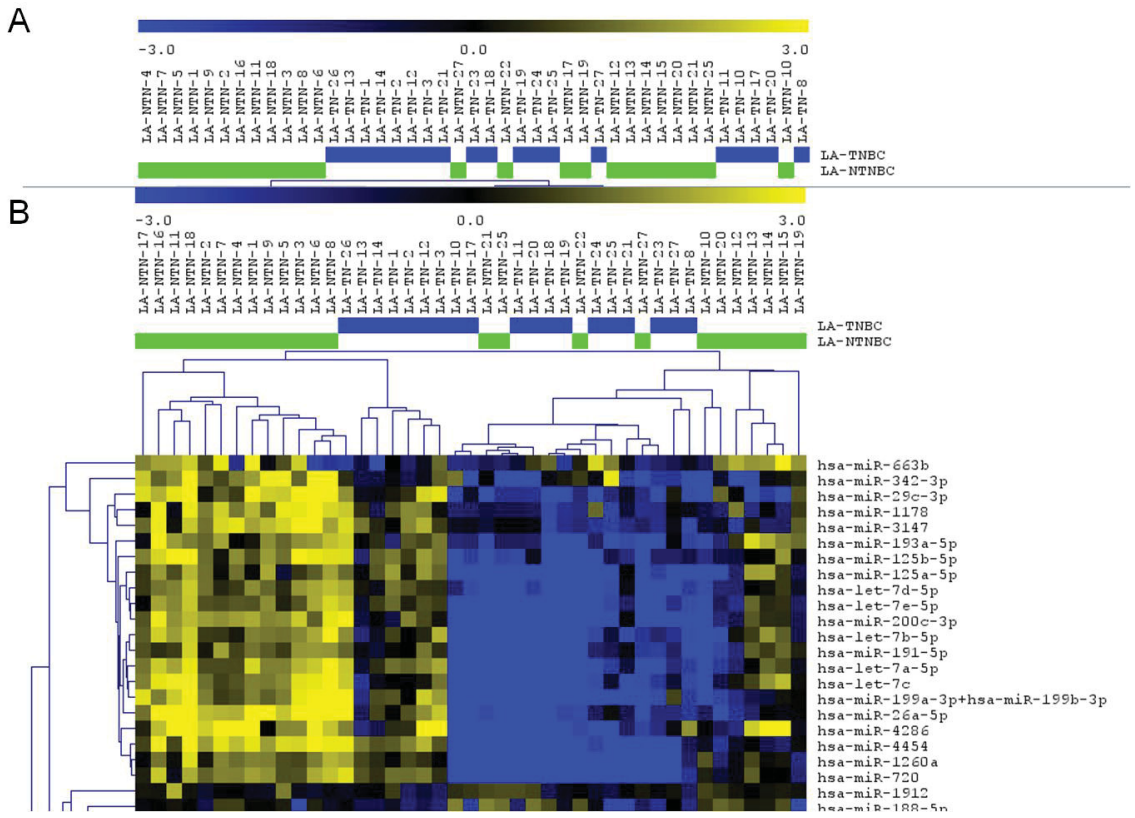


TABLE 3 - Top 15 miRNAs observed up and down-regulated in the TNBC group of patients compared to the non-TNBC group (ordered by Log2FC value)

miRNA	Log2FC	p	miRNA	Log2FC	p
hsa-miR-187-3p	3.64	0.009	hsa-miR-720	-3.98	0.012
hsa-miR-601	2.68	0.02	hsa-miR-1260a	-3.30	0.004
hsa-miR-663a	2.61	0.006	hsa-miR-4286	-3.09	0.008
hsa-miR-421	2.28	0.005	hsa-miR-4454	-2.72	0.047
hsa-miR-378b	2.26	0.003	hsa-miR-200c-3p	-2.45	0.013
hsa-miR-1305	2.22	0.019	hsa-let-7b-5p	-2.24	0.029
hsa-miR-135b-5p	1.93	0.001	hsa-miR-199a-3p	-2.19	0.032
hsa-miR-638	1.90	0.001	hsa-miR-199b-3p	-2.19	0.032
hsa-miR-708-5p	1.87	0.005	hsa-let-7c	-2.19	0.007
hsa-miR-1290	1.83	0.021	hsa-let-7a-5p	-2.09	0.041
hsa-miR-548z	1.68	0.036	hsa-miR-125b-5p	-1.87	0.023
hsa-miR-2117	1.63	0.002	hsa-miR-26a-5p	-1.85	0.034
hsa-miR-567	1.63	0	hsa-let-7e-5p	-1.80	0.028
hsa-miR-500b	1.60	0.01	hsa-miR-125a-5p	-1.73	0.048
hsa-miR-526a	1.55	0.005	hsa-miR-193a-5p	-1.72	0.007

FIGURE 3 - Penetrance plot of the array-CGH genomic profiling of the TNBC cases (n=25) from the patients analyzed with the corresponding mapping of the 17 miRNAs of the identified panel. Vertical lines represent each chromosome number. Red peaks indicate copy number gains and green peaks indicate copy number losses. The miRNAs expression levels up- and down regulated are in red and green color boxes, respectively (Agilent Genomic Workbench 7.0).

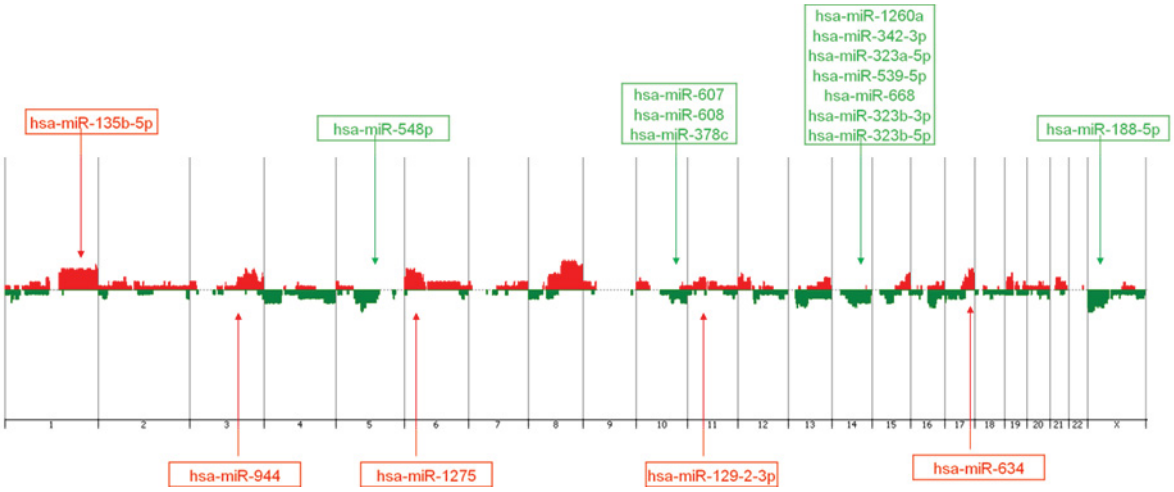


TABLE 4 - Seventeen differentially expressed miRNAs among the TNBC and non-TNBC group of patients, with expression levels in concordance with copy number alterations (CNAs) (presented by chromosome numerical order).

miRNAs	Cytoband	Start	Stop	CNA	miRNA expression	Log2FC	Adj p
hsa-miR-135b-5p	1q32.1	205448302	205448398	amp	up-regulated	1.93	0.001
hsa-miR-944	3q28	189829922	189830009	amp	up-regulated	1.13	0.004
hsa-miR-548p	5q21.1	100816482	100816565	del	down-regulated	-0.53	0.002
hsa-miR-1275	6p21.31	33999972	34000051	amp	up-regulated	0.61	0.049
hsa-miR-607	10q24.1	96828669	96828764	del	down-regulated	-0.50	0.006
hsa-miR-608	10q24.31	100974985	100975084	del	down-regulated	-0.52	0.04
hsa-miR-378c	10q26.3	130962588	130962668	del	down-regulated	-0.42	0.024
hsa-miR-129-2-3p	11p11.2	43581394	43581483	amp	up-regulated	0.77	0.036
hsa-miR-1260a	14q24.3	77266218	77266290	del	down-regulated	-3.30	0.004
hsa-miR-342-3p	14q32.2	100109655	100109753	del	down-regulated	-1.51	0.039
hsa-miR-323a-5p	14q32.31	101025732	101025817	del	down-regulated	-0.59	0.001
hsa-miR-539-5p	14q32.31	101047321	101047398	del	down-regulated	-0.68	>0.01
hsa-miR-668	14q32.31	101055258	101055323	del	down-regulated	-0.45	0.014
hsa-miR-323b-3p	14q32.31	101056219	101056300	del	down-regulated	-0.49	0.008
hsa-miR-323b-5p	14q32.31	101056219	101056300	del	down-regulated	-0.49	0.009
hsa-miR-634	17q24.2	66787072	66787168	amp	up-regulated	0.90	0.005
hsa-miR-188-5p	Xp11.23	50003503	50003588	del	down-regulated	-0.93	0.02

CNA: copy number alteration, FC: fold change

FIGURE 4.- Unsupervised Hierarchical Clustering Analysis applied to the TNBC (blue bars) and non-TNBC (green bars) cases analyzed using the selected 17miRNAs. Up-regulated miRNAs (yellow) and down-regulated miRNAs (blue). (MeV 4.9, Pearson Correlation,)

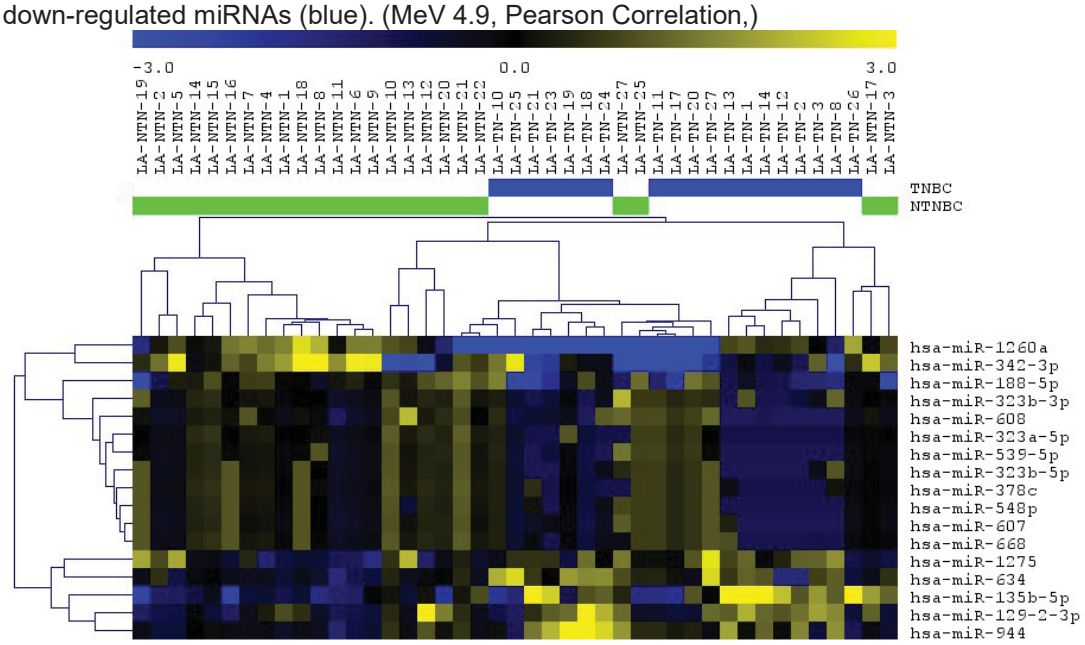


FIGURE 5 - Venn diagram showing the integration of genes located at the cytobands affected by CNAs and the corresponding target genes of miRNAs mapped in these regions. (VENNY 2.1)

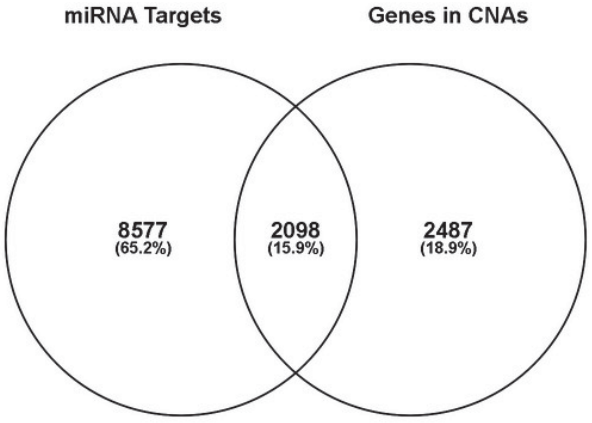


FIGURE 6 - miRNA dendogram of the 17 miRNA panel, based on miRNA-miRNA clustering (Diana miRpath, p<0.05, FDR<0.05)

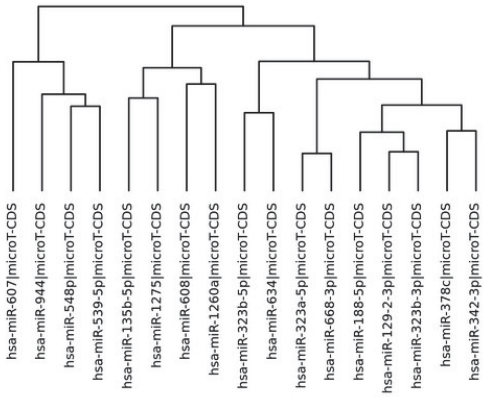


TABLE 5 - KEGG pathways potentially affected 7 miRNAs and selected target genes (DIANA mirPath 3.0)

Glycosaminoglycan biosynthesis – heparan sulfate / heparin (hsa00534, p=0,01967)	
miR-135b-5p	*EXT1*
miR-944	*EXT1*
miR-342-3p	*EXT1*
miR-1260a	*EXT2*
miR-539-5p	*EXT2*
Biosynthesis of unsaturated fatty acids (hsa01040, p=0,01967)	
miR-548p	*ACOX1*
miR-135b-5p	*ELOVL2*
miR-1275	*ELOVL2*
miR-129-2-3p	*ELOVL2*
miR-944	*HSD17B12, SCD*
miR-944	*SCD*
Hippo signaling pathway (hsa04390, p=0,01967)	
miR-1275	*DVL3, PPP2RD2*
miR-135b-5p	*LATS2, TCF7L2*
miR-323b-3p	*CCND2*
miR-342-3p	*FZD6*
miR-539-5p	*TCF7L2, WNT8B*
miR-548p	*BTRC, CCND2, CSNK1D, FRMD6, LATS2, STK3, YWHAZ*
miR-607	*CCND2, MPP5, SAV1, SOX2, YWHAZ*
miR-944	*CRB1, GDF6, SAV1, WWTR1*

FIGURE 7 - Hippo signaling pathway showing 7 out of the 17 miRNA panel and corresponding target genes. In yellow: genes targeted by one miRNA; in orange: genes targeted by more than one miRNA; in green: down-regulated miRNAs; in red: up-regulated miRNAs (Cytoscape 3.5.1)

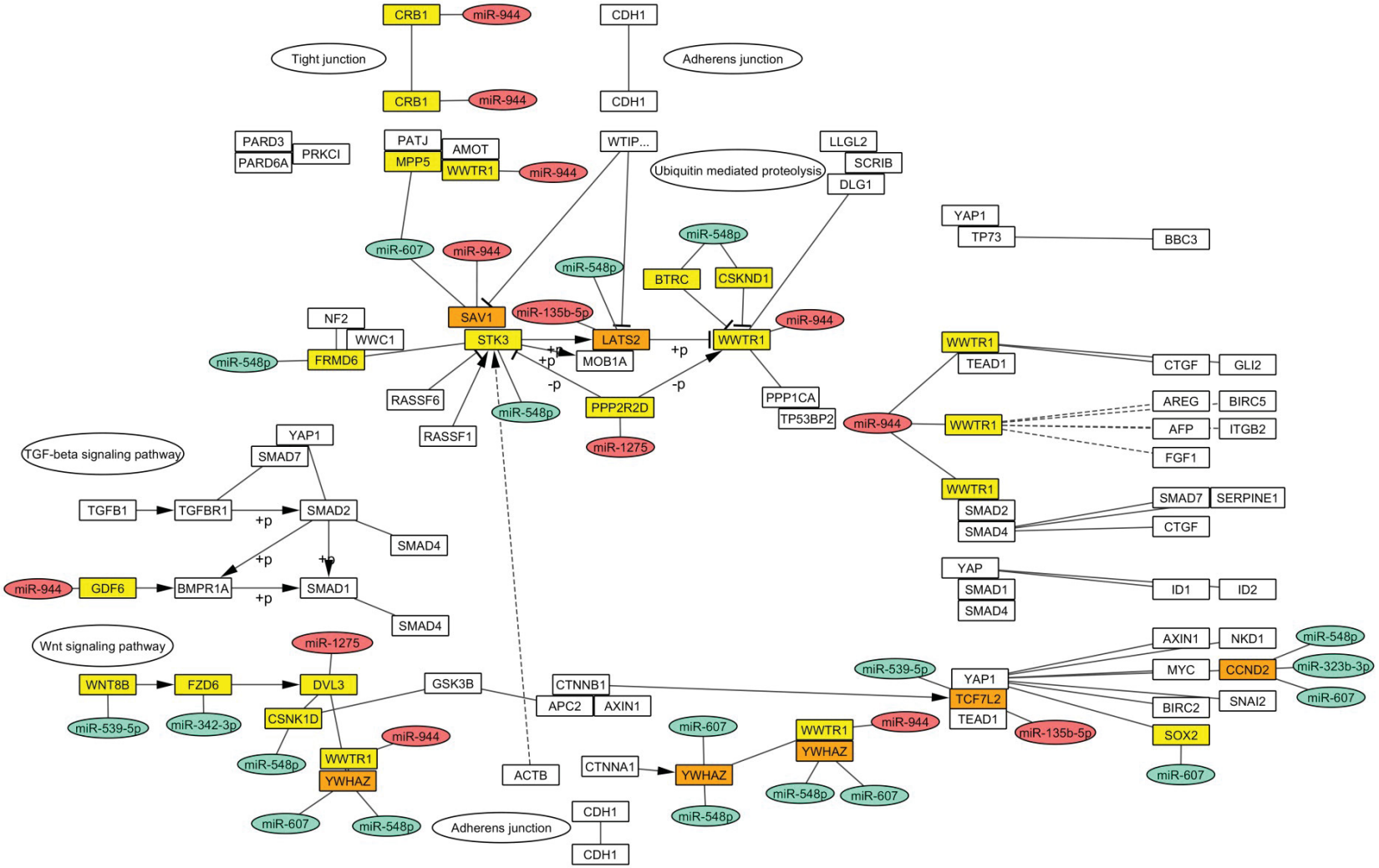


FIGURE 8 - Receiver operating characteristic (ROC) curve plots and Area Under the Curve (AUC) values of the combined and individual 17 miRNAs differentially expressed between the TNBC and NTNBC patients.

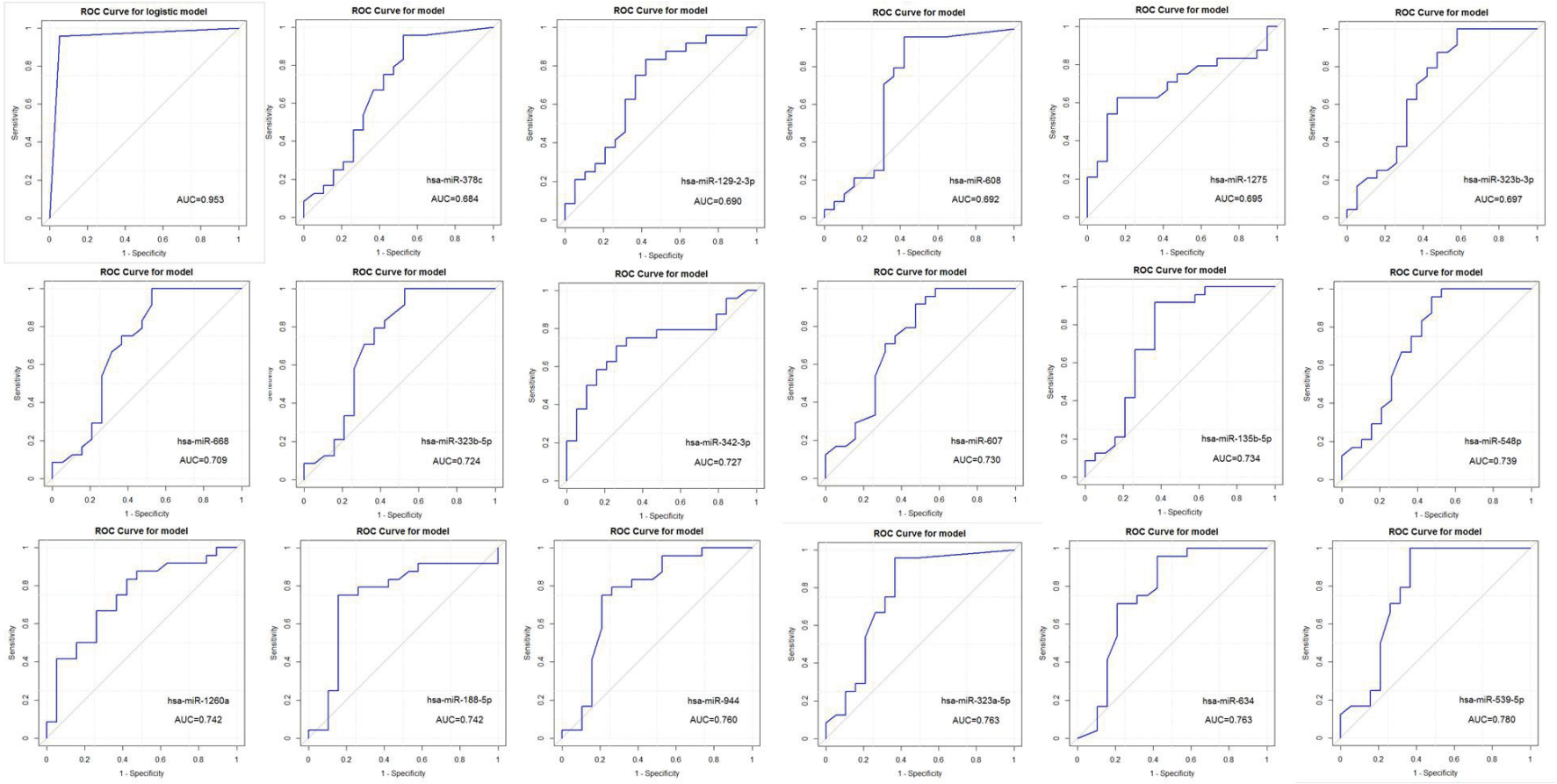


TABLE 6 - Kaplan-Meier data of 13 out of the 17 miRNA panel that discriminate the TNBC and non-TNBC samples of the METABRIC and TCGA data (KMPlot - miRpower).

miRNA	HR	IC 95%	P	Database
hsa-miR-129-2-3p	1.58	1.01 – 2.46	0.044	METABRIC
hsa-miR-135b-5p	1.56	1 – 2.44	0.049	METABRIC
hsa-miR-188-5p	1.63	0.93 – 2.86	0.087	METABRIC
hsa-miR-323a-5p	0.64	0.41 – 1	0.046	METABRIC
hsa-miR-323b-3p	0.55	0.16 – 1.83	0.32	TCGA
hsa-miR-342-3p	0.72	0.44 – 1.18	0.19	METABRIC
hsa-miR-378c	1.43	0.89 – 2.3	0.13	METABRIC
hsa-miR-539-5p	0.78	0.49 – 1.22	0.27	METABRIC
hsa-miR-548p	0.16	0.03 – 0.79	0.011	TCGA
hsa-miR-607	0.17	0.04 – 0.82	0.013	TCGA
hsa-miR-634	1.89	1.06 – 3.37	0.028	METABRIC
hsa-miR-668	0.23	0.06 – 0.89	0.021	TCGA
hsa-miR-1260a	1.26	0.79 – 2.01	0.33	METABRIC
hsa-miR-1275	0.64	0.38 – 1.07	0.084	METABRIC

FIGURE 9 - Kaplan-Meier plot for 13 out of the 17 miRNA panel (KMPlot – miRpower)

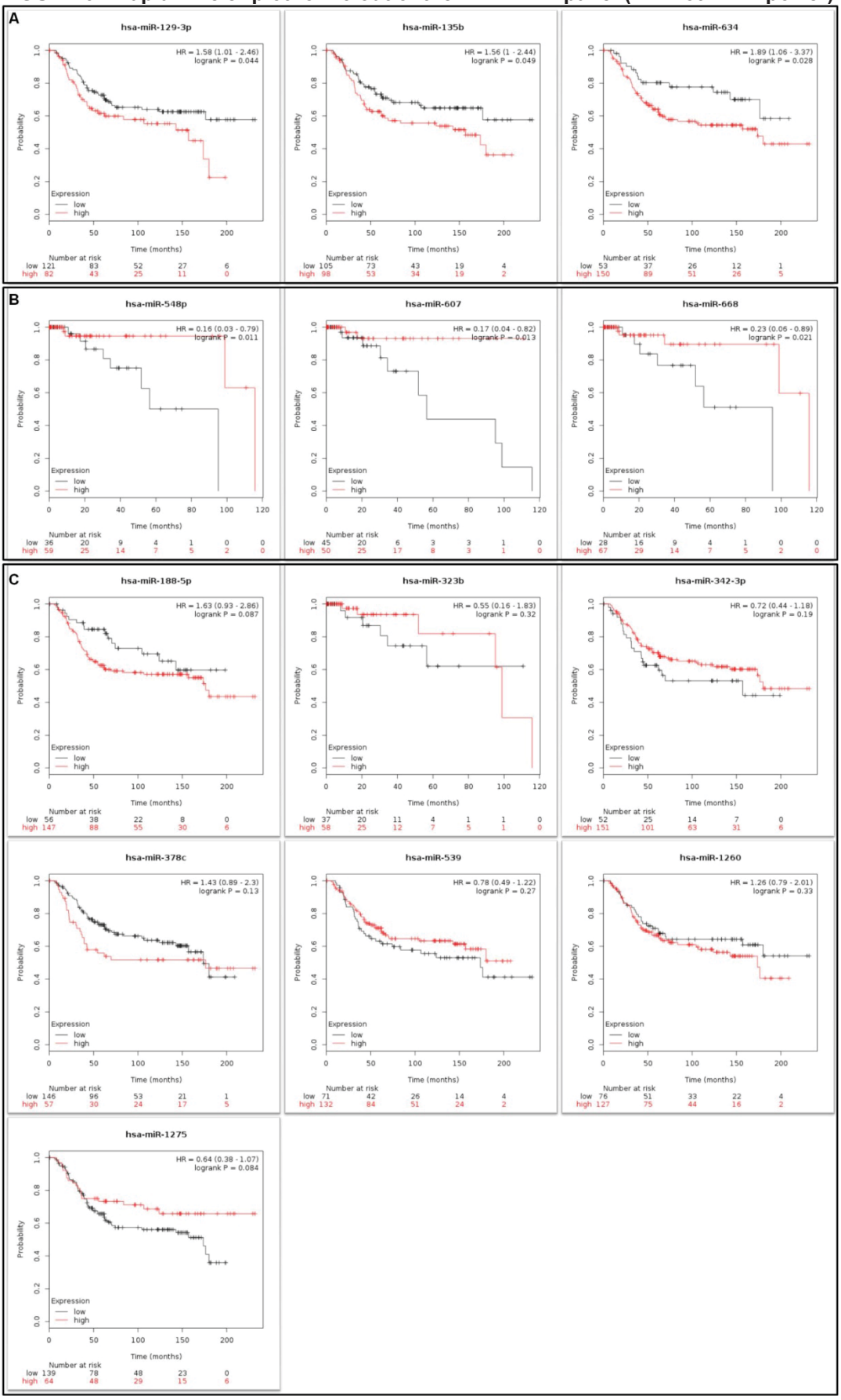


FIGURE S1 - Principal Component Analysis (PCA) plot of the cases genotyped for ancestral informative markers (AIMs). Note that the TNBC and non-TNBC Brazilian cases (orange) of this study (LABR) clustered mainly with the European (EUR) and Admixed Americans (AMR) reference groups of the 1000 Genomes Project.

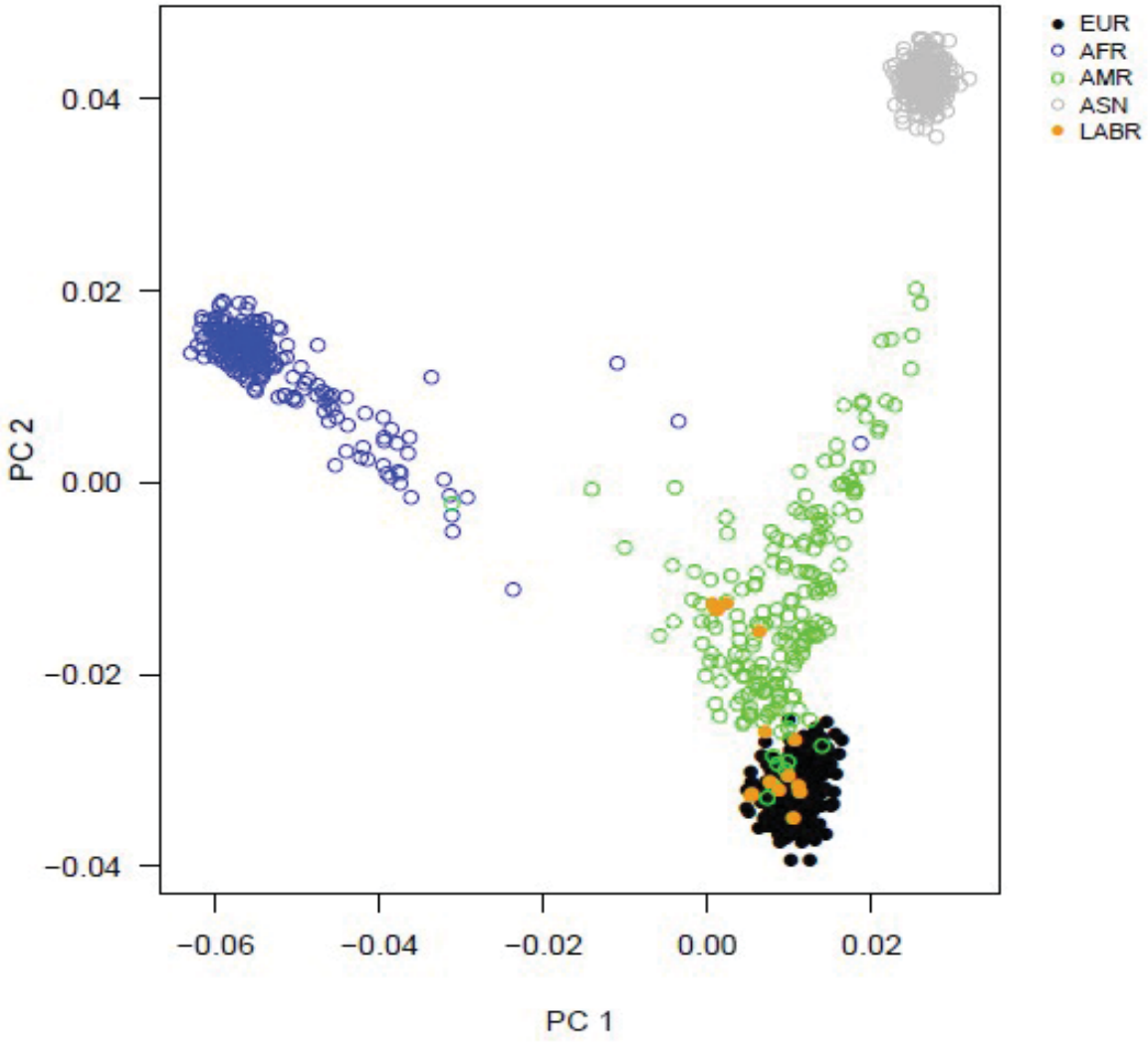


TABLE S1: The 163 differentially expressed miRNAs among the TNBC and non-TNBC groups.

miRNAs	Log2FC	p-value	miRNAs	Log2FC	p-value
hsa-miR-1299	1.241552	0.001	hsa-miR-548u	-0.50615	0.009
hsa-miR-135b-5p	1.928202	0.001	hsa-miR-577	0.968527	0.009
hsa-miR-3151	-0.56665	0.001	hsa-miR-1200	1.019496	0.01
hsa-miR-323a-5p	-0.59037	0.001	hsa-miR-184	-0.5105	0.01
hsa-miR-412	1.272243	0.001	hsa-miR-500b	1.601524	0.01
hsa-miR-555	-0.59636	0.001	hsa-miR-936	1.068307	0.01
hsa-miR-638	1.897654	0.001	hsa-miR-518c-3p	0.985269	0.011
hsa-miR-2117	1.634233	0.002	hsa-miR-943	-0.45362	0.011
hsa-miR-298	-0.6011	0.002	hsa-miR-1229	-0.43502	0.012
hsa-miR-299-3p	1.171332	0.002	hsa-miR-139-5p	-0.49376	0.012
hsa-miR-496	1.133951	0.002	hsa-miR-219-2-3p	-0.42312	0.012
hsa-miR-499b-5p	-0.56337	0.002	hsa-miR-3161	1.043253	0.012
hsa-miR-518a-3p	0.804529	0.002	hsa-miR-599	-0.516	0.012
hsa-miR-548p	-0.53097	0.002	hsa-miR-671-3p	-0.46463	0.012
hsa-miR-556-3p	-0.52457	0.002	hsa-miR-720	-3.97743	0.012
hsa-miR-610	0.811232	0.002	hsa-miR-200c-3p	-2.44742	0.013
hsa-miR-650	-0.55002	0.002	hsa-miR-323a-3p	0.628786	0.013
hsa-miR-378b	2.259158	0.003	hsa-miR-519b-5p	0.78908	0.013
hsa-miR-765	-0.56628	0.003	hsa-miR-519c-5p	0.78908	0.013
hsa-miR-1260a	-3.30145	0.004	hsa-miR-663b	-1.3021	0.013
hsa-miR-1286	1.241175	0.004	hsa-miR-1197	0.727561	0.014
hsa-miR-219-5p	1.167238	0.004	hsa-miR-548k	0.790092	0.014
hsa-miR-378g	1.050617	0.004	hsa-miR-563	-0.67788	0.014
hsa-miR-568	0.924021	0.004	hsa-miR-668	-0.44523	0.014
hsa-miR-802	0.959188	0.004	hsa-miR-762	1.056555	0.014
hsa-miR-944	1.133203	0.004	hsa-miR-1227	-0.45401	0.015
hsa-miR-1263	-0.47719	0.005	hsa-miR-637	-0.47987	0.015
hsa-miR-2276	0.826868	0.005	hsa-miR-378f	1.416524	0.016
hsa-miR-421	2.280377	0.005	hsa-miR-589-5p	-0.43799	0.016
hsa-miR-513a-3p	1.148336	0.005	hsa-miR-140-3p	-0.49958	0.017
hsa-miR-518d-5p	1.550398	0.005	hsa-miR-3144-5p	-0.47359	0.017
hsa-miR-520c-5p	1.550398	0.005	hsa-miR-379-5p	1.537607	0.017
hsa-miR-526a	1.550398	0.005	hsa-miR-514a-3p	0.70441	0.017
hsa-miR-634	0.903045	0.005	hsa-miR-876-5p	-0.43982	0.017
hsa-miR-708-5p	1.866813	0.005	hsa-miR-378e	-1.59139	0.018
hsa-miR-1269a	0.860137	0.006	hsa-miR-1305	2.21868	0.019
hsa-miR-1322	1.258168	0.006	hsa-miR-499a-3p	0.766943	0.019
hsa-miR-4425	-0.55469	0.006	hsa-miR-1296	-0.43391	0.02
hsa-miR-607	-0.4984	0.006	hsa-miR-188-5p	-0.92524	0.02
hsa-miR-620	-0.53163	0.006	hsa-miR-601	2.680344	0.02
hsa-miR-663a	2.60688	0.006	hsa-miR-1181	0.672413	0.021
hsa-let-7c	-2.18788	0.007	hsa-miR-1276	1.37538	0.021
hsa-miR-193a-5p	-1.71617	0.007	hsa-miR-1290	1.825019	0.021
hsa-miR-125a-3p	1.015675	0.008	hsa-miR-3147	-1.23049	0.021
hsa-miR-1910	-0.48273	0.008	hsa-miR-4792	0.60057	0.022
hsa-miR-323b-3p	-0.49473	0.008	hsa-miR-96-5p	0.875759	0.022
hsa-miR-4286	-3.08769	0.008	hsa-miR-125b-5p	-1.87456	0.023
hsa-miR-505-3p	0.90293	0.008	hsa-miR-18a-5p	1.053774	0.024
hsa-miR-576-5p	0.659678	0.008	hsa-miR-378c	-0.42007	0.024
hsa-miR-187-3p	3.642798	0.009	hsa-miR-519b-3p	-0.57555	0.024
hsa-miR-323b-5p	-0.48649	0.009	hsa-miR-302d-3p	-0.50028	0.026

TABLE S1: The 163 differentially expressed miRNAs among the TNBC and non-TNBC groups.

miRNAs	Log2FC	p-value	miRNAs	Log2FC	p-value
hsa-miR-548s	-0.73219	0.026	hsa-miR-1207-3p	-0.45121	0.041
hsa-miR-887	-0.44763	0.027	hsa-miR-1303	1.250427	0.041
hsa-let-7e-5p	-1.79934	0.028	hsa-miR-34c-5p	0.487047	0.041
hsa-miR-769-5p	1.013276	0.028	hsa-miR-513b	-0.46875	0.041
hsa-let-7b-5p	-2.24155	0.029	hsa-miR-1185-5p	0.689035	0.042
hsa-miR-1260b	0.788523	0.03	hsa-miR-300	0.509251	0.043
hsa-miR-548ak	0.724235	0.03	hsa-miR-19b-3p	1.210221	0.044
hsa-miR-1291	-0.43629	0.031	hsa-miR-324-3p	0.615431	0.045
hsa-miR-3182	-0.96625	0.031	hsa-miR-30c-5p	1.250616	0.046
hsa-miR-199a-3p	-2.19484	0.032	hsa-miR-593-3p	0.964414	0.046
hsa-miR-199b-3p	-2.19484	0.032	hsa-miR-760	0.45456	0.046
hsa-miR-367-3p	1.153571	0.032	hsa-miR-1178	-1.06657	0.047
hsa-miR-410	0.743609	0.033	hsa-miR-3190-5p	0.457707	0.047
hsa-miR-1277-3p	0.873712	0.034	hsa-miR-4454	-2.72352	0.047
hsa-miR-1284	0.660131	0.034	hsa-miR-125a-5p	-1.72745	0.048
hsa-miR-26a-5p	-1.85397	0.034	hsa-miR-571	-0.3322	0.048
hsa-miR-29c-3p	-1.54592	0.034	hsa-miR-1233	0.963333	0.049
hsa-miR-2052	0.704888	0.035	hsa-miR-1268a	1.379848	0.049
hsa-miR-129-2-3p	0.76581	0.036	hsa-miR-1275	0.610095	0.049
hsa-miR-370	0.620461	0.036	hsa-let-7d-5p	-1.28678	0.05
hsa-miR-548z	1.675082	0.036	hsa-miR-1273g-5p	-0.38748	0.05
hsa-miR-550a-5p	0.806703	0.036	hsa-miR-147b	0.502335	0.05
hsa-miR-191-5p	-1.683	0.037	hsa-miR-450b-5p	0.648325	0.05
hsa-miR-378d	-0.34324	0.037	hsa-miR-1225-5p	1.160807	<0.05
hsa-miR-4741	0.952086	0.038	hsa-miR-1258	1.331858	<0.05
hsa-miR-1912	-0.45077	0.039	hsa-miR-1302	-0.56988	<0.05
hsa-miR-342-3p	-1.50766	0.039	hsa-miR-488-3p	-0.62421	<0.05
hsa-miR-520e	1.140733	0.039	hsa-miR-539-5p	-0.68292	<0.05
hsa-miR-608	-0.52486	0.04	hsa-miR-567	1.63399	<0.05
hsa-let-7a-5p	-2.09134	0.041	hsa-miR-933	-0.56734	<0.05
hsa-miR-1203	-0.41278	0.041			

TABLE S2: The 38 KEGG pathways affected by the most significant 100 differentially expressed miRNAs (p-value) among the TNBC and non-TNBC groups.

#	KEGG pathway	p-value	#genes	#miRNAs
1	ECM-receptor interaction	4.16E-06	65	66
2	Adherens junction	4.16E-06	63	70
3	Mucin type O-Glycan biosynthesis	7.27E-06	25	33
4	Morphine addiction	1.34E-05	71	68
5	Proteoglycans in cancer	1.34E-05	148	78
6	Arrhythmogenic right ventricular cardiomyopathy (ARVC)	4.15E-05	56	65
7	GABAergic synapse	8.07E-05	66	67
8	Hippo signaling pathway	0.000422	110	69
9	Endocytosis	0.000937	153	78
10	Ras signaling pathway	0.001279	164	86
11	ErbB signaling pathway	0.001337	69	75
12	Rap1 signaling pathway	0.00197	155	82
13	PI3K-Akt signaling pathway	0.002122	245	86
14	Pathways in cancer	0.002122	288	87
15	Glutamatergic synapse	0.002125	86	67
16	Nicotine addiction	0.00534	30	54
17	Lysine degradation	0.00534	37	64
18	Thyroid hormone signaling pathway	0.00534	88	74
19	Focal adhesion	0.00534	152	79
20	mTOR signaling pathway	0.005391	49	67
21	Circadian rhythm	0.005785	27	57
22	Central carbon metabolism in cancer	0.011361	50	64
23	Retrograde endocannabinoid signaling	0.011361	76	69
24	TGF-beta signaling pathway	0.011361	57	74
25	Circadian entrainment	0.012166	73	76
26	Amoebiasis	0.014985	78	69
27	Renal cell carcinoma	0.018511	50	70
28	Wnt signaling pathway	0.018918	104	77
29	Pancreatic cancer	0.020395	49	70
30	Axon guidance	0.02064	91	74
31	FoxO signaling pathway	0.021891	98	73
32	Long-term depression	0.02446	44	58
33	Thyroid hormone synthesis	0.02446	52	60
34	Oxytocin signaling pathway	0.036428	113	86
35	Colorectal cancer	0.038055	48	66
36	Signaling pathways regulating pluripotency of stem cells	0.041672	100	76
37	Platelet activation	0.043618	92	70
38	Phosphatidylinositol signaling system	0.047876	56	73

TABLE S3. Top 10 KEGG pathways potentially affected by the 17 miRNAs.

KEGG pathway	p-value	#miRNAs	miRNAs	#genes	putative target genes
Axon guidance (hsa04360)	3.81E-08	17	miR-135b-5p, miR-944, miR-548p, miR-1275, miR-607, miR-608, miR-378c, miR-129-2-3p, miR-1260a, miR-342-3p, miR-323a-5p, miR-539-5p, miR-668, miR-323b-3p, miR-323b-5p, miR-634, miR-188-5p	66	*SEMA3G, SEMA6A, GSK3B, ABLIM3, PLXNB2, NTNG2, MET, ITGB1, ROCK1, SEMA4A, RAC2, CXCR4, NRAS, PAK2, ARHGEF12, EPHA5, ROCK2, PAK7, ROBO2, PPP3R1, SRGAP1, NTNG1, PLXNA1, GNAI3, RHPA, KRAS, PAK3, FYN, EFNA5, EPHA7, SEMA4G, NCK1, PPP3CA, PTK2, RASA1, EFNB3, SLIT2, DCC, PPP3CB, SRGAP3, EPHA3, DPYSL2, UNC5C, CFL2, SEMA3D, NRP1, RAC1, PLXNB1, SEMA4B, UNC5B, SEMA3A, PAK6, EDNB1, PLXNC1, LIMK1, SEMA4D, ROBO1, EPHA4, SEMA3E, UNC5D, LRRC4, EPHB1, GNAI1, PPPR32*
Glycosaminoglycan biosynthesis - chondroitin sulfate / dermatan sulfate (hsa00532)	1.57E-07	7	miR-135b-5p, miR-944, miR-1275, miR-548p, miR129-2-3p, miR342-3p, miR539-5p,	12	*UST, CHST3, CHST15, CHST11, CHSY3, DSE, CHPF, CHST14, CSGALNACT1, CHSY1, XYLT1*
Mucin type O-Glycan biosynthesis (hsa00512)	9.39E-06	5	miR-944, miR-548p, miR-607, miR-539-5p, miR342-3p	14	*POC1B-GALNT4, GALNT7, GALNT15, GALNTL5, GALNT13, GALNT4, GCNT3, GALNT1, GALNT10, CIGALT1, GCNT1, CIGALT1C1, GALNT5, GALNT16*
Thyroid hormone signaling pathway (hsa04919)	9.39E-06	15	miR-135b-5p, miR-944, miR-548p, miR-607, miR-608, miR-378c, miR-129-2-3p, miR-1260a, miR-342-3p, miR-323a-5p, miR-539-5p, miR-668, miR-323b-5p, miR-634, miR-1275	57	*ESR1, ACTB, GSK3B, PRKCA, ATP1B2, MED13L, NRAS, MED14, DIO2, PIK3R2, MED13, SLC16A10, ATP1B1, THRA, MED17, RCAN2, KRAS, MED12L, MED1, MED4, ITGAV, MED24, PLCB1, TBC1D4, PIK3CD, PIK3R3, HIF1A, PLN, PRKACA, PRKCG, NCOR1, ATP1A4, NOCA2, PIK3R1, PIK3CG, HDAC2, PRKX, PLCG2, THRB, KAT2B, EP300, AKT3, SLC2+A3, PLCE1, PIK3CA, SLC2A1, PDPK1, PFKFB2, ATP2A2, FOXO1, MDM2, PLCB4, BMP4, NCOA1, PRKACB, RXRA, PIK3R2*

TABLE S3. Top 10 KEGG pathways potentially affected by the 17 miRNAs.

KEGG pathway	p-value	#miRNAs	miRNAs	#genes	putative target genes
Signaling pathways regulating pluripotency of stem cells	2.32E-05	16	miR-135b-5p, miR-944, miR-548p, miR-1275, miR-607, miR-608, miR-378c, miR-129-2-3p, miR-1260a, miR-342-3p, miR-323a-5p, miR-539-5p, miR-668, miR-323b-3p, miR-323b-5p, miR-634	68	*BMI1, GSK3B, DVL3, FZD5, OTX1, KAT6A, PAX6, SMAD2, NRAS, NODAL, INHBB, APC, REST, ACVR1B, WNT5A, PIK3R2, MAPK14, TBX3, FZD6, SMARCAAD1, WNT2B, INHBA, WNT3, IGF1R, ZFHX3, ID4, FZD8, KRAS, FZD3, POU5F1B, ACVR2B, PCGF5, FZD4, RIF1, PIK3CD, PIK3R3, FZD10, JAK2, AXIN2, LIFR, SKIL, ZIC3, SMAD5, PIK3R1, PIK3CG, ACVR2A, FGF2, BMP2, FZD1, ACVR1C, IGF1, AKT3, BMPR1A, PIK3CA, WNT8B, IL6ST, ISL1, FGFR1, SOX2, PCGF3, WNT7B, KLF4, JAK1, BMP4, MEIS1, BMPR2, PIK3R2*
Proteoglycans in cancer (hsa05205)	2.46E-05	16	miR-135b-5p, miR-944, miR-548p, miR-1275, miR-607, miR-608, miR-378c, miR-129-2-3p, miR-342-3p, miR-323a-5p, miR-539-5p, miR-668, miR-323b-3p, miR-323b-5p, miR-634, miR-188-5p	88	*ESR1, CAMK2D, BRAF, ACTB, PRKCA, FZD5, MET, ITGB1, ROCK1, SMAD2, CBL, NRAS, THBS1, PTCH1, WNT5A, PIK3R2, ARHGEF12, PPP1CC, MAPK14, ROCK2, FRS2, FZD6, RDX, IQGAP1, ITGA5, WNT2B, TIAM1, WNT3, IGF1R, EGFR, RHOA, PPP1R12B, CAV2, FZD8, KRAS, FZD3, RRAS2, MSN, ANK2, CTTN, CAMK2A, PTK2, CBLB, ITGAV, ANK3, FZD4, PPP1R12A, PIK3CD, PIK3R3, FZD10, GPC1, HIF1A, IGF2, FLNB, PRKACA, PRKCG, EIF4B, TIMP3, FLNA, ITGA2, PIK3R1, SOS1, PIK3CG, PTPN11, PRKX, DDX5, RAC1, FGF2, PLCG2, PPP1R12C, FZD1, IGF1, AKT3, PLCE1, PIK3CA, WNT8B, SMP, PDPK1, VEGFA, FGFR1, ITPR2, WNT7B, MDM2, ERBB4, RPS6KB1, PRKACB, PPP1CB, PIK3R2*

TABLE S3. Top 10 KEGG pathways potentially affected by the 17 miRNAs.

KEGG pathway	p-value	#miRNAs	miRNAs	#genes	putative target genes
Wnt signaling pathway (hsa04310)	3.14E-05	17	miR-135b-5p, miR-944, miR-548p, miR-1275, miR-607, miR-608, miR-378c,miR-129-2-3p, miR-1260a, miR-342-3p, miR-323a-5p, miR-539-5p, miR-668,miR-323b-3p, miR-323b-5p, miR-634, miR-188-5p	66	*CAMK2D, GSK3B, PRKCA, DVL3, FZD5, LRP6, RAC2, BTRC, APC, VANGL1, TCF7L2, WNT5A, CCND2, ROCK2, FZD6, PPP3R1, CTBP2, WNT2B, WNT3, LRP5, RHOA, FZD8, FZD3, SKP1, FRAT2, CAMK2A, PPP3CA, PRICKLE1, NLK, PLCB1, FZD4, SENP2, FZD10, PPP3CB, AXIN2, MAPK8, CSNK1A1, NKD2, PRKACA, VANGL2, PRKCG, CSNK2B, CSNK2A1, PRKX, RAC1, SIAH1, PSEN1, FZD1, PRICKLE2, CSNK1E, EP300, CXXC4, RBX1, WNT8BM LEF1, DAAM1, MAP3K7, FBXW11, NFATC1, WNT7B, TLB1XR1, PPARD, PLCB4, MAPK10, PRKACB, PPP3R2*
Hippo signaling pathway (hsa04390)	7.29E-05	15	miR-135b-5p, miR-944, miR-548p, miR-1275, miR-607, miR-608, miR-378c,miR-129-2-3p, miR-1260a, miR-342-3p, miR-323a-5p, miR-539-5p, miR-668,miR-323b-3p, miR-634	67	*ACTB, GSK3B, DVL3, FZD5, TGFBR1, YWHAH, YAP1, SMAD2, YWHAE, BTRC, APC, TCF7L2, WNT5A, YWHAG, PPP1CC, CCND2, BMP5, FLI2, AREG, FZD6, WNT2B, MOB1B, WANT3, CRB1, PPP2REB, PPP2R2D, WWTR1, FZD8, TP53BP2, FZD3, BMP8B, LIMD1, MPP5, FZD4, CSNK1D, DLG4, FZD10, AXIN2, PPP2R2A, RASSF6, PPP2R1A, FRMD6, SAV1, TEAD1, YWHAZ, GDF6, STK3, BMP2, FZD1, CSNK1E, BMPR1A, CTNNA3, WNT8, DLG2, LEF1, PARD3, CTNNA2, LATS1, SOX2, BMP7, LATS2, FBXW11, PARD6B, WNT7B, BMP4, BMPR2, PPP1CB*

TABLE S3. Top 10 KEGG pathways potentially affected by the 17 miRNAs.

KEGG pathway	p-value	#miRNAs	miRNAs	#genes	putative target genes
Ras signaling pathway	0.0001079	17	miR-135b-5p, miR-944, miR-548p, miR-1275, miR-607, miR-608, miR-378c,miR-129-2-3p, miR-1260a, miR-342-3p, miR-323a-5p, miR-539-5p, miR-668,miR-323b-3p, miR-323b-5p, miR-634, miR-188-5p	95	*FGF12, PRKCA, KSR2, PDGFRA, GNG13, NFKB1, MET, FIGF, GNGT1, RAC2, NRAS, STK4, RASA2, CALM1, PAK2, PIK3R2, RAP1A, RASGRF2, ETS1, PAK7, GNG12, ANGPT2, RALA, CHUK, RALBP1, PLD1, TIAM1, RASGRP2, IGF1R, EGFR, RHOA, KRAS, PAK3, RRAS2, EFNA5, PLA2G2D, RAB5A, MLLT4, RASA1, PIK3CD, PIK3R3, PLA2G12A, RASGRP1, MAPK8, PLA2G16, REL, BRAP, PRKACA, PRKCG, FLT1, GNG2, PLA2G12B, PIK3R1, SOS1, PIK3CG, PTPN11, PRKX, RASA4, TBK1, KITLG, RAC1, FGF2, PLCG2, NF1, IGF1, SHC4, BCL2L1, AKT3, PDGFC, PAK6, ANGPT1, PLCE1, PDGFD, PIK3CA, SYNGAP1, GAB2, FGF8, GNG4, RASAL2, RGL1, TAB5C, FGFR2, VEGFA, FGFR1, ABL2, FGF7, CSF1, RAP1B, RAPGEF5, GNG3, MAPK10, PRKACB, PDGFA, GRIN2B, PIK3R2*

TABLE S3. Top 10 KEGG pathways potentially affected by the 17 miRNAs.

KEGG pathway	p-value	#miRNAs	miRNAs	#genes	putative target genes
Pathways in cancer (hsa05200)	0.0001787	17	miR-135b-5p, miR-944, miR-548p, miR-1275, miR-607, miR-608, miR-378c,miR-129-2-3p, miR-1260a, miR-342-3p, miR-323a-5p, miR-539-5p, miR-668, miR-323b-3p, miR-323b-5p, miR-634, miR-188-5p	169	*BRAF, FGF12, GSK3B, PRKCA, DVL3, PDGFRA, FZD5, E2F1, TGFBR1, GNG13, NFKB1, MET, ADCY1, ADCY5, SPI1, ITGB1, FIGF, PTGER4, ROCK1, GNGT1, CXCL8, RAC2, SMAD2, CBL, E2F2, CXCR4, ADCY7, NRAS, CRKL, STK4, LPAR3, APC, CRK, PTCH1, RAD51, MSH3, CUL2, TPR, TCF7L2, WNT5A, TGFA, HHIP, COL4A5, PIK3R2, ARHGEF12, GNA13, GLI2, ROCK2, ARNT, ETS1, ADCY3, GNG12, RALA, FZD6, CTBP2, CHUK, BCL2, CDKN1B, WNT2B, TRAF4, RALBP1, PLD1, CDKN2B, WNT3, RASGRP2, IGF1R, EGFR, GNAI3, PTCH2, F2RL3, RHOA, TRAF5, APPL1, FZD8, KRAS, CDK6, FZD3, PAX8, VHL, MITF, LPAR4, PTK2, GNA11, CBLB, ITGAV, PLCB1, FZD4, RUNX1T1, DCC, AR, PIK3CD, PIK3R3, FZD10, MSH6, AXIN2, RASGRP1, SKP2, HIF1A, E2F3, MAPK8, PTGER2, DAPK1, PRKACA, PRKCG, GNG2, COL4A3, ITGA2, PIK3R1, RB1, COL4A4, SOS1, PIK3CG, HDAC2, PRKX, KITLG, RAC1, FGF2, BMP2, PLCG2, FZD1, FLT3, IGF1, EP300, CASP8, GNAQ, NKX3-1, BCL2L1, AKT3, PIAS2, CTNNA3, PIK3CA, RBX1, WNT8B, LEF1, COL4A6, MECOM, FGF8, SMO, SLC2A1, GNG4, EDNRA, SUFU, FGFR2, ITGA6, SHH, CTNNA2, VEGFA, PTEN, FGFR1, FOXO1, FGF7, TRAF3, WNT7B, PPARD, JAK1, MDM2, PLCB4, BMP4, GNG3, CCDC6, MAPK10, GNAI1, PRKACB, XIAP, EGLN1, RXRA, PDGFA, LAMA4, PIK3R2*

REFERENCES

1. Siegel RL, Fedewa SA, Miller KD, Goding-Sauer A, Pinheiro PS, Martinez-Tyson D, Jemal A. Cancer statistics for Hispanics/Latinos, 2015. CA: a cancer journal for clinicians 2015;65: 457-80.

2. Society AC. Cancer Facts & FIGUREs for Hispanic/Latinos 2015-2017 Atlanta, 2015.

3. INCA/MS. Estimativa 2018 - Incidência de Câncer no Brasil Rio de Janeiro, 2018.

4. Hunt BR. Breast Cancer Prevalence and Mortality among Hispanic Subgroups in the United States, 2009-2013. Journal of cancer epidemiology 2016;2016: 8784040.

5. Justo N, Wilking N, Jonsson B, Luciani S, Cazap E. A review of breast cancer care and outcomes in Latin America. The oncologist 2013;18: 248-56.

6. Schinkel JK, Zahm SH, Jatoi I, McGlynn KA, Gallagher C, Schairer C, Shriver CD, Zhu K. Racial/ethnic differences in breast cancer survival by inflammatory status and hormonal receptor status: an analysis of the Surveillance, Epidemiology, and End Results data. Cancer causes & control : CCC 2014;25: 959-68.

7. Kohler BA, Sherman RL, Howlader N, Jemal A, Ryerson AB, Henry KA, Boscoe FP, Cronin KA, Lake A, Noone AM, Henley SJ, Eheman CR, et al. Annual Report to the Nation on the Status of Cancer, 1975-2011, Featuring Incidence of Breast Cancer Subtypes by Race/Ethnicity, Poverty, and State. Journal of the National Cancer Institute 2015;107: djv048.

8. Noone AM, Cronin KA, Altekruse SF, Howlader N, Lewis DR, Petkov VI, Penberthy L. Cancer Incidence and Survival Trends by Subtype Using Data from the Surveillance Epidemiology and End Results Program, 1992-2013. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology 2017;26: 632-41.

9. Brandzel S, Chang E, Tuzzio L, Campbell C, Coronado N, Bowles EJ, Bradford SC, Buist DS. Latina and Black/African American Women's Perspectives on Cancer Screening and Cancer Screening Reminders. Journal of racial and ethnic health disparities 2016.

10. Lynce F, Graves KD, Jandorf L, Ricker C, Castro E, Moreno L, Augusto B, Fejerman L, Vadaparampil ST. Genomic Disparities in Breast Cancer Among Latinas. Cancer control : journal of the Moffitt Cancer Center 2016;23: 359-72.

11. Connor AE, Visvanathan K, Baumgartner KB, Baumgartner RN, Boone SD, Hines LM, Wolff RK, John EM, Slattery ML. Ethnic differences in the relationships between diabetes, early age adiposity and mortality among breast cancer survivors: the Breast Cancer Health Disparities Study. Breast cancer research and treatment 2016;157: 167-78.

12. Arce-Salinas C, Aguilar-Ponce JL, Villarreal-Garza C, Lara-Medina FU, Olvera-Caraza D, Alvarado Miranda A, Flores-Diaz D, Mohar A. Overweight and obesity as poor prognostic factors in locally advanced breast cancer patients. Breast cancer research and treatment 2014;146: 183-8.

13. Bandera EV, Chandran U, Hong CC, Troester MA, Bethea TN, Adams-Campbell LL, Haiman CA, Park SY, Olshan AF, Ambrosone CB, Palmer JR, Rosenberg L. Obesity, body fat distribution, and risk of breast cancer subtypes in African American women participating in the AMBER Consortium. Breast cancer research and treatment 2015;150: 655-66.

14. Fejerman L, Stern MC, John EM, Torres-Mejia G, Hines LM, Wolff RK, Baumgartner KB, Giuliano AR, Ziv E, Perez-STABLE EJ, Slattery ML. Interaction between common breast cancer susceptibility variants, genetic ancestry, and nongenetic risk factors in Hispanic women. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology 2015;24: 1731-8.

15. Connor AE, Baumgartner RN, Pinkston CM, Boone SD, Baumgartner KB. Obesity, ethnicity, and quality of life among breast cancer survivors and women without breast cancer: the long-term quality of life follow-up study. Cancer causes & control : CCC 2016;27: 115-24.

16. Connor AE, Visvanathan K, Baumgartner KB, Baumgartner RN, Boone SD, Hines LM, Wolff RK, John EM, Slattery ML. Pre-diagnostic breastfeeding, adiposity, and mortality among parous Hispanic and non-Hispanic white women with invasive breast cancer: the Breast Cancer Health Disparities Study. Breast cancer research and treatment 2017;161: 321-31.

17. Tejeda S, Gallardo RI, Ferrans CE, Rauscher GH. Breast cancer delay in Latinas: the role of cultural beliefs and acculturation. Journal of behavioral medicine 2017;40: 343-51.

18. Molina Y, Thompson B, Espinoza N, Ceballos R. Breast cancer interventions serving US-based Latinas: current approaches and directions. Womens Health (Lond) 2013;9: 335-48; quiz 49-50.

19. Friese CR, Martinez KA, Abrahamse P, Hamilton AS, Graff JJ, Jagsi R, Griggs JJ, Hawley ST, Katz SJ. Providers of follow-up care in a population-based sample of breast cancer survivors. Breast cancer research and treatment 2014;144: 179-84.

20. O'Brien KM, Cole SR, Tse CK, Perou CM, Carey LA, Foulkes WD, Dressler LG, Geradts J, Millikan RC. Intrinsic breast tumor subtypes, race, and long-term survival in the Carolina Breast Cancer Study. Clinical cancer research : an official journal of the American Association for Cancer Research 2010;16: 6100-10.

21. Chavez-Macgregor M, Liu S, De Melo-Gagliato D, Chen H, Do KA, Pusztai L, Fraser Symmans W, Nair L, Hortobagyi GN, Mills GB, Meric-Bernstam F, Gonzalez-Angulo AM. Differences in gene and protein expression and the effects of race/ethnicity on breast cancer subtypes. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology 2014;23: 316-23.

22. Kroenke CH, Sweeney C, Kwan ML, Quesenberry CP, Weltzien EK, Habel LA, Castillo A, Bernard PS, Factor RE, Kushi LH, Caan BJ. Race and breast cancer survival by intrinsic subtype based on PAM50 gene expression. Breast cancer research and treatment 2014;144: 689-99.

23. Plasilova ML, Hayse B, Killelea BK, Horowitz NR, Chagpar AB, Lannin DR. Features of triple-negative breast cancer: Analysis of 38,813 cases from the national cancer database. Medicine 2016;95: e4614.

24. Patel TA, Colon-Otero G, Bueno Hume C, Copland JA, 3rd, Perez EA. Breast cancer in Latinas: gene expression, differential response to treatments, and differential toxicities in Latinas compared with other population groups. The oncologist 2010;15: 466-75.

25. Lara-Medina F, Perez-Sanchez V, Saavedra-Perez D, Blake-Cerda M, Arce C, Motola-Kuba D, Villarreal-Garza C, Gonzalez-Angulo AM, Bargallo E, Aguilar JL, Mohar A, Arrieta O. Triple-negative breast cancer in Hispanic patients: high prevalence, poor prognosis, and association with menopausal status, body mass index, and parity. Cancer 2011;117: 3658-69.

26. Killelea BK, Yang VQ, Wang SY, Hayse B, Mougalian S, Horowitz NR, Chagpar AB, Pusztai L, Lannin DR. Racial Differences in the Use and Outcome of Neoadjuvant Chemotherapy for Breast Cancer: Results From the National Cancer Data Base. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2015;33: 4267-76.

27. Sheppard VB, Hicks J, Makambi K, Hurtado-de-Mendoza A, Demark-Wahnefried W, Adams-Campbell L. The feasibility and acceptability of a diet and exercise trial in overweight and obese black breast cancer survivors: The Stepping STONE study. Contemporary clinical trials 2016;46: 106-13.

28. Warner ET, Ballman KV, Strand C, Boughey JC, Buzdar AU, Carey LA, Sikov WM, Partridge AH. Impact of race, ethnicity, and BMI on achievement of pathologic complete response following neoadjuvant chemotherapy for breast cancer: a pooled analysis of four prospective Alliance clinical trials (A151426). Breast cancer research and treatment 2016;159: 109-18.

29. Villarreal-Garza C, Soto-Perez-de-Celis E. Ethnicity, adverse events and outcomes in breast cancer. Breast 2015;24: 539-40.

30. Wu AH, Gomez SL, Vigen C, Kwan ML, Keegan TH, Lu Y, Shariff-Marco S, Monroe KR, Kurian AW, Cheng I, Caan BJ, Lee VS, et al. The California Breast Cancer Survivorship Consortium (CBCSC): prognostic factors associated with racial/ethnic differences in breast cancer survival. Cancer causes & control : CCC 2013;24: 1821-36.

31. Fejerman L, Ahmadiyeh N, Hu D, Huntsman S, Beckman KB, Caswell JL, Tsung K, John EM, Torres-Mejia G, Carvajal-Carmona L, Echeverry MM, Tuazon AM, et al. Genome-wide association study of breast cancer in Latinas identifies novel protective variants on 6q25. Nature communications 2014;5: 5260.

32. Kalinsky K, Lim EA, Andreopoulou E, Desai AM, Jin Z, Tu Y, Hibshoosh H, Wang A, Greenlee H, Crew KD, Maurer M, Sparano JA, et al. Increased expression of tumor proliferation genes in Hispanic women with early-stage breast cancer. Cancer investigation 2014;32: 439-44.

33. Sly J, Jandorf L, Erwin DO. Who's Missing? Predictors of Attrition Following Participation in Culturally Targeted Educational Breast and Cervical Cancer Outreach Programs for Latinas. Journal of health communication 2015;20: 851-8.

34. Hamel LM, Penner LA, Albrecht TL, Heath E, Gwede CK, Eggly S. Barriers to Clinical Trial Enrollment in Racial and Ethnic Minority Patients With Cancer. Cancer control : journal of the Moffitt Cancer Center 2016;23: 327-37.

35. Zhang W, Liu J, Wang G. The role of microRNAs in human breast cancer progression. Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine 2014;35: 6235-44.

36. Takahashi RU, Miyazaki H, Ochiya T. The Roles of MicroRNAs in Breast Cancer. Cancers 2015;7: 598-616.

37. Li JY, Jia S, Zhang WH, Zhang Y, Kang Y, Li PS. Differential distribution of microRNAs in breast cancer grouped by clinicopathological subtypes. Asian Pacific journal of cancer prevention : APJCP 2013;14: 3197-203.

38. Riaz M, van Jaarsveld MT, Hollestelle A, Prager-van der Smissen WJ, Heine AA, Boersma AW, Liu J, Helmijr J, Ozturk B, Smid M, Wiemer EA, Foekens JA, et al. miRNA expression profiling of 51 human breast cancer cell lines reveals subtype and driver mutation-specific miRNAs. Breast cancer research : BCR 2013;15: R33.

39. D'Ippolito E, Iorio MV. MicroRNAs and triple negative breast cancer. International journal of molecular sciences 2013;14: 22202-20.

40. Gyparaki MT, Basdra EK, Papavassiliou AG. MicroRNAs as regulatory elements in triple negative breast cancer. Cancer letters 2014;354: 1-4.

41. Yang F, Zhang W, Shen Y, Guan X. Identification of dysregulated microRNAs in triple-negative breast cancer (review). International journal of oncology 2015;46: 927-32.

42. Zhou L, Liu F, Wang X, Ouyang G. The roles of microRNAs in the regulation of tumor metastasis. Cell & bioscience 2015;5: 32.

43. Tian W, Chen J, He H, Deng Y. MicroRNAs and drug resistance of breast cancer: basic evidence and clinical applications. Clinical & translational oncology : official publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico 2013;15: 335-42.

44. Wang J, Yang M, Li Y, Han B. The Role of MicroRNAs in the Chemoresistance of Breast Cancer. Drug development research 2015;76: 368-74.

45. Rawlings-Goss RA, Campbell MC, Tishkoff SA. Global population-specific variation in miRNA associated with cancer risk and clinical biomarkers. BMC medical genomics 2014;7: 53.

46. Evans-Knowell A, LaRue AC, Findlay VJ. MicroRNAs and Their Impact on Breast Cancer, the Tumor Microenvironment, and Disparities. Advances in cancer research 2017;133: 51-76.

47. Bushel PR, McGovern R, Liu L, Hofmann O, Huda A, Lu J, Hide W, Lin X. Population differences in transcript-regulator expression quantitative trait loci. PloS one 2012;7: e34286.

48. Yao S, Graham K, Shen J, Campbell LE, Singh P, Zirpoli G, Roberts M, Ciupak G, Davis W, Hwang H, Khoury T, Bovbjerg DH, et al. Genetic variants in microRNAs and breast cancer risk in African American and European American women. Breast cancer research and treatment 2013;141: 447-59.

49. Wang PY, Gao ZH, Jiang ZH, Li XX, Jiang BF, Xie SY. The associations of single nucleotide polymorphisms in miR-146a, miR-196a and miR-499 with breast cancer susceptibility. PloS one 2013;8: e70656.

50. Khan S, Greco D, Michailidou K, Milne RL, Muranen TA, Heikkinen T, Aaltonen K, Dennis J, Bolla MK, Liu J, Hall P, Irwanto A, et al. MicroRNA related polymorphisms and breast cancer risk. PloS one 2014;9: e109973.

51. Bensen JT, Xu Z, Smith GJ, Mohler JL, Fontham ET, Taylor JA. Genetic polymorphism and prostate cancer aggressiveness: a case-only study of 1,536 GWAS and candidate SNPs in African-Americans and European-Americans. The Prostate 2013;73: 11-22.

52. Wang W, Ingles SA, Torres-Mejia G, Stern MC, Stanczyk FZ, Schwartz GG, Nelson DO, Fejerman L, Wolff RK, Slattery ML, John EM. Genetic variants and non-genetic factors predict circulating vitamin D levels in Hispanic and non-Hispanic White women: the Breast Cancer Health Disparities Study. International journal of molecular epidemiology and genetics 2014;5: 31-46.

53. Toyama T, Kondo N, Endo Y, Sugiura H, Yoshimoto N, Iwasa M, Takahashi S, Fujii Y, Yamashita H. High expression of microRNA-210 is an independent factor indicating a poor prognosis in Japanese triple-negative breast cancer patients. Japanese journal of clinical oncology 2012;42: 256-63.

54. Sugita B, Gill M, Mahajan A, Duttargi A, Kirolikar S, Almeida R, Regis K, Oluwasanmi OL, Marchi F, Marian C, Makambi K, Kallakury B, et al. Differentially expressed miRNAs in triple negative breast cancer between African-American and non-Hispanic white women. Oncotarget 2016;7: 79274-91.

55. Hammond ME, Hayes DF, Dowsett M, Allred DC, Hagerty KL, Badve S, Fitzgibbons PL, Francis G, Goldstein NS, Hayes M, Hicks DG, Lester S, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. Archives of pathology & laboratory medicine 2010;134: 907-22.

56. Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, Dowsett M, Fitzgibbons PL, Hanna WM, Langer A, McShane LM, Paik S, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2007;25: 118-45.

57. Ruiz-Linares A, Adhikari K, Acuna-Alonzo V, Quinto-Sanchez M, Jaramillo C, Arias W, Fuentes M, Pizarro M, Everardo P, de Avila F, Gomez-Valdes J, Leon-Mimila P, et al. Admixture in Latin America: geographic structure, phenotypic diversity and self-perception of ancestry based on 7,342 individuals. PLoS genetics 2014;10: e1004572.

58. Homburger JR, Moreno-Estrada A, Gignoux CR, Nelson D, Sanchez E, Ortiz-Tello P, Pons-Estel BA, Acevedo-Vasquez E, Miranda P, Langefeld CD, Gravel S, Alarcon-Riquelme ME, et al. Genomic Insights into the Ancestry and Demographic History of South America. PLoS genetics 2015;11: e1005602.

59. Conomos MP, Laurie CA, Stilp AM, Gogarten SM, McHugh CP, Nelson SC, Sofer T, Fernandez-Rhodes L, Justice AE, Graff M, Young KL, Seyerle AA, et al. Genetic Diversity and Association Studies in US Hispanic/Latino Populations: Applications in the Hispanic Community Health Study/Study of Latinos. American journal of human genetics 2016;98: 165-84.

60. Healy ME, Hill D, Berwick M, Edgar H, Gross J, Hunley K. Social-group identity and population substructure in admixed populations in New Mexico and Latin America. PloS one 2017;12: e0185503.

61. Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA. An integrated map of genetic variation from 1,092 human genomes. Nature 2012;491: 56-65.

62. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience 2015;4: 7.

63. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. PLoS genetics 2006;2: e190.

64. Torresan C, Oliveira MM, Pereira SR, Ribeiro EM, Marian C, Gusev Y, Lima RS, Urban CA, Berg PE, Haddad BR, Cavalli IJ, Cavalli LR. Increased copy number of the DLX4 homeobox gene in breast axillary lymph node metastasis. Cancer genetics 2014;207: 177-87.

65. Paraskevopoulou MD, Georgakilas G, Kostoulas N, Vlachos IS, Vergoulis T, Reczko M, Filippidis C, Dalamagas T, Hatzigeorgiou AG. DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows. Nucleic acids research 2013;41: W169-73.

66. Wong N, Wang X. miRDB: an online resource for microRNA target prediction and functional annotations. Nucleic acids research 2015;43: D146-52.

67. Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. eLife 2015;4.

68. Vlachos IS, Zagganas K, Paraskevopoulou MD, Georgakilas G, Karagkouni D, Vergoulis T, Dalamagas T, Hatzigeorgiou AG. DIANA-miRPath v3.0: deciphering microRNA function with experimental support. Nucleic acids research 2015;43: W460-6.

69. Nersisyan L, Samsonyan R, Arakelyan A. CyKEGGParser: tailoring KEGG pathways to fit into systems biology analysis workflows. F1000Research 2014;3: 145.

70. Montojo J, Zuberi K, Rodriguez H, Kazi F, Wright G, Donaldson SL, Morris Q, Bader GD. GeneMANIA Cytoscape plugin: fast gene function predictions on the desktop. Bioinformatics 2010;26: 2927-8.

71. Kutmon M, Kelder T, Mandaviya P, Evelo CT, Coort SL. CyTargetLinker: a cytoscape app to integrate regulatory interactions in network analysis. PloS one 2013;8: e82160.

72. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome research 2003;13: 2498-504.

73. Dai X, Chen A, Bai Z. Integrative investigation on breast cancer in ER, PR and HER2-defined subgroups using mRNA and miRNA expression profiling. Scientific reports 2014;4: 6566.

74. Bhattacharyya M, Nath J, Bandyopadhyay S. MicroRNA signatures highlight new breast cancer subtypes. Gene 2015;556: 192-8.

75. Iuliano R, Vismara MF, Dattilo V, Trapasso F, Baudi F, Perrotti N. The role of microRNAs in cancer susceptibility. BioMed research international 2013;2013: 591931.

76. Han W, Jung EM, Cho J, Lee JW, Hwang KT, Yang SJ, Kang JJ, Bae JY, Jeon YK, Park IA, Nicolau M, Jeffrey SS, et al. DNA copy number alterations and expression of relevant genes in triple-negative breast cancer. Genes, chromosomes & cancer 2008;47: 490-9.

77. Hua K, Jin J, Zhao J, Song J, Song H, Li D, Maskey N, Zhao B, Wu C, Xu H, Fang L. miR-135b, upregulated in breast cancer, promotes cell growth and disrupts the cell cycle by regulating LATS2. International journal of oncology 2016;48: 1997-2006.

78. Aakula A, Leivonen SK, Hintsanen P, Aittokallio T, Ceder Y, Borresen-Dale AL, Perala M, Ostling P, Kallioniemi O. MicroRNA-135b regulates ERalpha, AR and HIF1AN and affects breast and prostate cancer cell growth. Molecular oncology 2015;9: 1287-300.

79. Lowery AJ, Miller N, Devaney A, McNeill RE, Davoren PA, Lemetre C, Benes V, Schmidt S, Blake J, Ball G, Kerin MJ. MicroRNA signatures predict oestrogen receptor, progesterone receptor and HER2/neu receptor status in breast cancer. Breast cancer research : BCR 2009;11: R27.

80. He H, Tian W, Chen H, Jiang K. MiR-944 functions as a novel oncogene and regulates the chemoresistance in breast cancer. Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine 2016;37: 1599-607.

81. Pena-Chilet M, Martinez MT, Perez-Fidalgo JA, Peiro-Chova L, Oltra SS, Tormo E, Alonso-Yuste E, Martinez-Delgado B, Eroles P, Climent J, Burgues O, Ferrer-Lozano J, et al. MicroRNA profile in very young women with breast cancer. BMC cancer 2014;14: 529.

82. Cordero F, Ferrero G, Polidoro S, Fiorito G, Campanella G, Sacerdote C, Mattiello A, Masala G, Agnoli C, Frasca G, Panico S, Palli D, et al. Differentially methylated microRNAs in prediagnostic samples of subjects who developed breast cancer in the European Prospective Investigation into Nutrition and Cancer (EPIC-Italy) cohort. Carcinogenesis 2015;36: 1144-53.

83. Godbole M, Chandrani P, Gardi N, Dhamne H, Patel K, Yadav N, Gupta S, Badwe R, Dutt A. miR-129-2 mediates down-regulation of progesterone receptor in response to progesterone in breast cancer cells. Cancer biology & therapy 2017;18: 801-5.

84. Shi Y, Qiu M, Wu Y, Hai L. MiR-548-3p functions as an anti-oncogenic regulator in breast cancer. Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie 2015;75: 111-6.

85. Ke H, Zhao L, Feng X, Xu H, Zou L, Yang Q, Su X, Peng L, Jiao B. NEAT1 is Required for Survival of Breast Cancer Cells Through FUS and miR-548. Gene regulation and systems biology 2016;10: 11-7.

86. Guo J, Gong G, Zhang B. miR-539 acts as a tumor suppressor by targeting epidermal growth factor receptor in breast cancer. Scientific reports 2018;8: 2073.

87. Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, Mackowiak SD, Gregersen LH, Munschauer M, Loewer A, Ziebold U, et al. Circular RNAs are a large class of animal RNAs with regulatory potency. Nature 2013;495: 333-8.

8 CONCLUSÃO

Neste estudo demonstramos que carcinomas mamários triplo-negativos apresentam características moleculares (padrões de alterações de números de cópias de DNA e expressão de miRNAs) diferentes entre os grupos populacionais estudados: negras americanas, brancas americanas não-hispânicas e brasileiras. Assim, sugere-se que o fator biológico apresenta-se como um fator importante nas disparidades de saúde observadas no câncer de mama entre essas populações, uma vez que as análises de enriquecimento de vias e funções biológicas indicam o envolvimento de muitos desses miRNAs em processos frequentemente afetados na tumorigênese, como proliferação celular, migração e invasão.

A análise integrada realizada permitiu a seleção de microRNAs que podem ser utilizados como biomarcadores de diagnóstico (como os painéis de miRNAs), prognóstico e potenciais alvos no desenvolvimento de terapias.

9 CONSIDERAÇÕES FINAIS

Os miRNAs selecionados em cada comparação realizada neste estudo demonstram importância na tumorigênese mamária de TNBC dos grupos populacionais estudados, assim é recomendada a realização de estudos funcionais *in vitro* que permitam a elucidação das suas funções biológicas específicas, da mesma forma como foi realizado com o miR-150-5p. Esses estudos podem auxiliar na elucidação da função de cada um dos miRNAs,

Ainda, o painel de miRNAs obtido em cada um dos estudos, assim como os miRNAs individuais, devem ser posteriormente validados para serem utilizados como biomarcadores moleculares de diagnóstico e prognóstico de câncer de mama. Os miRNAs apresentam-se estavelmente em fluidos corporais como sangue, urina e secreções (incluindo fluído ductal mamilar) e assim a dosagem dos seus níveis nesses fluídos pode permitir o diagnóstico/prognóstico da doença por um método pouco invasivo, principalmente quando comparado a biópsias.

REFERÊNCIAS

ALIZART, M. et al. Molecular classification of breast carcinoma. **Diagnostic Histopathology,** v. 18, n. 3, p. 7, 2012.

AMBROS, V. The functions of animal microRNAs. **Nature,** v. 431, n. 7006, p. 350-5, Sep 16 2004.

BASTIEN, R. R. et al. PAM50 breast cancer subtyping by RT-qPCR and concordance with standard clinical molecular markers. **BMC Med Genomics,** v. 5, p. 44, Oct 4 2012.

BENTWICH, I. et al. Identification of hundreds of conserved and nonconserved human microRNAs. **Nat Genet,** v. 37, n. 7, p. 766-70, Jul 2005.

BHASKARAN, M.; MOHAN, M. MicroRNAs: history, biogenesis, and their evolving role in animal development and disease. **Vet Pathol,** v. 51, n. 4, p. 759-74, Jul 2014.

BHATTACHARYYA, S. N. et al. Relief of microRNA-mediated translational repression in human cells subjected to stress. **Cell,** v. 125, n. 6, p. 1111-24, Jun 16 2006.

CALIN, G. A. et al. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. **Proc Natl Acad Sci U S A,** v. 101, n. 9, p. 2999-3004, Mar 2 2004.

CDC, C. F. D. C. A. P. Health Equity. 2017. Disponível em: < https://www.cdc.gov/healthequity/index.html >. Acesso em: 10 de fevereiro de 2018.

CHAVEZ-MACGREGOR, M. et al. Male breast cancer according to tumor subtype and race: a population-based study. **Cancer,** v. 119, n. 9, p. 1611-7, May 01 2013.

CHEKULAEVA, M.; FILIPOWICZ, W. Mechanisms of miRNA-mediated post-transcriptional regulation in animal cells. **Curr Opin Cell Biol,** v. 21, n. 3, p. 452-60, Jun 2009.

CIRIELLO, G. et al. Emerging landscape of oncogenic signatures across human cancers. **Nat Genet,** v. 45, n. 10, p. 1127-33, Oct 2013.

DALY, B.; OLOPADE, O. I. A perfect storm: How tumor biology, genomics, and health care delivery patterns collide to create a racial survival disparity in breast cancer and proposed interventions for change. **CA Cancer J Clin,** v. 65, n. 3, p. 221-38, May-Jun 2015.

DI LEVA, G.; CROCE, C. M. Roles of small RNAs in tumor formation. **Trends Mol Med,** v. 16, n. 6, p. 257-67, Jun 2010.

DIDIANO, D.; HOBERT, O. Perfect seed pairing is not a generally reliable predictor for miRNA-target interactions. **Nat Struct Mol Biol,** v. 13, n. 9, p. 849-51, Sep 2006.

DOENCH, J. G.; SHARP, P. A. Specificity of microRNA target selection in translational repression. **Genes Dev,** v. 18, n. 5, p. 504-11, Mar 1 2004.

ELSTON, C. W.; ELLIS, I. O. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. **Histopathology,** v. 19, n. 5, p. 403-10, Nov 1991.

FABIAN, M. R.; SONENBERG, N.; FILIPOWICZ, W. Regulation of mRNA translation and stability by microRNAs. **Annu Rev Biochem,** v. 79, p. 351-79, 2010.

FARMER, H. et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. **Nature,** v. 434, n. 7035, p. 917-21, Apr 14 2005.

FARMER, P. et al. Identification of molecular apocrine breast tumours by microarray analysis. **Oncogene,** v. 24, n. 29, p. 4660-71, Jul 7 2005.

GARZON, R.; CALIN, G. A.; CROCE, C. M. MicroRNAs in Cancer. **Annu Rev Med,** v. 60, p. 167-79, 2009.

GIOLO, S. R. et al. Brazilian urban population genetic structure reveals a high degree of admixture. **Eur J Hum Genet,** v. 20, n. 1, p. 111-6, Jan 2012.

GIRALDEZ, A. J. et al. Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. **Science,** v. 312, n. 5770, p. 75-9, Apr 7 2006.

GOBBI, H. Classificação dos tumores de mama: atualização baseada na nova classificação da Organização Mundial da Saúde de 2012. **Jornal Brasileiro de Patologia e Medicina Laboratorial,** v. 48, n. 6, p. 12, 2012.

GOLDHIRSCH, A. et al. Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. **Ann Oncol,** v. 22, n. 8, p. 1736-47, Aug 2011.

GRAMANTIERI, L. et al. Cyclin G1 is a target of miR-122a, a microRNA frequently down-regulated in human hepatocellular carcinoma. **Cancer Res,** v. 67, n. 13, p. 6092-9, Jul 1 2007.

GREENUP, R. et al. Prevalence of BRCA mutations among women with triple-negative breast cancer (TNBC) in a genetic counseling cohort. **Ann Surg Oncol,** v. 20, n. 10, p. 3254-8, Oct 2013.

HEALTH, N. I. O. National Institute on Minority Health and Health Disparities. 2017. Disponível em: < https://www.nimhd.nih.gov/about/overview/ >. Acesso em 10 de fevereiro de 2018.

HERSCHKOWITZ, J. I. et al. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. **Genome Biol,** v. 8, n. 5, p. R76, 2007.

HHS. Healthy People 2020. 2010. Disponível em: < https://www.healthypeople.gov/ >. Acesso em: 10 de fevereiro de 2018.

HURTEAU, G. J. et al. Overexpression of the microRNA hsa-miR-200c leads to reduced expression of transcription factor 8 and increased expression of E-cadherin. **Cancer Res,** v. 67, n. 17, p. 7972-6, Sep 1 2007.

INCA/MS. **Estimativa 2018 - Incidência de Câncer no Brasil**. Rio de Janeiro 2018.

KAUFMAN, B. et al. Olaparib monotherapy in patients with advanced cancer and a germline BRCA1/2 mutation. **J Clin Oncol,** v. 33, n. 3, p. 244-50, Jan 20 2015.

KAVITHA, N. et al. MicroRNAs: biogenesis, roles for carcinogenesis and as potential biomarkers for cancer diagnosis and prognosis. **Asian Pac J Cancer Prev,** v. 15, n. 18, p. 7489-97, 2014.

KILLELEA, B. K. et al. Racial Differences in the Use and Outcome of Neoadjuvant Chemotherapy for Breast Cancer: Results From the National Cancer Data Base. **J Clin Oncol,** v. 33, n. 36, p. 4267-76, Dec 20 2015.

KOHLER, B. A. et al. Annual Report to the Nation on the Status of Cancer, 1975-2011, Featuring Incidence of Breast Cancer Subtypes by Race/Ethnicity, Poverty, and State. **J Natl Cancer Inst,** v. 107, n. 6, p. djv048, Jun 2015.

KROENKE, C. H. et al. Race and breast cancer survival by intrinsic subtype based on PAM50 gene expression. **Breast Cancer Res Treat,** v. 144, n. 3, p. 689-99, Apr 2014.

KUMAR, M. S. et al. Impaired microRNA processing enhances cellular transformation and tumorigenesis. **Nat Genet,** v. 39, n. 5, p. 673-7, May 2007.

LARA-MEDINA, F. et al. Triple-negative breast cancer in Hispanic patients: high prevalence, poor prognosis, and association with menopausal status, body mass index, and parity. **Cancer,** v. 117, n. 16, p. 3658-69, Aug 15 2011.

LEE, E. J. et al. Expression profiling identifies microRNA signature in pancreatic cancer. **Int J Cancer,** v. 120, n. 5, p. 1046-54, Mar 1 2007.

LEHMANN, B. D. et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. **J Clin Invest,** v. 121, n. 7, p. 2750-67, Jul 2011.

LEHMANN, B. D. et al. Refinement of Triple-Negative Breast Cancer Molecular Subtypes: Implications for Neoadjuvant Chemotherapy Selection. **PLoS One,** v. 11, n. 6, p. e0157368, 2016.

LEWIS, B. P. et al. Prediction of mammalian microRNA targets. **Cell,** v. 115, n. 7, p. 787-98, Dec 26 2003.

LI, M. et al. MicroRNAs: control and loss of control in human physiology and disease. **World J Surg,** v. 33, n. 4, p. 667-84, Apr 2009.

MARAGKAKIS, M. et al. DIANA-microT web server: elucidating microRNA functions through target prediction. **Nucleic Acids Res,** v. 37, n. Web Server issue, p. W273-6, Jul 2009.

MASUDA, H. et al. Differential response to neoadjuvant chemotherapy among 7 triple-negative breast cancer molecular subtypes. **Clin Cancer Res,** v. 19, n. 19, p. 5533-40, Oct 1 2013.

MICHELS, J. et al. Predictive biomarkers for cancer therapy with PARP inhibitors. **Oncogene,** v. 33, n. 30, p. 3894-907, Jul 24 2014.

MONTEMURRO, F.; DI COSIMO, S.; ARPINO, G. Human epidermal growth factor receptor 2 (HER2)-positive and hormone receptor-positive breast cancer: new insights into molecular interactions and clinical implications. **Ann Oncol,** v. 24, n. 11, p. 2715-24, Nov 2013.

MORTENSEN, R. D. et al. Posttranscriptional activation of gene expression in Xenopus laevis oocytes by microRNA-protein complexes (microRNPs). **Proc Natl Acad Sci U S A,** v. 108, n. 20, p. 8281-6, May 17 2011.

MS/SVS/CGIAE. Sistema de Informações sobre Mortalidade - SIM 2015. Disponível em: < http://tabnet.datasus.gov.br/cgi/tabcgi.exe?sim/cnv/obt10br.def >. Acesso em 2 de fevereiro de 2018.

NORUM, J. H.; ANDERSEN, K.; SORLIE, T. Lessons learned from the intrinsic subtypes of breast cancer in the quest for precision therapy. **Br J Surg,** v. 101, n. 8, p. 925-38, Jul 2014.

O'BRIEN, K. M. et al. Intrinsic breast tumor subtypes, race, and long-term survival in the Carolina Breast Cancer Study. **Clin Cancer Res,** v. 16, n. 24, p. 6100-10, Dec 15 2010.

O'SHAUGHNESSY, J. et al. Iniparib plus chemotherapy in metastatic triple-negative breast cancer. **N Engl J Med,** v. 364, n. 3, p. 205-14, Jan 20 2011.

PATEL, T. A. et al. Breast cancer in Latinas: gene expression, differential response to treatments, and differential toxicities in Latinas compared with other population groups. **Oncologist,** v. 15, n. 5, p. 466-75, 2010.

PEROU, C. M. et al. Molecular portraits of human breast tumours. **Nature,** v. 406, n. 6797, p. 747-52, Aug 17 2000.

PLASILOVA, M. L. et al. Features of triple-negative breast cancer: Analysis of 38,813 cases from the national cancer database. **Medicine (Baltimore),** v. 95, n. 35, p. e4614, Aug 2016.

PRAT, A.; BASELGA, J. The role of hormonal therapy in the management of hormonal-receptor-positive breast cancer with co-expression of HER2. **Nat Clin Pract Oncol,** v. 5, n. 9, p. 531-42, Sep 2008.

REIS-FILHO, J. S.; PUSZTAI, L. Gene expression profiling in breast cancer: classification, prognostication, and prediction. **Lancet,** v. 378, n. 9805, p. 1812-23, Nov 19 2011.

ROBSON, M. E. et al. OlympiAD: Phase III trial of olaparib monotherapy versus chemotherapy for patients (pts) with HER2-negative metastatic breast cancer (mBC) and a germline BRCA mutation (gBRCAm). **Journal of Clinical Oncology,** v. 35, n. 18_suppl, p. LBA4-LBA4, 2017.

RODRIGUEZ, A. et al. Identification of mammalian microRNA host genes and transcription units. **Genome Res,** v. 14, n. 10A, p. 1902-10, Oct 2004.

SALZANO, F. M.; SANS, M. Interethnic admixture and the evolution of Latin American populations. **Genet Mol Biol,** v. 37, n. 1 Suppl, p. 151-70, Mar 2014.

SARAIYA, A. A.; LI, W.; WANG, C. C. Transition of a microRNA from repressing to activating translation depending on the extent of base pairing with the target. **PLoS One,** v. 8, n. 2, p. e55672, 2013.

SCHINKEL, J. K. et al. Racial/ethnic differences in breast cancer survival by inflammatory status and hormonal receptor status: an analysis of the Surveillance, Epidemiology, and End Results data. **Cancer Causes Control,** v. 25, n. 8, p. 959-68, Aug 2014.

SHARMA, P. et al. Germline BRCA mutation evaluation in a prospective triple-negative breast cancer registry: implications for hereditary breast and/or ovarian cancer syndrome testing. **Breast Cancer Res Treat,** v. 145, n. 3, p. 707-14, Jun 2014.

SHEPPARD, V. B. et al. The feasibility and acceptability of a diet and exercise trial in overweight and obese black breast cancer survivors: The Stepping STONE study. **Contemp Clin Trials,** v. 46, p. 106-113, Jan 2016.

SHI, X. B. et al. An androgen-regulated miRNA suppresses Bak1 expression and induces androgen-independent growth of prostate cancer cells. **Proc Natl Acad Sci U S A,** v. 104, n. 50, p. 19983-8, Dec 11 2007.

SHIOBARA, M. et al. Enhanced polyadenosine diphosphate-ribosylation in cirrhotic liver and carcinoma tissues in patients with hepatocellular carcinoma. **J Gastroenterol Hepatol,** v. 16, n. 3, p. 338-44, Mar 2001.

SIEGEL, R. L.; MILLER, K. D.; JEMAL, A. Cancer Statistics, 2017. **CA Cancer J Clin,** v. 67, n. 1, p. 7-30, Jan 2017.

_____. Cancer statistics, 2018. **CA Cancer J Clin,** v. 68, n. 1, p. 7-30, Jan 2018.

SINN, H. P.; KREIPE, H. A Brief Overview of the WHO Classification of Breast Tumors, 4th Edition, Focusing on Issues and Updates from the 3rd Edition. **Breast Care (Basel),** v. 8, n. 2, p. 149-54, May 2013.

SORLIE, T. et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. **Proc Natl Acad Sci U S A,** v. 98, n. 19, p. 10869-74, Sep 11 2001.

SORLIE, T. et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. **Proc Natl Acad Sci U S A,** v. 100, n. 14, p. 8418-23, Jul 08 2003.

TAKAHASHI, R. U.; MIYAZAKI, H.; OCHIYA, T. The Roles of MicroRNAs in Breast Cancer. **Cancers (Basel),** v. 7, n. 2, p. 598-616, 2015.

United States Census Bureau. 2018. Disponível em: < https://www.census.gov/ >. Acesso em 2 de fevereiro de 2018.

VAN ROOIJ, E.; KAUPPINEN, S. Development of microRNA therapeutics is coming of age. **EMBO Mol Med,** v. 6, n. 7, p. 851-64, Jul 2014.

VASUDEVAN, S.; TONG, Y.; STEITZ, J. A. Cell-cycle control of microRNA-mediated translation regulation. **Cell Cycle,** v. 7, n. 11, p. 1545-9, Jun 1 2008.

WANG, X.; EL NAQA, I. M. Prediction of both conserved and nonconserved microRNA targets in animals. **Bioinformatics,** v. 24, n. 3, p. 325-32, Feb 01 2008.

WARNER, E. T. et al. Impact of race, ethnicity, and BMI on achievement of pathologic complete response following neoadjuvant chemotherapy for breast cancer: a pooled analysis of four prospective Alliance clinical trials (A151426). **Breast Cancer Res Treat,** v. 159, n. 1, p. 109-18, Aug 2016.

WEIL, M. K.; CHEN, A. P. PARP inhibitor treatment in ovarian and breast cancer. **Curr Probl Cancer,** v. 35, n. 1, p. 7-50, Jan-Feb 2011.

WHITEHEAD, M. The concepts and principles of equity and health. **Int J Health Serv,** v. 22, n. 3, p. 429-45, 1992.

YANG, H. et al. MicroRNA expression profiling in human ovarian cancer: miR-214 induces cell survival and cisplatin resistance by targeting PTEN. **Cancer Res,** v. 68, n. 2, p. 425-33, Jan 15 2008.

ZACK, T. I. et al. Pan-cancer patterns of somatic copy number alteration. **Nat Genet,** v. 45, n. 10, p. 1134-40, Oct 2013.

ZHANG, L. et al. microRNAs exhibit high frequency genomic alterations in human cancer. **Proc Natl Acad Sci U S A,** v. 103, n. 24, p. 9136-41, Jun 13 2006.

ZHANG, W.; LIU, J.; WANG, G. The role of microRNAs in human breast cancer progression. **Tumour Biol,** v. 35, n. 7, p. 6235-44, Jul 2014.