Evaluation of vitamin D biosynthesis and pathway target genes reveals \textit{UGT2A1/2} and \textit{EGFR} polymorphisms associated with epithelial ovarian cancer in African American Women

Delores J. Grant\textsuperscript{1} | Ani Manichaikul\textsuperscript{2} | Anthony J. Alberg\textsuperscript{3} | Elisa V. Bandera\textsuperscript{4} | Jill Barnholtz-Sloan\textsuperscript{5} | Melissa Bondy\textsuperscript{6} | Michele L. Cote\textsuperscript{7} | Ellen Funkhouser\textsuperscript{8} | Patricia G. Moorman\textsuperscript{9} | Lauren C. Peres\textsuperscript{2} | Edward S. Peters\textsuperscript{10} | Ann G. Schwartz\textsuperscript{7} | Paul D. Terry\textsuperscript{11} | Xin-Qun Wang\textsuperscript{12} | Temitope O. Keku\textsuperscript{13} | Cathrine Hoyo\textsuperscript{14} | Andrew Berchuck\textsuperscript{15} | Dale P. Sandler\textsuperscript{16} | Jack A. Taylor\textsuperscript{16} | Katie M. O’Brien\textsuperscript{16} | Digna R. Velez Edwards\textsuperscript{17} | Todd L. Edwards\textsuperscript{18} | Alicia Beeghly-Fadiel\textsuperscript{19} | Nicolas Wentzensen\textsuperscript{20} | Celeste Leigh Pearce\textsuperscript{21,22} | Anna H. Wu\textsuperscript{22} | Alice S. Whittemore\textsuperscript{23,24} | Valerie McGuire\textsuperscript{23} | Weiva Sieh\textsuperscript{25,26} | Joseph H. Rothstein\textsuperscript{25,26} | Francesmary Modugno\textsuperscript{27,28,29} | Roberta Ness\textsuperscript{30} | Kirsten Moysich\textsuperscript{31} | Mary Anne Rossing\textsuperscript{32,33} | Jennifer A. Doherty\textsuperscript{34} | Thomas A. Sellers\textsuperscript{35} | Jennifer B. Permuth-Way\textsuperscript{35} | Alvaro N. Monteiro\textsuperscript{35} | Douglas A. Levine\textsuperscript{36,37} | Veronica Wendy Setiawan\textsuperscript{38} | Christopher A. Haiman\textsuperscript{38} | Loic LeMarchand\textsuperscript{39} | Lynne R. Wilkens\textsuperscript{40} | Beth Y. Karlan\textsuperscript{41} | Usha Menon\textsuperscript{42} | Susan Ramus\textsuperscript{43,44} | Simon Gayther\textsuperscript{45,46} | Aleksandra Gentry-Maharaj\textsuperscript{42} | Kathryn L. Terry\textsuperscript{47,48} | Daniel W. Cramer\textsuperscript{47,48} | Ellen L. Goode\textsuperscript{49} | Melissa C. Larson\textsuperscript{50} | Scott H. Kaufmann\textsuperscript{51} | Rikki Cannioto\textsuperscript{52} | Kunle Odunsi\textsuperscript{53} | John L. Etter\textsuperscript{51} | Ruea-Yea Huang\textsuperscript{54} | Marcus Q. Bernardini\textsuperscript{55} | Alicia A. Tone\textsuperscript{55} | Taymaa May\textsuperscript{55} | Marc T. Goodman\textsuperscript{56,57} | Pamela J. Thompson\textsuperscript{56} | Michael E. Carney\textsuperscript{58} | Shelley S. Tworoger\textsuperscript{59} | Elizabeth M. Poole\textsuperscript{60} | Diether Lambrechts\textsuperscript{61,62} | Ignace Vergote\textsuperscript{63} | Adriaan Vanderstichele\textsuperscript{63} | Els Van Nieuwenhuysen\textsuperscript{63} | Hoda Anton-Culver\textsuperscript{64} | Argyrios Ziogas\textsuperscript{65} | James D. Brenton\textsuperscript{66} | Line Bjorge\textsuperscript{67,68} | Helga B. Salvesen\textsuperscript{67,68} | Lambertus A. Kiemene\textsuperscript{69} | Leon F. A. G. Massuger\textsuperscript{70} | Tanja Pejovic\textsuperscript{71,72} | Amanda Brueg\textsuperscript{71,72} | Melissa Moffitt\textsuperscript{71,72} | Linda Cook\textsuperscript{73} | Nhu D. Le\textsuperscript{74} | Angela Brooks-Wilson\textsuperscript{75,76} | Linda E. Kelemen\textsuperscript{77} | Paul D.P. Pharoah\textsuperscript{78} | Honglin Song\textsuperscript{79} | Ian Campbell\textsuperscript{80,81} | Diana Eccles\textsuperscript{82} | Anna DeFazio\textsuperscript{83,84} | Catherine J. Kennedy\textsuperscript{83,84} | Joellen M. Schildkraut\textsuperscript{12}
1Department of Biological and Biomedical Sciences, Cancer Research Program, JLC-Biomedical/Biotechnology Research Institute, North Carolina Central University, Durham, North Carolina
2Center for Public Health Genomics, University of Virginia, Charlottesville, Virginia
3Department of Epidemiology and Biostatistics, Arnold School of Public Health, University of South Carolina, Columbia, South Carolina
4Department of Population Science, Rutgers Cancer Institute of New Jersey, New Brunswick, New Jersey
5Case Comprehensive Cancer Center, Case Western Reserve University School of Medicine, Cleveland, Ohio
6Cancer Prevention and Population Sciences Program, Baylor College of Medicine, Houston, Texas
7Department of Oncology and the Karmanos Cancer Institute Population Studies and Disparities Research Program, Wayne State University School of Medicine, Detroit, Michigan
8Division of Preventive Medicine, University of Alabama at Birmingham, Birmingham, Alabama
9Department of Community and Family Medicine, Duke University Medical Center, Durham, North Carolina
10Epidemiology Program, Louisiana State University Health Sciences Center School of Public Health, New Orleans, Louisiana
11Department of Medicine, University of Tennessee Medical Center – Knoxville, Knoxville, Tennessee
12Department of Public Health Sciences, University of Virginia, Charlottesville, Virginia
13Departments of Medicine and Nutrition, Division of Gastroenterology and Hepatology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina
14Department of Biological Sciences, North Carolina State University, Raleigh, North Carolina
15Department of Obstetrics and Gynecology, Duke University Medical Center, Durham, North Carolina
16Epidemiology Branch, Division of Intramural Research, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North Carolina
17Vanderbilt Epidemiology Center, Center for Human Genetics Research, Department of Obstetrics and Gynecology, Vanderbilt University Medical Center, Nashville, Tennessee
18Division of Epidemiology, Center for Human Genetics Research, Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee
19Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Institute for Medicine and Public Health, Vanderbilt University Medical Center, Nashville, Tennessee
20Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland
21Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, Michigan
22Department of Preventive Medicine, Keck School of Medicine, University of Southern California Norris Comprehensive Cancer Center, Los Angeles, California
23Department of Health Research and Policy, Stanford University School of Medicine, Stanford, California
24Department of Biomedical Data Science, Stanford University School of Medicine, Stanford, California
25Department of Population Health Science and Policy, Icahn School of Medicine at Mount Sinai, New York, New York
26Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, New York
27Division of Gynecologic Oncology, Department of Obstetrics, Gynecology and Reproductive Sciences, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania
28Department of Epidemiology, University of Pittsburgh Graduate School of Public Health, Pittsburgh, Pennsylvania
29Ovarian Cancer Center of Excellence, Womens Cancer Research Program, Magee-Womens Research Institute and University of Pittsburgh Cancer Institute, Pittsburgh, Pennsylvania
30The University of Texas School of Public Health, Houston, Texas
31Department of Cancer Prevention and Control, Roswell Park Cancer Institute, Buffalo, New York
32Program in Epidemiology, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington
33Department of Epidemiology, University of Washington, Seattle, Washington
34Department of Population Health Sciences, Huntsman Cancer Institute, University of Utah, Salt Lake City, Utah
35Department of Cancer Epidemiology, Moffitt Cancer Center, Tampa, Florida
36Gynecology Service, Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, New York
37Gynecologic Oncology, Laura and Isaac Perlmutter Cancer Center, New York University Langone Medical Center, New York, New York
38University of Southern California Norris Comprehensive Cancer Center, Los Angeles, California
39University of Hawaii Cancer Center, Honolulu, Hawaii
40Cancer Epidemiology Program, University of Hawaii Cancer Center, Hawaii
Women's Cancer Program, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, California
MRC CTU at UCL, Institute of Clinical Trials and Methodology, University College London, London, UK
School of Women's and Children's Health, University of New South Wales, New South Wales, Australia
The Kinghorn Cancer Centre, Garvan Institute of Medical Research, Darlinghurst, New South Wales, Australia
Center for Cancer Prevention and Translational Genomics, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, California
Department of Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, California
Obstetrics and Gynecology Epidemiology Center, Brigham and Women's Hospital, Boston, Massachusetts
Harvard T. H. Chan School of Public Health, Boston, Massachusetts
Department of Health Science Research, Division of Epidemiology, Mayo Clinic, Rochester, Minnesota
Department of Health Science Research, Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, Minnesota
Departments of Medicine and Pharmacology, Mayo Clinic, Rochester, Minnesota
Cancer Pathology & Prevention, Division of Cancer Prevention and Population Sciences, Roswell Park Cancer Institute, Buffalo, New York
Department of Gynecological Oncology, Roswell Park Cancer Institute, Buffalo, New York
Center For Immunotherapy, Roswell Park Cancer Institute, Buffalo, New York
Division of Gynecologic Oncology, Princess Margaret Hospital, University Health Network, Toronto, Ontario, Canada
Cancer Prevention and Control, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, California
Department of Biomedical Sciences, Community and Population Health Research Institute, Cedars-Sinai Medical Center, Los Angeles, California
Department of Obstetrics and Gynecology, John A. Burns School of Medicine, University of Hawaii, Honolulu, Hawaii
Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts
Biostatistics, Sanofi Genzyme, Boston, Massachusetts
Vesalius Research Center, VIB, Leuven, Belgium
Laboratory for Translational Genetics, Department of Oncology, University of Leuven, Belgium
Division of Gynecologic Oncology, Department of Obstetrics and Gynaecology and Leuven Cancer Institute, University Hospitals Leuven, Leuven, Belgium
Department of Epidemiology, Director of Genetic Epidemiology Research Institute, Center for Cancer Genetics Research & Prevention, School of Medicine, University of California Irvine, Irvine, California
Department of Epidemiology, University of California Irvine, Irvine, California
Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, UK
Department of Gynecology and Obstetrics, Haukeland University Hospital, Bergen, Norway
Centre for Cancer Biomarkers, Department of Clinical Science, University of Bergen, Bergen, Norway
Radboud University Medical Center, Radboud Institute for Health Sciences, Nijmegen, Netherlands
Department of Gynaecology, Radboud University Medical Center, Radboud Institute for Molecular Life sciences, Nijmegen, The Netherlands
Department of Obstetrics & Gynaecology, Oregon Health & Science University, Portland, Oregon
Knight Cancer Institute, Oregon Health & Science University, Portland, Oregon
Division of Epidemiology and Biostatistics, Department of Internal Medicine, University of New Mexico, Albuquerque, New Mexico
Cancer Control Research, British Columbia Cancer Agency, Vancouver, British Columbia, Canada
Canada's Michael Smith Genome Sciences Centre, British Columbia Cancer Agency, Vancouver, British Columbia, Canada
Department of Biomedical Physiology and Kinesiology, Simon Fraser University, Burnaby, British Columbia, Canada
Hollings Cancer Center and Department of Public Health Sciences, Medical University of South Carolina, Charleston, South Carolina
Strangeways Research laboratory, Department of Oncology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK
Strangeways Research Laboratory, Department of Oncology, University of Cambridge, Cambridge, UK
Cancer Genetics Laboratory, Research Division, Peter MacCallum Cancer Centre, Victoria, Australia
Department of Pathology, University of Melbourne, Parkville, Victoria, Australia
Faculty of Medicine, University of Southampton, Southampton, UK
Centre for Cancer Research, The Westmead Institute for Medical Research, The University of Sydney, Sydney, New South Wales, Australia
Department of Gynaecological Oncology, Westmead Hospital, Sydney, New South Wales, Australia
Abstract
An association between genetic variants in the vitamin D receptor (VDR) gene and epithelial ovarian cancer (EOC) was previously reported in women of African ancestry (AA). We sought to examine associations between genetic variants in VDR and additional genes from vitamin D biosynthesis and pathway targets (EGFR, UGT1A, UGT2A1/2, UGT2B, CYP3A4/5, CYP2R1, CYP27B1, CYP24A1, CYP11A1, and GC).

Genotyping was performed using the custom-designed 533,631 SNP Illumina OncoArray with imputation to the 1,000 Genomes Phase 3 v5 reference set in 755 EOC cases, including 537 high-grade serous (HGSOC), and 1,235 controls. All subjects are of African ancestry (AA). Logistic regression was performed to estimate odds ratios (OR) and 95% confidence intervals (CI). We further evaluated statistical significance of selected SNPs using the Bayesian False Discovery Probability (BFDP).

A significant association with EOC was identified in the UGT2A1/2 region for the SNP rs10017134 (per allele OR = 1.4, 95% CI = 1.2-1.7, \( P = 1.2 \times 10^{-6}, \) BFDP = 0.02); and an association with HGSOC was identified in the EGFR region for the SNP rs114972508 (per allele OR = 2.3, 95% CI = 1.6-3.4, \( P = 1.6 \times 10^{-5}, \) BFDP = 0.29) and in the UGT2A1/2 region again for rs1017134 (per allele OR = 1.4, 95% CI = 1.2-1.7, \( P = 2.3 \times 10^{-5}, \) BFDP = 0.23). Genetic variants in the EGFR and UGT2A1/2 may increase susceptibility of EOC in AA women. Future studies to validate these findings are warranted. Alterations in EGFR and UGT2A1/2 could perturb enzyme efficacy, proliferation in ovaries, impact and mark susceptibility to EOC.

KEYWORDS
African ancestry risk, genetic association, ovarian cancer, vitamin D pathway
1 | INTRODUCTION

Women of African ancestry (AA) have the lowest incidence of ovarian cancer worldwide, but they tend to present with more advanced tumors and have lower 5-year survival (35%) compared to women of European descent (47%) in nearly every cancer subtype.\(^1\)\(^2\) Compared to Caucasian women, there have been fewer published studies investigating the association between common risk factors, such as tubal ligation, use of hormonal contraceptives, obesity, body powder and dietary patterns, and ovarian cancer risk in AA.\(^1\)\(^3\)\(^9\) Moreover, the investigation of genetic susceptibility to epithelial ovarian cancer (EOC) in AA has not been comprehensive. The limited assessment of genetic susceptibility among AA is in modest sized study populations of candidate genes including the repeat polymorphisms of the androgen receptor (AR), vitamin D receptor (VDR) and cellular transport genes, where an association with risk of ovarian cancer was observed.\(^10\)\(^12\)

The vitamin D receptor mediates the regulation of a pleiotropic cascade of physiological responses; including those involved in phase I and phase II detoxification and the epidermal growth factor receptor (EGFR) proliferation pathways in ovarian and other cancer cell lines; through VDR/DNA interactions and bioavailability of vitamin D.\(^13\)\(^17\) A VDR variant, rs7305032, was associated with ovarian cancer in 125 cases and 155 controls of AA but other observations were limited because of small sample size.\(^11\) Moreover, known genetic variations in the VDR/vitamin D biosynthesis and pathway target genes have been implicated in AA disease risk. Therefore an objective of this study was to assess those variants in ovarian cancer in women of African ancestry in a large sample.

Using a candidate gene approach, SNPs were selected from genes involved in vitamin D biosynthesis and metabolism; and putative targets of VDR regulation. Genes of the vitamin D biosynthesis pathways included cytochrome P450s: CYP2R1, CYP27B1, CYP24A1, CYP11A1, and group-specific component-vitamin D-binding protein (GC) which collectively are responsible for the homeostatic control and bioavailability of vitamin D.\(^18\)\(^23\) The candidate genes involved in vitamin D metabolic processes included CYP3A4/5 and UDP-glucuronosyltransferase 1A (UGT1A) locus members. They are responsible for glucuronidation and hydroxylation of the biologically active and circulatory forms of vitamin D. These genes are also inclusive of candidates regulated by vitamin D/VDR binding and included CYP3A4/5, UGT1A locus members, EGRF and UDP-glucuronosyltransferase 2 (UGT2) locus members; that are associated, in part, with other cancers in AA individuals.\(^24\)\(^39\) Thus, variants in VDR and additional genes from vitamin D biosynthesis and pathway targets are viable candidates to investigate the genetic underpinnings of ovarian cancer risk in women of African descent.

In this study, SNPs from 11 gene regions: VDR, EGRF, UGT1A, UGT2A1/2, UGT2B, CYP3A4/5, CYP2R1, CYP27B1, CYP24A1, CYP11A1, and GC, were genotyped, imputed then assessed for risk of EOC and high grade serous ovarian cancer (HGSOC) in cases and controls of AA from the African American Cancer Epidemiology Study (ACES)\(^40\) and the Ovarian Cancer Association Consortium (OCAC).\(^41\)

2 | MATERIALS AND METHODS

2.1 | Study populations

The Genetic Associations and Mechanisms in Oncology (GAME-ON) project comprised 63, mostly, case-control studies from four continents (North America, Europe, Asia and Australia). Only 32 studies contributed subjects of African Ancestry, including AACES and studies in OCAC, and were included in the current analysis (Supplemental Table S1). AACES, previously described elsewhere,\(^46\) is a multi-center population-based case-control study of newly diagnosed invasive EOC in African American women that enrolled study subjects between 2010 and 2015. Established in 2005, OCAC is an international consortium focused on genetic association and pooled risk factor analyses. The current analyses included 1990 samples: 1235 controls and 755 invasive EOC cases who passed quality control filters, all of whom were AA. The majority of the EOC cases were HGSOC (n = 537, 71%), followed by 49 mucinous cases (7%), 28 endometrioid cases (4%), 23 clear cell cases (3%), 12 mixed histology (2%) and 53 other (7%). All subjects included in this analysis provided written informed consent as well as data and blood samples under ethically approved protocols.

2.2 | Genotyping, ancestry analysis and quality control

Genotyping of AA women from OCAC was completed using the custom-designed 533,631 SNP array, the Illumina OncoArray. Sample level quality control included restriction to females, filter on call rate >95%, heterozygosity (either too big or too small), removal of ineligible samples, and relationship inference to check for unexpected first degree relatives. SNP level quality control included filter on call rate >95%, and Hardy–Weinberg Equilibrium p-value >1 x 10⁻⁵. After applying these procedures, 471,780 SNPs remained.

Intercontinental ancestry was calculated for the OCAC and AACES samples using the software package FastPop\(^42\) that was developed specifically for the OncoArray Consortium. Only the African ancestry samples defined as having >50% AA were used for the present analyses reported here. Seventy-seven cases and 120 controls were omitted due to African ancestry <50% and one gender mismatch. Principal
components computed using FastPop were further used to adjust for population structure in our analyses.

2.3 | Genotype imputation analysis

Using the genotyped SNPs that passed quality control, haplotypes were phased using SHAPEIT v2 followed by imputation to the 1,000 Genomes Phase 3 v5 reference set using Minimac3.

2.4 | Gene region and SNP selection

Eleven gene regions were defined based on human genome build 37. SNPs within the selected regions were filtered on imputation quality score (minimac imputation R-squared) >0.5 for imputed SNPs, or Hardy–Weinberg Equilibrium p-value >1.0 × 10−5 for genotyped SNPs. Quantile-quantile plots on the EOC and HGSOC dataset (Manichaikul et al, unpublished) have lambdas of 1.01 each within normal range. The imputation quality scores for significant SNPs are provided. We further applied filters on effective heterozygosity count (HC) > 30. After applying filters, the following number of SNPs remaining in each of the selected gene regions for EOC was: 288 in VDR, 433 in UGT2A1/2, 833 in UGT1A, 824 in EGFR, 15 in CYP2R1, 4 in CYP27B1, 106 in CYP24A1, 82 in CYP11A1, 375 in CYP3A4/5 and 282 in GC. For selected regions for HGSOC analysis, the number of SNPs was: 234 in VDR, 413 in UGT2A1/2, 833 in UGT1A, 824 in EGFR, 15 in CYP2R1, 4 in CYP27B1, 106 in CYP24A1, 82 in CYP11A1, 375 in CYP3A4/5 and 282 in GC.

2.5 | Statistical analysis

Genetic association testing was carried out with adjustment for two principal components (PCs) of ancestry using a logistic regression model that accounts for genotype uncertainty under a score test as implemented in SNPTEST v2.5.2 to estimate odds ratios (OR) and 95% confidence intervals (CI). For each gene region, we applied a gene-specific Bonferroni-threshold for statistical significance defined as 0.05/number of SNPs examined for that gene. We further assessed the main results with an alternative to the Bonferroni threshold using the Bayesian False Discovery Probability (BFDP) which provides the posterior probability of a false discovery based on a given prior probability of nonnull association at a given SNP. For this study we specified a prior probability of association at each SNP under investigation based on the total number of SNPs within each candidate gene region as 0.5 × 1/(N_{SNP}/3) where N_{SNP} represents the number of SNPs in the given candidate gene region. We considered N_{SNP}/3 to be an approximation of the effective number of independent SNPs within in each gene region, taking into account the fact that many SNPs will be correlated due to linkage disequilibrium. Accordingly, the specified prior indicates a 50% chance of true discovery within each gene region, with the prior probability of nonnull association distributed randomly among all SNPs within the region. In order to avoid spurious positive associations, we applied a filter on effective (HC) > 30 in each of cases and controls. Here, HC is defined as N × MAF × (1−MAF) for each SNP, N represents the sample size (either the number of cases or the number of controls), and MAF represents the SNP minor allele frequency. Based on 755 EOC cases and 537 HGSOC cases, respectively, applying this filter equates to applying a SNP MAF filter of 4.2% and 6% in analysis of EOC and HGSOC, respectively. Statistical power calculations for AA study participants and Caucasians are included in Supplemental Tables S2 and S3.

<table>
<thead>
<tr>
<th>SNP ID (Effect/other allele)</th>
<th>Nearest gene(s)</th>
<th>Effect Allele Frequency</th>
<th>N</th>
<th>OR</th>
<th>95% CI</th>
<th>P-value</th>
<th>Bayesian False Discovery Probability (BFDP)</th>
<th>Imputation quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs10017134 (C/T)</td>
<td>UGT2A1/2</td>
<td>0.73</td>
<td>1990</td>
<td>1.4</td>
<td>1.2-1.7</td>
<td>1.2 × 10^{-6}</td>
<td>0.020</td>
<td>0.998</td>
</tr>
<tr>
<td>rs2288741 (T/G)</td>
<td>UGT2A1/2</td>
<td>0.73</td>
<td>1990</td>
<td>1.4</td>
<td>1.2-1.6</td>
<td>1.9 × 10^{-6}</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>rs11939884 (T/G)</td>
<td>UGT2A1/2</td>
<td>0.14</td>
<td>1990</td>
<td>0.7</td>
<td>0.5-0.8</td>
<td>1.7 × 10^{-6}</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

\^Imputed.
\^Bonferroni correction was applied to adjust for multiple SNPs comparisons. There were 433 SNPs in UGT2A1/2 gene. BFDP is reported based on a prior probability of association (p0) equal to 0.6 × 1/(Number of SNPs/3).

3 | RESULTS

3.1 | VDR pathway gene regions and risk of EOC

SNPs from 11 gene regions (CYP3A4/5, CYP2R1, CYP27B1, CYP24A1, CYP11A1, EGFR, GC, UGT1A, UGT2A1/2, UGT2B and VDR) from VDR biosynthesis and pathway
targets were assessed for association with EOC (Supplemental Table S4). The top associations are reported in Table 1. Individuals carrying the major allele of SNP rs10017134 of the *UGT2A1/2* gene region had an increased odds of EOC when corrected for multiple comparisons (OR = 1.4, 95% CI = 1.2-1.7, \( P = 1.2 \times 10^{-5} \)). The BFDP for rs10017134 of 0.020 corresponds to 98% posterior probability of non-null association for this SNP. Significant associations with EOC were also observed for *UGT2A1/2* SNPs, rs2288741 and rs11939884. The variants are found in both *UGT2A1* and *UGT2A2* as the genes share common exons 2 through 6.45,49 Numerous splice variants of the *UGT* genes have shown using targeted RNAseq analysis that there are alternate splice sites, and cis-regulatory modules and transcription factors binding sites.46-48 In addition, recent studies have shown using targeted RNAseq analysis that there are numerous splice variants of the *UGT* genes.49

The *UGT2A1* and 2A genes are distinguished by unique first exons joined to common exons 2 through 6. Intron variants (rs10017134 and rs2288741) while the third intron variants (rs10017134 and rs2288741) while the third

### Table 2: Top SNP P-values from gene regions associated with HGSOC in African American OncoArray analysis

<table>
<thead>
<tr>
<th>SNP ID (Effect/other allele)</th>
<th>Effect Allele Frequency</th>
<th>N</th>
<th>OR</th>
<th>95% CI</th>
<th>P-value</th>
<th>Bayesian False Discovery Probability (BFDP)</th>
<th>Imputation quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs114972508 (T/C) <em>EGFR</em></td>
<td>0.04</td>
<td>1772</td>
<td>2.3</td>
<td>1.2-3.4</td>
<td>( 1.6 \times 10^{-5} )</td>
<td>0.293</td>
<td>0.890</td>
</tr>
<tr>
<td>rs10017134 (C/T) <em>UGT2A1/2</em></td>
<td>0.72</td>
<td>1772</td>
<td>1.4</td>
<td>1.2-1.7</td>
<td>( 2.3 \times 10^{-5} )</td>
<td>0.228</td>
<td>0.998</td>
</tr>
<tr>
<td>rs2288741 (T/G) <em>UGT2A1/2</em></td>
<td>0.72</td>
<td>1772</td>
<td>1.4</td>
<td>1.2-1.7</td>
<td>( 3.1 \times 10^{-5} )</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

*Bonferroni correction was applied to adjust for multiple SNPs comparisons. There were 824 SNPs in EGFR gene, and 413 SNPs in UGT2A1/2 gene. BFDP is reported based on a prior probability of association (pi0) equal to 0.5 * 1/(Number of SNPs/3).

### 3.2 | VDR pathway gene regions and risk of HGSOC

SNPs from the 10 gene regions from VDR biosynthesis and pathway targets were assessed for association with HGSOC (Supplemental Table S4). The top associations are in reported in Table 2. Individuals carrying the minor allele of *EGFR* SNP rs114972508 had more than twofold increased odds of HGSOC (OR = 2.3, 95% CI = 1.2-3.4, \( P = 1.6 \times 10^{-5} \)) (Table 2). The posterior BFDP is 29% for SNP rs114972508 corresponds to 71% posterior probability of non-null association. SNP rs10017134 of the *UGT2A1/2* gene region also showed association with HGSOC (OR = 1.4, 95% CI = 1.2-1.7, \( P = 2.3 \times 10^{-5} \)) (Table 2). The posterior BFDP is 22.8%. Supplemental Table S6 summarizes other notable (\( P < 0.01 \)) SNP associations with HGSOC in the OncoArray analysis.

### 4 | DISCUSSION

Few studies have investigated the genetic susceptibility for ovarian cancer among women of African descent. The assessment of candidate SNPs from chromosomal regions that contain genes regulated by VDR activity provides some evidence of association with EOC risk. The notable findings from this analysis show, for the first time, that risk assessments of variants in the *UGT2A1/2* and *EGFR* gene regions are suggestive of associations with EOC and HGSOC. The results also demonstrate evidence of associations for other SNPs from the candidate gene regions with EOC and HGSOC. Although the candidate SNPs are located in intronic regions there is ample evidence that many gene regulatory regions are present in those regions including encoded microRNAs, alternate splice sites, and cis-regulatory modules and transcription factors binding sites.46-48

The *UGT2A1* and 2A genes are distinguished by unique first exons joined to common exons 2-6 and are located downstream of *UGT2B4* on chromosome 4.45 *UGT2A* transcripts have been detected in several extrahepatic tissues such as the lung, trachea, larynx, intestine, pancreas, and kidney.50 *UGT2A1* is an extrahepatic enzyme that is expressed mainly in the nasal epithelium, catalyzing the glucuronidation of testosterone and epitestosterone at considerable rates and has similar kinetics as the *UGT2B* gene family members.51 There are reports that this enzyme also has activity toward estrogen metabolites epiestradiol and \( \beta \)-estradiol.52 *UGT2A1* has exhibited highest expression in the lung, followed by trachea, tonsil, larynx, colon, olfactory.53 *UGT2A2* mRNA expression was reported in fetal and adult nasal mucosa tissues.54 However, unlike UGT2A1, other expression analyses suggested that wild-type UGT2A2 had the highest expression in the breast, followed by trachea, larynx, and kidney.55

Neither the *UGT2A1* gene, nor UGT2A2 expression have been examined in ovarian tissue. However, VDR ChIPseq peak locations have been identified 430 kb downstream of the *UGT2A1/2* locus in experiments with THP-1 cells treated with 1a,25(OH)\(_2\)D\(_3\), the biologically active form of the vitamin D hormone, suggestive of a regulatory role for vitamin D.56 Splice variants found in *UGT2A1/2* that are highly conserved among both *UGT1A* and *UGT2* gene families have been implicated in altered glucuronidation activity against tobacco carcinogenesis.49,53,55,57 Two of the *UGT2A1* SNPs associated with EOC and HGSOC in this study are intron variants (rs10017134 and rs2288741) while the third
and proliferative function. It is probable that these variants alter enzyme function in target tissues including ovarian and/or alter risk in AA smokers. Of note, cigarette smoking has been found to be associated with the risk of mucinous EOC, but not HGSOC among Caucasian women. Moreover, providing some plausibility for the mechanism of the observed SNP association, a recent report suggests that cigarette smoking may be associated with serous EOC among African American women although a dose-response relationship was not observed. The association of genes from the UGT superfamily with ovarian cancer in AA is consistent with significant associations observed for Caucasian women for UGT1A. However in this study, no association was observed for AA samples with SNPs with a MAF of 0.42 for the risk allele while associations were observed in Caucasians with SNPs with a MAF of 0.07. Some but not all MAFs for the relationships observed in this study differ by race so it is unlikely to explain racial differences in risk.

The EGFR gene product has been a chemotherapeutic target for EOC since overexpression has been linked to poor prognosis in ovarian cancer patients. The signaling pathway for EGFR is mediated by ligands including the epidermal growth factor in the regulation of cell proliferation, differentiation and apoptosis in normal cells. Research into the mechanisms of EGFR overexpression has focused on mutations and amplifications in the coding region of the gene containing the receptor tyrosine kinase domain. However, few studies on SNP variants in this region have been linked to EOC or other ovarian cancer histologic subtypes. EGFR SNP rs114972508 is located in intron 1 of the EGFR gene. The location of the SNP is approximately 70 kb upstream of a VDR binding site also within EGFR intron 1 that has been shown experimentally to down regulate EGFR expression and proliferative function. Perhaps changes in the intron sequences may impact EGFR function and subsequently be as critical to cellular homeostasis as the receptor tyrosine function that has been extensively researched. Thus, EGFR SNPs could be abrogating vitamin D hormone regulation of ovarian cell proliferation and increasing susceptibility for the development of HGSOC in AA women.

Although we were unable to confirm the association between previously identified VDR variants and risk of EOC, a recent case-control study of women of European ancestry (10,065 cases, 21,654 controls) showed that SNPs associated with decreased circulating 25-hydroxyvitamin D were associated with ovarian cancer and HGSOC while another study showed that AA women exposed to increased sunlight had a decreased risk for ovarian cancer. These observations suggest that other mechanisms affecting vitamin D hormone activity independent of the VDR may be important in ovarian cancer etiology.

The main observations in the current study result from associations of imputations of genotyped SNPs but independent of VDR variant association with EOC and HGSOC. The VDR SNPs previously observed to be associated with the risk of EOC, including rs7975232 and rs7305032, were not associated with risk of EOC in the current study (Supplemental Table S7). A look up of the significant study SNPs in archived OCAC data on Caucasians shows no significant associations for the UGT2A1/2 SNPs. Data on the EGFR SNP were not available (Supplemental Table S8). Other VDR SNPs showed nominal (nonBonferroni corrected) associations with EOC but not with HGSOC (Supplemental Table S7). Although the largest study to date of genetic association with EOC in AA, the modest sample size remains a limitation of the current study and therefore some of the nominal SNP associations may be a result of inadequate power. The analyses are underpowered for discovery analysis across the selected gene regions and important associations may have been missed, nonetheless, we still found significantly associations with EOC and HGSOC. Several suggestive and nominal SNP associations (outside of Bonferroni significance) may provide some insight and consideration for future experimental studies to further explore the relevance of vitamin D biosynthesis and pathway target genes. Larger studies of AA are warranted to clarify these findings.

In summary, this study reports, for the first time, an association between EGFR and UGT2A1/2 variants with ovarian cancer risk in AA women. These gene variants could perturb cell proliferation and enzyme efficacy in ovaries and impact susceptibility to ovarian cancer by altering growth and intercellular hormone metabolism. Future studies are needed to validate the associations of the imputed SNPs and to determine their impact on cancer development. Currently, there are no published reports of population studies of UGT2A1/2 polymorphisms in Europeans or other racially distinct groups in larger sample sizes than this AA study that would allow intricate gene-environment analysis. At this present time, there is only limited evidence that UGT2B gene region variants may be associated with differences in nicotine metabolism across African American, Native Hawaiian, Caucasian, Latino, and Japanese American smokers. Analyses of the UGT2A1/2 variants across populations may reveal differential risk to ovarian disease. In addition, expression and functional analysis in ovarian tissue needs to be accomplished to elucidate the impact on tissue homeostasis. In spite of the limitations of this study, these results provide new insight into proliferative and hormone target pathways that may represent important opportunities for the development of chemotherapeutic targets and intervention strategies.

ACKNOWLEDGMENTS

The BEL study would like to thank Gilian Peuteman, Thomas Van Brussel, Annick Van den Broeck and Joke
De Roover for technical assistance. The CAM study was supported by Cancer Research UK, the University of Cambridge, and the National Institute for Health Research Cambridge Biomedical Research Centre. The MOF study would like to thank the Total Cancer Care™ Protocol and the Collaborative Data Services and Tissue Core Facilities at the H. Lee Moffitt Cancer Center & Research Institute, an NCI designated Comprehensive Cancer Center, Merck Pharmaceuticals and the state of Florida (MOF). The NHS/NHSII studies thank the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, and WY. The SEA study thanks the SEARCH team, Craig Luccarini, Caroline Baynes, Don Conroy. The UKO study thanks I. Jacobs, M. Widschwendter, E. Wozniak, A. Ryan, J. Ford and N. Balogun for their contribution to the study. The WMH study thanks the Gynaecological Oncology Biobank at Westmead, a member of the Australasian Biospecimen Network-Oncology group.

CONFLICT OF INTEREST
The authors have no conflicts of interest to disclose.

ORCID
Delores J. Grant https://orcid.org/0000-0001-8981-8930
Jill Barnholtz-Sloan https://orcid.org/0000-0001-6190-9304
Kirsten Moysich https://orcid.org/0000-0002-4678-2058
Jennifer B. Permuth-Way https://orcid.org/0000-0002-4726-9264
Daniel W. Cramer https://orcid.org/0000-0002-8024-3066

REFERENCES


**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section at the end of the article.

---