Type of the Paper (Original Article)

Whole Genome Transcriptome Analysis of the Association Between Obesity and Triple Negative Breast Cancer in Caucasian women

[1]Tarun Mamidi, [1]Jiande Wu, [2]Paul B. Tchounwou, [1]Lucio Miele and [1]Chindo Hicks
[1]Department of Genetics, Louisiana State University Health Sciences, School of Medicine, 533 Bolivar Street, New Orleans, LA 70112, USA
[2]NIH/NIMHD RCMI Center for Environmental Health, Jackson State University, 1400 Lynch Street, Box 18750, Jackson, MS 39217, USA

[1*]Chindo Hicks chick3@lsuhsc.edu
[1]Tarun Mamidi tmamid@lsuhsc.edu
[1]Jiande Wu jwu2@lsuhsc.edu
[1]Lucio Miele lmiele@lsuhsc.edu
[2]Paul B. Tchounwou paul.b.tchounwou@jsums.edu

*Corresponding author: chick3@lsuhsc.edu Tel: 01-504-568-2657

Abstract: Background: Triple negative breast cancer (TNBC) is the most aggressive form of breast cancer with poor outcomes. The molecular basis of TNBC remains poorly understood. The objective of this study was to explore the relationship between obesity and TNBC in premenopausal and postmenopausal Caucasian women using whole genome transcription profiling. **Methods:** We compared gene expression levels of tumor samples drawn from normal weight, overweight and obese in pre and postmenopausal women diagnosed with TNBC. We performed hierarchical clustering to assess similarity in patterns of gene expression profiles, and conducted network and pathway analysis to identify molecular networks and biological pathways. **Results:** We discovered gene signatures distinguishing normal weight from obese, normal weight from overweight and overweight from obese individuals in both premenopausal and postmenopausal women. The analysis revealed molecular networks and biological pathways dysregulated in response to obesity. Among the discovered pathways included the unfolded protein response, endoplasmatic reticulum stress, B cell receptor and the autophagy signaling pathways in obese premenopausal women and the integrin, axonal guidance, ERK/MAPK and Glutathione biosynthesis signaling pathways obese postmenopausal women. **Conclusions:** The results suggest that both overweight and obesity are associated with TNBC, highlighting the need for conformation of these results in independent studies.

Keywords: Gene expression obesity triple negative breast cancer

 doi:10.20944/preprints201803.0174.v1

1. Introduction

Triple negative breast cancer (TNBC) represents the breast cancers which lack expression of estrogen receptor (ER) and progesterone receptor (PR) and lack of amplification of the human epidermal growth factor receptor 2 (HER2) gene [1]. TNBC is a heterogeneous disease with a complex etiology. It is the most aggressive form of breast cancer with very poor clinical outcomes. Although TNBC represents only 15% of all breast cancers, it accounts for 25% of all breast cancer-related deaths [1-2]. Women with TNBC have a high frequency of metastasis to the lung, liver and brain, and survival is generally poor [3]. Even more concerning is that the median survival rate for women with metastatic TNBC is less than one year, and almost all patients die of their disease [3]. To date, there are no effective targeted therapies, chemotherapy remains the only effective therapeutic modality [1-3]. Therefore, there is a pressing need to understand the biological factors and pathways that driver these tumors and discovery of molecular markers for the development of targeted therapies.

Over the last decade, there has been considerable interest in investigating the role of modifiable factors such as obesity in women diagnosed with TNBC. This has been driven in part by the realization that factors such as socio-economic status and lifestyle may be associated with the disease [4-7]. To this end, epidemiologic studies have attempted to establish the relationship between obesity and TNBC. Several epidemiologic studies have associated overweight and obesity with TNBC [8-9]. Others have associated overweight and obesity with TNBC overall survival rate (OS) and disease free survival rate (DFS) [10-16]. In addition, several studies have shown that obesity is an independent prognostic factor of decreased pathological complete response to neoadjuvant chemotherapy in breast cancer patients [17-18]. However, there are no studies investigating the molecular mechanism association obesity and or overweight with TNBC. Moreover, other epidemiological studies did not find the association between overweight or obesity with TNBC [19-21]. Thus, taken together, these knowledge and seemingly contradictory results along with the need for biomarker discovery for TNBC point to the need for further research in this area.

TNBC affects both premenopausal and postmenopausal women. A recent epidemiologic study involving 326 TNBC patients treated with neoadjuvant chemotherapy identified BMI and menopausal status as two promising prognostic factors in TNBC [22]. However, to date little is known about the molecular mechanisms associating BMI with TNBC by menopausal status. Given the expanding obesity epidemic and the poor prognosis of the TNBC tumors, discovery of molecular markers associated with modifiable risk factors such as obesity may facilitate the development of novel prevention strategies and the realization of precision prevention. The objective of this exploratory study was we two-fold: (i) To investigate the association between obesity and/or overweight with TNBC in premenopausal and postmenopausal Caucasian women using whole genome transcription profiling and (ii) To discover the molecular networks and biological pathways dysregulated in response to obesity in premenopausal and postmenopausal women with TNBC. Our working hypothesis is that genomic alterations in tumors in premenopausal and postmenopausal women diagnosed with TNBC could lead to measurable changes associating obesity and overweight with

 doi:10.20944/preprints201803.0174.v1

TNBC. We further hypothesized that these changes in genomic alterations affect entire molecular networks and biological pathways which are dysregulated in response to increased weight or obesity. To test this hypothesis we used publicly available gene expression data derived from normal weight, overweight and obese premenopausal and postmenopausal Caucasian women diagnosed with TNBC to discover molecular signatures distinguishing patient groups and biological pathways dysregualated in response to overweight and or obesity.

2. Material and Methods

Gene expression data

We used publicly available gene expression data generated using tumor samples from premenopausal and postmenopausal Caucasian women of European ancestry diagnosed with TNBC. The data set was downloaded from the Gene Expression Omnibus (GEO) under accession number GSE76124. The data set consisted of 195 samples and was generated at Baylor University [23]. The experimental procedures have been fully described by the data originators [23]. The data set included body mass index (BMI), menopausal status defined as, either premenopausal or postmenopausal, age and tumor size. Samples without menopausal status, age and BMI were excluded from the final data set used in the analysis. The final data set included a total of 148 patients distributed according to menopausal status and BMI. The distribution of 54 patients with premenopausal status was: Normal weight (BMI ≤ 24; N = 21), overweight (BMI = 25 - 29; N = 21) and Obese (BMI = 30≥; N = 12). Similarly, the distribution of 94 patients with postmenopausal status was normal weight (BMI ≤ 24; N =25, over weight (BMI = 25 - 29; N = 31) and Obese (BMI = 30 ≥; N = 38). Because TNBC is a heterogeneous disease and its subtype classification has varied, we examined the literature to determine whether the classification used in this database represented the current reports in the literature. This evaluation revealed that the classification is consistent with other reports and meet the criteria of the current consensus on TNBC subtype classification [23-26]. The experimental procedures and methods of sample processing have been fully described by the data originators [23]. Clinico-pathological data from the patients used in the study include the tumor ER, PR and Her2/Neu status which were all negative. The data set was generated using the Affymetrix platform using the Human GeneChip U133Plus 2.0 which contains 54,614 probe sets). Expression values were calculated using the robust multi-array average (RMA) algorithm as implemented in the Affymetrix platform. All the expression values ware on a log scale (log2).

Data Analysis

We performed supervised analysis comparing gene expression levels among and between patient groups. We used analysis of variance (ANOVA) to

compare gene expression levels among the three patient groups: normal weight, overweight and obese by menopausal status. We performed supervised analysis using a t-test to compare gene expression levels between patient groups (normal weight versus overweight, normal weight versus obese and overweight versus obese), separately in premenopausal and postmenopausal women using Pomello and GenePattern Software packages [27-28]. Due to relatively small sample sizes for each patient group, we did not partition the data into test and validation sets as such an approach would lead to bias resulting from sampling errors. To address this issue, we used the leave-one-out cross-validation procedure as our prediction and validation model to identify genes with predictive power. This approach has been used successfully in gene expression data analysis to eliminate bias [29]. We used the false discovery rate (FDR) procedure to correct for multiple hypothesis testing [30]. Genes were ranked based on the p-values and the FDR, and highly significantly differentially expressed genes were selected for each comparison.

We performed unsupervised analysis using hierarchal clustering based on complete linkage model using the Pearson correlation coefficient as the measure of distance between pairs of genes. Prior to clustering, gene expression data was normalized using the median normalization, standardized and centered [31]. Hierarchical clustering was performed using GenePattern and Morpheus software [28,32]. We performed network and pathways analysis using Ingenuity Pathway Analysis (IPA) software [33]. Using IPA, the most highly significantly differentially expressed genes distinguishing patients with normal weight from obese patients in premenopausal and postmenopausal women were mapped onto networks and canonical pathways. The probability scores and the log *P*-values were calculated to assess the likelihood and reliability of correctly assigning the genes to the correct molecular networks and biological pathways. A false discovery rate was used to correct for multiple hypothesis testing in pathway analysis. The predicted molecular networks and biological pathways were ranked based on z-scores and log p-values; respectively. Gene ontology (GO) [34] analysis as implemented in IPA, was performed on the sets of differentially expressed genes to characterize the functional relationships among sets of genes associating overweight and obesity with TNBC and to identify the molecular, biological processes and cellular components in which the discovered genes are involved.

3. Results

Differences in gene expression levels among patient groups

To identify differentially expressed genes and assess variation in patterns of gene expression levels among the three patient groups, we performed analysis of variance by menopausal status. We hypothesized that the levels of gene expression differ and vary among patient groups in premenopausal and postmenopausal

 doi:10.20944/preprints201803.0174.v1

women. The analysis revealed significant differences in gene expression levels among patient groups. When we compared gene expression levels among patient groups in premenopausal women, we discovered a signature of 1034 significantly (P<0.05) differentially expressed genes. A subset of 242 genes were highly significantly (P<0.01) differentially expressed. The top 23 most highly significantly (P<0.001) differentially expressed genes included the genes *CD84, DUXAP8, NPC2, MAGEA5, PAWR, SNX29, IFNGR1, PRKXP1, WIPF1, ABCG1, DPY19L1, MGAT4A, KYNU, RNASET2, COX10-AS1, GPRIN3, MMD2, TMED10, FLVCR2, GABBR1, RPL32P3, RAPGEF1* and *LYST*.

Repeating the same analysis, but focusing on postmenopausal women revealed a signatures of 1551 significantly (P<0.05) differentially expressed genes. A subset of 376 genes were highly significantly (P<0.01) differentially expressed. The signature of the top 57 most highly significantly (P<0.001) differentially genes included the genes *IL4R, TAGLN2, ZNF92, MSX1, EPHA2, SERPINE1, PANX1, PDPN, KEAP1, STK10, KLF9, JRK, PLK3, ZNF138, PLEC, FPR1, CR1, ZNF85, TNC, SLC2A3, ANGPT2, MESDC1, ZEB2, CSGALNACT1, MUT, YWHAZ, ZNF140, PRKACA, MVP, COPS8, SEC23A, NNMT, YWHAH, SRSF7, CD163, HAS2, SCG2, LRRFIP1, PTPRE, EHD4, ZNF736, LOC101927523, DUSP1, PLP2, TWIST2, WWTR1-AS1, GRIA3, SEC62, IL6, KANSL1L, SCARF1, PROM2, RAPGEF2, BCAR3, OSMR, SMAD5* and *SPPL3.* There was no overlap between the two sets of highly significantly differentially expressed genes in premenopausal and postmenopausal women, suggesting that molecular perturbation in premenopausal and postmenopausal may regulated by different molecular mechanisms. As expected, there was significant variation in gene expression levels among patient groups in both premenopausal and postmenopausal women. This could be partially explained by the heterogeneity of the disease. TNBC is inherently a heterogeneous disease with many subtypes, therefore, such outcome should be expected. A list of all the significantly differentially expressed genes among patient groups by menopausal status is presented in supplementary Table SA, provided as supplementary data to this report.

Association of overweight and obesity with TNBC in premenopausal women

To address the hypothesis that overweight or obesity are associated with TNBC, we performed subclass mapping comparing gene expression levels between patients with normal weight and obese and between normal weight and overweight. We sought to discover gene signatures distinguishing individuals with normal weight from obese and or overweight.

Comparison of gene expression levels between normal and overweight patients revealed a signature of 1120 significantly (P<0.05) differentially expressed genes. The signature included 219 highly significantly (P<0.01) differentially expressed genes. A list of the top 32 most highly significantly genes distinguishing

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 March 2018 doi:10.20944/preprints201803.0174.v1

normal weight from overweight individuals is presented in Table 1. A complete list of the significantly differentially expressed genes distinguishing patients with normal weight from overweight individuals is presented in Table S1 provided as supplementary data to this report.

Analysis comparing gene expression levels between normal and obese individuals revealed a signature of 1218 significantly (P<0.05) differentially expressed genes. The signature included a set of 299 highly (P<0.01) significantly differentially expressed genes. A list of 32 most highly significantly (P<0.001) differentially expressed genes is presented in Table 1. A least of all the significantly differentially expressed genes distinguishing patients with normal weight from obese individuals is presented in Table S1 provided as supplementary data to this report.

To address the hypothesis that molecular perturbation in patients with overweight significantly differs from obese patients in premenopausal women, we compared gene expression levels between the two patient groups.

Table 1. List of the top 32 most highly significantly differentially expressed genes dysregulated in response to overweight and obese in premenopausal women

Normal weight versus Overweight				Normal weight versus Obese		
Gene Symbol	Cytoband	P-value		Gene Symbol	Cytoband	P-Value
COX10	17p12	0.000166		DUXAP8	22q11.1	0.0000984
DPY19L1	7p14.2	0.000236		ABCG1	21q22.3	0.000115
MYL6B	12q13.2	0.000242		IFNGR1	6q23.3	0.000187
CD84	1q23.3	0.000321		TPST2	22q12.1	0.000191
RAPGEF1	9q34.13	0.000325		CCDC32	15q15.1	0.000205
PAWR	12q21.2	0.000396		SFT2D1	6q27	0.00024
RPS21	20q13.33	0.000399		SMC2	9q31.1	0.00024
IL13RA1	Xq24	0.000496		NPC2	14q24.3	0.000251
GABBR1	6p22.1	0.000506		MGAT4A	2q11.2	0.000256
SLF2	10q24.31	0.000514		RNASET2	6q27	0.000261
ZNF621	3p22.1	0.000586		WAPL	10q23.2	0.000261
TOP1MT	8q24.3	0.000644		ATXN1	6p22.3	0.000267
MMD2	7p22.1	0.00068		ALG5	13q13.3	0.00027
N4BP2L2	13q13.1	0.000742		PTEN	10q23.31	0.000358
SNX29	16p13.13	0.000777		CREBL2	12p13.1	0.000429
CCDC34	11p14.1	0.000803		SUSD6	14q24.1	0.000447
EPB41L4A	5q22.1	0.000829		ZC3H7B	22q13.2	0.000461
MAP3K3	17q23.3	0.000887		ITFG1	16q12.1	0.000497
LYST	1q42.3	0.00093		FXN	9q21.11	0.000527
STC2	5q35.2	0.001075		BLVRA	7p13	0.000542
IQGAP1	15q26.1	0.001177		ICAM3	19p13.2	0.000554

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 March 2018 doi:10.20944/preprints201803.0174.v1

KYNU	2q22.2	0.001258		XBP1	22q12.1	0.000671
FBXO28	1q42.11	0.001358		CLN5	13q22.3	0.000682
ARSD	Xp22.3	0.001402		SEL1L	14q31	0.000687
NBN	8q21.3	0.001535		DNAJC3	13q32.1	0.000716
IFNGR1	6q23.3	0.001573		LAYN	11q23.1	0.000735
CASP9	1p36.21	0.001591		GPRIN3	4q22.1	0.000833
EHD4	15q15.1	0.00167		WIPF1	2q31.1	0.000838
SSH1	12q24.11	0.001679		SET	9q34.11	0.000896
SYN2	3p25.2	0.001725		FOXN3	14q31.3	0.00093
ZNF585A	19q13.13	0.001861		FBXO10	9p13.2	0.000936
FCGR3A	1q23.	0.001938		SNX29	16p13.13	0.000962

The analysis revealed a signature of 635 significantly (P<0.05) differentially expressed genes at a nominal p-value (P<0-05). A subset of 92 genes were highly significantly (P<0.01) differentially expressed genes. There was little overlap between genes associated with overweight and those associated with obesity, suggesting that overweight and obesity may be regulated by different biological mechanisms in premenopausal women.

Association of TNBC with obesity in postmenopausal women

Next we performed subclass mapping in postmenopausal women to address the hypothesis that obesity or weight are associated with TNBC in this the group of women. The goal was to discover gene signatures distinguishing women with normal weight from obese women and women with normal weight from overweight women.

Comparison of gene expression levels between individuals with normal weight and obese patients revealed a signature of 1556 significantly (P<0.05) differentially expressed genes at a nominal P-value (P<0.05). Among the significantly differentially expressed genes distinguishing women with normal from obese women, 401 genes were highly significantly (P<0.01) differentially expressed. A signature of the top 44 most highly (P<0.001) significantly differentially expressed genes is presented in Table 2. A complete list of all the significantly differentially expressed genes between normal and obese patients is presented in Table S2 provided as supplementary data to this report.

Analysis comparing patients with normal weight to overweight individuals produced a signature of 1327 significantly (P<0.05) differentially expressed genes. The signature included 560 highly significantly (P<0.01) differentially expressed genes. A signature of the top 32 most highly significantly (P<0.001) differentially expressed genes are presented in Table 2. A complete list of significantly differentially expressed genes distinguishing women with normal weight from women with overweight is presented in Table S2 provided as supplementary data to this report. There was little overlap between the two sets of genes.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 March 2018 doi:10.20944/preprints201803.0174.v1

To address the hypothesis that molecular perturbation differs between overweight and obese women in postmenopausal, we compared gene expression levels between the two patient groups. Comparison of gene expression levels between overweight and obese individuals produced a signature of 1438 significantly ($P<0.05$) differentially expressed genes. The signature included 367 highly significantly differentially expressed genes between the two patient groups. The top most highly significantly differentially expressed genes distinguishing overweight women from obese women in postmenopausal included the genes *ZNF230, PANX1, KLF9, EHD4, ACOT11, SPPL3, SEC23A, SEC62, CSGALNACT1, CCNY, WWC2, SNX19, WBP1L, COPS8, PPP2R2A, LRRFIP1, SMG7, ARF1, DUSP4, LOC101927523, LRCH3, BCAP29, PDPN, SMS, TRPC1, ANGPT2, ZNF140, PKD2, PLP2, CCDC7, SSBP2, CYP2U1, MGAT2, FOXP2, YWHAZ, IGBP1, STK17B, KCNQ3, DUSP1, TRIM32, SCARB1, PTGER4, PICALM, PSMF1* and *JRK*. A complete list of significantly differentially expressed genes distinguishing overweight from obese individuals in postmenopausal women is presented in Table S2 provided as supplementary data to this report.

Table 2. List of the top 32 most highly significantly differentially expressed genes dysregulated in response to obese and overweight postmenopausal women

	NW vs Obese				NW vs Over weight	
Gene Symbol	Cytoband	P-value		Gene Symbol	Cytoband	P-value
MSX1	4p16.2	0.000002		TAGLN2	1q23.2	0.0000002
IL4R	16p12.1	0.0000023		ZNF92	7q11.21	0.0000778
STK10	5q35.1	0.0000194		KEAP1	19p13.11	0.0000934
PLK3	1p34.1	0.0000257		ZNF253	19p13.11	0.0000988
EPHA2	1p36	0.0000278		YWHAH	22q12.3	0.000138
MUT	6p12.3	0.0000321		PLEC	8q24	0.000146
SERPINE1	7q22.1	0.0000412		TMC6	17q25.3	0.000161
MVP	16p11.2	0.0000842		CARHSP1	16p13.2	0.000196
TNC	9q33.1	0.000134		YWHAZ	8q22.3	0.000224
OSMR	5p13.1	0.000138		JRK	8q24.3	0.000255
MESDC1	15q25.1	0.000141		PROM2	2q11.1	0.000269
RBMS1	2q24.2	0.000196		SMAD5	5q31.1	0.000284
PDPN	1p36.21	0.000199		ZNF138	7q11.21	0.000293
ZEB2	2q22.3	0.000218		MESDC1	15q13	0.000314
CTDSP2	12q14.1	0.000243		ZNF85	19p12	0.000344
PTPRE	10q26.2	0.000304		ZNF736	7q11.21	0.00042
TWIST2	2q37.3	0.000335		NADSYN1	11q13.4	0.00048
PML	15q24.1	0.000349		SRSF9	12q24.31	0.000492
C6orf141	6p12.3	0.000386		BAHD1	15q15.1	0.000518
WWTR1-AS1	3q25.1	0.000441		ZBTB3	11q12.3	0.000528
MGAT1	5q35.3	0.000446		SRSF7	2p22.1	0.000554

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 March 2018 doi:10.20944/preprints201803.0174.v1

GCLC	6p12.1	0.000463		BBS9	7p14.3	0.000619
PELO	5q11.2	0.000495		FOSL2	2p23.2	0.000661
SLC2A3	12p13.31	0.000506		ACLY	17q21.2	0.000676
SSH1	12q24.11	0.000542		KANSL1L	2q34	0.000728
PRKACA	19p13.1	0.000627		TFAP2A-AS1	6p24.3	0.000766
PXN	12q24.23	0.000652		SMG7	1q25.3	0.000771
PLEC	8q24	0.000673		DUOXA1	15q21.1	0.000772
FAS	10q23.31	0.000737		HNF4A-AS1	20q13.12	0.0008
DUSP1	5q35.1	0.000777		EPHA2	1p36.13	0.000811
PTAFR	1p35.3	0.000786		C11orf54	11q23.1	0.00097
SCARF1	17p13.3	0.000807		ZNF506	19p13.11	0.000979
KLF7	2q3.3	0.000839		C11orf57	11q23.1	0.001022
FPR1	19q13.41	0.000848		HPCAL1	2p25.1	0.001038
CD163	12p13.31	0.000862		KLHDC7B	22q13.33	0.001083
RAP1B	12q15	0.000879		PLP2	Xp11.23	0.001178
CTSL	9q21.33	0.000896		ZNF592	15q25.2	0.001202
GGT5	22q11.23	0.000929		TTLL13P	15q26.1	0.001204
GLI2	2q14.2	0.000951		RAB34	17q11.2	0.001223
ITGA5	12q13.13	0.00097		ATM	11q22.3	0.001286
RBBP6	16p12.1	0.000979		STX8	17p13	0.001314
MKL1	22q13.1	0.000982		PAX8	2q14.1	0.001328
ST7L	1p13.2	0.000986		ODF3	11p15.5	0.001342
ANGPT2	8p23.1	0.000992		NEK1	4q33	0.001358

Premenopausal versus postmenopausal

Having established that obesity and overweight are both associated with TNBC in both premenopausal and postmenopausal women, we performed additional analysis to investigate whether molecular perturbation in overweight and obese premenopausal and postmenopausal women diagnosed with TNBC is likely modulated by different biological mechanisms. We hypothesized that genes associating obesity or overweight with TNBC in premenopausal women are not the same genes associating obesity or overweigh with TNBC in postmenopausal women. We addressed this hypothesis by evaluating genes found to be associated with obesity or overweight in premenopausal and postmenopausal women separately. To avoid any spurious discoveries, we focused on highly differentially expressed sets of genes.

The results showing highly significantly differentially expressed genes which overlap between premenopausal and postmenopausal are presented in Venn diagrams in Figure 1(1A for obese and 1B for overweight). Evaluation of differentially expressed genes for obesity revealed 11 overlapping genes between the two patient groups. A similar evaluation of differentially expressed genes for overweight between

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 March 2018 doi:10.20944/preprints201803.0174.v1

the two patient groups revealed only 2 overlapping genes between the two patient groups. This suggests that obesity could potentially have divergent impacts on risk of TNBC in premenopausal versus postmenopausal women.

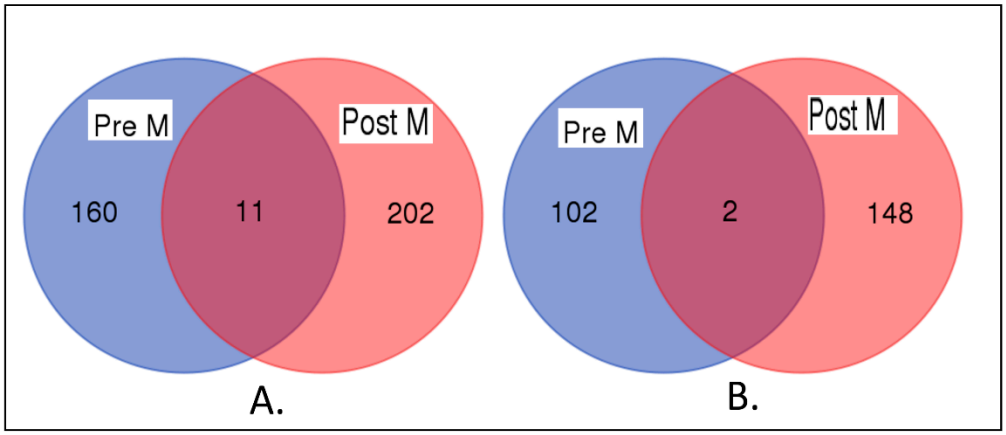

Figure 1. (A) Venn diagram showing overlap between genes associated with obesity in premenopausal and postmenopausal women. (B) Venn diagram showing overlap between genes associated with overweight in premenopausal and postmenopausal women diagnosed with TNBC. Pre M and Post M denote premenopausal and postmenopausal, respectively.

Similarity in patterns of gene expression profiles

To determine whether genes dysregulated in response to overweight and obesity in premenopausal and postmenopasaul women are co-regulated and have similar patterns of expression profiles, we performed unsupervised analysis using hierarchical clustering by menopausal status. To ensure reliability of clustering we focused on genes that were highly significantly ($P<0.01$) associated with obesity and or overweight. The results are presented in Figure 2A for 171 highly significantly differentially expressed genes associating obesity with TNBC and Figure 2B for 102 significantly differentially expressed genes associating overweight with TNBC in premenopausal women. In both cases, the genes were co-expressed and had similar patterns of expression profiles. Similarly for postmenopausal women Figure 3A for the 213 significantly differentially expressed genes associating obese with TNBC and Figure 3B for the 146 significantly differentially expressed genes associating overweight with TNBC were co-expressed and had similar patterns of expression profile. As expected there were significant variations in patterns of expression profiles. TNBC is inherently a heterogeneous disease consisting of different subtypes, thus under such conditions, the observed outcome was expected.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 March 2018 doi:10.20944/preprints201803.0174.v1

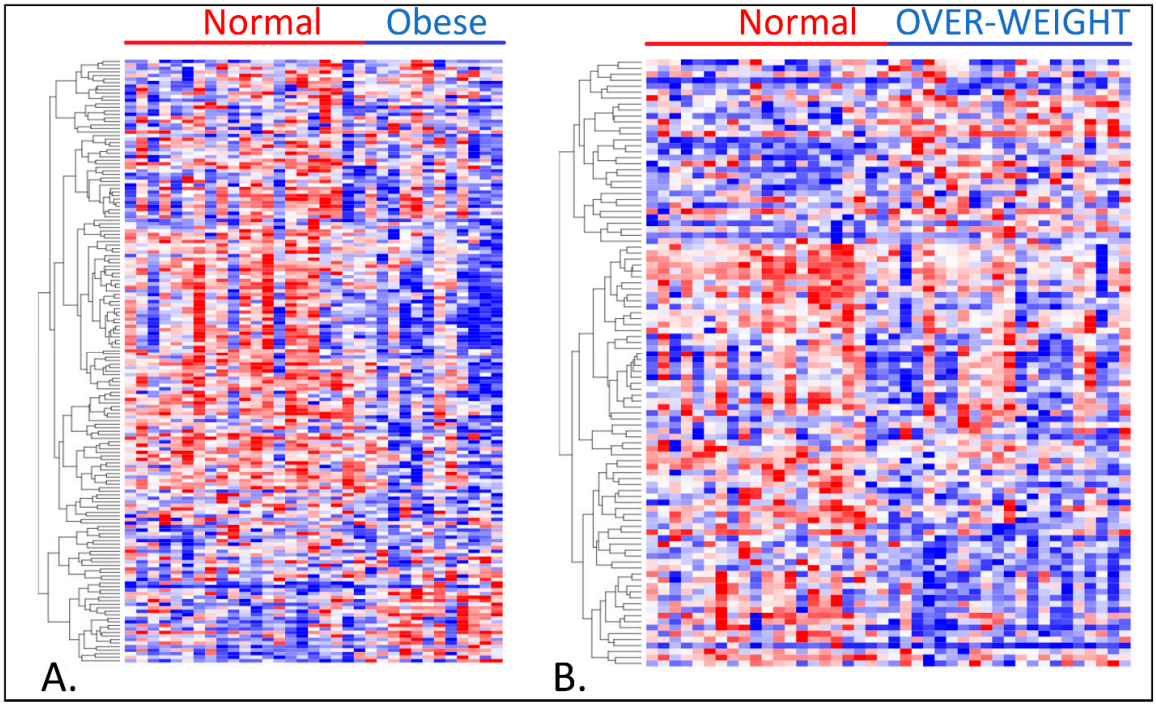

Figure 2. (A) Patterns of gene expression profiles for the 171 genes differentially expressed between women with normal weight and obese women with premenopausal status. (B) Patterns of gene expression profiles for the 102 genes differentially expressed between women with normal weight and obese women with premenopausal status. Genes in rows and patients in columns. Red font indicates up regulated and blue font indicates down regulated.

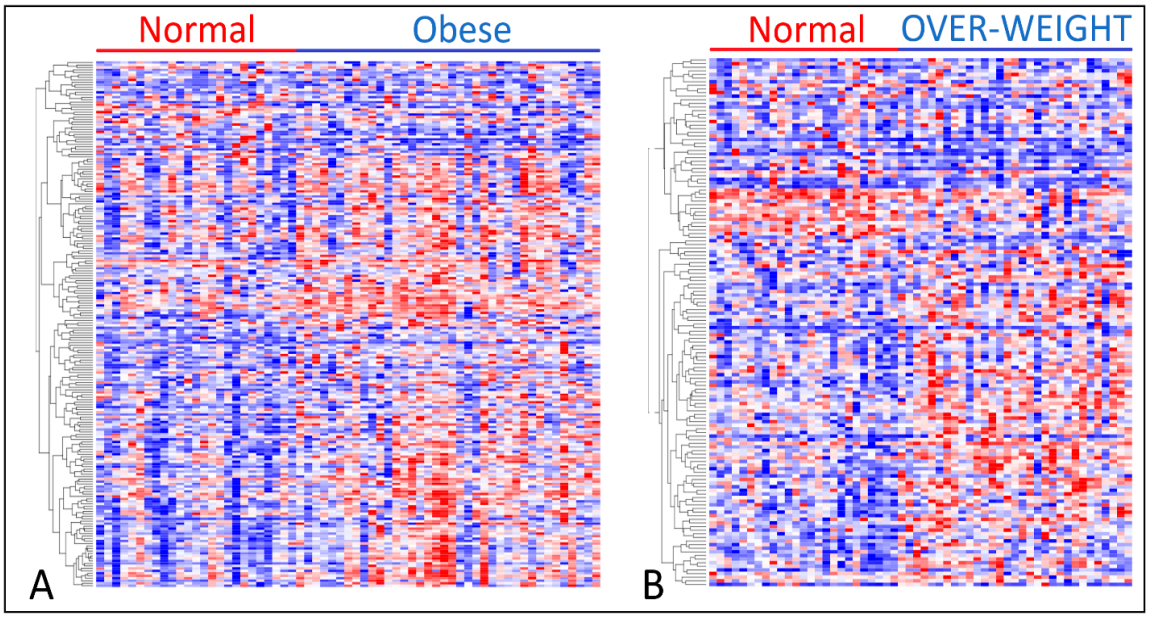

Figure 3. (A) Patterns of gene expression profiles for the 213 genes differentially expressed between women with normal weight and obese women with postmenopausal status. (B) Patterns of gene expression profiles for the 146 genes differentially expressed between women with normal weight and overweight women

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 March 2018 doi:10.20944/preprints201803.0174.v1

with postmenopausal status. Genes in rows and patients in columns. Red font indicates up regulated and blue font indicates down regulated.

Molecular networks and biological pathways

To gain insights about the broader biological context in which genes associating obesity with TNBC operate in premenopausal and postmenopausal women, we performed network and pathway analysis by menopausal status using IPA and focusing on highly significantly differentially expressed genes associating obesity with TNBC. We hypothesized that genes involved in obesity interact with one another in molecular networks and biological pathways.

The results of network and pathway analysis in premenopausal women are presented in Figure 4 and Figure 5; respectively. Network analysis revealed genes predicted to be significantly involved in cell cycle, cell death and survival, cellular development, cellular growth and proliferation, cell morphology and cellular function and maintenance. The most significant genes in the network included *PTPRE, PTPRF, PTEN, ATXN1, MAP3K5, FAS, FOXO1, E2F1* and *ATG7* (Figure 4). Pathway analysis using the same set of genes, revealed pathways predicted to be highly significantly associated with unfolded protein response, endoplasmic reticulum stress pathway, the B cell receptor signaling pathway, production of nitric oxide and reactive oxygen species in macrophages and the autophagy signaling pathways (Figure 5). The analysis also revealed top upstream regulators including *CD3, SEL1L, TGFB1* and *TNFSF11*.

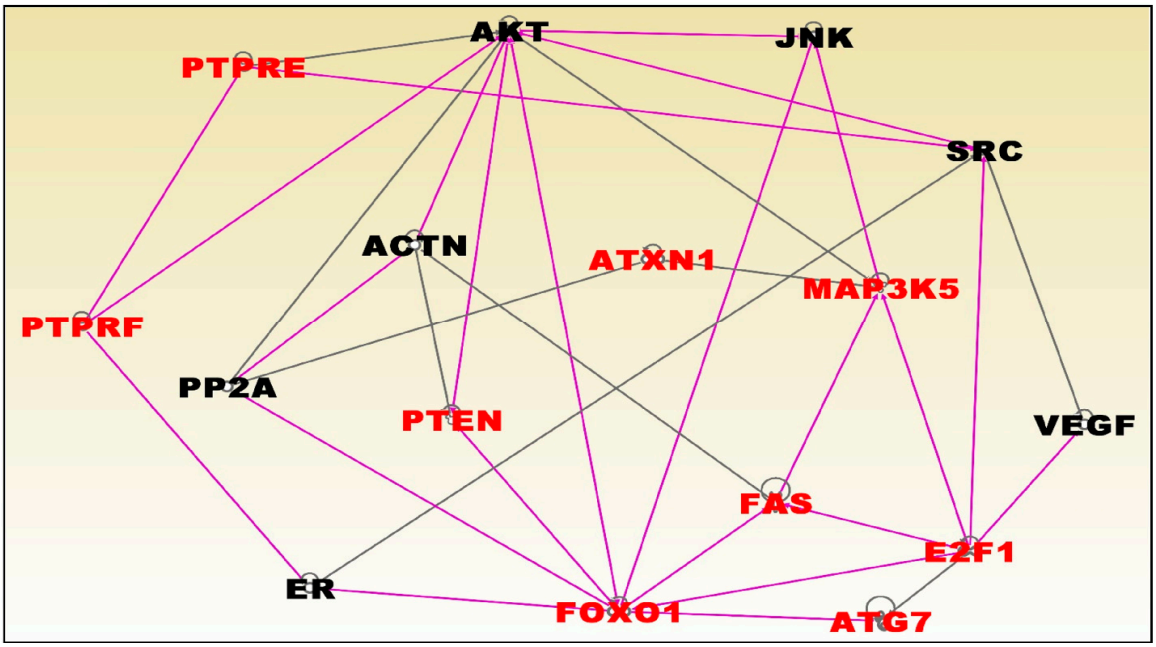

Figure 4. Molecular networks containing genes predicted to be significantly associated with obesity in premenopausal women. Network analysis was based on higly significantly differentially expressed genes (P<0.01) in red fonts. Gene symbols in red fonts were predicted to be highly significantly associated with obesity. Genes

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 March 2018 doi:10.20944/preprints201803.0174.v1

in black symbols are predicted to be functionally related. The pink and black lines denote the relationships between merged networks.

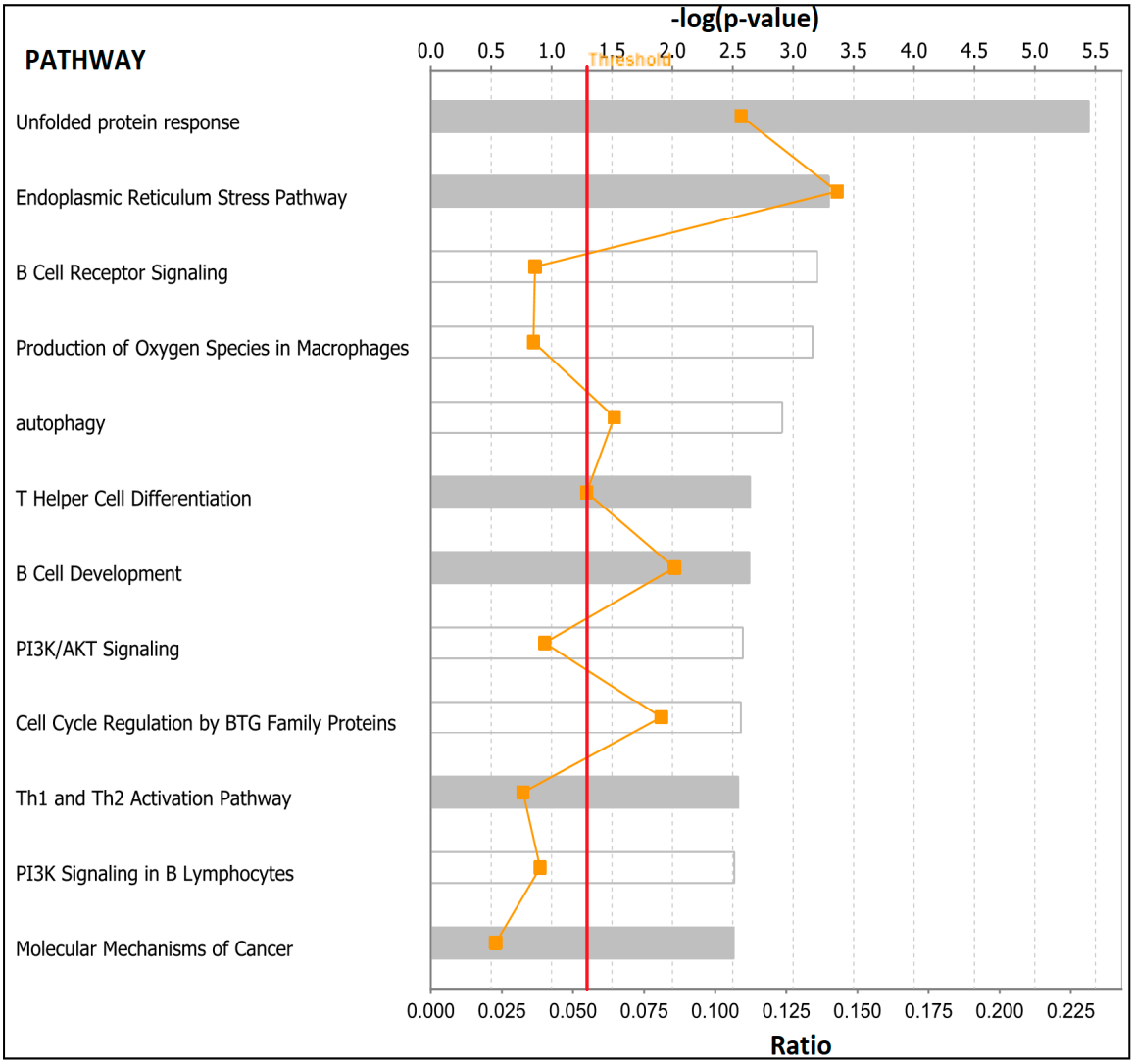

Figure 5. Biological pathways predicted to be highly significantly associated with obesity in premenopausal women. Pathway analysis was based on the most significantly differentially expressed genes. The red line indicates the threshold level above which significance is declared. The zigzagging orange line denotes the ratio of the number of genes predicted to map to that pathway to the original number of genes in that pathway.

Repeating network and pathways analysis focusing on the set of genes associating obesity with TNBC in postmenopausal women revealed molecular networks containing genes predicted to be significantly involved in cellular movement, cell-to-cell signaling and interactions, cell death and survival, cellular function and maintenance, cell development, drug metabolism and cellular growth and proliferation (Figure 6).The most significant genes in the networks were *CSF1R, SHC1, IQGAP1, PXN, HMOX1, CXCL8, COL1A1, ITGA5, CYRG1, JUNB, PDPN,*

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 March 2018 doi:10.20944/preprints201803.0174.v1

PAK2 and *NR3C1*(Figure 6). Pathway analysis revealed the integrin, axonal guidance, hepatic fibrosis, ERK/MAPK and the signaling, pathways predicted glutathione biosynthesis signaling pathways (Figure 7). The top upstream regulators included *TNF, TGFB1, Cycloheximide, lipopolysaccharide and the IL1*.

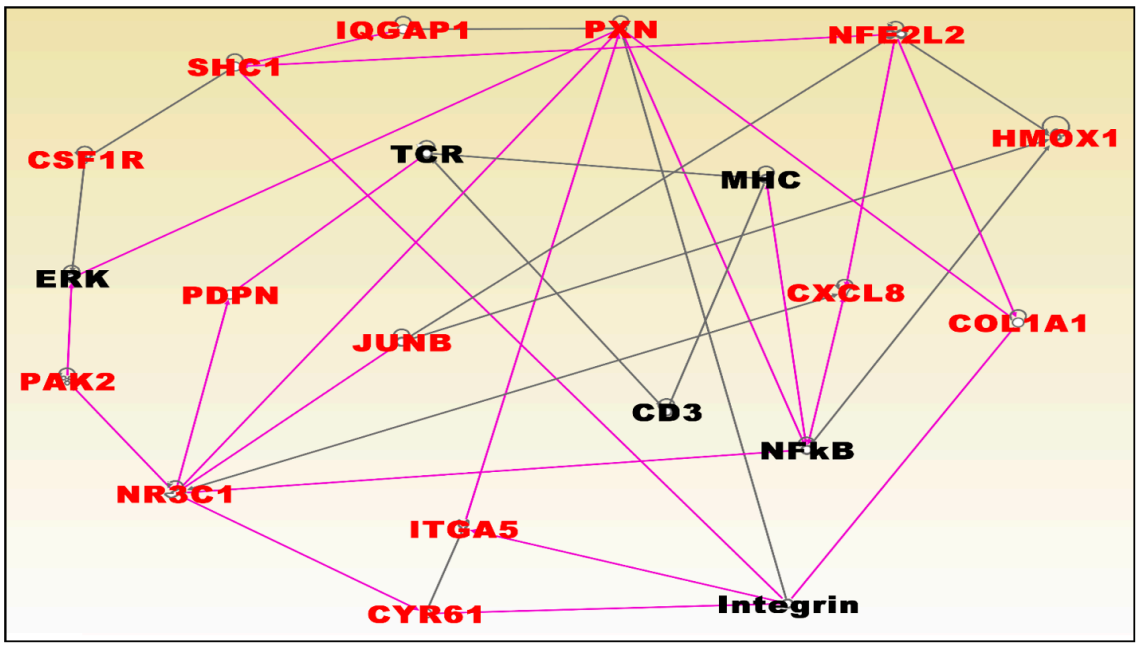

Figure 6. Molecular networks predicted to be significantly associated with obesity in postmenopausal women. Network analysis was based on the most significantly differentially expressed genes. Gene symbols in red fonts were predicted to be highly significantly associated with obesity. Genes in black symbols are predicted to be functionally related.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 March 2018 doi:10.20944/preprints201803.0174.v1

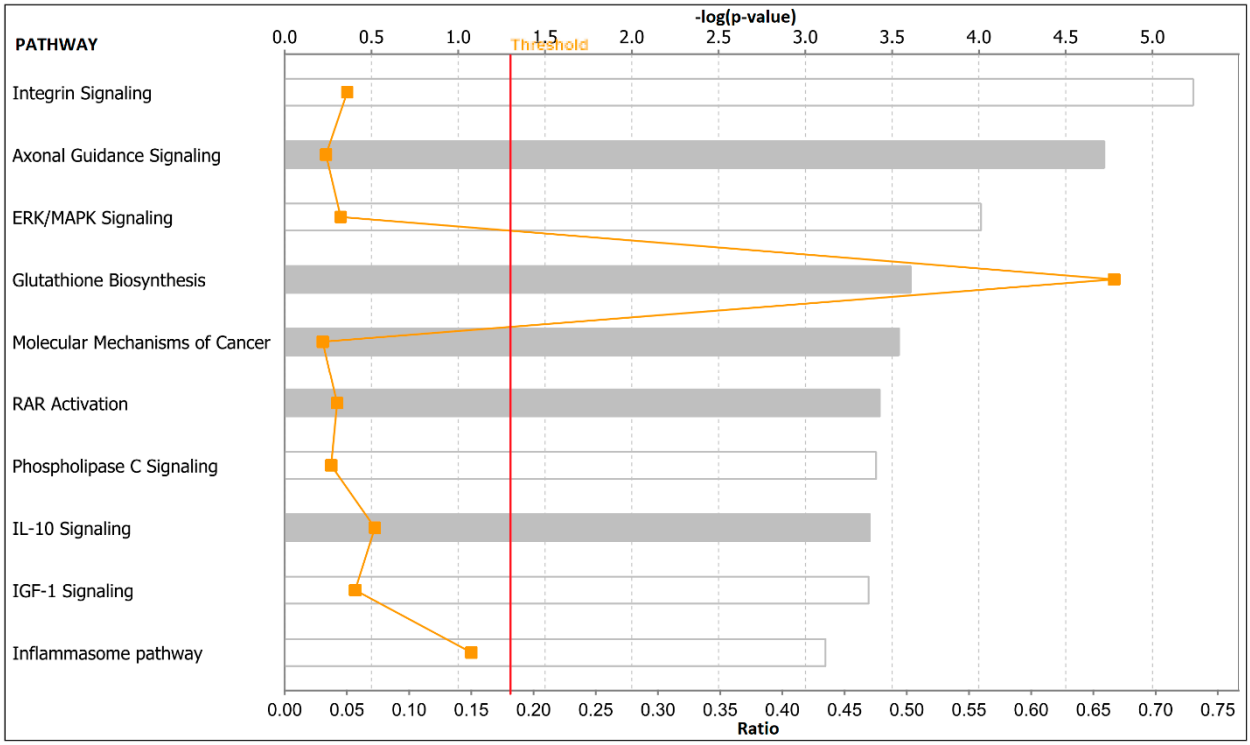

Figure 7. Top biological pathways predicted to be highly significantly associated with obesity in postmenopausal women. Pathway analysis was based on the most significantly differentially expressed genes. The red line indicates the threshold level above which significance is declared.

4. Discussion

We performed whole genome transcriptome analysis to determine whether obesity and overweight are associated with TNBC in premenopausal and postmenopausal Caucasian women. In both patient groups, obesity and overweight were associated with TNBC suggesting that overweight and obesity likely play a role in the etiology of TNBC. These results are consistent with several epidemiological studies which have associated obesity or overweight with TNBC [8-9]. The novel aspect of this study is that it delineates the molecular mechanisms associating overweight and obesity with TNBC in both premenopausal and postmenopausal women. To our knowledge this is the first study to use transcription profiling to investigate the relationship between obesity and TNBC in both premenopausal and postmenopausal women. The clinical significance of the results is that, given the expanding obesity epidemic in the US and its increase world-wide and the lack of targeted therapies for the discovered biomarkers could be used for precision prevention and the development of novel therapeutics. Importantly, molecular markers associating overweight and obesity with TNBC could potentially serve prognostic markers. Epidemiological studies have shown that overweight is an independent prognostic factor for overall survival and disease free survival [10].

 doi:10.20944/preprints201803.0174.v1

The little overlap or lack thereof in molecular perturbations for overweight and obese women between premenopausal and postmenopausal women is of particular interest. This along with differences in biological pathways modulating obesity in the two groups of women tends to suggest that molecular perturbation between premenopausal and postmenopausal women with TNBC may be controlled by different biological mechanisms. This is particularly interesting because epidemiological studies have shown that premenopausal women with overweight are at greater risk of death and progression than women with normal weight [10]. It is conceivable that molecular perturbation in obese individuals with TNBC may be related to metabolism and inflammation [35-37]. Although we did not investigate this relationship, previous epidemiological studies have shown that before menopause, triple-negative cancers were related to obesity and chronic inflammation, and that after menopause, in women aged <65 these latter subtypes were related to metabolic syndrome [22,35,36,37,38].

This study was focused on women of European ancestry. The results are consistent with recent reports of epidemiological studies in Caucasian women [39]. For example, a recent epidemiology study on obesity and TNBC involving socio-economically deprived Caucasian women in the Apalacian in West Virginia revealed the occurrence of TNBC in younger women with later stage of diagnosis [39].

Limitation of the study: Although this exploratory study provides some insights about the relationship between obesity and overweight with TNBC in premenopausal and postmenopausal women, limitations must be acknowledged. This study used samples from women of European ancestry diagnosed with TNBC and therefore cannot be generalized to other ethnic populations. Studies have reported that the incidences and mortality rate of TNBC are significantly higher in African American (AA) women and that the disease tends to have a higher impact in premenopausal AA women [2,40,41]. This study did not include AA women. However, although AA women have the highest incidence of TNBC [2,40,41], published reports of survival outcomes for African-American women with TNBC, relative to European-American women, are conflicting [2]. Recent reports have shown that there are no differences in survival and outcomes between AA and EA women after adjusting for disparities in access to health-care treatment, co-morbid disease and income [42-43]. Thus, although we did not use the AA women in this study, the significance of this exploratory study is that both obesity and overweight are modifiable risk factors, and affect both AA and EA women with TNBC [40,41,44], therefore they provide new opportunities for the development of risk reduction strategies that may decrease mortality by preventing the development of TNBC in both AA and EA women. Lastly but not the least, in this study we used BMI as the surrogate measure of obesity. However, different mechanisms such as adiposity may be more reliable measures of obesity than BMI [45] and are worth exploring.

 doi:10.20944/preprints201803.0174.v1

5. CONCLUSIONS

The results of this exploratory study show that overweight and obesity are associated with TNBC in premenopausal and postmenopausal Caucasian women. The results further demonstrate that obesity and overweight could potentially have divergent impact in premenopausal and postmenopausal women. More research involving larger sample sizes from different race/ethnic populations is needed to confirm these results.

Supplementary Materials: The following are available online at www.mdpi.com/link:

Table SA. List of significantly differentially expressed genes showing evidence of association with TNBC at P<0.05 among the three patient groups in women with premenopausal status.

Table SB. List of significantly differentially expressed genes showing evidence of association with TNBC at P<0.05 among the three patient groups in women with postmenopausal status.

Table S1. List of significantly differentially expressed genes showing evidence of association between TNBC and Obesity or overweight in women with premenopausal status. Also shown in this table is a list of significantly differentially expressed genes distinguishing obese from non-obese women.

Table S2. List of significantly differentially expressed genes showing evidence of association between TNBC and Obesity or overweight in women with postmenopausal status. Also shown in this table is a list of significantly differentially expressed genes distinguishing obese from non-obese women.

Acknowledgments

This research was supported by Louisiana State University School of Medicine in New Orleans. National Institutes of Health Grant # G12MD007581 is also acknowledged. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author Contributions

Chindo Hicks, Tarun Mamidi, Jiande Wu, conceived, designed, and drafted the manuscript; and Lucio Miele and Paul B. Tchounwou participated in the implementation of the study, interpretation of data and writing of the manuscript. All authors read and approved the final draft of the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Foulkes WD, Smith IE, Reis-Filho JS. Triple-negative breast cancer. N Engl J Med. 2010;363:1938–1948.
2. 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 Apr;15(4):248-54.
3. Carey LA, Perou CM, Livasy CA et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. JAMA. 2006 Jun 7;295(21):2492-502.
4. Turkman YE, Sakibia Opong A, Harris LN, Knobf MT.Biologic, demographic, and social factors affecting triple negative breast cancer outcomes. Clin J Oncol Nurs. 2015 Feb;19(1):62-7.
5. Bradley CJ, Given CW, Roberts C. Race, socioeconomic status, and breast cancer treatment and survival. J Natl Cancer Inst. 2002;94:490–496.
6. . Bauer KR, Brown M, Cress RD, Parise CA, Caggiano V. Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype: a population-based study from the California cancer registry. Cancer. 2007;109:1721–1728.
7. Bao PP, Zhao GM, Shu XO, Peng P, Cai H, Lu W, Zheng Y. Modifiable Lifestyle Factors and Triple-negative Breast Cancer Survival: A Population-based Prospective Study. Epidemiology. 2015 Nov;26(6):909-16.
8. Pierobon M, Frankenfeld CL. Obesity as a risk factor for triple-negative breast cancers: a systematic review and meta-analysis. Breast Cancer Res Treat. 2013 Jan;137(1):307-14.
9. Turkoz FP, Solak M, Petekkaya I et al. Association between common risk factors and molecular subtypes in breast cancer patients. Breast. 2013 Jun;22(3):344-50.
10. Al Jarroudi O1, Abda N, Seddik, Brahmi SA, Afqir S. Overweight: Is It a Prognostic Factor in Women with Triple-Negative Breast Cancer? Asian Pac J Cancer Prev. 2017 Jun 25;18(6):1519-1523.
11. Hao S, Liu Y, Yu KD, Chen S, Yang WT, Shao ZM. Overweight as a Prognostic Factor for Triple-Negative Breast Cancers in Chinese Women. PLoS One. 2015 Jun 24;10(6):e0129741.
12. Chen HL, Ding A, Wang ML. Impact of central obesity on prognostic outcome of triple negative breast cancer in Chinese women. Springerplus. 2016 May 11;5:594.
13. Sahin S, Erdem GU, Karatas F et al. The association between body mass index and immunohistochemical subtypes in breast cancer. Breast. 2017 Apr;32:227-236.
14. Cakar B, Muslu U, Erdogan AP, Ozisik M, Ozisik H et al. The Role of Body Mass Index in Triple Negative Breast Cancer. Oncol Res Treat. 2015;38(10):518-22.

15. Widschwendter P, Friedl TW, Schwentner L et al. The influence of obesity on survival in early, high-risk breast cancer: results from the randomized SUCCESS A trial. Breast Cancer Res. 2015 Sep 18;17:129.
16. Phipps AI, Malone KE, Porter PL, Daling JR, Li CI. Body size and risk of luminal, HER2-overexpressing, and triple-negative breast cancer in postmenopausal women. Cancer Epidemiol Biomarkers Prev. 2008 Aug;17(8):2078-86.
17. Fontanella C, Lederer B, Gade S et al. Impact of body mass index on neoadjuvant treatment outcome: a pooled analysis of eight prospective neoadjuvant breast cancer trials. Breast Cancer Res Treat. 2015 Feb;150(1):127-39.
18. Karatas F, Erdem GU, Sahin S et al. Obesity is an independent prognostic factor of decreased pathological complete response to neoadjuvant chemotherapy in breast cancer patients. Breast. 2017 Apr;32:237-244.
19. Tait S, Pacheco JM, Gao F, Bumb C, Ellis MJ, Ma CX. Body mass index, diabetes, and triple-negative breast cancer prognosis. Breast Cancer Res Treat. 2014 Jul;146(1):189-97.
20. Ademuyiwa FO, Groman A, O'Connor T, Ambrosone C, Watroba N, Edge SB. Impact of body mass index on clinical outcomes in triple-negative breast cancer Cancer. 2011 Sep 15;117(18):4132-40.
21. Mowad R, Chu QD, Li BD, Burton GV, Ampil FL, Kim RH. Does obesity have an effect on outcomes in triple-negative breast cancer? J Surg Res. 2013 Sep;184(1):253-9.
22. Bonsang-Kitzis H, Chaltier L, Belin L et al. Beyond Axillary Lymph Node Metastasis, BMI and Menopausal Status Are Prognostic Determinants for Triple-Negative Breast Cancer Treated by Neoadjuvant Chemotherapy. PLoS One. 2015 Dec 18;10(12):e0144359.
23. Burstein MD, Tsimelzon A, Poage GM, Covington KR et al. Comprehensive genomic analysis identifies novel subtypes and targets of triple-negative breast cancer.
24. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, Pietenpol JA. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. J Clin Invest. 2011 Jul;121(7):2750-67.
25. Perou CM. Molecular stratification of triple-negative breast cancers. Oncologist. 2011;16 Suppl 1:61-70.
26. Perou CM. Oncologist. 2010;15 Suppl 5:39-48. Molecular stratification of triple-negative breast cancers.
27. Morrissey ER, Diaz-Uriarte R. Pomelo II: finding differentially expressed genes. *Nucleic Acids Res*. 2009;37:W581–6.
28. Reich M, Liefeld T, Gould J, Lerner J, Tamayo P. GenePattern 2.0. *Nat Genet*. 2006;38(5):500–1.

29. Ramacher MD, Mcshane LM, Simon R. A paradigm for class prediction using gene expression profiles. *J Comput Biol*. 2002;9(3):505–11.
30. Benjamini Y, Hochberg Y. Controlling the false discovery rat: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol*. 1995;57(1):289–300.
31. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Nat Acad Sci U S A*. 1998;95: 14863–8.
32. Mi B, Liu G, Zhou W, Lv H, Liu Y, Liu J. Identification of genes and pathways in the synovia of women with osteoarthritis by bioinformatics analysis. Mol Med Rep. 2018 Mar;17(3):4467-4473. doi: 10.3892/mmr.2018.8429. Epub 2018 Jan
33. Ingenuity Pathways Analysis (IPA). *Ingenuity Pathways Analysis (IPA) system*. Redwood, CA: Ingenuity Systems, Inc. Available at http://www.ingenuity.com/
34. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. 2000;25(1):25–9.
35. Maiti B, Kundranda MN, Spiro TP, Daw HA. The association of metabolic syndrome with triple-negative breast cancer. Breast Cancer Res Treat. 2010;121(2):479-83.
36. Davis AA, Kaklamani VG.Metabolic syndrome and triple-negative breast cancer: a new paradigm. Int J Breast Cancer. 2012;2012:809291.
37. Dietze EC, Chavez TA, Seewaldt VL. Obesity and Triple-Negative Breast Cancer: Disparities, Controversies, and Biology. Am J Pathol. 2018 Feb;188(2):280-290.
38. Agresti R, Meneghini E, Baili P, Minicozzi P et al. Association of adiposity, dysmetabolisms, and inflammation with aggressive breast cancer subtypes: a cross-sectional study. Breast Cancer Res Treat. 2016;157(1):179-89.
39. Vona-Davis L, Rose DP, Hazard H, Howard-McNatt M, Adkins F, Partin J, Hobbs G. Triple-negative breast cancer and obesity in a rural Appalachian population. Cancer Epidemiol Biomarkers Prev. 2008 Dec;17(12):3319-24.
40. Stead LA, et al. Triple-negative breast cancers are increased in black women regardless of age or body mass index. Breast Cancer Res. 2009;11:R18.
41. Tayyari F, Gowda GAN, Olopade OF et al. 2018. Metabolic profiles of triple-negative and luminal A breast cancer subtypes in African-American identify key metabolic differences. Oncotarget. 2018 Feb 7;9(14):11677-11690.
42. 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.

 doi:10.20944/preprints201803.0174.v1

655 43. Pacheco JM1, Gao F, Bumb C, Ellis MJ, Ma CX. Racial differences in
656 outcomes of triple-negative breast cancer. Breast Cancer Res Treat.
657 2013;138(1):281-9.
658 44. Chen L, Cook LS, Tang MT et al., Body mass index and risk of luminal, HER2-
659 overexpressing, and triple negative breast cancer. Breast Cancer Res Treat.
660 2016; 157(3):545-54.
661 45. Bandera EV, Chandran U, Hong CC, Troester MA et al. Obesity, body fat
662 distribution, and risk of breast cancer subtypes in African American women
663 participating in the AMBER Consortium. Breast Cancer Res Treat. 2015
664 Apr;150(3):655-66.