RESEARCH

Prevalence of the *UGT1A1*28* promoter polymorphism and breast cancer risk among African American women in Memphis, TN

Alana Smith[1], Cheryl D. Cropp[2,3], Gregory Vidal[4,5], Elizabeth Pritchard[5], Jennifer Cordero[1], Claire Simpson[1], and Athena Starlard-Davenport[1*]

[*1]Department of Genetics, Genomics and Informatics, University of Tennessee Health Science Center, Memphis, TN 38163, USA, [2]Integrated Cancer Genomics Division, Translational Genomics Research Institute (TGen), Phoenix, AZ, 85004, USA, [3]Department of Pharmaceutical, Social and Administrative Sciences, McWhorter School of Pharmacy, Samford University, Birmingham, AL, 35229, USA, [4]Division of Hematology/Oncology, Department of Medicine, University of Tennessee Health Science Center, Memphis, TN 38163, [5]The University of Tennessee West Cancer Center; Memphis, TN 38163

***Corresponding author's e-mail**: astarlar@uthsc.edu

ABSTRACT

Inherited variations in UDP-glucuronosyltransferase 1A1 (UGT1A1) are associated with an increased breast cancer risk in women of African ancestry. The *UGT1A1*28* promoter polymorphism is characterized by the presence of 7 TA repeats in the TATA box sequence and results in reduced UGT1A1 gene expression and enzymatic activity. In this study, we investigated associations between the *UGT1A1*28* polymorphism and breast cancer risk among African American (AA) women in Memphis, Tennessee, a city with increased breast cancer mortality rates among AA women. Saliva was collected from 352 AA women, including breast cancer cases (n=82) and controls (n=270) between June 2016 to June 2017. DNA was isolated and sequenced for the *UGT1A1*28* polymorphism. The odds ratio for cases with the low *UGT1A1* activity alleles (TA)7/8 repeat genotypes versus 5/5, 5/6, and 6/6 genotypes was 1.46 [95% CI, 0.65-3.31; P = 0.36] in premenopausal women and 1.10 (95% CI, 0.52-2.38; P = 0.79) in postmenopausal women. Further analysis of TCGA RNA-seq data showed that UGT1A1 mRNA was significantly lower among estrogen receptor (ER)-negative breast cancers from AA as compared to non-Hispanic white women with ER-negative breast cancer. Larger epidemiological studies are needed to determine the functional consequence of the *UGT1A1*28* polymorphism on breast cancer risk in AA women.

KEYWORDS: UGT1A1, UGT1A1*28, breast cancer disparities, African American, Memphis, TN

Citation: Smith A et al (2019) Prevalence of the UGT1A1*28 promoter polymorphism and breast cancer risk among African American women in Memphis, TN. Cancer Health Disparities 3:e1-e12. doi:10.9777/chd.2019.1015.

INTRODUCTION

Breast cancer is a leading cause of cancer death among African American (AA) women (DeSantis et al., 2017). African American (AA) women under the age of 45 are most likely to develop breast cancer, and AA women of all ages are most likely to die from breast cancer (DeSantis et al., 2017). Recent data showed that AA women in Memphis, Tennessee suffer disproportionately from higher breast cancer mortality rates compared to non-Hispanic white women (Hunt et al., 2014; Vidal et al., 2017; Whitman et al., 2012). Epidemiological evidence highlights the influence of estrogen exposures with breast cancer development in women of African ancestry (Ambrosone et al., 2015; Corsini et al., 2017). Established breast cancer risk factors, e.g. increased body mass index (BMI), obesity, and postmenopausal status (Dietze et al., 2018; Gerard and Brown, 2018), are thought to increase breast cancer risk in women through modulation of estrogens. In support of these findings, AA women tend to have higher BMI and obesity rates compared to non-Hispanic white women (Tan et al., 2017). These factors, along with genetic variation in key genes involved in estrogen conjugation and metabolism, may contribute to increased breast cancer risk among AA women.

A major pathway involved in estrogen conjugation and detoxification from circulation is through the process of glucuronidation, catalyzed by UDP-Glucuronosyltransferases (UGTs) (Bock et al., 1987; Starlard-Davenport et al., 2007). Human UGT1A1 is a major UGT enzyme involved in estrogen conjugation (Lepine et al., 2004). Variation in the number of thymine-adenine (TA) repeats in the TATA-box of the *UGT1A1* promoter (rs8175347) significantly alters UGT1A1 expression and catalytic activity towards estrogens and other compounds (Boyd et al., 2006; Iyer et al., 1999). The number of TA repeats vary from 5 to 8, with 6 TA repeats representing the most common (wild-type) number of repeats (Beutler et al., 1998). Individuals with more than 6 TA repeats (i.e., 7 TA repeats; *UGT1A1*28*) have markedly reduced enzymatic activity, resulting in increased estrogen and bilirubin levels, and increased risk for diseases, including breast cancer development (Adegoke et al., 2004; Guillemette et al., 2001; Huo et al., 2008). Studies have shown that the frequency of the low activity *UGT1A1*28* promoter polymorphism varies significantly between ethnic groups, with prevalence being highest among populations of African descent (Beutler et al., 1998; Guillemette et al., 2000; Huo et al., 2008).

Considering recent reports of higher breast cancer mortality rates among AA women in Memphis, TN as compared to 49 of the largest US cities (Hunt et al., 2014; Vidal et al., 2017), the present study was conducted to determine the possible association between the *UGT1A1* TA repeat polymorphism and breast cancer risk in a sample of AA women in Memphis, TN. We further explored the relationship between UGT1A1 expression and breast cancer risk in AA and non-Hispanic white women using RNA sequencing data from The Cancer Genome Atlas (TCGA).

MATERIALS AND METHODS

Study Population

A total of 352 AA women participants, comprising 82 breast cancer cases and 270 healthy women controls age 18 and older were included in this pilot study. Breast cancer patients treated at the West Cancer Center (WCC) in Memphis, TN were recruited. Normal healthy controls and cancer survivors were also recruited at local community outreach events in Memphis, TN between June

2016 and June 2017. All recruited breast cancer patients in this study were either being treated for breast cancer or had a prior diagnosis of breast cancer. The control group included healthy volunteers without a prior diagnosis of breast cancer and who were recruited at community outreach events. All recruited participants completed a 3-page health questionnaire that included information on participants' reproductive history, diet and lifestyle factors, and family history of cancer. Race was classified according to self-report. A 2 mL saliva sample was also collected using Oragene® OG-500 DNA Self-Collection Kit (Ottawa, Ontario, Canada) at the same time the participants completed the health questionnaire.

Ethics

Written informed consent was obtained from each study participant and the study protocol (protocol # 16-04551-XP and 16-04502-XP) was approved by the Institutional Review Board of the University of Tennessee Health Science Center, Memphis, TN and was carried out in accordance with the guidelines of the Declaration of Helsinki.

UGT1A1 Promoter Genotyping Analysis. DNA was isolated from saliva using DNA Genotek's Prep-It L2P DNA isolation kit (Ottawa, Ontario, Canada). DNA was purified using a ZymoResearch DNA clean and concentrator kit (Irvine, California). The *UGT1A1*28* polymorphism was identified from DNA that was PCR amplified as previously described (Massacesi et al., 2006). Briefly, polymerase chain reactions (PCR) were conducted in a total volume of 50 µL containing 100 ng of genomic DNA, 25 pmoles of the *UGT1A1* primer (*UGT1A1* forward: 5'-GAT TTG AGT ATG AAA TTC CAG CCA G-3' and *UGT1A1* reverse: 5'-CCA GTG GCT GCC ATC CAC T-3') (Massacesi et al., 2006) and the following reagents (Promega, Madison, WI): 0.1 mM each of dCTP, dGTP, dATP and dTTP; 2.5 mM of MgCl2; 50 mM of potassium chloride; 10 mM of Tris (pH 9.0); 0.1 % Triton-X; and 2.5 U of Taq polymerase. After 35 cycles of amplification (denaturation at 94 °C for 30 sec, annealing at 60 °C for 1 min, and extension at 72 °C for 1 min), the amplification products were electrophoresed in 3% agarose gel and visualized after staining with ethidium bromide.

Genotypes of the A(TA)6-8TAA polymorphism were determined by bidirectional sequencing with forward or reverse primers. PCR products were purified using the Qiagen DNA Purification Kit (Valencia, CA). DNA sequencing of the *UGT1A1* product (351 bp) containing the polymorphic TA repeats was performed using the Sanger method employing Applied Biosystems Big Dye v3.1 Reaction Mix at 1/10X strength combined with an appropriate amount of PCR product and 5 pmol of the relevant primer. All sequencing data was generated in the Molecular Resource Center (MRC) of Excellence at The University of Tennessee Health Science Center, Memphis, TN. All PCR protocols and template/primer ratios were those suggested by the manufacturer. The ThermoFisher (Applied Biosystems) BDXterminator system was used to remove salts, primers and unincorporated nucleotides from the labeling reaction. Labelled samples were analyzed with an Applied Biosystems 3130xL Genetic Analyzer using Sequencing Analysis v5.2.0 software employing the KB.bcp program using the KB_3130_POP7_BDTv3 mobility file. All electropherograms were visually examined in the MRC to ensure the highest data quality.

Sample quality was determined prior to dNTP incorporation by assaying the A260, A280 and A230 values using a Nanodrop Spectrophotometer: the preferred A260/A280 and

A260/A230 ratios were 1.7 or better. *UGT1A1* genotypes were assigned based on the number of TA repeats for each allele (i.e., 6/6, 6/7, 7/7, or 7/8). The percentage of samples with successful calls was ≥98%.

RNA sequencing analysis. RNA sequencing analysis was done with Cufflinks (Trapnell et al., 2013). This algorithm reports changes at the transcript and gene level as fragment reads per kilobase of exon per million reads mapped (FPKM) values. Breast carcinoma RNA-Seq data, made available by Nationwide's Children, was queried through the National Cancer Institute GCD Data Portal using the following criteria: female, 18-90 years old at age of diagnosis, alive or dead, African American or European American, and not Hispanic. The raw FPKM values from this RNA-Seq data were used for the analysis. No differential gene expression analysis, Cuffdiff, was performed and thus no log2 (Fold Change) values were obtained. The raw FPKM values from Cufflinks were converted to log scale for graphing purposes.

Statistical Analysis

Demographic characteristics and selected risk factors for breast cancer were compared between cases and controls using T-tests for continuous variables and Chi-square tests for categorical data. Data were expressed as the mean ± standard deviation. ORs and 95% CIs for breast cancer were calculated from unconditional logistic regression models and used to estimate relative risk. To estimate the allele frequencies, data were analyzed using $\chi 2$ analysis and two-tailed Fisher's exact test. The observed genotype frequencies of *UGT1A1* were compared with the expected frequencies, according to the Hardy-Weinberg equation. A p-value of <0.05 was considered statistically significant. All statistical analyses were performed using GraphPad Prism, version 7 software (La Jolla, CA).

RESULTS

Participant characteristics

This study was conducted on a total of 352 AA women participants, consisting of 82 breast cancer cases and 270 healthy women volunteers (Table 1). The mean ± SD age of the breast cancer cases and the comparison group was 54.5 ± 13.8 and 48.2 ± 13.8 years, respectively. The minimum and maximum age among breast cancer cases was 20 and 88 years. The minimum and maximum age among healthy volunteers was 18 and 79 years. As can be seen from Table 1, the group of breast cancer patients was matched to the control group on age ($p > 0.05$). There was no significant difference between cases and controls for body mass index (BMI), smoking status, alcohol consumption, family history of cancer or menopausal status. The majority of participants, both cases and controls, in the cohort had a BMI greater than 30, did not smoke, had a family history of breast cancer, or were postmenopausal. Specifically, more than 90% of cases and 89% of controls reported that they do not smoke. However, more than 73% of breast cancer cases reported that they do not consume alcoholic beverages while greater than 52% of healthy volunteers reported that they do consume alcohol. Overall, no differences were found between the groups regarding the characteristics shown in Table 1.

Table 1. Characteristics of study participants.

Characteristic	Cases (n = 82) [a]	Controls (n = 270)	p-value
Age in years, mean ± SD	54.5 ± 13.8	48.2 ± 13.8	0.170
Minimum	20.0	18.0	
25% Percentile	44.0	38.0	
Median	52.5	50.0	
75% Percentile	64.3	58.0	
Maximum	88.0	79.0	
Body Mass Index (BMI), n (%)			0.920
Optimal, BMI 18.5 – 25	17 (20.7)	47 (17.4)	
Overweight, BMI 25-30	21 (25.6)	75 (27.8)	
Obese, BMI > 30	43 (52.4)	146 (54.1)	
Missing	1	2	
Smoking status, n (%)			0.886
Yes	3 (3.66)	23 (8.52)	
No	79 (96.3)	241 (89.3)	
Missing	0	6	
Alcohol consumption, n (%)			0.965
Yes	22 (26.8)	141 (52.2)	
No	60 (73.2)	121 (44.8)	
Missing	0	8	
Family history of cancer, n (%)			0.958
Yes	52 (63.4)	141 (52.2)	
No	29 (35.4)	122 (45.2)	
Missing	1	5	
Menopause Status, n (%)			>0.999
Pre-menopausal	36 (43.9)	132 (48.9)	
Post-menopausal	46 (56.1)	138 (51.1)	

Distribution of *UGT1A1* allele and genotype frequencies in our study population

Results of genotyping analysis for the *UGT1A1* promoter TATA box variants among breast cancer cases and controls in our study population are presented in Table 2. The number of (TA) repeats in the *UGT1A1* promoter TATA box included 5TA, 6TA, 7TA and 8TA. The normal or wild-type allele is characterized by 6 TA repeats, while 5TA, 7TA, and 8TA repeats are variants. The polymorphism with the highest allele frequency among breast cancer cases was 7TA (*UGT1A1*28*) at 43%. However, among controls the normal or wild-type allele, consisting of 6TA, was most common at 47%. The allele frequency of the *UGT1A1*36* (5TA) and *UGT1A1*37* (8TA) was least common among both breast cancer cases and controls.

Table 2 also shows the *UGT1A1* (TA) genotype frequencies. The 7/7 homozygous genotype of the UGT1A1 gene (*UGT1A1*28*) was identified in 17.8 % of breast cancer cases and 17.1% of controls. The 6/7 heterozygous genotype was the most prevalent in both breast cancer cases (35.6%) and controls (34.1%). There was no departure from Hardy-Weinberg equilibrium and no significant

association in allele and genotype frequencies between cases and controls (χ2 test, p>0.05). The frequency distribution of the *UGT1A1* promoter variants in our population and the DNA sequence lengths of the promoter variants are shown in supplementary Figure 1.

Table 2. *UGT1A1* promoter allele and genotype frequencies of study population.

Allele [a]	Cases [b]	Controls [b]	p-value
5	0.07	0.10	0.33
6	0.42	0.47	
7	0.43	0.40	
8	0.08	0.03	
Genotype [c]	Cases, n (%) [d]	Controls, n (%) [d]	
5/5	4 (5.48)	7 (2.78)	0.63
5/6	2 (2.74)	17 (6.75)	
6/6	17 (23.3)	68 (27.0)	
6/7	26 (35.6)	86 (34.1)	
7/7	13 (17.8)	43 (17.1)	
7/8	7 (15.1)	31 (12.3)	
8/8	0 (0)	0 (0)	

Adapted from Guillemette et al, 2000 (Guillemette et al., 2001)

[a] χ2 test comparing distribution in cases versus controls for trend: p = 0.33

[b] Number of alleles/number of chromosomes (unweighted)

[c] χ2 test comparing distribution in cases versus controls for trend: p = 0.63

[d] Number of participants with genotype/total participants (unweighted)

Association between *UGT1A1* genotypes and obesity and menopausal status

We further determined the effect of *UGT1A1* genotypes on menopausal status and obesity status using odds ratios (ORs) (Table 3). Participants were divided into two categories based on *UGT1A1* promoter activity. Participants with high *UGT1A1* (TA) activity alleles carried 5/5, 5/6, and 6/6 TA repeats and those considered to have low activity alleles had the presence of 6/7, 7/7 or 7/8 TA repeats. All comparisons were achieved with the reference group including 5 or 6 allele-containing genotypes (5/5, 5/6, and 6/6). On stratification by menopausal status, association between *UGT1A1* high-risk genotypes and risk of breast cancer showed an OR of 1.46 in premenopausal women (95% CI, 0.65-3.31; P = 0.361) and an OR of 1.10 in postmenopausal women (95% CI, 0.52-2.38; P = 0.793). Although not significant, overweight and obese status among women with the low activity 7/8 TA repeat alleles had a 92% increased risk of breast cancer compared to wild-type (5TA) and high activity (6TA) alleles.

Table 3. *UGT1A1* genotypes and risk of breast cancer by menopause and obesity status.

Genotypes	Cases, n (%)	Controls, n (%)	OR (95% CI)	p-value
Premenopausal women				
5/5, 5/6, 6/6	10 (27.8)	45 (36.0)	1.46 (0.65-3.31)	0.361
6/7, 7/7, 7/8	26 (72.2)	80 (64.0)		
Postmenopausal women				
5/5, 5/6, 6/6	13 (35.1)	48 (37.5)	1.10 (0.52-2.38)	0.793
6/7, 7/7, 7/8	24 (64.9)	80 (62.5)		
Optimal weight				
5/5, 5/6, 6/6	6 (40.0)	14 (34.1)	0.78 (0.23-2.63)	0.686
6/7, 7/7, 7/8		27 (65.9)		
Overweight/Obese				
5/5, 5/6, 6/6	12 (24.0)	75 (37.7)	1.92 (0.94-3.89)	0.073
6/7, 7/7, 7/8	38 (76.0)	124 (62.3)		

Investigation of UGT1A1 mRNA expression in breast tumor tissues from AA and non-Hispanic white women in the TCGA

We previously showed that UGT1A1 mRNA levels was significantly lower in breast tumors as compared to normal breast tissues from AA and non-Hispanic white women (Starlard-Davenport et al., 2012). We demonstrated that low UGT1A1 mRNA expression in breast tumors correlated with *UGT1A1*28* low activity alleles. Guillemette et al also showed that the *UGT1A1*28* was associated with Estrogen Receptor (ER)-breast cancers in premenopausal AA women (Guillemette et al., 2000). Therefore, we explored The Cancer Genome Atlas (TCGA), a comprehensive database that catalogues genomic alterations responsible for cancer from 500 patients, to determine whether UGT1A1 RNA expression was altered in a group of human breast cancer samples from AA (n = 94) and non-Hispanic white (n = 244) women by estrogen receptor (ER) status. TCGA RNA-sequencing data and the log2 (average raw FPKM) value was plotted per UGT1A1 gene for AA tumor and non-Hispanic white tumor samples by ER status (Figure 1). We observed that UGT1A1 mRNA expression was lower in breast tumors from AA as compared to non-Hispanic white women although not significant (Figure 1). We further showed that UGT1A1 gene expression was significantly higher in ER negative breast tumors from non-Hispanic white women as compared to AA women with ER negative and ER positive breast tumors.

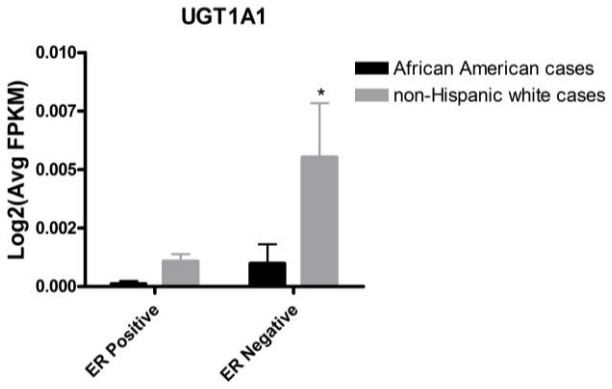

Figure 1. UGT1A1 RNA expression from TCGA database. The log2(average raw FPKM) value is plotted per UGT1A1 gene expression for AA tumor and non-Hispanic white tumor samples by Estrogen Receptor (ER) status. The average raw FPKM values for UGT1A1 were zero or close to zero.

DISCUSSION

In the present study, we investigated the associations between the *UGT1A1*28* TA repeat polymorphism located in the promoter region of UGT1A1 among AA breast cancer cases and healthy AA women volunteers. Our study is the first to investigate the association between the *UGT1A1*28* polymorphism and breast cancer risk among women in Memphis, TN, a city with one of the highest breast cancer mortality disparity rates among AA women compared to non-Hispanic white women in the US.

The human UGT1A1 gene is a member of the UGT1A subfamily of enzymes that are encoded by the UGT1A gene locus on chromosome 2q37.1. The UGT1A isoforms share 4 common exons from exon 2 to exon 5 and have a unique exon 1 and individual promoter pairs (Tukey and Strassburg, 2000). UGT1A1 is the most abundant member of the UGT1A family in human liver and is also the major isoform responsible for the metabolism and clearance of many endogenous and exogenous compounds, including β-estradiol and its metabolites (Lepine et al., 2004). It is well known that prolonged exposure to β-estradiol is a major risk factor for breast cancer development in women. Elevated levels of circulating estrogens are associated with increased breast cancer risk in postmenopausal women. The carcinogenic potential of estrogens is not only mediated by estrogen receptor signaling but also mediated by increased cellular oxidative metabolism of estrogens to genotoxic estrogen metabolites (Cavalieri et al., 1997).

Several studies have investigated the impact of inherited genetic variations in *UGT1A1* on the incidence risk of various cancers including breast cancer (Adegoke et al., 2004; Guillemette et al., 2000; Huo et al., 2008; Shatalova et al., 2006). The scientific premise for the association of *UGT1A1* genetic variants and breast cancer risk stems from the hypothesis that reduced UGT1A1 estrogen conjugating activity results in decreased cellular oxidative estrogen metabolism and increased estrogen bioavailability in circulation. The most studied *UGT1A1* polymorphism is *UGT1A1*28* (rs8175347) polymorphism is characterized by an extra TA repeat (TA-7) in the TATA box region in the *UGT1A1* promoter resulting in decreased gene transcription and glucuronidation activity (Beutler et al., 1998). *UGT1A1*28* has been previously linked to increased risk of breast cancer. For instance, a study by Guillemette et al showed for the first time that 200 premenopausal AA women with ER negative breast cancer have a higher prevalence of low *UGT1A1* activity allele-containing genotypes than 200 female controls of African ancestry (Guillemette et al., 2000). Although a limitation of our genetic study was the small population sample size and lack of information regarding hormone receptor breast tumor status among breast cancer patients, we still observed the low activity *UGT1A1*28* polymorphism among premenopausal AA women with breast cancer but not in postmenopausal AA breast cancer patients, which is in agreement with findings previously reported (Guillemette et al., 2000). Additionally, in our study, the gene frequency of all *UGT1A1* alleles in African American women was similar to those reported previously by Guillemette et al (Guillemette et al., 2000). A similar study was conducted on women in the Shanghai Breast Cancer Study, a population-based case-control study that found similar results (Adegoke et al., 2004). In that study, Adegoke et al also found that women under the age of 40 who carried the *UGT1A1*28* risk allele had an increased risk of breast cancer with an OR = 1.7; 95% CI = 1.0-2.7) but not among women 40 years old and

over (OR = 0.8; 0.7-1.1) (Adegoke et al., 2004). Likewise, similar findings have been found in Russian women (Shatalova et al., 2006). By contrast, Huo et al genotyped DNA for the *UGT1A1*28* polymorphism among 512 Nigerian breast cancer cases and 226 community controls and found that premenopausal indigenous Nigerian African women with the low-activity *UGT1A1* TA repeat alleles were protected against breast cancer (Huo et al., 2008). These findings show that the frequency of the *UGT1A1*28* polymorphism varies between populations by race and ethnicity.

In our previous published study, we showed that the *UGT1A1*28* promoter polymorphism was associated with decreased UGT1A1 gene expression in AA and non-Hispanic white women with breast cancer (Starlard-Davenport et al., 2012). In that study, we observed a significant decrease in UGT1A1 mRNA expression levels among breast tissues isolated from AA and non-Hispanic white women with breast cancer (N = 17) and normal healthy controls (N = 30) ($p = 0.02$). Based on these findings and those of Guillemette et al (Guillemette et al., 2000) who found an association between the *UGT1A1*28* polymorphism and ER negative breast cancer risk in AA women, we interrogated existing TCGA RNA sequencing datasets to determine whether UGT1A1 gene expression was differentially expressed among AA and non-Hispanic white breast cancer patients by ER status. In our analysis, we were unable to compare UGT1A1 mRNA expression between AA breast cancer cases and AA healthy controls since there was only one healthy control available for comparison. Thus, it was impossible to conduct statistical analysis of UGT1A1 mRNA between normal and breast cancer AA patients in the TCGA. However, further analysis of UGT1A1 mRNA expression in the TCGA database between AA breast cancers and non-Hispanic white breast cancer cases showed that UGT1A1 expression was downregulated in breast cancers and was lower among AA breast cancers compared to non-Hispanic white breast cancer cases, especially among ER negative breast tumors from AA women.

In conclusion, this genetic study supports previous findings highlighting the association between the *UGT1A1*28* polymorphism and premenopausal breast cancer risk among AA women. This study also highlights the importance of investigating genetic studies in an underserved minority AA population such as Memphis, TN where the breast cancer mortality rates is significantly higher among AA women compared to non-Hispanic white women. Larger epidemiological studies are required to replicate these findings in other populations of African ancestry and to delineate the exact biological mechanisms underlying the effects of the *UGT1A1*28* promoter polymorphism in AA women.

Acknowledgements

The authors wish to thank all participants in this study and volunteers who assisted in consenting and collecting biospecimens. Special thanks to Mrs. Gwendolyn Brown with the Carin and Sharin Breast Cancer Support Group and Dr. Osborne Burks at Southwest Tennessee Community College. The authors wish to acknowledge support from the University of Tennessee Health Science Center (UTHSC) Molecular Resource Center and the West Cancer Center.

Funding Sources

This work was supported by generous funds from the UTHSC Center for Integrative Genetics (CITG) and Methodist Mission to AS-D).

Conflict of interest

The authors declare that no competing or conflict of interests exists. The funders had no role in study design, writing of the manuscript, or decision to publish.

Authors' contributions

Conceptualization: AS-D, Formal analysis: AS-D, CDC, CS. Methodology: AS, JC, GV, EP. Supervision: AS-D. Writing ± original draft: AS-D, AS, CDC, GV. Writing ± review & editing: AS-D, AS, CDC, GV, CS

REFERENCES

Adegoke, O.J., Shu, X.O., Gao, Y.T., Cai, Q., Breyer, J., Smith, J., and Zheng, W. (2004). Genetic polymorphisms in uridine diphospho-glucuronosyltransferase 1A1 (UGT1A1) and risk of breast cancer. Breast cancer research and treatment *85*, 239-245.

Ambrosone, C.B., Zirpoli, G., Hong, C.C., Yao, S., Troester, M.A., Bandera, E.V., Schedin, P., Bethea, T.N., Borges, V., Park, S.Y., *et al.* (2015). Important Role of Menarche in Development of Estrogen Receptor-Negative Breast Cancer in African American Women. Journal of the National Cancer Institute *107*.

Beutler, E., Gelbart, T., and Demina, A. (1998). Racial variability in the UDP-glucuronosyltransferase 1 (UGT1A1) promoter: a balanced polymorphism for regulation of bilirubin metabolism? Proceedings of the National Academy of Sciences of the United States of America *95*, 8170-8174.

Bock, K.W., Lilienblum, W., Fischer, G., Schirmer, G., and Bock-Henning, B.S. (1987). The role of conjugation reactions in detoxication. Archives of toxicology *60*, 22-29.

Boyd, M.A., Srasuebkul, P., Ruxrungtham, K., Mackenzie, P.I., Uchaipichat, V., Stek, M., Jr., Lange, J.M., Phanuphak, P., Cooper, D.A., Udomuksorn, W., *et al.* (2006). Relationship between hyperbilirubinaemia and UDP-glucuronosyltransferase 1A1 (UGT1A1) polymorphism in adult HIV-infected Thai patients treated with indinavir. Pharmacogenetics and genomics *16*, 321-329.

Cavalieri, E.L., Stack, D.E., Devanesan, P.D., Todorovic, R., Dwivedy, I., Higginbotham, S., Johansson, S.L., Patil, K.D., Gross, M.L., Gooden, J.K., *et al.* (1997). Molecular origin of cancer: catechol estrogen-3,4-quinones as endogenous tumor initiators. Proceedings of the National Academy of Sciences of the United States of America *94*, 10937-10942.

Corsini, C., Henouda, S., Nejima, D.B., Bertet, H., Toledano, A., Boussen, H., Habib, F., Mouhout, A., Gaballah, A., Ghazaly, H.E., *et al.* (2017). Early onset breast cancer: differences in risk factors, tumor phenotype, and genotype between North African and South European women. Breast cancer research and treatment *166*, 631-639.

DeSantis, C.E., Ma, J., Goding Sauer, A., Newman, L.A., and Jemal, A. (2017). Breast cancer statistics, 2017, racial disparity in mortality by state. CA: a cancer journal for clinicians *67*, 439-448.

Dietze, E.C., Chavez, T.A., and Seewaldt, V.L. (2018). Obesity and Triple-Negative Breast Cancer: Disparities, Controversies, and Biology. The American journal of pathology *188*, 280-290.

Gerard, C., and Brown, K.A. (2018). Obesity and breast cancer - Role of estrogens and the molecular underpinnings of aromatase regulation in breast adipose tissue. Molecular and cellular endocrinology *466*, 15-30.

Guillemette, C., De Vivo, I., Hankinson, S.E., Haiman, C.A., Spiegelman, D., Housman, D.E., and Hunter, D.J. (2001). Association of genetic polymorphisms in UGT1A1 with breast cancer and plasma hormone levels. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology *10*, 711-714.

Guillemette, C., Millikan, R.C., Newman, B., and Housman, D.E. (2000). Genetic polymorphisms in uridine diphospho-glucuronosyltransferase 1A1 and association with breast cancer among African Americans. Cancer research *60*, 950-956.

Hunt, B.R., Whitman, S., and Hurlbert, M.S. (2014). Increasing Black:White disparities in breast cancer mortality in the 50 largest cities in the United States. Cancer epidemiology *38*, 118-123.

Huo, D., Kim, H.J., Adebamowo, C.A., Ogundiran, T.O., Akang, E.E., Campbell, O., Adenipekun, A., Niu, Q., Sveen, L., Fackenthal, J.D., *et al.* (2008). Genetic polymorphisms in uridine diphospho-glucuronosyltransferase 1A1 and breast cancer risk in Africans. Breast cancer research and treatment *110*, 367-376.

Iyer, L., Hall, D., Das, S., Mortell, M.A., Ramirez, J., Kim, S., Di Rienzo, A., and Ratain, M.J. (1999). Phenotype-genotype correlation of in vitro SN-38 (active metabolite of irinotecan) and bilirubin glucuronidation in human liver tissue with UGT1A1 promoter polymorphism. Clinical pharmacology and therapeutics *65*, 576-582.

Lepine, J., Bernard, O., Plante, M., Tetu, B., Pelletier, G., Labrie, F., Belanger, A., and Guillemette, C. (2004). Specificity and regioselectivity of the conjugation of estradiol, estrone, and their catecholestrogen and

methoxyestrogen metabolites by human uridine diphospho-glucuronosyltransferases expressed in endometrium. The Journal of clinical endocrinology and metabolism *89*, 5222-5232.

Massacesi, C., Terrazzino, S., Marcucci, F., Rocchi, M.B., Lippe, P., Bisonni, R., Lombardo, M., Pilone, A., Mattioli, R., and Leon, A. (2006). Uridine diphosphate glucuronosyl transferase 1A1 promoter polymorphism predicts the risk of gastrointestinal toxicity and fatigue induced by irinotecan-based chemotherapy. Cancer *106*, 1007-1016.

Shatalova, E.G., Loginov, V.I., Braga, E.A., Kazubskaia, T.P., Sudomoina, M.A., Blanchard, R.L., and Favorova, O.O. (2006). [Association of polymorphisms in SULT1A1 and UGT1A1 Genes with breast cancer risk and phenotypes in Russian women]. Molekuliarnaia biologiia *40*, 263-270.

Starlard-Davenport, A., Word, B., and Lyn-Cook, B. (2012). Characterization of UDP-glucuronosyltransferase (UGT1A1) Promoter Polymorphisms and Gene Expression on Ethnicity, Stage of Disease, and Menopausal Status in Breast Cancer. Drug Metabolism and Toxicology *S4*, 1-5.

Starlard-Davenport, A., Xiong, Y., Bratton, S., Gallus-Zawada, A., Finel, M., and Radominska-Pandya, A. (2007). Phenylalanine(90) and phenylalanine(93) are crucial amino acids within the estrogen binding site of the human UDP-glucuronosyltransferase 1A10. Steroids *72*, 85-94.

Tan, M., Mamun, A., Kitzman, H., Mandapati, S.R., and Dodgen, L. (2017). Neighborhood Disadvantage and Allostatic Load in African American Women at Risk for Obesity-Related Diseases. Preventing chronic disease *14*, E119.

Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. Nature biotechnology *31*, 46-53.

Tukey, R.H., and Strassburg, C.P. (2000). Human UDP-glucuronosyltransferases: metabolism, expression, and disease. Annual review of pharmacology and toxicology *40*, 581-616.

Vidal, G., Bursac, Z., Miranda-Carboni, G., White-Means, S., and Starlard-Davenport, A. (2017). Racial disparities in survival outcomes by breast tumor subtype among African American women in Memphis, Tennessee. Cancer medicine *6*, 1776-1786.

Whitman, S., Orsi, J., and Hurlbert, M. (2012). The racial disparity in breast cancer mortality in the 25 largest cities in the United States. Cancer epidemiology *36*, e147-151.

Supplemental Figure 1