Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Madeline Davis 04/28/2025

Finding Ancestry-Associated Genes Among Black Women Diagnosed with High-Grade Serous Ovarian Cancer (HGSOC)

By Madeline Davis
MPH

Epidemiology

Joellen Schildkraut, PhD

Finding Ancestry-Associated Genes Among Black Women Diagnosed with High-Grade Serous Ovarian Cancer (HGSOC)

By

Madeline Davis
Emory University
2025

Thesis Committee Chair: Joellen Schildkraut, PhD

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health in Epidemiology
2025

Abstract

Finding Ancestry-Associated Genes Among Black Women Diagnosed with High-Grade Serous Ovarian Cancer (HGSOC)
By Madeline Davis

Although the incidence of epithelial ovarian cancer (EOC) is lower among Black women compared to White women, Black women diagnosed with EOC experience poorer prognoses and overall survival (OS). While social determinants and unequal access to care partly explain these disparities, biological factors may also contribute. This study aimed to identify ancestry-associated differential gene expression in ovarian tumors of Black women to explore potential differences in the tumor microenvironment influencing outcomes. Our study sample included 211 Black women with the high-grade serous carcinoma (HGSC) subtype of EOC enrolled in the African American Cancer Epidemiology Study (AACES) and the North Carolina Ovarian Cancer Study (NCOCS). Risk factor and tumor characteristic data were collected through surveys and pathology reports. RNA sequencing (RNA-Seq) assessed expression of 18,699 genes, and RFMIX estimated global and subcontinental ancestry proportions, including African, European, Asian, and West African ancestries. We found that the downregulation of the gene *SNX27* was significantly associated with higher African ancestry proportions ($p = 2.90 \times 10^{-6}$), below the Bonferroni-adjusted threshold ($p < 5.35 \times 10^{-6}$). Significant differential expression by BMI (≥40 kg/m² vs. <30 kg/m²) was observed for *FOXD4L5*, *TMEM106B*, *GDF10*, and *G3BP1P1*, with *FOXD4L5* also differing by type 2 diabetes status.
Several differentially expressed genes linked to African ancestry were associated with immune evasion, enhanced tumor cell migration, and proliferation, suggesting that ancestry-related biological factors may contribute to worse EOC outcomes among Black women.

Finding Ancestry-Associated Genes Among Black Women Diagnosed with High-Grade Serous Ovarian Cancer (HGSOC)

By

Madeline Davis

Emory University
2025

Thesis Committee Chair: Joellen Schildkraut, PhD

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health in Epidemiology
2025

Background

Ovarian Cancer Characteristics

Ovarian cancer is a rare yet lethal disease with a 50.9% overall five-year relative survival, establishing its role as the most deadly gynecological cancer.[1] Patients who present with early-stage fare better with 90-95% five-year relative survival following surgery, while patients who present with advanced-stage ovarian cancer have a 10-30% survival rate.[2] The high fatality of ovarian disease is presumably due to nonspecific early symptoms of the disease and lack of effective screening methods, often leading to a delay in diagnosis until the disease has reached an advanced stage (Stage III or IV).[3] In a large prospective cohort study, following a large sample of UK women for cancer incidence and death, it was found that ovarian cancer stage was a major prognostic indicator of survival (stage IV vs I, RR=10.54, 95% CI: 9.16–12.13).[4] Nearly 75% of incident cases present as advanced stages (stage III and IV).[5] Despite high response rates to primary treatments, the recurrence rate of the disease remains high. Among patients with FIGO III or IV stage ovarian cancer, greater than 70% will experience recurrence within 5 years, necessitating further treatment of the advanced disease.[6]

A multifaceted approach, contingent on the stage and characteristics of the disease, is implemented as the standard of care for ovarian cancer. First-line therapy commonly consists of tumor debulking surgery followed by adjuvant chemotherapy if advanced stages have been reached.[7] Occasionally, targeted drugs are used alone, in conjugation with chemotherapy, or as maintenance therapy. Among these treatments, the most robust prognostic indicator of survival is the size of residual disease following tumor debulking surgery; residual disease < 1 cm in diameter is considered an optimistic

status.[8,9] Further, pathological examination of the debulked tissues identifies critical clinical and molecular characteristics of the malignancy: tumor stage, ovarian cancer subtype, histology, and tumor grade are frequently included on pathology reports, aiding in determining prognosis and patient therapeutic course.

The etiology of ovarian cancer remains poorly understood due to its heterogeneous nature, composed of several distinct subtypes and complex molecular signatures. Ovarian cancer, commonly referred to as a single entity, is more appropriately classified as a collection of multiple distinguishable diseases marked by uncontrolled cell proliferation. Epithelial ovarian cancer (EOC), the malignancy of the outermost tissue of the ovary, is the most prevalent type of ovarian cancer, at 95%, and will be the subject of this paper.[10] EOC consists of five principal histotypes: low-grade serous carcinoma (LGSC), high-grade serous carcinoma (HGSC), mucinous carcinoma (MC), endometrioid carcinoma (EC), and clear cell carcinoma (CCC).[11] All EOC histotypes markedly differ by pathogenesis, molecular attributes, epidemiology, and clinical behavior. This analysis focuses on HGSC cases exclusively. Accounting for approximately 75% of EOC cases, the HGSC histotype is characterized by severe nuclear atypia and comprehensive TP53 mutations, and it is the most fatal histotype, often presenting as metastatic at diagnosis.[12] From a pathology perspective, the tumor may present with solid, granular, cribriform, or papillary architecture with substantial necrosis frequently observed at the site of solid growth.[10] Tothill et al. has further differentiated the HGSC histotype into subtypes of molecular and clinical significance: proliferative, immunoreactive, differentiated, and mesenchymal.[13] Variations in the tumor microenvironment and distinctive responses to treatments result in histotype-specific survival patterns.[14] To illustrate, EC and LGSC are associated with better survival

outcomes compared to HGSC, regardless of tumor stage.[14] HGSC has the highest mortality rates among EOC histotypes, particularly four or more years after diagnosis, making it the most lethal subtype.

EOC Fatality in Black Women

Irrespective of histotype, EOC prognosis is frequently poor among all women diagnosed. However, survival and incidence rates of EOC significantly differ when stratified by race.[15] The 5-year relative survival of EOC is approximately 52% overall, yet declines to 42% for Black women, despite the incidence of EOC being 25% lower in Black women compared to White women.[16,17] Limited access to quality healthcare among Black women in the United States could lead to the presentation of advanced disease at diagnosis and reduced possibility of receiving guideline-recommended treatment.[18] This inequity places Black women at a greater risk of a poor prognosis and death due to EOC.

Additionally, the disproportionate distribution of race-specific population-attributable (PAR) risks and prognostic factors pertaining to EOC, including body mass index (BMI), aspirin use, and body powder, may also partially explain the discrepancy in survival.[19] In a study collecting data on established EOC determinants, 51% of Black women with EOC were classified as obese (BMI ≥ 30 kg/m2) as compared to 21% of non-Hispanic White women (NHW).[20] Inflammation-related exposures, including obesity, nutritional influences, NSAID use, history of pelvic inflammatory disease (PID), history of endometriosis, history of PCOS, and hormone therapy duration, are theorized to be associated with EOC survival, as they may modify the immunological composition of the tumor microenvironment.[21,22] An analysis assessing the association between a comprehensive

inflammation-related risk score (IRRS), incorporating 11 of the 12 pro-inflammatory and anti-inflammatory components, and overall survival among a cohort of Black women with EOC determined that IRRS was associated with an increased risk of mortality risk of mortality (HR: 1.37, 95% CI: 1.02–1.85 for Q4 vs Q1).[22] A statistically significant dose-response relationship of an 11% increase in mortality per quartile of IRRS was also found. Additional evaluation of the imbalance of inflammatory prognostic variables and their corresponding effects on ovarian tumor microenvironments would be valuable to characterize racial disparities in EOC fatality further.

Type 2 Diabetes, which is disproportionately prevalent in Black women compared to White women, is another prognostic variable that has been linked to worse overall survival (OS) and progression-free survival (PFS) of EOC.[23, 24] A retrospective cohort study with 250 EOC patients found the HR for the association between diabetes and OS time was 3.93 (p-value < 0.001), following adjustment for BMI, parity, stage, histology, debulking status, HTN, metformin intake, and neoadjuvant chemotherapy.[24] The mechanism linking type 2 diabetes and EOC prognoses is not well understood, but this phenomenon may partly contribute to racial disparities in EOC survival along with the variables previously mentioned.

Genetic association among self-identified Black women with EOC

In addition to several inflammatory exposures, genetic ancestry may further drive differential gene expression in the tumor, shaping the immunological landscape and influencing overall survival.[25] The first genome-wide association study (GWAS) of EOC in women of African ancestry largely included data pulled from the African American Cancer

Epidemiology Study (AACES), the largest cohort to date of Black women diagnosed with EOC. This analysis aimed to determine whether women of African descent carry identical EOC susceptible genes as those previously identified in European-ancestry women. Four loci associated with EOC and six loci associated with HGSC histotype were detected among women with African ancestry at a suggestive threshold of $p < 1 \times 10-6$.[25] Of the novel SNPs identified, only one had been previously associated with EOC in GWAS studies of women with European ancestry, and the direction of the association was not congruent with that of African ancestry women in this study population. The functions of some newly identified variants in this analysis suggest biologically plausible involvement in EOC risk for women of African descent. Two SNPs (AKR1C3 and FST) are located near genes noted for regulating hormones and diseases of the ovary, and two are linked to cancer (AKR1C3 and MAGEC1).

Admixture and Ancestry-Associated Gene Signatures

Discovering differential ancestry-associated tumor expression among self-reported Black Americans is more complex, as the genetic makeup of African American populations in the United States is highly admixed or comprised of varying frequencies of previously isolated populations. African ancestry ranges from 16.64% to 99.9% among self-identified Black Americans, and European ancestry may vary regionally.[25] Due to this ancestral admixture of many African American populations, using self-reported race (SRR) alone has limited utility in genomic studies. Race is a valuable proxy when examining social determinants of health, such as structural racism, health care access, and environmental exposures. While ancestry does not capture social determinants of health, it does provide

crucial insights into genetic predisposition to diseases. Previous work identifying African ancestry-associated expression in Triple-Negative Breast Cancer (TNBC) tumors determined that 48.1% of the African ancestry-associated genes were distinguishable from SRR-associated genes.[26] Quantifying admixture by estimating global ancestry proportions for each individual can reveal ancestry-specific expression. Such an analysis has not been conducted in EOC tumors.

The first comparative bulk RNA-sequencing study comprising a cohort of African women and African American women with TNBC uncovered 613 African ancestry-associated gene signatures reported by Martini et al.[26] Pathway enrichment analysis indicates the predominant function of these genes is related to TNBC tumor growth and inflammatory response, including mechanisms of immune cell trafficking. These findings are consistent with preceding studies indicating that African ancestry may be linked with heightened inflammatory responses and immune cell enrichment in the tumor, potentially affecting prognosis and tumor response among women of African descent.[27]

Further evidence of increased inflammatory response at the tumor site among women with African ancestry is demonstrated by the differential expression of atypical chemokine receptor 1 (ACKR1/DARC) which encodes a seven-transmembrane G-protein-coupled receptor. Chemokine receptors such as ACKR1/DARC play a vital role in cell proliferation, metastasis, and overall cancer progression.[28] In a breast cancer study examining the relevance of DARC/ACKR1 concerning tumor immune response, they found that expression is significantly different across race.[29] African Americans have the greatest proportion of DARC/ ACKR1-low tumors (40.1%), and the DARC gene contains many single nucleotide variants (SNVs) specific to Sub-Saharan African ancestry. Higher

DARC/ACKR1 tumor expression is linked to longer overall survival and relapse-free survival. Functioning as a decoy receptor for several chemokines, including CCL2 and CXCL8, the DARC/ACKR1 receptor terminates their signaling mechanisms involved in pro-inflammatory pathways and tumor proliferation.[30] DARC/ACKR1 is expressed on various inflammatory, endothelial cells, and occasionally tumor-derived epithelial cells.[31] DARC/ACKR1 may modify the tumor microenvironment, leading to variation in prognosis, treatment response, and survival, yet, the relevance of ancestry-associated DARC/ACKR1 expression pertaining to EOC has not been extensively investigated.

Implications

Improved comprehension of variance in the tumor microenvironment and the clinical significance of differential gene expression may lead to improved outcomes in cancer patients. Identification of differentially expressed tumor genes linked to pro-inflammatory, or cell proliferation pathways may aid in prognosis prediction, earlier disease detection, and more precisely targeted therapies.

The purpose of this analysis is to identify ancestry-associated genes among Black women diagnosed with high-grade serous EOC, which may provide crucial information necessary to improve survival rates and help attenuate race disparities in survival. There has been little assessment of differential expression among Black women with EOC, therefore, we will take advantage of a unique opportunity to study the patterns of expression and resulting outcomes.

Methods

Study Population

The study population represents a subset of women enrolled in the African American Cancer Epidemiology Study (AACES), the largest population-based cohort of Black women diagnosed with EOC, and the North Carolina Ovarian Cancer Study (NCOCS), a population-based case-control study also among women diagnosed with EOC.[32] Collectively, these studies include a total of 747 Black women diagnosed with EOC, encompassing a comprehensive range of histotypes and disease stages. However, cases involving aggressive forms of EOC were underrepresented due to some participants succumbing to the disease prior to study enrollment. Participants in AACES provided verbal consent and signed forms authorizing pathology release and access to medical records, while written informed consent was obtained from participants in NCOCS.

Among the 747 Black EOC cases, 464 were classified as high-grade serous carcinoma (HGSC). Of this subset, 325 participants provided written consent for biospecimen-based analysis and had sufficient tumor tissue available for RNA extraction. Fifty-three participants who recently underwent neoadjuvant therapy were excluded from RNA sequencing due to the creation of specific signatures by chemotherapy in the tumor, rendering it incomparable to tumors from women without neoadjuvant chemotherapy. Furthermore, this study population includes participants from both AACES and NCOCS with available germline data, which was used to estimate global ancestry proportions. The final analytical sample consists of 211 self-identified African American women diagnosed with the HGSC histotype of EOC.

Demographic characteristics and information on established EOC risk factors were previously obtained from surveys for all cases in this analytical sample. Tumor attributes, including stage and debulking status, were abstracted from both pathology and medical records.

RNA Sequencing

To evaluate the tissue-specific expression among participants, RNA sequencing (RNA-seq) was executed using formalin-fixed paraffin-embedded (FFPE) tumor tissue. This technique utilizes a high-throughput technology that produces a comprehensive overview of the complete transcriptome, assessing both the arrangement and quantity of RNA. RNA-seq has proven to be a highly effective tool for identifying functional genes and pathways associated with the development and progression of cancer. Compiling RNA-seq findings, the Cancer Genome Atlas (TCGA) project generates a transcriptome database of 33 cancer types, including EOC, to identify potential therapeutic targets.[33] To illustrate, HGSC has been previously identified as having *TP53* mutations in nearly all tumors, along with a low frequency but statistically significant recurrence of somatic mutations in nine other genes, including *NF1, BRCA1, BRCA2, RB1, and CDK12*.[34] These RNA-seq findings depicted molecular subtypes and allowed for pathway analysis, revealing signaling pathways associated with tumor pathophysiology.

Following extraction from formalin-fixed paraffin-embedded (FFPE) tissue, RNA was stored at -80 degrees Celsius. Repurification was conducted using deoxyribonuclease (DNase) enzyme treatment and spin column purification in order to minimize the quantity of degraded RNA present (RNA < 200 nucleotides). Total RNA samples were used to

prepare RNA libraries using reagents from the Illumina Stranded mRNA Prep and the Illumina RNA UD Indexes Set (20091657). Then, to produce strand-specific libraries enriched for coding regions of RNA, amplified libraries were hybridized to biotin-labeled probes from the Illumina Exome Panel (cat# 20020183). Sequencing libraries were chemically denatured and subsequently administered to an Illumina NovaSeq flow cell using the NovaSeq XP workflow. Following the transfer of the flow cell to an Illumina NovaSeq 6000 instrument, a 150 x 150 cycle paired-end sequence run was conducted.

Utilizing fastp, a pre-processing and quality control tool, reads were filtered to a PHRED score of 15 at a minimum and a length of at least 20 base pairs. Although, most base pairs were high quality with a score greater than 30. The program Salmon was used to quantify paired-end reads using GRCh38 release 95. Sequence-specific biases were corrected using seqBias and gcBias flags. Additionally, to ensure quantification accuracy, the rangeFactorizationBins parameter was set to 4. Low-expression genes were filtered out if they contained a median expression of 0 within a dataset. To ensure estimates were not biased by the total number of reads per sample, an upper quantile normalization technique was enacted by matching the 85th percentile across samples. Further samples were excluded where the median normalized read count was under 925. Lastly, samples were removed that were flagged as having very similar expression patterns by doppelgangR. Further detail on the RNA-seq processing of this data can be found in the paper by Davidson & Barnard et. al.[35]

Statistical Analysis

Global Ancestry Proportions

DNA extracted from participant blood or saliva samples was genotyped using the OncoArray platform. Single nucleotide polymorphisms (SNPs) were included in the analysis if they existed in the pre-imputation OncoArray dataset, the 1000 Genomes Project phased reference data, and centimorgan coordinate maps. The percentage of SNPs assigned to African, European, and Asian ancestries was estimated across chromosomes 1 through 22. To approach regional ancestry, five African subpopulations were incorporated: Esan in Nigeria (ESN), Mende in Sierra Leone (MSL), Yoruba in Nigeria (YRI), Gambian in Western Divisions in the Gambia (GWD), and Luhya in Webuye, Kenya (LWK). Notably, LWK represents the sole East African subpopulation in the 1000 Genomes Project, while the other subpopulations are classified as West African. Reference samples selected from the 1000 Genomes Project were evenly distributed across the previously mentioned populations and incorporated all women to ensure comparability.

Chromosomes 1 through 22 were analyzed individually using RFMix, a robust discriminative algorithm for estimating ancestry proportions.[36] RFMix assigns each SNP to its most likely ancestry of origin by leveraging random forest methodology, which integrates multiple decision trees to generate a final prediction. The algorithm uses genetic information collected from small segments, referred to as windows, to manipulate the random forest classifier.

To quantify ancestry proportions, the total number of SNPs assigned to each ancestry was summed and divided by the total number of SNPs analyzed, estimating the

percentage of SNPs attributable to each ancestry. These assignments included African ancestry and West African ancestry in addition to the aforementioned populations.

Ancestry-Expression Linear Regression

Quantified genetic ancestry proportions were used to identify ancestry-associated tumor-specific genes in participants, using linear regression analysis comparing gene expression with the continuous ancestry variable. This analysis was conducted independently for each ancestry group (African, European, Asian, Mende in Sierra Leone (MSL), Yoruba in Nigeria (YRI), Esan in Nigeria (ESN), Yoruba in Nigeria (YRI), Gambian in Western Divisions in the Gambia (GWD), and Luhya in Kenya (LWK). We reported the top five most significant genes for each ancestry. A Bonferroni correction for the exploratory p-value of 0.10 was implemented to address the increased risk of false positive results due to multiple comparisons. The number of comparisons for each linear regression analysis corresponded to the number of genes (18,699), bringing the resulting p-value threshold to 5.35×10^{-6}.

Covariates

Some risk factors were identified in the literature and assessed to examine the potential influence on the association between ancestry and tumor-specific expression, including body mass index (BMI). Higher BMI is of notable interest due to its established association with obesity-related chronic low-grade inflammation, which may promote tumor growth and influence the composition of the tumor microenvironment. To evaluate potential differential gene expression by BMI, violin plots illustrating the distribution of

expression and summary statistics were generated for the top five most significant genes of each ancestry group, stratified by BMI (BMI ≥40 kg/m^2 versus BMI <30 kg/m^2). And two-sample t-tests were performed to determine whether there were statistically significant differences in gene expression between the stratified BMI groups. Participants classified as overweight (BMI 30–39.9 kg/m^2) were excluded from this analysis to focus on the comparison between individuals with obesity, who are more likely to exhibit increased pro-inflammatory mechanisms, and those within a healthy weight range.

Additionally, type 2 diabetes may influence the tumor microenvironment through the differential expression of genes and pathways associated with chronic inflammation, potentially forming a tumor microenvironment more conducive to tumor progression and metastasis. Previous studies have found poorer prognosis in EOC patients diagnosed with diabetes compared to those without, controlling for obesity and metformin use.[37] Violin plots displaying the distribution and summary of expression for the top five most significant genes of each ancestry group were created, stratified by the presence or absence of a type 2 diabetes diagnosis. Two-sample t-tests were executed to identify statistically significant differences in expression by type 2 diabetes status.

Results

Characteristics of Participants and Cancer Diagnosis

Variables comprised of demographic information, ovarian cancer risk factors and prognostic factors, protective factors, and ovarian tumor pathology data were recorded for all study participants (Table 1). Established ovarian cancer risk factors which were documented for participants include body mass index (BMI), smoking status, history of type 2 diabetes mellitus (T2DM), history of endometriosis, nulliparity, first-degree family history of ovarian cancer, and first-degree family history of breast cancer.

BMI is categorized by the following variables: (<18.5 kg/m^2), normal weight (18.5 - 24.9 kg/m^2), overweight (25 - 29.9 kg/m^2), class I obesity (30 - 34.9 kg/m^2), class II obesity (35 - 39.9 kg/m^2), and class III obesity (> 40 kg/m^2). The majority of participants were categorized as overweight or obese, as only 16.3% of the cohort fell within the “normal weight” category. The most prevalent BMI categories among study participants were overweight and class I obesity, which encompassed 28.4% and 25.5% of participants, respectively. Overall, 83.3% of participants were classified as overweight or obese.

Type 2 diabetes (T2DM), associated with both obesity and an increased risk of ovarian cancer, was diagnosed in 14% of participants- higher than the national average for self-report Black women of approximately 12.7% in 2021.[38] Notably, the proportion of participants with T2DM who were obese (BMI ≥ 30) was nearly equivalent to the proportion of those with T2DM and a BMI below 30.

In our study, 29.6% of participants reported a first-degree family history of breast cancer- substantially higher than the 11% prevalence observed in a representative cohort of cancer-free women in the United States.[39] Additionally, 9.4% of subjects reported a first-

degree family history of ovarian cancer, which significantly increases the risk of being diagnosed with the disease.

A history of endometriosis was reported by 8.7% of participants, which is comparable to the estimated 10% prevalence among women in general. However, it is important to acknowledge that Black women have historically been underdiagnosed with this condition.[40]

Smoking was common among participants: 20.9% identified as current smokers, 38.4% as former smokers, and 40.8% had never smoked. Oral contraceptive use- known to provide protective effects against ovarian cancer- was reported by 69% of participants, with a mean duration of use of approximately 43 months. This statistic is generally consistent with CDC data showing that 64.9% of U.S. women aged 15 to 49 use oral contraceptives.[41]

23.1% of participants reported general use of NSAIDs, while 14.7% specifically indicated aspirin use in the past.

Approximately half of the participants in this analysis reported an education level of high school graduate, GED equivalent, or lower. Married individuals comprised 40.8% of the study population, representing the most common marital status. The majority (83.3%) of participants had given birth, with an average of 2.33 pregnancies across the cohort.

Lastly, regarding tumor characteristics, approximately 75% of high-grade serous carcinoma (HGSC) cases were diagnosed at stage III. Suboptimal debulking status post-surgery was reported in 33.6% of participants. 88% of participants underwent a hysterectomy, which is commonly executed in ovarian cancer treatment. Among molecular

subtypes, the immunoreactive subtype was most prevalent, accounting for 36.5% of tumors, followed by the mesenchymal, proliferative, and differentiated subtypes.

Table 1 summarizes the demographic and clinical characteristics of the study population and highlights potential patterns associated with HGSC cases in Black women.

Table 1: Demographics and clinical characteristics of the study population (n = 211)	
EOC subtype (%)	
Differentiated	17 (8.1%)
Immunoreactive	77 (36.5%)
Mesenchymal	56 (26.5%)
Proliferative	51 (24.2%)
Unknown	10 (4.7%)
FIGO stage (%)	
I	21 (10.1%)
II	22 (10.6%)
III	156 (75.0%)
IV	9 (4.3%)
Age at diagnosis (mean(SD))	42.8 (64.8)
Debulking status (%)	
Optimal	89 (66.4%)
Suboptimal	45 (33.6%)
Family history of breast cancer (%)	58 (29.6%)
Family history of ovarian cancer (%)	18 (9.4%)
Endometriosis diagnosis (%)	18 (8.7%)
Type 2 diabetes diagnosis (%)	28 (13.3%)
BMI category (%)	
Underweight (< 18.5 kg/m2)	1 (0.5%)
Normal weight (18.5 - 24.9 kg/m2)	34 (16.3%)
Overweight (25 - 29.9 kg/m2)	59 (28.4%)
Class I obesity (30 - 34.9 kg/m2)	53 (25.5%)
Class II obesity (35 - 39.9 kg/m2)	28 (13.5%)
Class III obesity (=> 40 kg/m2)	33 (15.9%)
History of hysterectomy (%)	185 (88%)
Age at hysterectomy (mean(SD))	50.7 (12.9)
Smoking status (%)	
Current smoker	44 (20.9%)
Never smoker	86 (40.8%)
Former smoker	81 (38.4%)
History of oral contraceptive use (%)	145 (69.0%)
Oral contraceptive duration in months (mean(SD))	42.8 (64.8)
Number of pregnancies (mean(SD))	2.33 (1.82)
Nulliparous (%)	35 (16.7%)
Education level (%)	
HS graduate/ged or less	109 (51.9%)
Some college	40 (19.0%)
College grad	35 (16.7%)
Grad/professional school	26 (12.4%)
Marital status (%)	
Single/never married	30 (14.9%)
Married	82 (40.8%)
Divorced/separated	69 (34.3%)
Widowed	20 (10.0%)
NSAID use (%)	46 (23.1%)
Aspirin use (%)	27 (14.7%)
Talc use on genital areas (%)	79 (38.0%)

Global Ancestry Proportions

Figure 1.1 presents the distribution of global ancestry proportions among the study population (n = 211). Overall, the distribution of Asian ancestry proportions displayed low variance and held a few higher outliers with a median of 1.6% across participants and range of 0.01% to 1.3%. As the representation of Asian descent was extremely low among this study group, Asian ancestry was excluded from the ancestry-expression linear regression analysis. Alternatively, increased variability with a relatively even distribution was observed for European ancestry percentages, ranging from nearly 0% to nearly 44%. The median European ancestry proportion was 16.1%. Among the African subpopulations selected for analysis, the highest ancestry proportions were attributed to

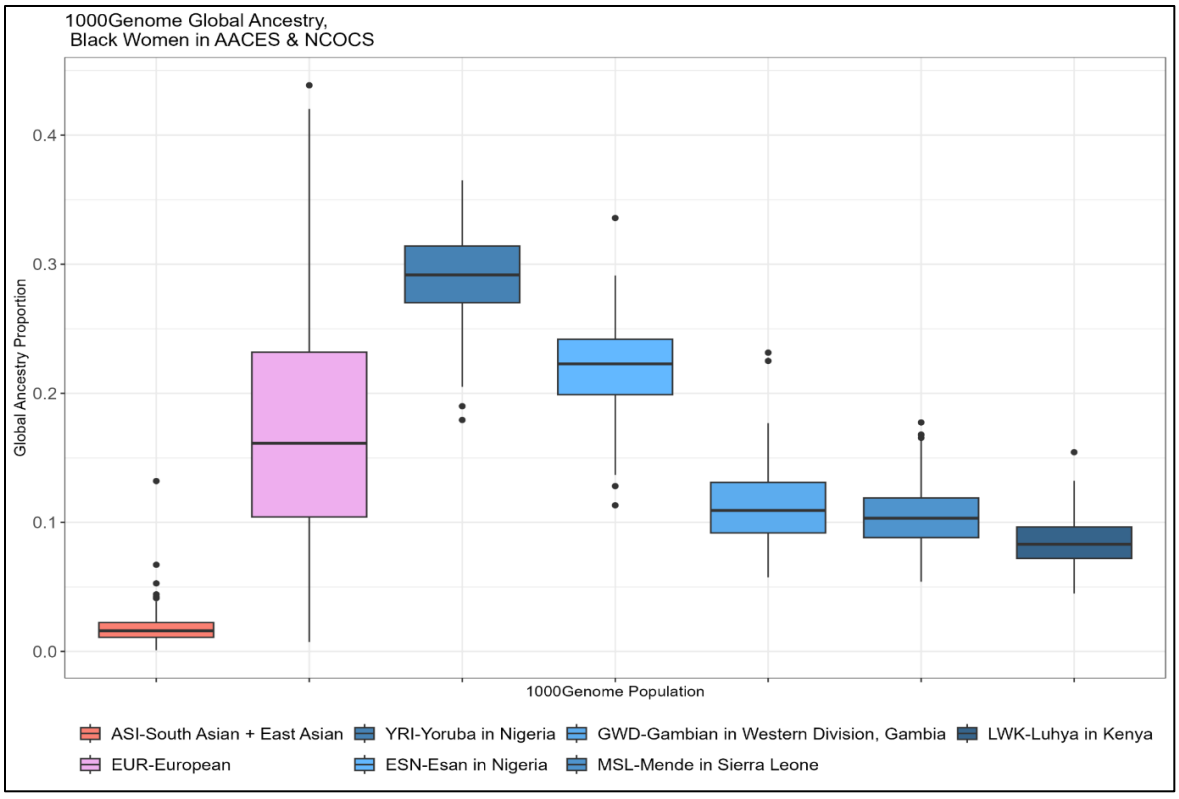


Figure 1.1: Study population global ancestry percentages.

Yoruba in Nigeria (YRI) and Esan in Nigeria (ESN) subpopulations, with respective medians of 29.2% and 22.3%, and ranges of 1.8-36.5% and 11.3-33.6%. The sole East African population assessed, Luhya in Kenya (LWK), contributed to the smallest proportion of African ancestry overall with a median of 8.3% (range: 0.5%-15.4%). The genetic contributions the of Gambian in Western Division (GWD) and Mende in Sierra Leone (MSL) subpopulations fell in between the Nigerian subpopulations (YRI and ESN) and East African (LWK).

Figure 1.2: Global ancestry proportions stratified by state of residence.

To evaluate potential regional variances in ancestry proportions among participants, ancestry plots were stratified by state of residence (Figure 1.2). It is worth noting that the study population is not evenly distributed across represented states with 92 individuals residing in North Carolina and a majority living in southeastern states overall. The mean proportion of European ancestry slightly varied across states, ranging from 12.5% in South Carolina up to 20.1% in Louisiana. All proportions of African subpopulations (YRI, ESN, GWD, MSL, and LWK) remained nearly equivalent across the states depicted.

Ancestry-Associated Gene Expression

Table 2 presents the top five most statistically significant genes identified for each ancestry group, along with their unadjusted linear regression coefficients and corresponding p-values. This section provides a brief overview of these genes, while a more in-depth discussion of their biological functions and potential roles in cancer progression is included in the discussion section.

Referring to the Bonferroni p-value threshold of $5.35 \times 10^{-6,}$ only one resulting gene is considered statistically significant in their relationship with ancestry proportions. The expression of *SNX27* is significantly and inversely associated with the proportion of African ancestry.

African, West African, and European ancestry share the same first and second most significant genes, *SNX27* and *TRBV29.1,* respectively. The association between *SNX27* expression and African ancestry displayed the singular statistically significant relationship out of the whole dataset with a p-value of 2.90×10^{-6}. The expression of *SNX27* and *TRBV29.1* are both upregulated with European ancestry, having a positive LR

coefficient, and they are downregulated with African ancestry and West African ancestry, resulting in negative coefficients. *SNX27*, also known as sorting nexin 27, is involved in retrograde transport from the endosome to the plasma membrane, promoting the recycling of internalized transmembrane proteins, rather than the destruction by lysosomes. The protein product of *TRBV29.1* is the variable domain of the T-cell receptor complex which participates in antigen recognition. Following protein recognition, signaling cascades lead to activation of transcription factors which promote T-cell growth and differentiation. Therefore, the expression of *TRBV29.1* is crucial for a successful adaptive immune response.

The remaining African ancestry-associated genes of the top five include *VPS45*, *TXNIP*, and *RASGEF1B*, all of which, are downregulated relative to higher proportions of African ancestry. The function of *VPS45* is largely unknown but evidence suggests that the gene plays a role in inflammatory mediator transportation. *TXNIP* encodes a thioredoxin-binding protein, inhibiting the antioxidative properties of the thioredoxin, and exposing the cell to oxidative stress. The encoding protein is also thought to act as a tumor suppressor in certain cases. *RASGEF1B* functions as a guanine-nucleotide exchange factor (GEF) proposed to be involved with the GTP-binding proteins RAS signal transduction pathway crucial to cell growth and differentiation.

The genes *IGHGP* and *HEBP2P1* are unique to European ancestry, while West African and European ancestry both contain the *YWHAH* gene in their resulting 5 most significant lists, with *YWHAH* being upregulated for European and downregulated for West African ancestry. *IGHGP* is upregulated for European ancestry, while *HEBP2P1* is downregulated. Both *IGHGP* and *HEBP2P1* are classified as pseudogenes, and while they

may contribute to gene transcription or cellular functions, very little is known about their contribution. The product of *YWHAH* belongs to the 14-3-3 family of proteins, which bind to phosphorylated proteins, mediating several signaling pathways including MAPK, WNT, and AKT. 14-3-3 proteins are instrumental to cell metabolism, protein trafficking, apoptosis, and cell cycle regulation.

Top West African genes *SNX27*, *TRBV29.1*, and *YWHAH* have been previously noted due to the overlap with top African and European genes. The pseudogene *RPL21P32* is upregulated by both West African and MSL ancestries, while *CEP104* was downregulated and unique to West African ancestry. *CEP104* encodes a centrosomal protein required for cilia formation and maintenance. The protein also plays a role in microtubule regulation, influencing cell division.

Two forkhead box genes, *FOXD4L5* and *FOXD4L6*, were upregulated for LWK ancestry. Both *FOXD4L5* and *FOXD4L6* are predicted to participate in cell differentiation and transcription though RNA polymerase II regulation, although, the details of its mechanisms are largely unknown. The most significant gene for LWK ancestry, *FUT2*, encodes a protein crucial to the final step of secretor-type antigen synthesis and has implications for disease susceptibility and immune response. *KCNJ1* facilitates potassium homeostasis in cells by encoding an inwardly-rectifying potassium channel embedded in the cell membrane. Last among LWK genes, *NPM1P29*, is a pseudogene with limited research concerning its involvement in cellular processes. However, its parent gene *NPM1* facilitates cell proliferation and protein chaperoning. The top five most significant genes for LWK ancestry, which were mentioned above, were all upregulated.

The most statistically significant genes found for MSL ancestry were *RAB11B, TPH2, UPF3AP2, RPL21P32,* and *CCDC42,* all of which were upregulated. The function of *RPL21P32* was previously mentioned, as the gene is also significant for West African ancestry. The next gene, *RAB11B* plays a critical role in the regulation of endocytic and exocytotic pathways, facilitating transport processes of numerous proteins. *TPH2* is primarily known for its role in serotonin and melatonin biosynthesis pathways. *UPF3AP2* is a non-functional gene derived from *UPF3A* which regulates mRNA stability and quality control. Lastly, *CCDC42* plays a critical role in the formation of motile cilia, acting either on or upstream of the centrosome cycle.

With respect to YRI ancestry proportions, the most significant genes were *GDAP2, RXFP1, PPP2R2C, LTBP1,* and *GDF10*. The top five most significant genes were all downregulated. *GDAP2* is predicted to play a role in retinoic acid (vitamin A derivative) pathways, otherwise, little is known regarding the gene function. The gene *RXFP1* encodes for a part of a subgroup of the G protein-coupled 7-transmembrane receptor superfamily. As a receptor for the hormone relaxin, the encoded protein is crucial for pregnancy and parturition. Reduced expression of *RXFP1* has been associated with endometriosis. Next, the product of *PPP2R2C* is a regulatory subunit for protein phosphatase 2a (PP2A), an enzyme that modulates cell cycle progression. PP2A activity is commonly suppressed with cancer, leading to sustained cell proliferation. The protein encoded by *LTBP1* regulates and activates the transforming growth factor beta (TGF-β) signaling pathway which influences cell growth and immune responses. Interestingly, the final gene for YRI, *GDF10*, encodes the ligand that binds to the TGF-β receptor. Subsequent to ligand induction, SMAD transcription factors are activated and bind to

promoter regions of DNA, influencing expression. *GDF10* also acts as a tumor suppressor gene in certain cancers, applying its effects through the TGF-β signaling pathway's control of cell proliferation and differentiation.

In relation to GWD ancestry proportions, the expressions of *MAGEH1, TRO, COL25A1, HMGB1,* and *HCLS1* were upregulated and the most statistically significant genes. *MAGEH1* belongs to of a non-CT (cancer/testis) subgroup of the melanoma-associated antigen (MAGE) protein family, indicating that its expression is not necessarily restricted to germ cells or cancer cells, unlike most MAGE genes. The protein encoded by *MAGEH*1 is predicted to be involved with cell cycle arrest and apoptosis. Similarly, *TRO* also belongs to the MAGE protein family. The protein encoded by *TRO* mediates cell adhesion between endometrium epithelial cells and trophoblastic cells during embryo implantation, and it is presumed to play a role in ectopic pregnancy and potential cancer formation. *COL25A1* is highly expressed in brain cells and has no known function outside of the nervous system. *HMGB1* produces a protein belonging to the High Mobility Group-box superfamily. The encoded protein has several functions, including regulation of tumor cell migration, inflammatory processes, and cell differentiation. The final gene *HCLS1* is primarily involved in cell migration and immune cell signaling.

The top five genes identified for ESN ancestry were *G3BP1P1, YWHAH, TMEM106B, CEP164,* and *SNCA. YWHAH* function has been previously noted due to its overlap with West African and European ancestries. Similar to West African ancestry, it is downregulated. *G3BP1P1* is an upregulated pseudogene derived from *G3BP1*, which is involved in innate immunity processes and stress granule formation. *TMEM106B,* downregulated relative to higher ESN ancestry proportions, is highly expressed in neurons

and controls endosomal and lysosomal trafficking. *CEP164* is upregulated and modulates microtubule organization, chromosome segregation, and DNA damage response through the production of a centrosomal protein. Finally, *SNCA* is involved in several neuronal processes including synaptic vesicle trafficking and subsequent neuron release. *SNCA* is downregulated with higher proportions of ESN ancestry.

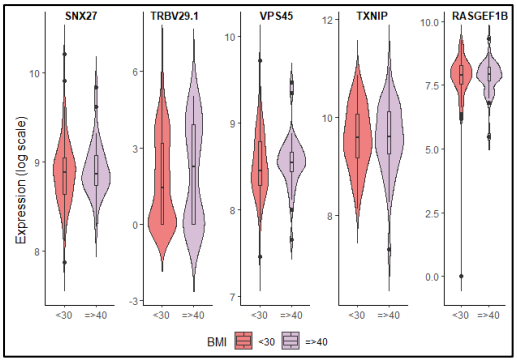
ancestry	gene	coefficient	pvalue
African	SNX27	-13436.0	2.90E-06
	TRBV29.1	-122.6	2.80E-04
	VPS45	-6576.6	3.29E-04
	TXNIP	-30260.1	5.28E-04
	RASGEF1B	-5099.1	6.49E-04
European	SNX27	12421.2	2.37E-05
	TRBV29.1	140.8	3.61E-05
	YWHAH	9610.3	1.19E-04
	IGHGP	376.0	4.57E-04
	HEBP2P1	-151.7	4.76E-04
West African	SNX27	-12184.7	1.04E-04
	TRBV29.1	-138.9	1.35E-04
	RPL21P32	92.0	4.33E-04
	CEP104	-8659.1	4.72E-04
	YWHAH	-9136.7	6.18E-04
LWK	FUT2	1.39E+04	2.52E-05
	FOXD4L5	2.22E+03	4.67E-05
	KCNJ1	1.07E+03	9.46E-05
	NPM1P29	1.65E+02	1.12E-04
	FOXD4L6	2.81E+03	1.53E-04
MSL	RAB11B	49877.2	1.52E-05
	TPH2	439.9	7.85E-05
	UPF3AP2	2126.9	1.73E-04
	RPL21P32	350.4	1.75E-04
	CCDC42	1472.4	2.13E-04
YRI	GDAP2	-9443.5	2.04E-04
	RXFP1	-3297.5	2.18E-04
	PPP2R2C	-8390.2	2.64E-04
	LTBP1	-87196.5	2.75E-04
	GDF10	-1414.0	3.12E-04
GWD	MAGEH1	4417.7	3.60E-05
	TRO	52830.4	1.01E-04
	COL25A1	11307.4	3.04E-04
	HMGB1	29342.7	4.99E-04
	HCLS1	26243.3	5.74E-04
ESN	G3BP1P1	826.9	2.89E-05
	YWHAH	-25259.8	1.60E-04
	TMEM106B	-12071.3	1.94E-04
	CEP164	33844.1	2.68E-04
	SNCA	-2732.3	3.12E-04

Table 2: Top 5 genes of each ancestry with respective LR coefficients and raw p-values

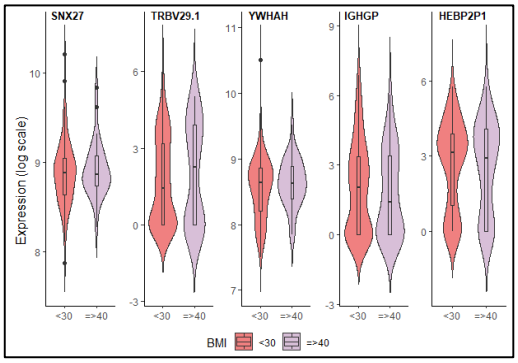
Expression stratified by BMI and Type 2 Diabetes

Of the top five genes for each ancestry, four were found to have statistically significant differential expression by the selected BMI categories (BMI ≥40 kg/m^2 versus BMI <30 kg/m^2) when conducting 2 sample t-tests (Figure 2). This consists of the genes *FOXD4L5* of LWK (p-value = 0.0212), *G3BP1P1* of ESN (p-value = 0.0028), *TMEM106B* of ESN (p-value = 0.0129), and *GDF10* of YRI (p-value = 0.0471). *FOXD4L5*, *G3BP1P1*, *GDF10*, and *TMEM106B* were all significantly downregulated among participants with a BMI of 40 kg/m^2 compared to expression in participants with a BMI of less than 30 kg/m^2.

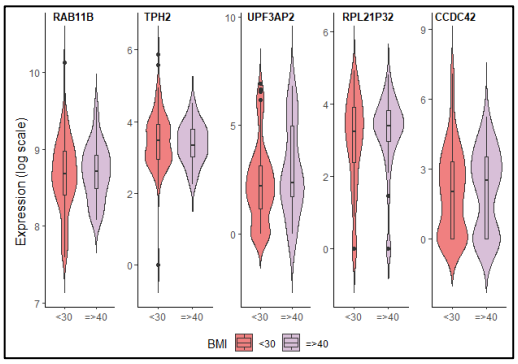
As for type 2 diabetes diagnosis, *FOXD4L5* of LWK ancestry was the only statistically significant gene found to differ in expression by type 2 diabetes status (p-value = 0.0211), which is displayed in Figure 3. Similar to its differential expression by BMI, *FOXD4L5* is significantly downregulated in participants diagnosed with type 2 diabetes, compared to expression in participants with no history of type 2 diabetes.



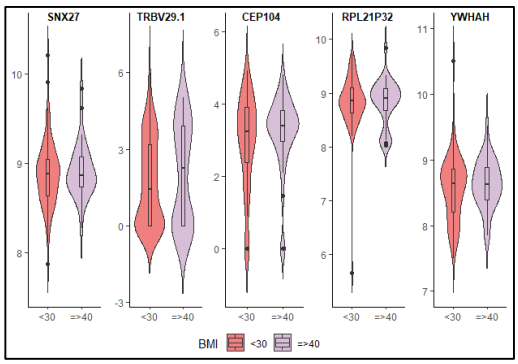
A. African



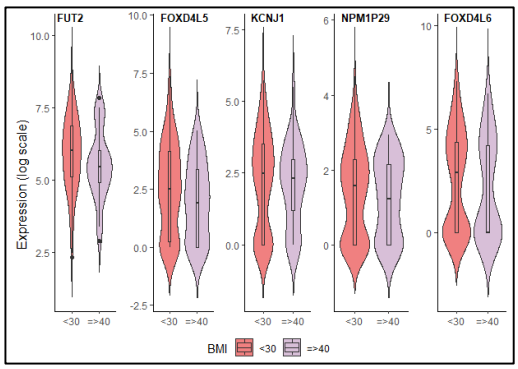
B. European



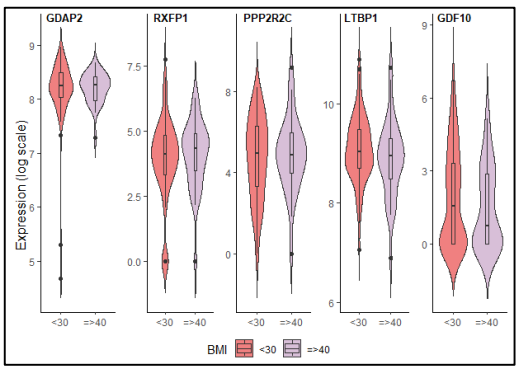
C. MSL



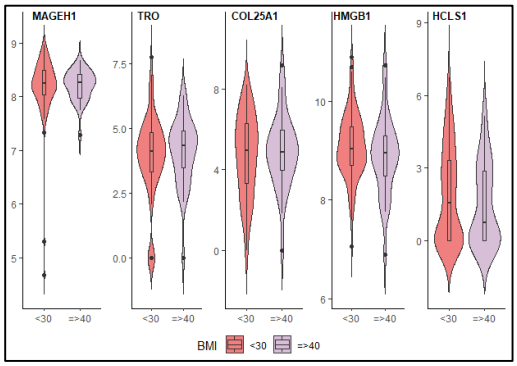
D. West African



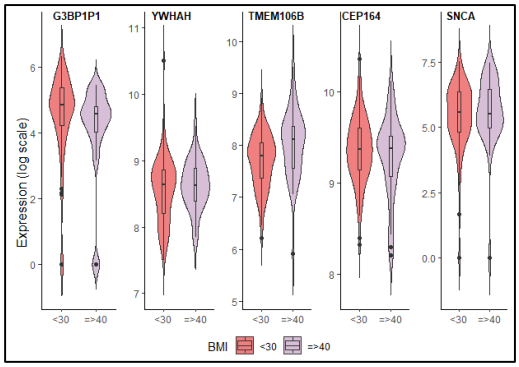
E. LWK



F. YRI

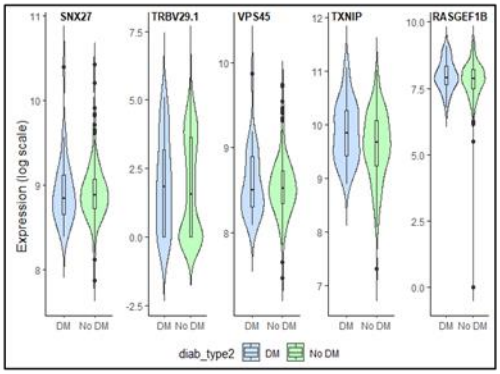


G. GWD

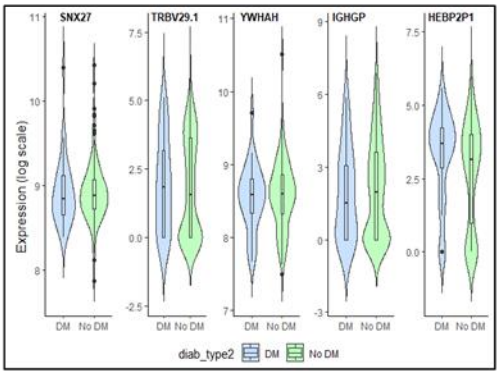


H. ESN

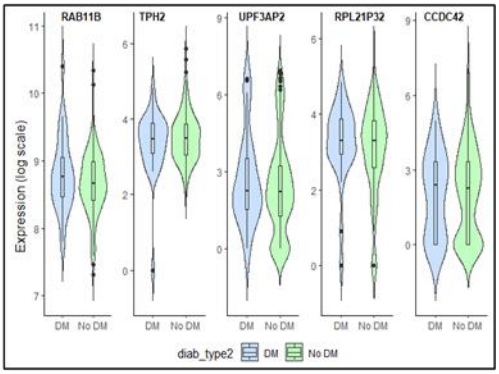
Figure 2: Expression of the 5 most significant genes of each ancestry. Stratified by BMI (<30 versus =>40).



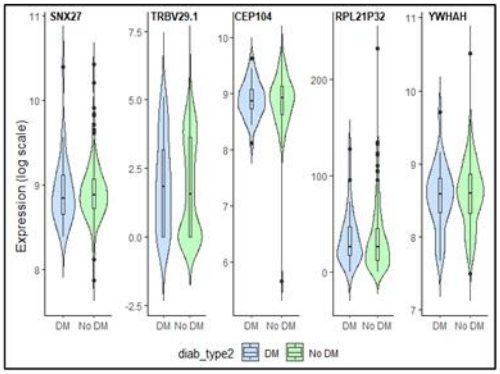
A. African



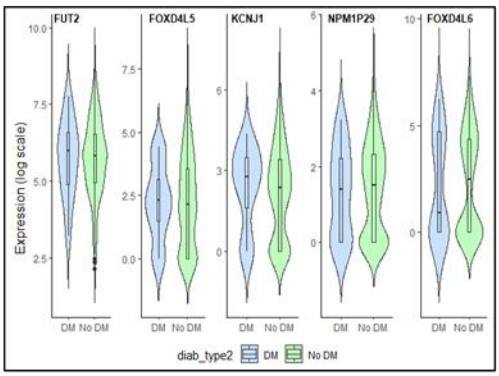
B. European



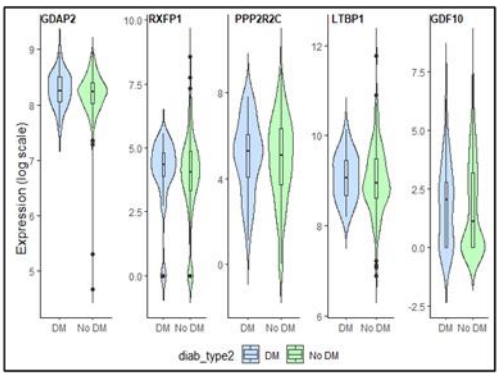
C. MSL



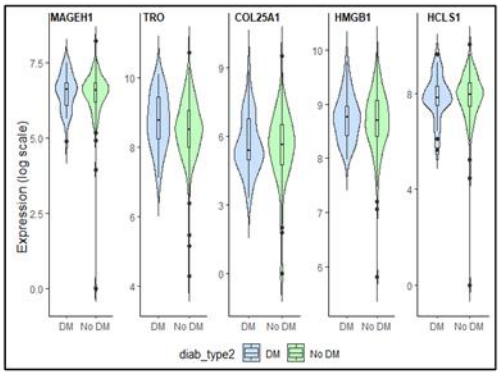
D. West African



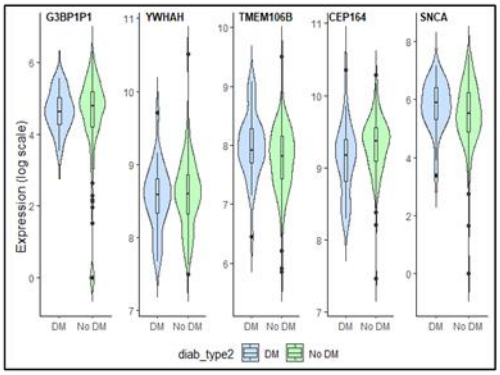
E. LWK



F. YRI



G. GWD



H. ESN

Figure 3: Expression of the 5 most significant genes of each ancestry. Stratified by type 2 diabetes diagnosis.

	Cell proliferation	Immune response + inflammation	Transcription regulation	Cell differentiation	Cell homeostasis	Protein transport	Cell migration	Cell metabolism	Cell-cell interaction	DNA damage response
SNX27		X				X				
TRBV29.1		X								
VPS45						X				
TXNIP		X			X			X		
RASGEF1B	X		X	X						
YWHAH	X					X		X		
IGHGP										
HEBP2P1										
RPL21P32										
CEP104										
FUT2	X								X	
FOXD4L5			X	X						
KCNJ1					X					
NPM1P29										
FOXD4L6			X	X						
RAB11B						X				
TPH2										
UPF3AP2										
CCDC42	X		X				X			
GDAP2					X					
RXFP1										
PPP2R2C	X				X					
LTBP1	X	X		X			X			
GDF10	X			X						
MAGEH1	X	X								
TRO									X	
COL25A1										
HMGB1		X					X	X		
HCLS1			X				X			
G3BP1P1										
TMEM106B					X					
CEP164	X									X
SNCA										

Figure 4: Common functions of resulting genes.

Discussion

Implications of Ancestry-Associated Gene Expression

Overall, the top five genes identified as differentially expressed for each ancestry exhibit diverse functions and play varied roles in vital cellular processes. Some of these genes have been repeatedly linked to cancer metastasis and tumor progression, given their established involvement in cell proliferation, migration, immune evasion, and other processes that promote a tumor-supportive environment. Conversely, some identified genes are either non-functional or have limited information available in the literature regarding their potential association with cancer in general or EOC specifically. This section explores the possible implications of the five most significant genes for each ancestry in relation to EOC.

SNX27 and *TRBV29.1* upregulation is associated with higher proportions of European ancestry, while SNX27 and *TRBV29.1* downregulation is associated with higher proportions of African and West African descent. Further, the genes have correlated functions, suggesting that *SNX27* and *TRBV29.*1 may function in conjunction to promote tumorigenesis. Specifically, the overexpression or underexpression of *SNX27* and *TRBV29.1* may impact tumor cell growth and proliferation through T-cell regulation. *SNX27* facilitates the recycling of the T-cell receptor (TCR) from endosomes to the plasma membrane, while *TRBV29.1* encodes the variable domain of the TCR receptor, enabling cancer cell antigen recognition.[42,43] Upon antigen recognition, downstream signaling cascades stimulate T-cell activation and proliferation. Insufficient expression of either *SNX27* or *TRBV29.1* may impair T-cell presence, reducing the immune system's ability to identify and eliminate cancer cells.[44] This weakened adaptive immune response to cancer cells could contribute to uncontrolled cell proliferation.

VPS45, another downregulated gene in individuals with African ancestry, is thought to be involved in leukocyte intracellular protein transport, similar to SNX27, although its precise function remains unclear.[45] Given the characterizations and expressions of *VPS45*, *SNX27*, and *TRBV29.1*, individuals with higher African and West African ancestries may be susceptible to decreased T-cell production in response to EOC cells, potentially leading to poorer prognoses.

TXNIP, a key regulator of cellular oxidation and metabolism, plays a complex role in cancer cells, as it has been suggested that its function may have varying implications contingent on cancer type and stage. The consequence of *TXNIP* upregulation relative to higher African ancestry is unclear until its mechanistic role in EOC cells is better understood. Adding to the complexity, when stratifying early- vs. late-stage cases of ovarian cancer in one study, *TXNIP* expression was found to be associated with differential clinical outcomes, specifically, poor survival in late-stage disease but better survival in early-stage disease.[46] The mechanism responsible for this phenomenon remains unknown.

Additionally, it has been proposed that *TXNIP* may play a key role in cancer and diabetes, as it targets both thioredoxin (TRX)-mediated redox regulation and glucose transporter (GLUT)-mediated metabolism.[47] *TXNIP* expression is associated with a significant reduction in cellular glucose uptake. Increased glucose uptake and hyperactive glycolysis are typical metabolic changes observed in EOC to fuel cell growth and proliferation, aligning with the downregulation of *TXNIP* associated with African ancestry. Further, increased expression of GLUT receptors is linked with poorer prognosis and

survival times across EOC patients, illustrating *TXNIP* as a potential prognostic marker for EOC.[48]

RASGEF1B is thought to activate RAP2 (RAP2A, RAP2B, RAP2C) proteins from their inactive-GDF-bound state to active GTP-bound state.[49] RAP2 proteins act as intracellular signal transducers responding to the stiffness and composition of the extracellular matrix (ECM).[50] The rigidity of the ECM is a vital cell property that can drastically alter cell signaling and has the capacity to promote a tumor microenvironment conducive to cancer metastasis. RAP2 activation has been previously implicated in cancer progression, due to its promotion of cell migration and negative regulation of cell adhesion.[51] This relationship between RAP2 protein and cancer progression is probable but unconfirmed. *RASGEF1B* is downregulated relative to higher proportions of African ancestry, which, in theory, would decrease the quantity of RAP2 proteins and slow tumor cell migration, protecting against EOC progression. With this logic, African ancestry would be associated with improved EOC prognoses.

YWHAH, upregulated for higher proportions of European ancestry and downregulated for higher proportions of West African ancestry, has been demonstrated to function as an oncogene in thyroid cancer cells and as a tumor-suppressor gene in hepatocellular cancer cells through the mediations of several signaling pathways.[52] It is poorly understood how YWHAH-related pathways impact ovarian cancer cells specifically. Further research is needed to provide insight into the differences in the expression of *YWHAH* for European and West African ancestries, as *YWHAH* is involved in numerous cellular processes such as metabolism, cell cycle regulation, and protein trafficking.

CEP104, which is downregulated relative to higher proportions of West African ancestry, facilitates microtubule formation and ciliogenesis, impacting cell division and motility.[53] Errors in *CEP104* may lead to dysfunctional mitotic spindles that are essential to mitosis and consequently trigger abnormal cell division.[54] *CEP104* knockout has been demonstrated to cause ciliopathy, a group of disorders resulting from defective cilia. The gene's potential connection to cancer is not established, although underexpression may be related to abnormal cell division or dysfunctional cell signaling because of cilia abnormalities. One study found that the percentage of ciliated cells was lower in HGSOC metastasis cases compared to patient-matched primary tumors.[55] Further work must be done to comprehend the implications of these cilia patterns and how the downregulation of *CEP104* for West African ancestry may affect cancer progression.

Expressed differentially by LWK ancestry, *FUT2* controls the secretion and expression of ABO blood group antigens in bodily fluids.[56] *FUT1* and *FUT2* catalyze the last step of Globo-H antigen biosynthesis, adding fructose to the precursor. A ceramide portion is synthesized in the ER and Golgi apparatus, allowing the newly formed Globo-H ceramide to anchor to the cell membrane. Globo-H ceramides are overexpressed on the cell surfaces in a variety of epithelial cancers, contributing to immune evasion and tumorigenesis.[57] A study assessing the levels of anti-Globo H antibodies in participants diagnosed with serous ovarian cancer in comparison to women without gynecological malignancies, found that participants with serous ovarian cancer displayed significantly higher levels of the antibody when measured with both flow cytometry and suspension array techniques.[58] This finding aligns with the upregulated expression of *FUT2* in

association with LWK ancestry among HGSC cases. Globo-H has even been proposed as a potential tumor marker and therapeutic target.

FOXD4L5 and *FOXD4L6* are both slightly upregulated among women with greater proportions of LWK ancestry, and the genes may have coregulated expression as they are located in the same gene cluster on chromosome 9.[59] Notably, *FOXD4L5* was differentially expressed in our analysis when stratified by BMI and when stratified by type 2 diabetes diagnosis. However, existing literature has not implicated *FOXD4L5* in either obesity or type 2 diabetes. Both genes belong to the Forkhead Box (FOX) transcription factor family, composed of an evolutionarily conserved group of transcription regulators predicted to be involved in several cellular processes and human diseases.[60] All genes in the FOX family have close mouse orthologs with one exception. The mouse has a single *FOXD4* gene whereas the human gene has undergone a recent replication to a total of 7 genes (*FOXD4* and *FOXD4L1* to *FOXD4L6*).[61] The FOXD4 subgroup genes are not well understood relative to the rest of the FOX family genes. Further research is needed to better comprehend the functions of *FOXD4L5* and FOXD4L6, particularly in relation to obesity, diabetes, and cancer.

KCNJ1, which is upregulated with higher proportions of LWK ancestry, encodes an inwardly-rectifying potassium channel protein that plays a vital role in cellular ion homeostasis.[62] There is little information in the literature about the potential role of *KCNJ1* in cancer. However, an analysis examining *KCNJ1* expression in clear cell renal cell carcinoma (ccRCC) found that expression was significantly associated with tumor pathology grade and clinical stage.[63] KCNJ1 gene knockout enhanced cancer cell proliferation, while the re-expression of KCNJ1 in ccRCC cell lines resulted in increased

cancer cell apoptosis and inhibited cancer cell growth. The influence of *KCNJ1* distinctive to EOC cells is unknown, but it is possible that it could function similarly to what has been observed in ccRCC cells. If so, then higher proportions of LWK ancestry may play a protective role against EOC growth and metastasis.

Upregulated with higher proportions of MSL ancestry, *RAB11B* plays a crucial role in intracellular membrane trafficking through the activation of proteins from their inactive GDF-bound state to their active GTP-bound state.[64] *RAB11B* controls the expression and recycling of proteins back to the cell membrane surface, including integrin β1.[65] Integrin β1 is a transmembrane protein that facilitates cell-cell adhesion and cell-ECM interactions. The protein has previously been implicated in cancer progression due to its involvement in cell adhesion, migration, proliferation, and chemotherapy resistance.[66] In ovarian cancer specifically, integrin β1 was found to be frequently upregulated, enhancing ovarian tumorigenesis.[67] More details regarding the mechanistic relationship between *RAB11B* and integrin β1 expression on the cell surface are needed to meaningfully interpret the upregulation of *RAB11B* in relation to MSL ancestry.

TPH2, which is upregulated for greater proportions of MSL ancestry, is primarily expressed in the brain due to its importance in serotonin biosynthesis.[68] This gene has no other known functions and appears unrelated to cellular processes involved in cancer progression, making it difficult to draw conclusions about its upregulation.

CCDC42 expression increases with higher proportions of MSL ancestry. There is very little information concerning the function of *CCDC42* in ovarian cells, and several studies highlight the gene's importance in male fertility and sperm development.[69] *CCDC42* is also predicted to influence the centrosome cycle and cilium assembly. More concrete

information concerning the gene's function is required to evaluate its potential role in EOC cells and cancer progression.

GDAP2 is downregulated relative to higher proportions of YRI ancestry. Although the vast majority of research on this gene investigates its role in neuronal differentiation and development, it may also be implicated in cellular stress responses. The knockdown of *GDAP2* in flies resulted in increased susceptibility to oxidative stress and nutrient deprivation.[70] The gene has no known relationship to cancer development or progression, and its role specific to EOC is unknown. Further research into *GDAP2* function in non-neuronal cells is needed to draw any conclusions about its downregulation with YRI ancestry.

The gene *RXFP1*, which is downregulated with higher proportions of YRI ancestry, encodes the receptor of the hormone relaxin.[71] Ovarian cancer cells are characterized by low levels of *RXFP1* expression, which aligns with our results. Complete inhibition of *RXFP1* or relaxin has been demonstrated to increase cisplatin sensitivity of OC cell lines and suppress in vivo tumor formation.[72] Another analysis found that the relaxin-2/RXFP signaling pathway may contribute to ovarian cancer progression by modifying biological properties of ovarian cancer cells. For instance, lower expression of *RXFP1* reduces ovarian cancer cell and ECM adhesion and increases their inclination towards apoptosis.[73] Conversely, increased expression of *RXFP1* enhances cell-ECM adhesion and improves tumor cell migration capacity. With this logic, very low levels or complete knockout of *RXFP1* may contribute to reduced metastasis and improved prognoses among EOC patients. In our results, expression of *RXFP1* is decreased relative to higher YRI proportions, but it is unclear if expression is low enough to have protective effects.

The gene *PPP2R2C* is downregulated relative to higher proportions of YRI ancestry. Regulating cell division and growth, *PPP2R2C* is considered a tumor suppressor in the brain, as its downregulation is associated with greater cell proliferation of glioma cells.[74] The gene's downregulation in relation to YRI ancestry may lead to greater ovarian tumor growth and worse prognoses, although the gene's roles in tissues apart from the brain are understudied.

The gene *LTBP1* is downregulated relative to higher proportions of YRI ancestry. *LTBP1* regulates the activation of transforming growth factor-beta (TGF-β) proteins, impacting numerous cellular processes relevant to cancer.[75] TGF-β signaling typically suppresses tumorigenesis and promotes differentiation of normal cells in early stage cancers.[76] In advanced cancer stages, which pertains to the majority of cases in this analysis, TGF-β signaling is frequently modified to promote tumor progression and metastasis. *LTBP1* gene expression may be related to TGF-β overactivation, inducing epithelial-mesenchymal transition (EMT) to optimize cancer cell migration and ECM remodeling to facilitate growth and metastasis. A study on cervical cancer cells found that *LTBP1* knockdown resulted in increased cell proliferation and heightened in vitro migration and in vivo metastasis.[77] Additionally, researchers found that the quantity of myeloid-derived suppressive cells in vitro increased and T cells decreased. Although *LTBP1* expression is not well researched in terms of ovarian cancer cells, decreased expression may be associated with worse outcomes and metastasis for EOC as well.

Interestingly, *GDF10*, which is also downregulated with higher proportions of YRI ancestry, encodes a secreted ligand belonging to the transforming growth factor-beta superfamily.[78] An analysis exploring the underexpression and overexpression of *GDF10* in

TNBC tumors found the gene to be downregulated in tumors compared to matched controls, promoting cell proliferation and invasion. Alternatively, *GDF10* overexpression inhibited proliferation, invasion, and EMT. Additional findings indicated that *GDF10* may act as a tumor suppressor gene in epithelial cells, as its overexpression reduced tumor progression and enhanced apoptosis in a TNBC xenograft mouse model. If *GDF10* expression functions similarly in EOC cells as in TNBC cells, then downregulation in YRI ancestry would correlate with increased tumor invasiveness and worse outcomes. Although the previously discussed gene, *LTBP1*, primarily regulates TGF-β isoforms, *LTBP*-mediated regulation of TGF-β signaling indirectly affects *GDF10*, as it belongs to the transforming growth factor beta superfamily. Both *LTBP1* and *GDF10* are downregulated in correlation to higher proportions of YRI ancestry, so further research into their pathways is necessary to understand their impact on EOC progression.

Additionally, *GDF10* was significantly differentially expressed by BMI stratification, downregulated in participants with BMI of 40 kg/m^2 or greater compared to those with a BMI less than 30 $kg/m^{2.}$ The gene plays a crucial role in inhibiting adipogenesis through the expression of key transcriptional factors. A prospective case-control study examining the role of *GDF10* in childhood obesity found that children with an increased BMI displayed significantly lower serum concentrations of *GDF10* compared to children with normal BMI.[79] Although our analysis consists of adults exclusively, the findings align with those of this study.

Trophinin, the protein product of the *TRO* gene, facilitates cell adhesion between epithelial endometrial cells and trophoblastic cells, during the initial process of embryo implantation.[80] Additionally, *TRO* is one of the genes that is most downregulated through

RAS pathway induction, a pathway promoting cell growth and differentiation which is frequently enhanced in ovarian cancer.[81] The knockdown of *TRO* leads to cisplatin resistance, a common chemotherapy drug used to treat ovarian cancer and greater invasiveness of ovarian cancer cells. Because of this, overexpression of *TRO* would be linked to improved prognoses. *TRO* was found to be upregulated in association with GWD ancestry, suggesting increased protection against ovarian cancer invasiveness.

In another overexpressed GWD gene, *MAGEH1*, its protein product is linked with cell cycle arrest, apoptosis, and growth inhibition.[82] In an analysis examining *MAGEH1* expression, researchers found that the gene was downregulated in hepatocellular carcinoma (HCC) tumor tissues compared to normal liver tissue, with upregulation resulting in reduced HCC cell proliferation and migration.[83] Also in this analysis, low MAGEH1 expression was significantly associated with poor HCC prognosis. These findings suggest that the upregulation of MAGEH1 for GWD ancestry may have protective effects for EOC patients. However, it is unknown if the gene functions similarly across different types of cancers.

The MAGE family genes TRO and MAGEH1, both of which are upregulated relative to higher proportions of GWD ancestry, are located in close proximity on the X chromosome. *MAGEH1* is located at Xp11.21 and TRO is located at the bands Xp11.21-22, suggesting that *TRO* and *MAGEH1* may be linked genes and jointly impact ovarian cancer progression.[84, 85]

COL25A1 is upregulated relative to higher proportions of GWD ancestry. The gene encodes a neuron-specific membrane-associated collagen and plays a vital role in Alzheimer's disease pathogenesis.[86] The literature available on the gene's function

excluding neuronal expression is limited; however the collagen encoded by *COL25A1* has recently been proposed to modify the ECM of various tissues aside from neurons, as ECM is primarily composed of collagen.[87] Modifications to the ECM may influence cancer cells' capacity for invasion, adhesion, and migration, ultimately creating a more conducive environment for tumor progression. Notably, *COL25A1* has been found to be dysregulated and expressed in several cancers. It would be meaningful to further explore the specific implications of *COL25A1* expression in cancerous tissues, as many of the genes discussed have been implicated in ECM regulation and cancer progression.

HMGB1, belonging to the High Mobility Group-Box superfamily, is upregulated with higher proportions of GWD ancestry. It regulates numerous cellular processes, including cell migration, inflammatory response, and cell differentiation.[88] Its pathways are complex and poorly understood, considering the gene has been demonstrated to act both as an oncogene and tumor suppressor under various circumstances.[89] For instance, extracellular *HMGB1* promotes the secretion of cytokines including IL-6 and IL-8, which subsequently stimulate tumor cell invasion, metastasis, angiogenesis, and EMT. However, *HMGB1* expression in the cell nucleus assists in the regulation of telomeres and the maintenance of DNA repair, so the loss of *HMGB1* leads to genome instability and amplified tumorigenesis. It is difficult to draw conclusions about the role of *HMGB1* expression in EOC progression due to its intricate role as an oncogene and a tumor suppressor gene.

HCLS1, commonly referred to as *HS1*, is upregulated relative to higher proportions of GWD ancestry. *HS1* has been shown to be overexpressed in several cancers, with its elevated expression associated with enhanced cell migration, metastasis, and poor

prognoses.[90] In a study examining *HS1* expression, Kaplan–Meier analysis revealed that higher HS1 expression was significantly linked to shorter OS among patients with stage II–IV ovarian cancer, while no statistically significant association was observed for stage I ovarian cancer cases. Furthermore, the downregulation of *HS1* suppressed the migration and invasion of ovarian cancer cells, whereas overexpression of exogenous HS1 enhanced these aggressive processes. These findings suggest that the upregulation of HS1 may contribute to worse outcomes, potentially placing EOC patients with higher GWD ancestry at greater risk.

CEP164 is upregulated relative to higher proportions of ESN ancestry, influencing microtubule organization and DNA damage response. *CEP164* downregulation has been found to be associated with the deletion of primary cilia and the proliferation of pancreatic cancer cells.[91] However, the literature concerning the characterization of *CEP164* and its potential impact on cancer progression is very limited.

TMEM106 is downregulated relative to higher proportions of ESN ancestry. The gene encodes a transmembrane protein contributing to the physiological regulation of the lysosome. Changes in *TMEM106B* expression, whether decreased or increased, can lead to lysosome abnormalities.[92] Further, *TMEM106B* has been identified as a robust driver of lung cancer metastasis, with its overexpression linked to poor OS and progression-free survival. However, the gene's role in lung cancer metastasis has not been extensively studied or implicated in any other types of cancer. More research is required to corroborate its role in lung cancer progression and clarify the gene's role specific to EOC cells.

Additionally, we found that *TMEM106*B expression is significantly downregulated among participants with a BMI of 40 kg/m^2 or greater compared to those with a BMI of less than 30 kg/m^2. An analysis of *TMEM106B* in mouse models found that *TMEM106B* gene knockout resulted in significant body weight gain and increased fat deposition.[93] This finding aligns with the pattern of *TMEM106B* expression found in our BMI stratification. However, the mechanisms responsible for this trend remain unclear.

SNCA is downregulated relative to higher proportions of ESN ancestry. The encoded protein product, alpha-synuclein, is primarily known for its regulation of synaptic activity and role in Parkinson's Disease pathogenesis. Some literature suggests *SNCA* could influence tumorigenesis, however, it remains uncertain. One analysis on lung adenocarcinoma (LUAD) found that *SNCA* expression was lower in LUAD tissue compared to normal tissue, and higher expression in LUAD tissue was associated with improved prognosis.[94] Additionally, the researchers discovered statistically significant positive associations between *SNCA* expression and immune infiltrations, including CD8+ T cells, CD4+ T cells, B cells, neutrophils, and macrophages. Enrichment analysis confirmed the gene's involvement in the PI3K-AKT signaling pathway and other key tumor signaling pathways. *SNCA* may function similarly in EOC cells, and if so, downregulation relative to ESN ancestry may reduce the count of infiltrating immune cells and potentially worsen prognosis.

In this analysis, several pseudogenes were found to be differentially expressed by ancestry, including *IGHGP*, *HEBP2P1*, *RPL21P32*, *NPM1P26*, *UPF3AP2*, and *G3BP1P1*. Historically, pseudogenes were thought to be inconsequential, as they do not encode proteins. As a result, information on their function and influence on cellular processes

remains limited. More recent research suggests that pseudogenes may modulate gene expression and potentially contribute to the pathogenesis of cancer. Ancestry-differentiated pseudogene expression is highly prevalent in our results, with six of the resulting genes being implicated, suggesting a possible connection between the pseudogenes' function and involvement in EOC.

Strengths + Limitations

This analysis has a few limitations that should be considered when interpreting the findings. First, the study population was relatively small, which may limit the statistical power of the results. A larger sample size would provide greater confidence in the observed associations and improve the ability to detect subtle differences in ancestry-associated differential gene expression.

Second, potential confounding factors may influence the results. While efforts were made to account for known confounders, unmeasured variables may have influenced the results and affected the strength of the associations observed.

Despite these limitations, this study offers valuable insights into ancestry-associated gene expression patterns among Black women with EOC, including the identification of differentially expressed genes that may contribute to tumor metastasis and poor prognoses.

A key strength of this exploratory analysis is its novel focus. This is one of the first studies to investigate the relationship between genetic ancestry and EOC tumor gene expression in Black women. These findings may help advance our understanding of survival disparities and disease progression in this population.

Summary + Conclusion

Overall, the top five genes identified for each ancestry demonstrated distinct biological implications. Many of the differentially expressed genes associated with African ancestry or its subpopulations appear to promote cancer progression. Several of these genes play key roles in processes such as cell adhesion and migration, immune evasion, and enhanced cell proliferation, mechanisms that are likely to contribute to EOC metastasis and poor clinical outcomes. However, it remains challenging to definitively determine the involvement of these genes in EOC, as many are implicated in multiple cellular pathways and exhibit diverse effects across different cancer types.

Alternatively, some differentially expressed genes appear to have limited or no known association with cancer progression or EOC specifically. Several factors may account for this. For some genes, particularly pseudogenes, there is a general lack of functional information in the existing literature. In other cases, the known function of a gene is primarily characterized within a specific tissue type unrelated to EOC, such as neurons, with little evidence available regarding its role in other contexts.

The genes identified as significantly differentially expressed by BMI categories or type 2 diabetes diagnosis included *FOXD4L5*, *TMEM106B*, *GDF10*, and *G3BP1P1*. The amount of data available on these genes widely varies. For example, *FOXD4L5*, which was differentially expressed by both BMI and type 2 diabetes status, has limited functional characterization in the literature, offering little insight into its potential role in obesity or diabetes. Similarly, *G3BP1P1*, a pseudogene, is also poorly understood. In contrast, previous studies have associated *TMEM106B* and *GDF10,* both of which were found to be downregulated with increasing BMI, with obesity-related mechanisms.

The results of this analysis include several pseudogenes, under-researched genes, and genes involved with a wide variety of cellular processes. To better comprehend the implications of these results, further research is necessary to investigate functions of key genes that currently lack sufficient characterization. Strengthening our knowledge in this area will enhance our understanding of the differentially expressed pathways and the mechanisms shaping the EOC tumor microenvironment. For genes with more compelling links to cancer progression, future gene knockout studies in EOC cell models may help clarify how gene expression influences the tumor microenvironment and potentially drives metastasis.

References:

1. Cancer of the Ovary - Cancer Stat Facts. SEER. Accessed August 28, 2024. https://seer.cancer.gov/statfacts/html/ovary.html

2. Wei W, Li N, Sun Y, Li B, Xu L, Wu L. Clinical outcome and prognostic factors of patients with early-stage epithelial ovarian cancer. *Oncotarget*. 2016;8(14):23862-23870. doi:10.18632/oncotarget.13317

3. Liberto JM, Chen SY, Shih IM, Wang TH, Wang TL, Pisanic TR. Current and Emerging Methods for Ovarian Cancer Screening and Diagnostics: A Comprehensive Review. *Cancers*. 2022;14(12):2885. doi:10.3390/cancers14122885

4. Gaitskell K, Hermon C, Barnes I, et al. Ovarian cancer survival by stage, histotype, and pre-diagnostic lifestyle factors, in the prospective UK Million Women Study. *Cancer Epidemiol*. 2022;76:102074. doi:10.1016/j.canep.2021.102074

5. Epithelial ovarian cancer: Review article - ScienceDirect. Accessed November 7, 2024. https://www.sciencedirect.com/science/article/pii/S2468294222001204

6. du Bois A, Reuss A, Pujade-Lauraine E, Harter P, Ray-Coquard I, Pfisterer J. Role of surgical outcome as prognostic factor in advanced epithelial ovarian cancer: a combined exploratory analysis of 3 prospectively randomized phase 3 multicenter trials: by the Arbeitsgemeinschaft Gynaekologische Onkologie Studiengruppe Ovarialkarzinom (AGO-OVAR) and the Groupe d'Investigateurs Nationaux Pour les Etudes des Cancers de l'Ovaire (GINECO). *Cancer*. 2009;115(6):1234-1244. doi:10.1002/cncr.24149

7. Garrido MP, Fredes AN, Lobos-González L, Valenzuela-Valderrama M, Vera DB, Romero C. Current Treatments and New Possible Complementary Therapies for Epithelial Ovarian Cancer. *Biomedicines*. 2021;10(1):77. doi:10.3390/biomedicines10010077

8. Elattar A, Bryant A, Winter-Roach BA, Hatem M, Naik R. Optimal primary surgical treatment for advanced epithelial ovarian cancer. *Cochrane Database Syst Rev*. 2011;2011(8):CD007565. doi:10.1002/14651858.CD007565.pub2

9. Ataseven B, Grimm C, Harter P, et al. Prognostic impact of debulking surgery and residual tumor in patients with epithelial ovarian cancer FIGO stage IV. *Gynecol Oncol*. 2016;140(2):215-220. doi:10.1016/j.ygyno.2015.12.007

10. Schildkraut JM, Iversen ES, Akushevich L, et al. Molecular signatures of epithelial ovarian cancer: analysis of associations with tumor characteristics and epidemiologic risk factors. *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol*. 2013;22(10):1709-1721. doi:10.1158/1055-9965.EPI-13-0192

11. Lisio MA, Fu L, Goyeneche A, Gao Z hua, Telleria C. High-Grade Serous Ovarian Cancer: Basic Sciences, Clinical and Therapeutic Standpoints. *Int J Mol Sci*. 2019;20(4):952. doi:10.3390/ijms20040952

12. Chui MH, Song Q, Zhu J, et al. Early genetic divergence of high-grade carcinomas originating from low-grade ovarian serous neoplasms. *Mod Pathol*. Published online October 9, 2024:100629. doi:10.1016/j.modpat.2024.100629

13. Tothill RW, Tinker AV, George J, et al. Novel Molecular Subtypes of Serous and Endometrioid Ovarian Cancer Linked to Clinical Outcome. *Clin Cancer Res*. 2008;14(16):5198-5208. doi:10.1158/1078-0432.CCR-08-0196

14. Peres LC, Risch H, Terry KL, et al. Racial/ethnic differences in the epidemiology of ovarian cancer: a pooled analysis of 12 case-control studies. *Int J Epidemiol*. 2018;47(2):460-472. doi:10.1093/ije/dyx252

15. Stenzel AE, Buas MF, Moysich KB. Survival disparities among racial/ethnic groups of women with ovarian cancer: An update on data from the Surveillance, Epidemiology and End Results (SEER) registry. *Cancer Epidemiol*. 2019;62:101580. doi:10.1016/j.canep.2019.101580

16. Cancer Facts & Figures 2024. Published online 1930.

17. Pierson C, Mooney SJ, Lawson AB, Schildkraut JM, Peters ES. Abstract B089: Understanding the relationship between neighborhood disorder, racial segregation, and ovarian cancer survival in African American women. *Cancer Epidemiol Biomarkers Prev*. 2023;32(12_Supplement):B089. doi:10.1158/1538-7755.DISP23-B089

18. Kaufman M, Cruz A, Thompson J, et al. A review of the effects of healthcare disparities on the experience and survival of ovarian cancer patients of different racial and ethnic backgrounds. *J Cancer Metastasis Treat*. 2019;5:13. doi:10.20517/2394-4722.2018.25

19. Peres LC, Bethea TN, Camacho TF, et al. Racial Differences in Population Attributable Risk for Epithelial Ovarian Cancer in the OCWAA Consortium. *JNCI J Natl Cancer Inst*. 2020;113(6):710-718. doi:10.1093/jnci/djaa188

20. Ochs-Balcom HM, Johnson C, Guertin KA, et al. Racial differences in the association of body mass index and ovarian cancer risk in the OCWAA Consortium. *Br J Cancer*. 2022;127(11):1983-1990. doi:10.1038/s41416-022-01981-6

21. Schildkraut JM. Invited Commentary: Relationship Between Ovulation and Markers of Systemic Inflammation Versus Markers of Localized Inflammation. *Am J Epidemiol*. 2020;189(7):671-673. doi:10.1093/aje/kwz265

22. Association of inflammation-related exposures and ovarian cancer survival in a multi-site cohort study of Black women - PubMed. Accessed October 14, 2024. https://pubmed.ncbi.nlm.nih.gov/37537254/

23. Factors contributing to higher incidence of diabetes for black Americans. National Institutes of Health (NIH). January 8, 2018. Accessed April 9, 2025. https://www.nih.gov/news-events/nih-research-matters/factors-contributing-higher-incidence-diabetes-black-americans

24. Akhavan S, Ghahghaei-Nezamabadi A, Modaresgilani M, et al. Impact of diabetes mellitus on epithelial ovarian cancer survival. *BMC Cancer*. 2018;18(1):1246. doi:10.1186/s12885-018-5162-3

25. Manichaikul A, Peres LC, Wang XQ, et al. Identification of novel epithelial ovarian cancer loci in women of African ancestry. *Int J Cancer*. 2020;146(11):2987-2998. doi:10.1002/ijc.32653

26. Martini R, Delpe P, Chu TR, et al. African Ancestry–Associated Gene Expression Profiles in Triple-Negative Breast Cancer Underlie Altered Tumor Biology and Clinical Outcome in Women of African Descent. *Cancer Discov*. 2022;12(11):2530-2551. doi:10.1158/2159-8290.CD-22-0138

27. Davis M, Tripathi S, Hughley R, et al. AR negative triple negative or “quadruple negative” breast cancers in African American women have an enriched basal and immune signature. *PloS One*. 2018;13(6):e0196909. doi:10.1371/journal.pone.0196909

28. Liu H, Yang Z, Lu W, et al. Chemokines and chemokine receptors: A new strategy for breast cancer therapy. *Cancer Med*. 2020;9(11):3786. doi:10.1002/cam4.3014

29. Jenkins BD, Martini RN, Hire R, et al. Atypical Chemokine Receptor 1 (DARC/ACKR1) in Breast Tumors Is Associated with Survival, Circulating Chemokines, Tumor-Infiltrating Immune Cells, and African Ancestry. *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol*. 2019;28(4):690-700. doi:10.1158/1055-9965.EPI-18-0955

30. Luo H, Chaudhuri A, Johnson KR, et al. Cloning, Characterization, and Mapping of a Murine Promiscuous Chemokine Receptor Gene: Homolog of the Human Duffy Gene. *Genome Res*. 1997;7(9):932-941.

31. Zlotnik A, Yoshie O, Nomiyama H. The chemokine and chemokine receptor superfamilies and their molecular evolution. *Genome Biol*. 2006;7(12):243. doi:10.1186/gb-2006-7-12-243

32. Schildkraut J. Abstract IA42: Risk factors and ovarian cancer in African American women: Contributors to disparities. *Cancer Epidemiol Biomarkers Prev*. 2015;24(10_Supplement):IA42. doi:10.1158/1538-7755.DISP14-IA42

33. Wang J, Dean DC, Hornicek FJ, Shi H, Duan Z. RNA sequencing (RNA-Seq) and its application in ovarian cancer. *Gynecol Oncol*. 2019;152(1):194-201. doi:10.1016/j.ygyno.2018.10.002

34. Bell D, Berchuck A, Birrer M, et al. Integrated genomic analyses of ovarian carcinoma. *Nature*. 2011;474(7353):609-615. doi:10.1038/nature10166

35. Davidson NR, Barnard ME, Hippen AA, et al. Molecular subtypes of high-grade serous ovarian cancer across racial groups and gene expression platforms. *BioRxiv Prepr Serv Biol*. Published online December 2, 2023:2023.11.01.565179. doi:10.1101/2023.11.01.565179

36. Maples BK, Gravel S, Kenny EE, Bustamante CD. RFMix: A Discriminative Modeling Approach for Rapid and Robust Local-Ancestry Inference. *Am J Hum Genet*. 2013;93(2):278-288. doi:10.1016/j.ajhg.2013.06.020

37. Shah MM, Erickson BK, Matin T, et al. Diabetes mellitus and ovarian cancer: More complex than just increasing risk. *Gynecol Oncol*. 2014;135(2):273-277. doi:10.1016/j.ygyno.2014.09.004

38. CDC. National Diabetes Statistics Report. Diabetes. July 23, 2024. Accessed April 10, 2025. https://www.cdc.gov/diabetes/php/data-research/index.html

39. Haber G, Ahmed NU, Pekovic V. Family History of Cancer and Its Association With Breast Cancer Risk Perception and Repeat Mammography. *Am J Public Health*. 2012;102(12):2322-2329. doi:10.2105/AJPH.2012.300786

40. Farland LV, Horne AW. Disparity in endometriosis diagnoses between racial/ethnic groups. *Bjog*. 2019;126(9):1115. doi:10.1111/1471-0528.15805

41. Products - Data Briefs - Number 327 - December 2018. October 31, 2022. Accessed April 24, 2025. https://www.cdc.gov/nchs/products/databriefs/db327.htm

42. Deb S, Sun J. Endosomal Sorting Protein SNX27 and Its Emerging Roles in Human Cancers. *Cancers*. 2022;15(1):70. doi:10.3390/cancers15010070

43. Human Gene TRBV29-1 (ENST00000634198.1) from GENCODE V47. Accessed April 18, 2025. https://www.genome.ucsc.edu/cgi-bin/hgGene?hgg_gene=ENST00000634198.1&hgg_chrom=chr7_KI270803v1_alt&hgg_start=742798&hgg_end=743487&hgg_type=knownGene&db=hg38

44. Rodriguez-Rodriguez C, González-Mancha N, Ochoa-Echeverría A, Mérida I. Sorting nexin 27–dependent regulation of Lck and CD4 tunes the initial stages of T-cell activation. Accessed April 18, 2025. https://dx.doi.org/10.1093/jleuko/qiae086

45. VPS45 vacuolar protein sorting 45 homolog [Homo sapiens (human)] - Gene - NCBI. Accessed April 18, 2025. https://www.ncbi.nlm.nih.gov/gene/11311

46. Spaeth-Cook D, Burch M, Belton R, et al. Loss of TXNIP enhances peritoneal metastasis and can be abrogated by dual TORC1/2 inhibition. *Oncotarget*. 2018;9(86):35676-35686. doi:10.18632/oncotarget.26281

47. Kim S, Ge J, Kim D, et al. TXNIP-mediated crosstalk between oxidative stress and glucose metabolism. *PloS One*. 2024;19(2):e0292655. doi:10.1371/journal.pone.0292655

48. Ben Ali F, Qmichou Z, Oukabli M, et al. Alteration of glucose metabolism and expression of glucose transporters in ovarian cancer. *Explor Target Anti-Tumor Ther*. 2024;5(2):384-399. doi:10.37349/etat.2024.00224

49. Yaman E, Gasper R, Koerner C, Wittinghofer A, Tazebay UH. RasGEF1A and RasGEF1B are guanine nucleotide exchange factors that discriminate between Rap GTP-binding proteins and mediate Rap2-specific nucleotide exchange. *FEBS J*. 2009;276(16):4607-4616. doi:10.1111/j.1742-4658.2009.07166.x

50. Meng Z, Qiu Y, Lin KC, et al. RAP2 mediates mechanoresponses of the Hippo pathway. *Nature*. 2018;560(7720):655-660. doi:10.1038/s41586-018-0444-0

51. Di J, Huang H, Qu D, et al. Rap2B promotes proliferation, migration and invasion of human breast cancer through calcium-related ERK1/2 signaling pathway. *Sci Rep*. 2015;5(1):12363. doi:10.1038/srep12363

52. Liu B, Li H, Zhou J, et al. LINC00943 regulates miR-1252-5p/YWHAH axis to promote tumor proliferation and metastasis in lung adenocarcinoma. *Heliyon*. 2023;9(6):e16736. doi:10.1016/j.heliyon.2023.e16736

53. Rezabkova L, Kraatz SHW, Akhmanova A, Steinmetz MO, Kammerer RA. Biophysical and Structural Characterization of the Centriolar Protein Cep104 Interaction Network. *J Biol Chem*. 2016;291(35):18496-18504. doi:10.1074/jbc.M116.739771

54. Rezabkova L, Kraatz SHW, Akhmanova A, Steinmetz MO, Kammerer RA. Biophysical and Structural Characterization of the Centriolar Protein Cep104 Interaction Network. *J Biol Chem*. 2016;291(35):18496-18504. doi:10.1074/jbc.M116.739771

55. Richardson MT, Recouvreux MS, Karlan BY, Walts AE, Orsulic S. Ciliated Cells in Ovarian Cancer Decrease with Increasing Tumor Grade and Disease Progression. *Cells*. 2022;11(24):4009. doi:10.3390/cells11244009

56. FUT2 fucosyltransferase 2 (H blood group) [Homo sapiens (human)] - Gene - NCBI. Accessed April 18, 2025. https://www.ncbi.nlm.nih.gov/gene/2524

57. Sigal DS, Hermel ,David J, Hsu ,Pei, and Pearce T. The role of Globo H and SSEA-4 in the Development and Progression of Cancer, and their Potential as Therapeutic Targets. *Future Oncol*. 2022;18(1):117-134. doi:10.2217/fon-2021-1110

58. Pochechueva T, Alam S, Schötzau A, et al. Naturally occurring anti-glycan antibodies binding to Globo H-expressing cells identify ovarian cancer patients. *J Ovarian Res*. 2017;10:8. doi:10.1186/s13048-017-0305-8

59. Human Gene FOXD4L4 (ENST00000377413.3) from GENCODE V47. Accessed April 18, 2025. https://genome.ucsc.edu/cgi-bin/hgGene?db=hg38&hgg_chrom=chr9&hgg_gene=ENST00000377413.2&hgg_start=65737145&hgg_end=65738395&hgg_type=knownGene

60. Laissue P. The forkhead-box family of transcription factors: key molecular players in colorectal cancer pathogenesis. *Mol Cancer*. 2019;18(1):5. doi:10.1186/s12943-019-0938-x

61. Jackson BC, Carpenter C, Nebert DW, Vasiliou V. Update of human and mouse forkhead box (FOX) gene families. *Hum Genomics*. 2010;4(5):345. doi:10.1186/1479-7364-4-5-345

62. KCNJ1 potassium inwardly rectifying channel subfamily J member 1 [Homo sapiens (human)] - Gene - NCBI. Accessed April 19, 2025. https://www.ncbi.nlm.nih.gov/gene/3758

63. Guo Z, Liu J, Zhang L, et al. KCNJ1 inhibits tumor proliferation and metastasis and is a prognostic factor in clear cell renal cell carcinoma. *Tumour Biol J Int Soc Oncodevelopmental Biol Med*. 2015;36(2):1251-1259. doi:10.1007/s13277-014-2746-7

64. RAB11B - Ras-related protein Rab-11B - Homo sapiens (Human) | UniProtKB | UniProt. Accessed April 19, 2025. https://www.uniprot.org/uniprotkb/Q15907/entry

65. Howe EN, Burnette MD, Justice ME, et al. Rab11b-mediated integrin recycling promotes brain metastatic adaptation and outgrowth. *Nat Commun*. 2020;11:3017. doi:10.1038/s41467-020-16832-2

66. Desgrosellier JS, Cheresh DA. Integrins in cancer: biological implications and therapeutic opportunities. *Nat Rev Cancer*. 2010;10(1):9-22. doi:10.1038/nrc2748

67. Yang Z, Zhou X, Liu Y, et al. Activation of integrin β1 mediates the increased malignant potential of ovarian cancer cells exerted by inflammatory cytokines. *Anticancer Agents Med Chem*. 2014;14(7):955-962. doi:10.2174/1871520614666140613123108

68. Tph2 tryptophan hydroxylase 2 [Mus musculus (house mouse)] - Gene - NCBI. Accessed April 19, 2025. https://www.ncbi.nlm.nih.gov/gene/216343

69. CCDC42 coiled-coil domain containing 42 [Homo sapiens (human)] - Gene - NCBI. Accessed April 19, 2025. https://www.ncbi.nlm.nih.gov/gene/146849

70. Eidhof I, Baets J, Kamsteeg EJ, et al. GDAP2 mutations implicate susceptibility to cellular stress in a new form of cerebellar ataxia. *Brain J Neurol*. 2018;141(9):2592-2604. doi:10.1093/brain/awy198

71. RXFP1 relaxin family peptide receptor 1 [Homo sapiens (human)] - Gene - NCBI. Accessed April 20, 2025. https://www.ncbi.nlm.nih.gov/gene/59350

72. Burston HE, Kent OA, Communal L, et al. Inhibition of relaxin autocrine signaling confers therapeutic vulnerability in ovarian cancer. *J Clin Invest*. 131(7):e142677. doi:10.1172/JCI142677

73. Domińska K, Urbanek KA, Kowalska K, et al. The consequences of manipulating relaxin family peptide receptor 1 (RXFP1) level in ovarian cancer cells. *Reprod Biol*. 2024;24(2):100864. doi:10.1016/j.repbio.2024.100864

74. Fan Y ling, Chen L, Wang J, Yao Q, Wan J qing. Over expression of PPP2R2C inhibits human glioma cells growth through the suppression of mTOR pathway. *FEBS Lett*. 2013;587(24):3892-3897. doi:10.1016/j.febslet.2013.09.029

75. Ltbp1 latent transforming growth factor beta binding protein 1 [Mus musculus (house mouse)] - Gene - NCBI. Accessed April 20, 2025. https://www.ncbi.nlm.nih.gov/gene/268977

76. Seoane J, Gomis RR. TGF-β Family Signaling in Tumor Suppression and Cancer Progression. *Cold Spring Harb Perspect Biol*. 2017;9(12):a022277. doi:10.1101/cshperspect.a022277

77. Gu H, Wang W, Sun C, et al. Immune suppressive signaling regulated by latent transforming growth factor beta binding protein 1 promotes metastasis in cervical cancer. *Braz J Med Biol Res Rev Bras Pesqui Medicas E Biol*. 2023;55:e12206. doi:10.1590/1414-431X2022e12206

78. GDF10 growth differentiation factor 10 [Homo sapiens (human)] - Gene - NCBI. Accessed April 20, 2025. https://www.ncbi.nlm.nih.gov/gene/2662

79. Yousof TR, Mejia-Benitez A, Morrison KM, Austin RC. Reduced plasma GDF10 levels are positively associated with cholesterol impairment and childhood obesity. *Sci Rep*. 2024;14(1):1805. doi:10.1038/s41598-024-51635-1

80. TRO trophinin [Homo sapiens (human)] - Gene - NCBI. Accessed April 20, 2025. https://www.ncbi.nlm.nih.gov/gene/7216

81. Baba T, Mori S, Matsumura N, et al. Trophinin is a potent prognostic marker of ovarian cancer involved in platinum sensitivity. *Biochem Biophys Res Commun*. 2007;360(2):363-369. doi:10.1016/j.bbrc.2007.06.070

82. PubChem. MAGEH1 - MAGE family member H1 (human). Accessed April 27, 2025. https://pubchem.ncbi.nlm.nih.gov/gene/MAGEH1/human

83. Wang PC, Hu ZQ, Zhou SL, et al. Downregulation of MAGE family member H1 enhances hepatocellular carcinoma progression and serves as a biomarker for patient prognosis. *Future Oncol Lond Engl*. 2018;14(12):1177-1186. doi:10.2217/fon-2017-0672

84. MAGEH1 MAGE family member H1 [Homo sapiens (human)] - Gene - NCBI. Accessed April 27, 2025. https://www.ncbi.nlm.nih.gov/gene?Db=gene&Cmd=DetailsSearch&Term=28986

85. TRO trophinin [Homo sapiens (human)] - Gene - NCBI. Accessed April 27, 2025. https://www.ncbi.nlm.nih.gov/gene/7216

86. portal.gdc.cancer.gov/genes/ENSG00000188517. Accessed April 20, 2025. https://portal.gdc.cancer.gov/genes/ENSG00000188517

87. Guo KS, Brodsky AS. Tumor collagens predict genetic features and patient outcomes. *Npj Genomic Med*. 2023;8(1):1-14. doi:10.1038/s41525-023-00358-9

88. Yuan S, Liu Z, Xu Z, Liu J, Zhang J. High mobility group box 1 (HMGB1): a pivotal regulator of hematopoietic malignancies. *J Hematol OncolJ Hematol Oncol*. 2020;13(1):91. doi:10.1186/s13045-020-00920-3

89. Wang S, Zhang Y. HMGB1 in inflammation and cancer. *J Hematol OncolJ Hematol Oncol*. 2020;13:116. doi:10.1186/s13045-020-00950-x

90. Koya Y, Liu W, Yamakita Y, et al. Hematopoietic lineage cell-specific protein 1 (HS1), a hidden player in migration, invasion, and tumor formation, is over-expressed in ovarian carcinoma cells. *Oncotarget*. 2018;9(66):32609-32623. doi:10.18632/oncotarget.25975

91. Kobayashi T, Tanaka K, Mashima Y, Shoda A, Tokuda M, Itoh H. CEP164 Deficiency Causes Hyperproliferation of Pancreatic Cancer Cells. *Front Cell Dev Biol*. 2020;8:587691. doi:10.3389/fcell.2020.587691

92. Feng T, Lacrampe A, Hu F. Physiological and pathological functions of TMEM106B: A gene associated with brain aging and multiple brain disorders. *Acta Neuropathol (Berl)*. 2021;141(3):327-339. doi:10.1007/s00401-020-02246-3

93. Yang C, Lee GB, Hao L, Hu F. TMEM106B deficiency leads to alterations in lipid metabolism and obesity in the TDP-43Q331K knock-in mouse model. *Commun Biol*. 2025;8(1):1-19. doi:10.1038/s42003-025-07752-2

94. Zhang X, Wu Z, Ma K. SNCA correlates with immune infiltration and serves as a prognostic biomarker in lung adenocarcinoma. *BMC Cancer*. 2022;22:406. doi:10.1186/s12885-022-09289-7