A Transcriptome-Wide Association Study Among 97,898 Women to Identify Candidate Susceptibility Genes for Epithelial Ovarian Cancer Risk

Yingchang Lu[1], Alicia Beeghly-Fadiel[1], Lang Wu[1], Xingyi Guo[1], Bingshan Li[2], Joellen M. Schildkraut[3], HaeKyung Im[4],Yian A.Chen[5], Jennifer B. Permuth[6], Brett M. Reid6, Jamie K. Teer[5], Kirsten B. Moysich[7], Irene L. Andrulis[8,9], Hoda Anton-Culver[10], Banu K. Arun[11], Elisa V. Bandera1[2], Rosa B. Barkardottir[13,14], Daniel R. Barnes[15], Javier Benitez[16,17], Line Bjorge[18,19], James Brenton[20], Ralf Butzow[21], Trinidad Caldes[22], Maria A. Caligo[23], Ian Campbell[24,25,26], Jenny Chang-Claude[27,28], Kathleen B.M. Claes[29], Fergus J. Couch[30], Daniel W. Cramer[31,32], Mary B. Daly[33], Anna DeFazio[34,35], Joe Dennis[15], Orland Diez[36], Susan M. Domchek[37], Thilo Dörk[38], Douglas F. Easton[15,39], Diana M. Eccles[40], Peter A. Fasching [41,42], Renée T. Fortner[27], George Fountzilas[43], Eitan Friedman [44,45], Patricia A. Ganz[46], Judy Garber[47], Graham G. Giles[48,49,50], Andrew K. Godwin[51], David E. Goldgar[52], Marc T. Goodman[53,54], Mark H. Greene[55], Jacek Gronwald[56], Ute Hamann[57], Florian Heitz[58,59], Michelle A.T. Hildebrandt[60], Claus K. Høgdall[61], Antoinette Hollestelle[62], Peter J. Hulick[63,64], David G. Huntsman[65,66,67], Evgeny N. Imyanitov[68], Claudine Isaacs[69], Anna Jakubowska[56], Paul James[25,70], Beth Y. Karlan[71], Linda E. Kelemen[72], Lambertus A. Kiemeney[73], Susanne K. Kjaer[74,75], Ava Kwong[76,77,78], Nhu D. Le[79], Goska Leslie[15], Fabienne Lesueur[80,81,82,83], Douglas A. Levine[84,85], Amalia Mattiello[86], Taymaa May[87], Lesley McGuffog[15], Iain A. McNeish[88], Melissa A. Merritt[89,90], Francesmary Modugno[91,92,93], Marco Montagna[94], Susan L. Neuhausen[95], Heli Nevanlinna[96], Finn C. Nielsen[97], Liene Nikitina-Zake[98], Robert L. Nussbaum[99], Kenneth Offit[100,101], Edith Olah[102], Olufunmilayo I. Olopade[103], Sara H. Olson[104], Håkan Olsson[105], Ana Osorio[16,106], Sue K. Park[107,108,109], Michael T. Parsons[110], Petra H.M. Peeters[111], Tanja Pejovic[112,113], Paolo Peterlongo[114], Catherine M. Phelan[6], Miquel Angel Pujana[115], Susan J. Ramus[116,117], Gad Rennert[118], Harvey Risch[119], Gustavo C. Rodriguez[120], Cristina Rodríguez-Antona[121], Isabelle Romieu[122], Matti A. Rookus[123], Mary Anne Rossing[124,125], Iwona K. Rzepecka[126], Dale P. Sandler[127], Rita K. Schmutzler[128,129], Veronica W. Setiawan[130], Priyanka Sharma[131], Weiva Sieh[132], Jacques Simard[133], Christian F. Singer[134], Honglin Song[39], Melissa C. Southey[26], Amanda B. Spurdle[110], Rebecca Sutphen[135], Anthony J. Swerdlow[136,137], Manuel R. Teixeira[138,139], Soo H. Teo[140,141], Mads Thomassen[142], Marc Tischkowitz[143,144], Amanda E.Toland[145], Antonia Trichopoulou[146,147], Nadine Tung[148], Shelley S. Tworoger[6,149], Elizabeth J. van Rensburg[150], Adriaan Vanderstichele[151], Ana Vega[152], Digna Velez Edwards[153], Penelope M.Webb[154], Jeffrey N.Weitzel[155], Nicolas Wentzensen[156], Emily White[157,158], Alicja Wolk[159], Anna H.Wu[130], Drakoulis Yannoukakos[160], Kristin K. Zorn[161], Simon A. Gayther[130,162,163], Antonis C. Antoniou[15], Andrew Berchuck[164], Ellen L.Goode[165],Georgia Chenevix-Trench[110], Thomas A. Sellers[6], Paul D.P. Pharoah[15,39],Wei Zheng[1], and Jirong Long[1,*]

[1]Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, Tennessee.

[2]Department of Molecular Physiology and Biophysics, Vanderbilt University Medical Center, Nashville, Tennessee.

[3]Department of Public Health Sciences, University of Virginia, Charlottesville, Virginia.

[4]Section of Genetic Medicine, Department of Medicine, University of Chicago, Chicago, Illinois.

[5]Department of Biostatistics, Moffitt Cancer Center, Tampa, Florida.

[6]Department of Cancer Epidemiology, Moffitt Cancer Center, Tampa, Florida.

[7]*Division of Cancer Prevention and Control, Roswell Park Cancer Institute, Buffalo, New York.*
[8]*Fred A. Litwin Center for Cancer Genetics, Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, Ontario, Canada.*
[9]*Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada.*
[10]*Department of Epidemiology, Genetic Epidemiology Research Institute, University of California Irvine, Irvine, California.*
[11]*Department of Breast Medical Oncology, University of Texas MD Anderson Cancer Center, Houston, Texas.*
[12]*Cancer Prevention and Control Program, Rutgers Cancer Institute of New Jersey, New Brunswick, New Jersey.*
[13]*Department of Pathology, Landspitali University Hospital, Reykjavik, Iceland.*
[14]*BMC (Biomedical Centre), Faculty of Medicine, University of Iceland, Reykjavik, Iceland.*
[15]*Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, United Kingdom.*
[16]*Human Cancer Genetics Program, Spanish National Cancer Research Centre, Madrid, Spain.*
[17]*Centro de Investigación en Red de Enfermedades Raras (CIBERER), Valencia, Spain.*
[18]*Department of Gynecology and Obstetrics, Haukeland University Horpital, Bergen, Norway.*
[19]*Centre for Cancer Biomarkers, Department of Clinical Science, University of Bergen, Bergen, Norway.*
[20]*Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, United Kingdom.*
[21]*Department of Pathology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland.*
[22]*Medical Oncology Department, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria San Carlos (IdISSC), Centro Investigación Biomédica en Red de Cáncer (CIBERONC), Madrid, Spain.*
[23]*Section of Genetic Oncology, Department of Laboratory Medicine, University and University Hospital of Pisa, Pisa, Italy.*
[24]*Peter MacCallum Cancer Center, Melbourne, Victoria, Australia.*
[25]*Sir Peter MacCallum Department of Oncology, The University of Melbourne, Melbourne, Victoria, Australia.*
[26]*Department of Pathology, The University of Melbourne, Melbourne, Victoria, Australia.*
[27]*Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany.*
[28]*Research Group Genetic Cancer Epidemiology, University Cancer Center Hamburg (UCCH), University Medical Center Hamburg-Eppendorf, Hamburg, Germany.*
[29]*Centre for Medical Genetics, Ghent University, Gent, Belgium.*
[30]*Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota.*
[31]*Obstetrics and Gynecology Epidemiology Center, Brigham and Women's Hospital, Boston, Massachusetts.*
[32]*Harvard T.H. Chan School of Public Health, Boston, Massachusetts.*
[33]*Department of Clinical Genetics, Fox Chase Cancer Center, Philadelphia, Pennsylvania.*
[34]*Centre for Cancer Research, The Westmead Institute for Medical Research, The University of Sydney, Sydney, New South Wales, Australia.*
[35]*Department of Gynaecological Oncology, Westmead Hospital, Sydney, New South Wales, Australia.*
[36]*Oncogenetics Group, Clinical and Molecular Genetics Area, Vall d'Hebron Institute of Oncology (VHIO), University Hospital, Vall d'Hebron, Barcelona, Spain.*
[37]*Department of Medicine, Abramson Cancer Center, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania.*
[38]*Gynaecology Research Unit, Hannover Medical School, Hannover, Germany.*
[39]*Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Cambridge, United Kingdom.*
[40]*Cancer Sciences Academic Unit, Faculty of Medicine, University of Southampton, Southampton, United Kingdom.*
[41]*Department of Gynaecology and Obstetrics, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nuremberg, Comprehensive Cancer Center Erlangen-EMN, Erlangen, Germany.*
[42]*David Geffen School of Medicine, Department of Medicine Division of Hematology and Oncology, University of California at Los Angeles, Los Angeles, California.*
[43]*Department of Medical Oncology, "Papageorgiou" Hospital, Aristotle University of Thessaloniki School of Medicine, Thessaloniki, Greece.*
[44]*The Susanne Levy Gertner Oncogenetics Unit, Chaim Sheba Medical Center, Ramat Gan, Israel.*
[45]*Sackler Faculty of Medicine, Tel Aviv University, Ramat Aviv, Israel.*
[46]*Schools of Medicine and Public Health, Division of Cancer Prevention & Control Research, Jonsson Comprehensive Cancer Centre, UCLA, Los Angeles, California.*
[47]*Cancer Risk and Prevention Clinic, Dana-Farber Cancer Institute, Boston, Massachusetts.*
[48]*Cancer Epidemiology & Intelligence Division, Cancer Council Victoria, Melbourne, Victoria, Australia.*

[49]*Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Victoria, Australia.*
[50]*Department of Epidemiology and Preventive Medicine, Monash University, Melbourne, Victoria, Australia.*
[51]*Department of Pathology and Laboratory Medicine, Kansas University Medical Center, Kansas City, Kansas.*
[52]*Department of Dermatology, Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, Utah.*
[53]*Cancer Prevention and Control, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, California.*
[54]*Community and Population Health Research Institute, Department of Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, California.*
[55]*Clinical Genetics Branch, DCEG, National Cancer Institute, Bethesda, Maryland.*
[56]*Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland.*
[57]*Molecular Genetics of Breast Cancer, German Cancer Research Center (DKFZ), Heidelberg, Germany.*
[58]*Department of Gynecology and Gynecologic Oncology, Dr. Horst Schmidt Kliniken Wiesbaden, Wiesbaden, Germany.*
[59]*Department of Gynecology and Gynecologic Oncology, Kliniken Essen-Mitte/Evang. Huyssens-Stiftung/Knappschaft GmbH, Essen, Germany.*
[60]*Department of Epidemiology, University of Texas MD Anderson Cancer Center, Houston, Texas.*
[61]*Department of Gynecology, The Juliane Marie Centre, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark.*
[62]*Department of Medical Oncology, Family Cancer Clinic, Erasmus MC Cancer Institute, Rotterdam, The Netherlands.*
[63]*Center for Medical Genetics, NorthShore University HealthSystem, Evanston, Illinois.*
[64]*The University of Chicago Pritzker School of Medicine, Chicago, Illinois.*
[65]*British Columbia's Ovarian Cancer Research (OVCARE) Program, Vancouver General Hospital, BC Cancer Agency and University of British Columbia, Vancouver, British Columbia, Canada.*
[66]*Department of Molecular Oncology, BC Cancer Agency Research Centre, Vancouver, British Columbia, Canada.*
[67]*Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, British Columbia, Canada.*
[68]*N.N. Petrov Institute of Oncology, St. Petersburg, Russia.*
[69]*Lombardi Comprehensive Cancer Center, Georgetown University, Washington, District of Columbia.*
[70]*Parkville Familial Cancer Centre, Peter MacCallum Cancer Center, Melbourne, Victoria, Australia.*
[71]*Women's Cancer Program at the Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, California.*
[72]*Department of Public Health Sciences, Medical University of South Carolina, Charleston, South Carolina.*
[73]*Radboud Institute for Health Sciences, Radboud University Medical Center, Nijmegen, The Netherlands.*
[74]*Department of Virus, Lifestyle and Genes, Danish Cancer Society Research Center, Copenhagen, Denmark.*
[75]*Department of Gynaecology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark.*
[76]*Hong Kong Hereditary Breast Cancer Family Registry, Happy Valley, Hong Kong.*
[77]*Department of Surgery, The University of Hong Kong, Pok Fu Lam, Hong Kong.*
[78]*Department of Surgery, Hong Kong Sanatorium and Hospital, Happy Valley, Hong Kong.*
[79]*Cancer Control Research, BC Cancer Agency, Vancouver, British Columbia, Canada.*
[80]*Genetic Epidemiology of Cancer team, Inserm U900, Paris, France.*
[81]*Institut Curie, Paris, France.*
[82]*PSL University, Paris, France.*
[83]*Mines ParisTech, Fontainebleau, France.*
[84]*Gynecology Service, Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, New York.*
[85]*Gynecologic Oncology, Laura and Isaac Pearlmutter Cancer Center, NYU Langone Medical Center, New York, New York.*
[86]*Dipertimento Di Medicina Clinica E Chirurgia, Federico II University, Naples, Italy.*
[87]*Division of Gynecologic Oncology, University Health Network, Princess Margaret Hospital, Toronto, Ontario, Canada.*
[88]*Institute of Cancer Sciences, University of Glasgow, Wolfson Wohl Cancer Research Centre, Glasgow, United Kingdom.*
[89]*Cancer Epidemiology Program, University of Hawaii Cancer Center, Honolulu, Hawaii.*

[90]*Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, United Kingdom.*
[91]*Ovarian Cancer Center of Excellence, Womens Cancer Research Program, Magee-Womens Research Institute and University of Pittsburgh Cancer Institute, Pittsburgh, Pennsylvania.*
[92]*Department of Epidemiology, University of Pittsburgh Graduate School of Public Health, Pittsburgh, Pennsylvania.*
[93]*Division of Gynecologic Oncology, Department of Obstetrics, Gynecology and Reproductive Sciences, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania.*
[94]*Immunology and Molecular Oncology Unit, Veneto Institute of Oncology IOV - IRCCS, Padua, Italy.*
[95]*Department of Population Sciences, Beckman Research Institute of City of Hope, Duarte, California.*
[96]*Department of Obstetrics and Gynecology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland.*
[97]*Center for Genomic Medicine, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark.*
[98]*Latvian Biomedical Research and Study Centre, Riga, Latvia.*
[99]*Cancer Genetics and Prevention Program, University of California San Francisco, San Francisco, California.*
[100]*Clinical Genetics Research Lab, Department of Cancer Biology and Genetics, Memorial Sloan-Kettering Cancer Center, New York, New York.*
[101]*Clinical Genetics Service, Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, New York.*
[102]*Department of Molecular Genetics, National Institute of Oncology, Budapest, Hungary.*
[103]*Center for Clinical Cancer Genetics, The University of Chicago, Chicago, Illinois.*
[104]*Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, New York.*
[105]*Department of Cancer Epidemiology, Clinical Sciences, Lund University, Lund, Sweden.*
[106]*Biomedical Network on Rare Diseases (CIBERER), Madrid, Spain.*
[107]*Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Korea.*
[108]*Department of Biomedical Sciences, Seoul National University Graduate School, Seoul, Korea.*
[109]*Cancer Research Institute, Seoul National University, Seoul, Korea.*
[110]*Department of Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia.*
[111]*Julius Center for Health Sciences and Primary Care, UMC Utrecht, University Utrecht, Utrecht, The Netherlands.*
[112]*Department of Obstetrics and Gynecology, Oregon Health & Science University, Portland, Oregon.*
[113]*Knight Cancer Institute, Oregon Health & Science University, Portland, Oregon.*
[114]*IFOM, the FIRC (Italian Foundation for Cancer Research) Institute of Molecular Oncology, Milan, Italy.*
[115]*Catalan Institute of Oncology, ProCURE, Oncobell, Bellvitge Biomedical Research Institute (IDIBELL). Barcelona, Spain.*
[116]*School of Women's and Children's Health, University of NSW Sydney, Sydney, New South Wales, Australia.*
[117]*The Kinghorn Cancer Centre, Garvan Institute of Medical Research, Sydney, New South Wales, Australia.*
[118]*Clalit National Cancer Control Center, Carmel Medical Center and Technion Faculty of Medicine, Haifa, Israel.*
[119]*School of Public Health, Yale University, New Haven, Connecticut.*
[120]*Division of Gynecologic Oncology, NorthShore University HealthSystem, University of Chicago, Evanston, Illinois.*
[121]*Hereditary Endocrine Cancer group, Spanish National Cancer Research Center (CNIO), Madrid, Spain.*
[122]*Nutrition and Metabolism Section, International Agency for Research on Cancer (IARC-WHO), Lyon, France.*
[123]*Department of Epidemiology, The Netherlands Cancer Institute, Amsterdam, The Netherlands.*
[124]*Program in Epidemiology, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington.*
[125]*Department of Epidemiology, University of Washington, Seattle, Washington.*
[126]*Department of Pathology and Laboratory Diagnostics, Maria Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland.*
[127]*Epidemiology Branch, National Institute of Environmental Health Sciences, NIH, Research Triangle Park, Maryland.*
[128]*Center for Hereditary Breast and Ovarian Cancer, University Hospital of Cologne, Cologne, Germany.*
[129]*Center for Molecular Medicine Cologne (CMMC), University of Cologne, Cologne, Germany.*
[130]*Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California.*

[131]*Department of Medicine, University of Kansas Medical Center, Kansas City, Kansas.*
[132]*Department of Genetics and Genomic Sciences, Department of Population Health Science and Policy, Icahn School of Medicine at Mount Sinai, New York, New York.*
[133]*Genomics Center, Centre Hospitalier Universitaire de Québec Research Center, Laval University, Québec City, Québec, Canada.*
[134]*Dept of OB/GYN and Comprehensive Cancer Center, Medical University of Vienna, Vienna, Austria.*
[135]*Epidemiology Center, College of Medicine, University of South Florida, Tampa, Florida.*
[136]*Division of Genetics and Epidemiology, The Institute of Cancer Research, London, United Kingdom.*
[137]*Division of Breast Cancer Research, The Institute of Cancer Research, London, United Kingdom.*
[138]*Department of Genetics, Portuguese Oncology Institute, Porto, Portugal.*
[139]*Biomedical Sciences Institute (ICBAS), University of Porto, Porto, Portugal.*
[140]*Cancer Research Malaysia, Subang Jaya, Selangor, Malaysia.*
[141]*Breast Cancer Research Unit, Cancer Research Institute, University Malaya Medical Centre, Kuala Lumpur, Malaysia.*
[142]*Department of Clinical Genetics, Odense University Hospital, Odence C, Denmark.*
[143]*Program in Cancer Genetics, Departments of Human Genetics and Oncology, McGill University, Montréal, Québec, Canada.*
[144]*Department of Medical Genetics, Cambridge University, Cambridge, United Kingdom.*
[145]*Department of Cancer Biology and Genetics, Comprehensive Cancer Center, The Ohio State University, Columbus, Ohio.*
[146]*Hellenic Health Foundation, Athens, Greece.*
[147]*WHO Collaborating Center for Nutrition and Health, Unit of Nutritional Epidemiology and Nutrition in Public Health, Department of Hygiene, Epidemiology and Medical Statistics, University of Athens Medical School, Athens, Greece.*
[148]*Department of Medical Oncology, Beth Israel Deaconess Medical Center, Boston, Massachusetts.*
[149]*Research Institute and Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts.*\
[150]*Department of Genetics, University of Pretoria, Arcadia, South Africa.*
[151]*Division of Gynecologic Oncology, Department of Obstetrics and Gynaecology and Leuven Cancer Institute, University Hospitals Leuven, Leuven, Belgium.*
[152]*Fundación Pública Galega Medicina Xenómica, Santiago De Compostela, Spain.*
[153]*Vanderbilt Epidemiology Center, Vanderbilt Genetics Institute, Department of Obstetrics and Gynecology, Vanderbilt University Medical Center, Nashville, Tennessee.*
[154]*Population Health Department, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia.*
[155]*Clinical Cancer Genetics, City of Hope, Duarte, California.*
[156]*Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland.*
[157]*Fred Hutchinson Cancer Research Center, Seattle, Washington.*
[158]*Department of Epidemiology, University of Washington, Seattle, Washington.*
[159]*Department of Environmental Medicine, Division of Nutritional Epidemiology, Karolinska Institutet, Stockholm, Sweden.*
[160]*Molecular Diagnostics Laboratory, INRASTES, National Centre for Scientific Research 'Demokritos', Athens, Greece.*
[161]*Magee-Womens Hospital, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania.*
[162]*Center for Cancer Prevention and Translational Genomics, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, California.*
[163]*Department of Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, California.*
[164]*Department of Obstetrics and Gynecology, Duke University Medical Center, Durham, North Carolina.*
[165]*Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota.*

***Corresponding Author:**
Jirong Long, Vanderbilt Epidemiology Center, Vanderbilt University School of Medicine, 2525 West End Avenue, Eight Floor, Nashville, TN 37203-1738. Phone: 615-343-6741; Fax: 615-936-8241; E-mail: jirong.long@vanderbilt.edu

Abstract

Large-scale genome-wide association studies (GWAS) have identified approximately 35 loci associated with epithelial ovarian cancer (EOC) risk. The majority of GWAS-identified disease susceptibility variants are located in noncoding regions, and causal genes underlying these associations remain largely unknown. Here, we performed a transcriptome-wide association study to search for novel genetic loci and plausible causal genes at known GWAS loci. We used RNA sequencing data (68 normal ovarian tissue samples from 68 individuals and 6,124 cross-tissue samples from 369 individuals) and high-density genotyping data from European descendants of the Genotype-Tissue Expression (GTEx V6) project to build ovarian and cross-tissue models of genetically regulated expression using elastic net methods. We evaluated 17,121 genes for their *cis*-predicted gene expression in relation to EOC risk using summary statistics data from GWAS of 97,898 women, including 29,396 EOC cases. With a Bonferroni-corrected significance level of $P < 2.2 \times 10^{-6}$, we identified 35 genes, including *FZD4* at 11q14.2 ($Z = 5.08$, $P = 3.83 \times 10^{-7}$, the cross-tissue model; 1 Mb away from any GWAS-identified EOC risk variant), a potential novel locus for EOC risk. All other 34 significantly associated genes were located within 1 Mb of known GWAS-identified loci, including 23 genes at 6 loci not previously linked to EOC risk. Upon conditioning on nearby known EOC GWAS-identified variants, the associations for 31 genes disappeared and three genes remained ($P < 1.47 \times 10^{-3}$). These data identify one novel locus *(FZD4*) and 34 genes at 13 known EOC risk loci associated with EOC risk, providing new insights into EOC carcinogenesis.

Significance: Transcriptomic analysis of a large cohort confirms earlier GWAS loci and reveals FZD4 as a novel locus associated with EOC risk.

Introduction

Epithelial ovarian cancer (EOC) has a substantial heritable component with a heritability estimated to be 22% (1). Genome-wide association studies (GWAS) have identified approximately 35 loci associated with EOC risk (2–12). Most reported associations are specific to the most common histologic subtype, serous EOC (2–7, 9–12). Together, known GWAS-identified variants account for approximately 6.4% of EOC risk in the general population (12), indicating that additional susceptibility variants remain to be identified. In addition, genes that underlie the large majority of GWAS-identified risk loci remain unknown; most GWAS-identified variants are located in noncoding genomic regions that may be involved in regulation of gene expression. Recent mechanistic studies have demonstrated that GWAS-identified variants are more frequently located in active chromatin regions, and highly-enriched with expression quantitative trait loci (eQTL; refs. 13, 14). This evidence underscores the importance of transcriptional regulation in influencing human traits and disease susceptibility.

Prior studies on genetically regulated gene expression were largely limited to easily accessible sources, such as adipose tissue and peripheral blood cells (15). Although the sample size in eQTL studies of peripheral blood cells recently reached the thousands, a relatively small number of genes are expressed in blood cells compared with other tissue types (14). Conclusions from eQTL studies in tumor tissue (e.g., TCGA) should also be

interpreted with caution due to the inherent complexity of transcriptional regulation caused by acquired somatic alterations (16). The Genotype-Tissue Expression (GTEx) project provides high-density genotype data and RNA sequencing (RNA-seq) transcriptome data from 53 tissues (14). We used these data to build models of genetically regulated expression for 17,121 genes. We investigated the association between these genetically predicted gene expressions and EOC risk using data from 97,898 women including 29,396 EOC cases. We identified 35 genes at 14 loci associated with EOC risk, and provide additional evidence of a potential role for dysregulated ovarian function and imbalanced ovarian hormone production in ovarian carcinogenesis.

Materials and Methods

Genomic and transcriptomic data

The GTEx preliminary cleaned genome-wide genotype data and RNA-seq transcriptome data across 53 unique tissues (released on 2015-01-12) were downloaded from dbGaP (accession phs000424.GTEx.v6.p1). It included 183 GTEx donors genotyped on Illumina's Omni 5M and 267 GTEx donors genotyped on Omni 2.5M. Genomic and transcriptomic data were processed according to the GTEx protocol (http://www.gtexportal.org/home/documentationPage). The Omni 2.5M portion of hard-called genotypes from the Omni 2.5M or Omni 5M across all 450 donors were extracted and merged for analysis. We excluded variants with a genotyping call rate < 98%, with differential missingness between Omni 2.5M and Omni 5M arrays, with Hardy–Weinberg equilibrium $P < 10^{-6}$ (for subjects of European ancestry), or with batch effects. Genotype data were imputed to the Haplotype Reference Consortium reference panel using minimac3 for imputation and SHAPEIT for prephasing (17). Variants with high imputation quality ($R^2 \geq 0.8$), minor allele frequency (MAF) ≥ 0.05, and inclusion in the HapMap Phase 2 project were used to build predicted expression models.

We used gene level expression in Reads Per Kilobase of transcript per Million mapped reads (RPKM) from RNA-SeQC for gene expression data. For ovarian transcriptomic data, genes were required to have expression in ≥10 individuals with >0.1 RPKM and raw counts >6. For our analysis of cross-tissue derived transcriptomic data (below), genes were filtered on mean expression levels with >0.1 RPKM and RPKM >0 required in at least 3 individuals (18). We performed quantile normalization to transform the expression profile of each sample to the same scale, and performed inverse quantile normalization for each gene to map each set of expression values to a normal distribution. Residual expression was calculated by regressing transformed expression data against three top principal components (PC) derived from common genetic variants (MAF ≥0.05), top 15 or 35 probabilistic estimation of expression residuals (PEER) factors, respectively, for ovarian tissue and cross-tissue derived models (below; ref. 19), sex (for cross-tissue only) to correct for batch effects and other potential experimental confounders.

European ancestry analysis of GTEx subjects

The ancestral analysis was conducted with 2,836 ancestry informative markers for 450 GTEx individuals and 1,092 individuals included in the 1000 Genome project (Phase 1; ref. 20). Of the individuals with both genotype and transcriptome data available, 369 were clustered together with EUR populations (CEU, FIN, GBR, IBS, and TSI) on the multidimensional scaling plot of the pairwise Identity-By-State distance and were included in the analysis, 68 of whom had transcriptome data available for ovarian tissue.

Orthogonal tissue decomposition–derived cross tissue estimation

Mixed effect models were used to decompose gene expression levels into subject-specific and subject-by-tissue–specific components (18). GTEx data consisted of expression measurements from multiple tissues for each subject. The expression level of a gene at a given tissue for individual *i* was considered to be composed of a cross-tissue component represented as Y_i^{CT} and a tissue-specific component that was estimated as the difference between the expression levels and cross-tissue components given the lack of replicated measurement for a specific tissue/subject pair (18). Z'_i represents a vector of covariates that have effects of β on the expression levels of the gene, such as PEER factors, ancestry information derived from the principal component analysis, and sex. The expression of a gene for individual *i* in tissue *t,* $Y_{i,t}$, is modeled as $Y_{i,t} = Y_i^{CT} + Z_i'\beta + \epsilon_{i,t}$

The mixed effect model parameters were estimated using the lme4 package in R. Posterior modes of the subject level random intercepts were used as estimates of the cross-tissue components (18). Cross-tissue model included gene expression from 6,124 GTEx tissue samples from 369 unique European individuals who had genome-wide genotype data available.

Ovarian-specific and cross-tissue genetically regulated expression model building

We built an expression prediction model for each gene using the elastic net method as implemented in the glmnet R package, with a ridge-lasso mixing parameter of $\alpha = 0.5$ and a penalty parameter lambda chosen through 10-fold cross-validation (18, 21, 22). The elastic net method with $\alpha = 0.5$ is a compromise between the ridge-regression penalty ($\alpha = 0$) for solutions with many parameters (each of small effects) and the lasso penalty ($\alpha = 1$) for solutions with fewer parameters (each of large effects; ref. 18). The genetically regulated expression for each gene was estimated by including SNPs within 1 Mb of the gene start or end, as defined by GENCODE V19 gene annotations. Expression prediction models were built for protein-coding genes, long noncoding RNAs (lncRNA), miRNAs, processed transcripts, immunoglobulin genes, and T-cell receptor genes, according to categories described in the GENCODE V19 gene annotation file. Pseudogenes were not included in this study because of potential concerns of inaccurate calling (23). Prediction r^2 values (the square of the correlation between predicted and observed expression) were generated to estimate the prediction performance for each gene in our prediction models.

With genome-wide genomic data and RNAseq-based tissue transcriptome data, we built an ovarian tissue *cis* genetically regulated expression model for 8,580 genes that had predicted performance of $r^2 > 0.01$ and a cross-tissue *cis* genetically regulated expression model for 14,085 genes that had predicted performance of $r^2 > 0.01$.

Association analysis of predicted gene expression with EOC risk

Associations between predicted gene expression levels and EOC risk were evaluated using MetaXcan (22). Briefly, the formula: $Z_g \approx \sum_{l \in \text{Model}_g} w_{lg} \frac{\hat{\sigma}_l}{\hat{\sigma}_g} \frac{\hat{\beta}_l}{\text{se}(\hat{\beta}_l)}$

was used to estimate the Z-score of the association between predicted gene expression and ovarian cancer risk. Here w_{lg} is the weight of SNP l for predicting the expression of gene g, $\hat{\beta}_l$, and $\text{se}(\hat{\beta}_l)$ are the association regression coefficient and its standard error for SNP l in GWAS, and $\hat{\sigma}_l$ and $\hat{\sigma}_g$ are the estimated variances of SNP l and the predicted expression of gene g, respectively. The input variables for the MetaXcan analyses include the weights for gene expression predicting SNPs, GWAS summary statistics results, and correlations between predictor SNPs. We integrated prediction models of gene expression levels with summary statistics from GWAS of EOC risk for 97,898 European women with 29,396 EOC cases from the Ovarian Cancer Association Consortium (OCAC) and Consortium of Investigators of Modifiers of *BRCA1/2* (CIMBA; ref. 12) based on the variance and covariance matrix of genetic variants derived from 1000 Genome phase 3 EUR population ($N = 503$). The performance of MetaXcan has been shown to be similar to PrediXcan that uses individual-level genetic data for the identification of genes with expression that is associated with disease risk (21, 22).

Details of individual contributing studies were previously reported (12). Briefly, the OCAC summary statistics were based on analysis of 40,941 controls and 25,509 population-based EOC cases (22,406 invasive cases and 3,103 borderline cases). OCAC cases included 1,954 serous borderline ovarian cancers, 1,149 mucinous borderline ovarian cancers, 1,417 mucinous invasive ovarian cancer, 1,012 low-grade serous ovarian cancers, 13,037 high-grade serous ovarian cancers, 2,810 endometrioid ovarian cancers, 1,366 clear-cell ovarian cancer and 2,764 other EOC cases. The CIMBA summary statistics were based on the analysis of 19,036 *BRCA1* and 12,412 *BRCA2* mutation carriers, of whom 2,933 and 954, respectively, were diagnosed with EOC. Details of the genotyping procedure and QC have been described elsewhere (12). In brief, samples were excluded if they had a genotyping call rate < 95%, excessively low or high heterozygosity, if they were not female or had ambiguous sex, or were duplicates (cryptic or intended) (12). SNPs were excluded for a call rate <95%, deviating from Hardy–Weinberg equilibrium ($P < 10^{-7}$ in controls or unrelated samples in CIMBA and $P < 10^{-12}$ in cases) and concordance <98% among 5,280 duplicate pairs (12). All participants provided written informed consent and each contributing study was approved by the appropriate local institutional ethical review board. The studies were conducted in accordance with Declaration of Helsinki.

We used a Bonferroni-corrected *P* value threshold of 2.21×10^{-6} (adjusting for 22,665 gene–tissue pairs) to determine a statistically significant association in our analysis. This threshold was conservative as 5,544 genes appeared in both ovarian and cross-tissue models. We did

the primary analysis for high-grade serous EOC, as this had the largest sample size. In our secondary analyses, we also evaluated other histotypes or the combined histotypes, even though power to discover novel gene associations was relatively low for some (i.e., clear-cell, endometrioid, or low-grade serous). To determine whether associations identified between genetically predicted gene expression and EOC risk were influenced by variants previously identified by GWAS, we conducted conditional analyses adjusting for index SNPs. Briefly, we performed conditional analyses developed by Yang and colleagues (24) (GCTA-COJO) to calculate association betas and standard errors of SNPs with ovarian cancer risk after adjusting for index SNPs of interest. This was followed by reperforming MetaXcan analyses using updated summary statistics.

Results

Gene expression prediction model building

We constructed genetically regulated expression models based on genome-wide genotype data and RNA-seq transcriptome data from the GTEx project (Supplementary Fig. S1; ref. 14). Ovarian transcriptome data were available for 68 European individuals, and 8,580 genes achieved a prediction performance of $r^2 \geq 0.01$ in the ovarian model (Supplementary Table S1). Because a large portion of *cis* expression regulation is shared across multiple tissues (14, 18), we also used transcriptome data for 6,124 tissue samples from 369 European individuals to build cross-tissue models for 14,085 genes with a prediction performance of $r^2 \geq 0.01$ (Supplementary Table S1).

Association analyses between predicted gene expression and EOC risk

We evaluated associations between predicted gene expression levels and EOC risk using MetaXcan (22) with summary statistics for individual GWAS SNPs from 97,898 European women including 29,396 EOC cases from OCAC and CIMBA (Supplementary Fig. S1; ref. 12). Our primary analysis focused on high-grade serous EOC; secondary analyses included other EOC histotypes (Supplementary Fig. S1).

In total, we identified 35 genes with genetically predicted expression that were associated with EOC risk at a Bonferroni-corrected threshold of $P < 2.21 \times 10^{-6}$ (Fig. 1; Supplementary Figs. S2 and S3; Tables 1 and 2; Supplementary Table S2). One gene at 11q14.2 (*FZD4*), was more than 1 Mb away from any GWAS-identified EOC susceptibility variant (Fig. 1), suggesting a potential novel risk locus for this disease. High predicted *FZD4* expression was associated with increased risk of high-grade serous EOC ($Z = 5.08$, $P = 3.83 \times 10^{-7}$; Fig. 1). The remaining 34 genes were located within 1 Mb of previously identified EOC susceptibility variants (Tables 1 and 2; Supplementary Tables S2 and S3), including 11 genes (at 8 loci) that were previously implicated in EOC risk using functional annotation, bioinformatic prediction, *in vitro* cellular models or known gene biology. Our study provides additional evidence to support these previous findings (Table 2; Supplementary Table S3). However, 23 genes (at 6 known risk loci) had not been reported to be associated with EOC risk in prior studies (Table 1; Supplementary Table S3). For 31 of these 34 genes, the associations were no longer statistically significant at $P < 1.47 \times 10^{-3}$ (multiple comparisons correction of 0.05/34) after adjustment for the nearest SNP identified by EOC GWAS (Supplementary Table S4),

indicating that the previously identified GWAS SNPs for EOC at these 31 regions might regulate the expression of these associated gene to affect EOC risk. Associations for three genes (Z = 6.84 vs. 3.27 for *DNALI1*, Z = 5.16 vs. 3.81 for *HOXD3* and Z = −8.60 vs. −4.18 for *CCDC171*; Tables 1 and 2; Supplementary Table S4) remained statistically significant at $P < 1.47 \times 10^{-3}$ after adjusting for the nearest EOC risk SNP, although the strength of the association was attenuated. Four loci (2q31.1, 9p22.3, 17q21.31, and 17q21.32) had multiple nearby genes associated with EOC risk (Tables 1 and 2). This may be partially due to coregulated gene expression in these chromosomal regions (Supplementary Table S5 and Supplementary Material).

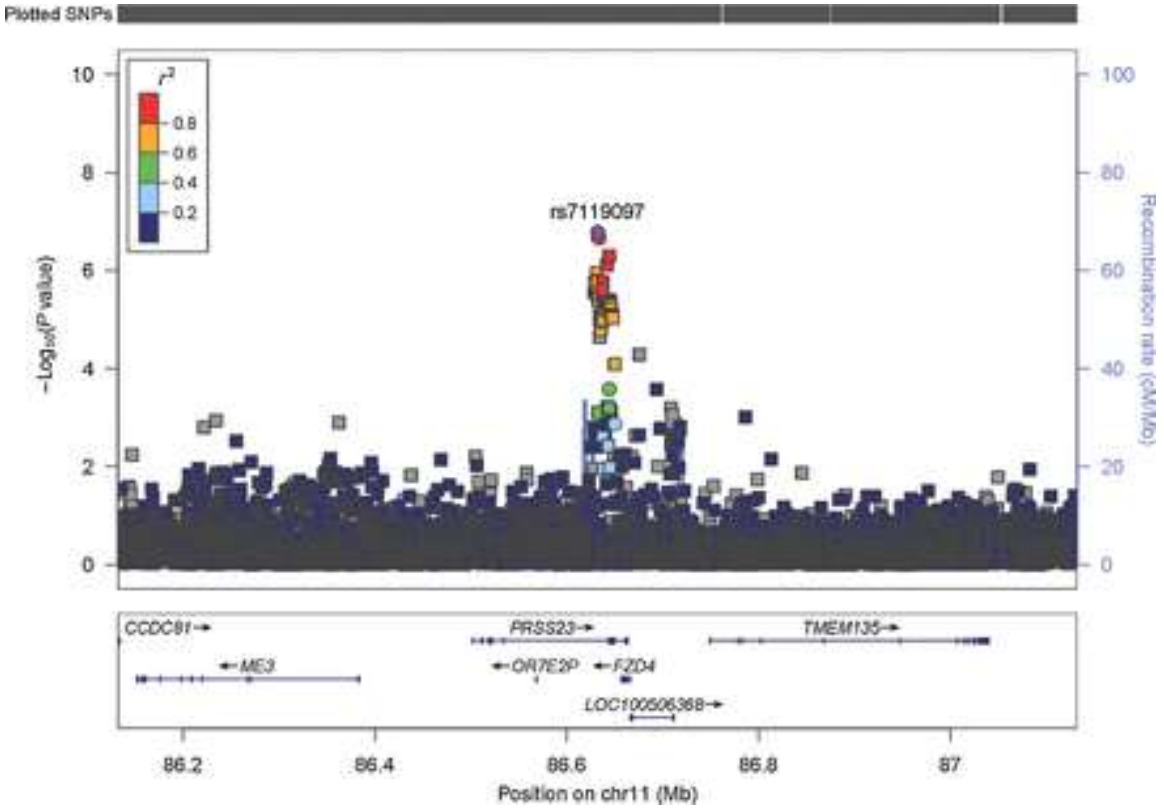

Figure 1. Regional plot of OCAC and CIMBA GWAS summary statistics around the *FZD4* gene associated with high-grade serous EOC risk (Z = 5.08; $P = 3.83 \times 10^{-7}$ based on the cross-tissue model of $r^2 = 0.07$; see Supplementary Table S2 for details). Each symbol represents the significance (*P* value on a $\log_{10}$ scale) of a SNP with invasive EOC risk as a function of the SNP's genomic position (NCBI Build 37). The most significantly associated SNP is represented in purple. The color of all other SNPs indicates LD with this SNP (estimated by EUR r^2 from the 1000 Genome Project data). Recombination rates were also estimated from 1000 Genome Project data, and gene annotations were obtained from the UCSC Genome Browser. The circle denotes the SNPs included in the model construction of genetically regulated *FZD4* expression and the square denotes the SNPs not included in the model construction. The gene model was constructed including SNPs within 1 Mb of the gene start or end, and one SNP included in the model construction was located outside the 1 Mb window size of the locus zoom plot (rs7944482 at chr11:86091532, $P = 0.52$ for association with high-grade serous EOC risk).

Table 1. Association results for genes in known loci not previously reported in association with epithelial ovarian cancer risk

Region	Gene[a]	Z-score	*P*	r^2[b]	Histotype	Model	GWAS Index SNP[c]	Distance to the index SNP (kb)[d]
1p34.3	*DNALI1*	6.84	7.84E-12	0.29	High-grade serous[e]	Cross-tissue	rs58722170	64
9p22.3	*CCDC171*	−8.60	8.08E-18	0.02	High-grade serous[e]	Ovary	rs10962692	854
9p22.3	*C9orf92*	−5.16	2.45E-07	0.15	High-grade serous[e]	Ovary	rs10962692	640
17q21.31	*ADAM11*	−4.86	1.19E-06	0.05	High-grade serous[e]	Ovary	rs1879586	708
17q21.31	*AC091132.1*	−7.18	7.02E-13	0.03	High-grade serous[e]	Cross-tissue	rs1879586	26
17q21.31	*RP11–798G7.8*	6.58	4.77E-11	0.05	High-grade serous[e]	Ovary	rs1879586	42
17q21.31	*CRHR1*	8.61	7.23E-18	0.60	High-grade serous[e]	Cross-tissue	rs1879586	132
17q21.31	*RP11–105N13.4*	6.77	1.33E-11	0.05	High-grade serous[e]	Ovary	rs1879586	132
17q21.31	*MAPT-AS1*	7.74	9.60E-15	0.10	High-grade serous[e]	Cross-tissue	rs1879586	354
17q21.31	*RP11–669E14.6*	−8.35	6.64E-17	0.30	High-grade serous[e]	Cross-tissue	rs1879586	545
17q21.31	*KANSL1-AS1*	8.26	1.48E-16	0.85	High-grade serous[e]	Cross-tissue	rs1879586	704
17q21.31	*LRRC37A*	8.38	5.08E-17	0.54	High-grade serous[e]	Ovary	rs1879586	803
17q21.31	*LRRC37A2*	8.26	1.44E-16	0.55	High-grade serous[e]	Ovary	rs1879586	1,022
17q21.31	*NSF*	−5.55	2.78E-08	0.02	High-grade serous[e]	Ovary	rs1879586	1,101
17q21.32	*RP11–138C9.1*	5.54	3.04E-08	0.02	High-grade serous[e]	Cross-tissue	rs7207826	741
17q21.32	*RP11–6N17.6*	5.93	3.00E-09	0.19	High-grade serous[e]	Cross-tissue	rs7207826	475
17q21.32	*PNPO*	5.34	9.38E-08	0.30	High-grade serous[e]	Cross-tissue	rs7207826	475
17q21.32	*PRR15L*	−4.91	9.18E-07	0.04	High-grade serous[e]	Cross-tissue	rs7207826	465
17q21.32	*HOXB2*	−5.48	4.28E-08	0.40	High-grade serous[e]	Cross-tissue	rs7207826	118
17q21.32	*HOXB-AS1*	−5.15	2.59E-07	0.29	High-grade serous[e]	Cross-tissue	rs7207826	120
17q21.32	*HOXB3*	−5.59	2.30E-08	0.12	High-grade serous[e]	cross-tissue	rs7207826	126
18q11.2	*RP11–403A21.1*	−5.53	3.13E-08	0.11	Low grade/borderline serous[f]	cross-tissue	rs8098244	132
19q13.2	*ZNF546*	7.14	9.07E-13	0.01	Invasive/borderline mucinous[f]	Ovary	rs688187	757

[a] *ARHGAP27* and *PLEKHM1* were previously considered as potential EOC candidate susceptibility genes by Permuth-Wey and colleagues (10) with an integrated molecular analysis of multiple genes at 17q21.31 locus (see Table 2 and Supplementary Table S3).
[b] r^2 of tissue model's correlation to gene's measured transcriptome (prediction performance).
[c] See Supplementary Table S4 for detailed information in selecting the GWAS index SNPs.
[d] If the GWAS index SNP is located upstream of the gene, the gene start position is used; otherwise, the gene end position was used; *LRRC37A2* and *NSF* are within 1M of reported GWAS SNPs considering the association of all variants with EOC risk at $P < 5 \times 10^{-8}$ at this locus (See text and Supplementary Table S4 for details).
[e] The analyses were based on summary statistics for high-grade serous ovarian cancers from Ovarian Cancer Association Consortium (OCAC) and Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA).
[f] The analyses were based on summary statistics from OCAC.

Table 2. Association results for genes in known loci previously reported in association with ovarian cancer risk

Region	Gene	Z-score	*P*	r^2[a]	Histotype	Model	GWAS Index SNP[b]	Distance to the index SNP (kb)[c]
2q31.1	*HOXD3*	5.16	2.42E-07	0.04	Borderline mucinous[e]	Cross-tissue	rs711830	0
2q31.1	*HOXD1*	6.07	1.31E-09	0.04	High-grade serous[f]	Cross-tissue	rs711830	16
3q25.31	*LEKR1*	−5.81	6.24E-09	0.46	High-grade serous[f]	Cross-tissue	rs62274041	108
8q21.13	*CHMP4C*	−6.69	2.24E-11	0.47	High-grade serous[e]	Cross-tissue	rs11782652	0
9q34.2	*ABO*	5.44	5.37E-08	0.49	High-grade serous[f]	Ovary	rs635634	4
10q24.33	*OBFC1*	−5.09	3.66E-07	0.01	Borderline serous[e]	Cross-tissue	rs7902587	16
15q26.1	*RCCD1*	−5.46	4.64E-08	0.59	High-grade serous[e]	Cross-tissue	rs8037137	0
17q21.31[d]	*PLEKHM1*	4.80	1.59E-06	0.01	High-grade serous[f]	Cross-tissue	rs1879586	0
17q21.31[d]	*KANSL1*	4.74	2.15E-06	0.18	High-grade serous[f]	Ovary	rs1879586	540
17q21.31[d]	*WNT3*	6.81	9.82E-12	0.40	High-grade serous[f]	Cross-tissue	rs1879586	1,273
19p13.11	*ABHD8*	4.79	1.69E-06	0.23	High-grade serous[f]	Cross-tissue	rs4808075	13

[a] r^2 of tissue model's correlation to gene's measured transcriptome (prediction performance).
[b] See Supplementary Table S4 for detailed information in selecting the GWAS index SNPs.
[c] If the GWAS index SNP is located upstream of the gene, the gene start position is used; otherwise, the gene end position was used; *WNT3* is within 1M of reported GWAS SNPs considering the association of all variants with EOC risk at $P < 5 \times 10^{-8}$ at this locus (See text and Supplementary Table S4 for details).
[d] Eleven novel genes associated with EOC risk at this locus were presented in Table 1.
[e] The analyses were based on summary statistics from OCAC.
[f] The analyses were based on summary statistics for high-grade serous ovarian cancers from Ovarian Cancer Association Consortium (OCAC) and Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA).

Consistent with the etiologic heterogeneity of EOC (25), GWAS-identified risk variants differed across histologic subtypes (12). Therefore, we investigated associations between genes with $P < 2.21 \times 10^{-6}$ across all major histotypes of EOC (Supplementary Table S6). The majority of identified genes were associated with high-grade serous EOC risk, likely due to the large number of cases in our primary analysis. A few additional histotype specific associations were identified from secondary analyses. *HOXD3* at 2q31.1 was associated with borderline mucinous EOC risk (Table 2; Supplementary Table S6: Z = 5.16, $P = 2.42 \times 10^{-7}$). *RP11-403A21.1* at 18q11.2 was associated with low-grade or borderline serous EOC risk (Table 1; Supplementary Table S6: Z = −5.53, $P = 3.13 \times 10^{-8}$). *ZNF546* at 19q13.2 was associated with mucinous EOC risk (Table 1; Supplementary Table S6: Z = 7.14, $P = 9.07 \times 10^{-13}$ for invasive/borderline mucinous EOC combined; Z = 5.99 and $P = 2.14 \times 10^{-9}$ for borderline mucinous EOC only). *HOXD1* at 2q31.1 was associated with both invasive serous (Supplementary Table S6: Z = 4.92, $P = 8.55 \times 10^{-7}$) and borderline mucinous (Supplementary Table S6: Z = 5.24, $P = 1.59 \times 10^{-7}$) EOC risk.

Evidence from previous eQTL analyses of identified EOC susceptibility risk variants supports several currently identified gene associations (Table 2; Supplementary Table S3). Reduced *OBFC1* expression was associated with risk allele of GWAS identified EOC SNP at 10q24.33 (12), and we found that higher predicted *OBFC1* expression was associated with lower EOC risk. Similarly, reduced *RCCD1* expression was associated with risk allele of GWAS identified EOC SNP (11), and we found that higher predicted *RCCD1* expression was associated with reduced EOC risk at 15q26.1. In addition, multiple lines of evidence support our finding between higher predicted *ABHD8* and increased EOC risk at 19p13.11. Increased *ABHD8* expression was associated with risk allele of GWAS identified EOC SNP (26). Copy number variant analysis indicated that 46% of high-grade serous EOC had amplification at 19p13.11 that contains *ABHD8* (3).

Discussion

In this large transcriptome-wide association study (TWAS) among 97,898 women of European ancestry, we identified 35 genes with genetically predicted expression levels associated with EOC risk. One of these genes (*FZD4*) is located more than 1 Mb away from any previously identified GWAS EOC variant (25 Mb away from the nearest reported EOC risk variant; ref. 11), suggesting it is a potential novel risk locus. All other 34 genes identified were located within 1 Mb of known GWAS loci, including 23 genes at 6 loci that had not previously been associated with EOC risk. After adjustment for nearby known EOC GWAS-identified variants, the associations for 3 of the 34 genes retained.

FZD4 is a member of the frizzled gene family that encodes seven-transmembrane domain proteins (Fz) as the receptors for the secreted Wnts signaling ligands. Several Wnts and Fzs (including Fzd4 and Wnt4), as well as downstream targets of the canonical WNT signaling pathway, are expressed at different stages of ovarian follicular development, ovulation, and luteinization, suggesting specific functions for these signaling molecules in the mature ovary (27). Recent studies using transgenic mouse models demonstrated that Wnt4, Fzd4, and Ctnnb1 are required for normal folliculogenesis, luteogenesis and steroidogenesis, and that dysregulated WNT signaling leads to granulosa cell tumor development (27, 28). *FZD4*-null female mice are infertile and exhibit reduced progesterone production, reduced

luteinization-associated gene expression, impaired corpora lutea formation and function, and impaired vascular development (28). Interestingly, *WNT4* (1p36.12) encodes a potential Fzd4-binding ligand, which was also recently identified as a potentially causal gene underlying EOC risk by GWAS (Supplementary Table S3; ref. 7). Aberrant activation of WNT signaling in adult tissues has been implicated in the pathogenesis of several types of cancer, including colorectal cancer (29). The positive association between *FZD4* expression and invasive serous EOC risk suggests that dysregulated corpus luteum function and/or progesterone production may contribute to EOC pathogenesis.

A locus 17q21.31 was previously identified by GWAS as associated with EOC risk (10, 30). This region contains a 900-kb inversion in Europeans that has extensive linkage disequilibrium likely due to restriction from crossovers in individuals who are heterozygous with respect to inversion (31). The H2 haplotype is less frequent (20% in Europeans) and is associated with higher number of children born to women (31). Interestingly, minor alleles of genetic variants in this region were almost universally associated with reduced breast cancer risk but increased EOC risk at genome-wide significance levels (Supplementary Table S7; Supplementary Material; refs. 10, 30). Permuth-Wey and colleagues (10) investigated several of these genes, including *KIF18B, C1QL1, DCAKD, NMT1, PLCD3, ACBD4, HEXIM1, HEXIM2, FMNL1, C17orf46, MAP3K14, ARHGAP27, PLEKHM1, CRHR1, IMP5*, and *MAPT*; extensive functional analysis suggested that *ARHGAP27* and *PLEKHM1* may be EOC susceptibility genes (10). One of the other candidate genes at this region, *CRHR1*, is involved in regulating ovarian function; it is expressed in ovarian thecal cells, granulosa cells and luteal cells (32), and upregulated in EOC (10). High *CRHR1* expression was almost universally associated with minor alleles of multiple genetic variants in this chromosome 17 region (Supplementary Table S8; Supplementary Material; ref. 33). Enhanced *CRHR1* activation in the ovary leads to reduced production of testosterone (32) and estrogen (32, 34–36), but increased progesterone accumulation and production (32). This may explain the lower breast cancer risk associated with variants in this region from lower estrogen exposure and higher progesterone exposure associated with multiparity (31, 37). Similarly, this also suggests that imbalanced estrogen and/or progesterone production contributes to EOC pathogenesis.

Two of the candidate genes at the 17q21.32 locus, *HOXB2* and *HOXB3*, belong to the homeobox gene family, which is important for normal vertebrate limb and organ development. This gene family was also recently shown to be enriched for genes underlying serous EOC risk by GWAS (38). Inconsistent tumorigenic effects of *HOXB2* and *HOXB3* were reported across several types of cancers (breast, pancreatic, lung, cervical cancer, and acute myeloid leukemia; refs. 39–43). This may be due to context-dependent effects from specific tumor microenvironments (39, 43). With regard to ovarian cancer, increased *HOXB2* and *HOXB3* expression were associated with reduced EOC risk; potential molecular mechanisms underlying *HOXB* suppressive effect on EOC warrant further investigation.

Several additional findings from this study are noteworthy. The precise function of *DNALI1* at 1p34.3 is not known. It is a potential candidate gene for primary cilia syndrome or Kartagener syndrome, in which the action of cilia lining the respiratory tract and fallopian tube is compromised (44). A marked reduction in fertility was observed in female patients with Kartagener syndrome due to dysfunction of the oviductal cilia (45). The predicted

expression of *CCDC171* at 9p22.3 was associated with reduced EOC risk. *CCDC171* was shown to interact with *KRAS* by a stringent screening for Ras synthetic lethal genes (46). Several lncRNAs were associated with EOC risk, including *RP11-403A21.1* at 18q11.2 (Table 1). Little is known about their particular function in either tumor initiation or tumor development, but lncRNAs have been increasingly implicated in many classic cancer biology pathways (47). In addition to *HOXD3* and *HOXD1* at 2q31.1 (Table 2; Supplementary Table S3; refs. 4, 8), *ZNF546* at 19q13.2 was identified as a novel candidate gene for mucinous EOC. Enrichment for expression in gonadal tissues (14) supports a potential role in EOC pathogenesis. Because of the complexity of mucinous EOC, and undetermined cell/tissue of origin, identification of associated genetic variants and/or genes is particularly important (8, 25).

The tissue samples used in building gene expression models in GTEx (V6) came most from the people who recently died of traumatic injury (for these young donors) or cardio-cerebrovascular diseases (for the old donors). There were no overlaps between the tissues used in building gene expression models and the samples used in EOC GWAS in OCAC or CIMBA. Our ability to detect genes significantly associated with EOC risk is affected by tissue specificity and the sample size of the dataset used to build genetic prediction models for gene expression. Four genes were identified from both ovarian and cross-tissue models; 8 genes were only identified on the basis of ovarian models, and 23 genes were only identified from cross-tissue models (Supplementary Table S2). The ovarian tissue transcriptome that we used to model gene expression was potentially derived from multiple ovarian cell types, including surface epithelial cells, oocytes, granulosa cells, Theca cells, luteal cells, and other interstitial cells. Because of the importance of tissue or cell-specific regulators (i.e., transcription factors or epigenomic features) in governing development and function, the ovarian-specific model should best capture transcriptional regulatory mechanisms of the ovary. However, in light of abundant shared *cis* regulation of expression across multiple tissues (14, 18), we also pooled constitutive variant-dependent regulatory information across tissues and built cross-tissue gene expression models. We would expect this model to yield greater power as the number of tissues in which a variant is functional increases. By coupling both tissue-specific and cross-tissue models, we aimed to robustly capture genetically regulated genes expression using a large sample size. Because of insufficient samples in the GTEx project, we did not build Fallopian tube–specific models.

In summary, we identified one novel locus (*FZD4*) and 34 genes at 13 known EOC risk loci associated with EOC risk, and these findings may help improve our mechanistic understanding of EOC pathogenesis. In line with tentative observations of increased borderline EOC risk from ovarian hormone dysregulation for women who received fertility drug treatment with *in vitro* fertilization (48–50), the known biology of *FZD4* and *CRHR1* in the ovary implicates the potential of long-term dysregulated ovarian function or imbalanced ovarian hormone production as a possible mechanism underlying EOC pathogenesis.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: Y. Lu, J.M. Schildkraut, H. Anton-Culver, B.K. Arun, P.A. Fasching, R.L. Nussbaum, P.H.M. Peeters, V.W. Setiawan, M.C. Southey, S.A. Gayther, A. Berchuck, W. Zheng, J. Long

Development of methodology: Y. Lu, X. Guo, H.K. Im, C.F. Singer, W. Zheng

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): Y. Lu, L. Wu, J.M. Schildkraut, B.M. Reid, K.B. Moysich, I.L. Andrulis, H. Anton-Culver, B.K. Arun, E.V. Bandera, R.B. Barkardottir, J. Benitez, L. Bjorge, J. Brenton, R. Butzow, T. Caldes, M.A. Caligo, I.G. Campbell, J. Chang-Claude, K.B.M. Claes, F.J. Couch, D.W. Cramer, M.B. Daly, A. DeFazio, O. Diez, S.M. Domchek, T. Dork, D.F. Easton, D.M. Eccles, P.A. Fasching, R.T. Fortner, G. Fountzilas, E. Friedman, P.A. Ganz, J. Garber, G.G. Giles, A.K. Godwin, D.E. Goldgar, M.T. Goodman, M.H. Greene, J. Gronwald, U. Hamann, F. Heitz, M.A.T. Hildebrandt, C.K. Høgdall, A. Hollestelle, P.J. Hulick, E.N. Imyanitov, C. Isaacs, A. Jakubowska, B.Y. Karlan, L.E. Kelemen, L.A. Kiemeney, S.K. Kjaer, A. Kwong, N.D. Le, F. Lesueur, D.A. Levine, A. Mattiello, T. May, I.A. McNeish, F. Modugno, M. Montagna, S.L. Neuhausen, H. Nevanlinna, L. Nikitina-Zake, R.L. Nussbaum, K. Offit, E. Olah, O.I. Olopade, S.H. Olson, H. Olsson, A. Osorio, S.K. Park, P.H.M. Peeters, T. Pejovic, P. Peterlongo, C.M. Phelan, M.A. Pujana, S.J. Ramus, G. Rennert, H. Risch, G.C. Rodriguez, C. Rodríguez-Antona, M.A. Rookus, M.A. Rossing, I.K. Rzepecka, D.P. Sandler, R.K. Schmutzler, V.W. Setiawan, P. Sharma, W. Sieh, J. Simard, C.F. Singer, H. Song, M.C. Southey, A.B. Spurdle, R. Sutphen, A.J. Swerdlow, M.R. Teixeira, S.H. Teo, M. Thomassen, M. Tischkowitz, A. Trichopoulou, N. Tung, S.S. Tworoger, E.J. van Rensburg, A. Vega, D.V. Edwards, P.M. Webb, J.N. Weitzel, N. Wentzensen, E. White, A. Wolk, A.H. Wu, D. Yannoukakos, K.K. Zorn, S.A. Gayther, A.C. Antoniou, E.L. Goode, G. Chenevix-Trench, T.A. Sellers, P.D.P. Pharoah, W. Zheng

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): Y. Lu, A. Beeghly-Fadiel, L. Wu, X. Guo, B. Li, Y.A. Chen, J.K. Teer, K.B. Moysich, D.R. Barnes, D.R. Barnes, S.M. Domchek, R.T. Fortner, J. Gronwald, N.D. Le, K. Offit, H. Risch, G.C. Rodriguez, I. Romieu, W. Sieh, C.F. Singer, M.C. Southey, A.E. Toland, N. Tung, A. Vanderstichele, N. Wentzensen, A.C. Antoniou, A. Berchuck, E.L. Goode, W. Zheng

Writing, review, and/or revision of the manuscript: Y. Lu, A. Beeghly-Fadiel, L. Wu, B. Li, H.K. Im, J.B. Permuth, B.M. Reid, J.K. Teer, K.B. Moysich, I.L. Andrulis, H. Anton-Culver, B.K. Arun, E.V. Bandera, R.B. Barkardottir, J. Benitez, L. Bjorge, J. Brenton, I.G. Campbell, J. Chang-Claude, K.B.M. Claes, F.J. Couch, D.W. Cramer, M.B. Daly, A. DeFazio, O. Diez, S.M. Domchek, T. Dork, D.F. Easton, D.M. Eccles, P.A. Fasching, R.T. Fortner, G. Fountzilas, E. Friedman, P.A. Ganz, J. Garber, G.G. Giles, A.K. Godwin, M.T. Goodman, M.H. Greene, J. Gronwald, U. Hamann, F. Heitz, C.K. Høgdall, A. Hollestelle, P.J. Hulick, E.N. Imyanitov, C. Isaacs, A. Jakubowska, P. James, B.Y. Karlan, L.E. Kelemen, L.A. Kiemeney, S.K. Kjaer, A. Kwong, N.D. Le, A. Mattiello, T. May, M.A. Merritt, F. Modugno, S.L. Neuhausen, R.L. Nussbaum, E. Olah, O.I. Olopade, H. Olsson, A. Osorio, S.K. Park, P.H.M. Peeters, S.J. Ramus, G. Rennert, H. Risch, G.C. Rodriguez, I. Romieu, M.A. Rossing, D.P. Sandler, R.K. Schmutzler, V.W. Setiawan, P. Sharma, W. Sieh, J. Simard, C.F. Singer, H. Song, M.C. Southey, A.B. Spurdle, R. Sutphen, A.J. Swerdlow, S.H. Teo, M. Tischkowitz, A.E. Toland, A. Trichopoulou,

N. Tung, S.S. Tworoger, E.J. van Rensburg, A. Vega, P.M. Webb, J.N. Weitzel, N. Wentzensen, E. White, A. Wolk, A.H. Wu, D. Yannoukakos, S.A. Gayther, A.C. Antoniou, A. Berchuck, E.L. Goode, G. Chenevix-Trench, T.A. Sellers, P.D.P. Pharoah, W. Zheng, J. Long

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): Y. Lu, L. Wu, K.B. Moysich, J. Dennis, T. Dork, M.A.T. Hildebrandt, B.Y. Karlan, L.E. Kelemen, L.A. Kiemeney, G. Leslie, F. Lesueur, L. McGuffog, F. Modugno, K. Offit, O.I. Olopade, M.T. Parsons, S.J. Ramus, M.A. Rookus, P. Sharma, M.C. Southey, J.N. Weitzel, E. White, A.C. Antoniou, A. Berchuck, W. Zheng

Study supervision: P.J. Hulick, D.G. Huntsman, L.A. Kiemeney, C.F. Singer, M.C. Southey, R. Sutphen, S.A. Gayther, W. Zheng, J. Long

Other (variant classification): A.B. Spurdle

Acknowledgments

The Ovarian Cancer Association Consortium is supported by a grant from the Ovarian Cancer Research Fund thanks to donations by the family and friends of Kathryn Sladek Smith (PPD/RPCI.07). The scientific development and funding for this project were in part supported by the US National Cancer Institute GAME-ON Post-GWAS Initiative (U19-CA148112). This study made use of data generated by the Wellcome Trust Case Control consortium that was funded by the Wellcome Trust under award 076113. The results published here are in part based upon data generated by The Cancer Genome Atlas Pilot Project established by the National Cancer Institute and National Human Genome Research Institute (dbGap accession number phs000178.v8.p7). The OCAC OncoArray genotyping project was funded through grants from the NIH [U19-CA148112 (to T.A. Sellers), R01-CA149429 (to C.M. Phelan), and R01-CA058598 (to M.T. Goodman); Canadian Institutes of Health Research (MOP-86727 (to L.E. Kelemen)] and the Ovarian Cancer Research Fund (to A. Berchuck). The COGS project was funded through a European Commission's Seventh Framework Programme grant (agreement number 223175 - HEALTH-F2-2009-223175). Funding for individual studies: AAS: NIH (RO1-CA142081); AOV: The Canadian Institutes for Health Research (MOP-86727); **AUS:** The Australian Ovarian Cancer Study Group was supported by the U.S. Army Medical Research and Materiel Command (DAMD17-01-1-0729), National Health & Medical Research Council of Australia (199600, 400413 and 400281), Cancer Councils of New South Wales, Victoria, Queensland, South Australia and Tasmania and Cancer Foundation of Western Australia (Multi-State Applications 191, 211, and 182). The Australian Ovarian Cancer Study gratefully acknowledges additional support from Ovarian Cancer Australia and the Peter MacCallum Foundation; **BAV:** ELAN Funds of the University of Erlangen-Nuremberg; **BEL:** National Kankerplan; **BGS:** Breast Cancer Now, Institute of Cancer Research; **BVU:** Vanderbilt CTSA grant from the National Institutes of Health (NIH)/National Center for Advancing Translational Sciences (NCATS) (ULTR000445); **CAM:** National Institutes of Health Research Cambridge Biomedical Research Centre and Cancer Research UK Cambridge Cancer Centre; **CHA:** Innovative Research Team in University (PCSIRT) in China (IRT1076); **CNI:** Instituto de Salud Carlos III (PI 12/01319); Ministerio de Economía y Competitividad (SAF2012); **COE:** Department of Defense (W81XWH-11-2-0131); **CON:** National Institutes of Health (R01-CA063678, R01-CA074850; R01-CA080742); **DKE:**

Ovarian Cancer Research Fund and National Institutes of Health 1R01CA211574 (J.M. Schildkraut); **DOV:** National Institutes of Health R01-CA112523 and R01-CA87538; **EMC:** Dutch Cancer Society (EMC 2014-6699); **EPC:** The coordination of EPIC is financially supported by the European Commission (DG-SANCO) and the International Agency for Research on Cancer. The national cohorts are supported by Danish Cancer Society (Denmark); Ligue Contre le Cancer, Institut Gustave Roussy, Mutuelle Générale de l'Education Nationale, Institut National de la Santé et de la Recherche Médicale (INSERM) (France); German Cancer Aid, German Cancer Research Center (DKFZ), Federal Ministry of Education and Research (BMBF) (Germany); the Hellenic Health Foundation (Greece); Associazione Italiana per la Ricerca sul Cancro-AIRC-Italy and National Research Council (Italy); Dutch Ministry of Public Health, Welfare and Sports (VWS), Netherlands Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF), Statistics Netherlands (The Netherlands); ERC-2009-AdG 232997 and Nordforsk, Nordic Centre of Excellence programme on Food, Nutrition and Health (Norway); Health Research Fund (FIS), PI13/00061 to Granada, PI13/01162 to EPIC-Murcia, Regional Governments of Andalucía, Asturias, Basque Country, Murcia and Navarra, ISCIII RETIC (RD06/0020) (Spain); Swedish Cancer Society, Swedish Research Council and County Councils of Skåne and Västerbotten (Sweden); Cancer Research UK (14136 to EPIC-Norfolk; C570/A16491 and C8221/A19170 to EPIC-Oxford), Medical Research Council (1000143 to EPIC-Norfolk, MR/M012190/1 to EPIC-Oxford) (United Kingdom); **GER:** German Federal Ministry of Education and Research, Programme of Clinical Biomedical Research (01 GB 9401) and the German Cancer Research Center (DKFZ); **GRC:** This research has been co-financed by the European Union (European Social Fund - ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program of the General Secretariat for Research & Technology: SYN11_10_19 NBCA. Investing in knowledge society through the European Social Fund; **GRR:** Roswell Park Cancer Institute Alliance Foundation, P30 CA016056; **HAW:** U.S. National Institutes of Health (R01-CA58598, N01-CN-55424 and N01-PC-67001); **HJO:** Intramural funding; Rudolf-Bartling Foundation; **HMO:** Intramural funding; Rudolf-Bartling Foundation; **HOC:** Helsinki University Research Fund; **HOP:** Department of Defense (DAMD17-02-1-0669) and NCI (K07-CA080668, R01-CA95023, P50-CA159981 MO1-RR000056 R01-CA126841); **HUO:** Intramural funding; Rudolf-Bartling Foundation; **JGO:** JSPS KAKENHI grant; **JPN:** Grant-in-Aid for the Third Term Comprehensive 10-Year Strategy for Cancer Control from the Ministry of Health, Labour and Welfare; **KRA:** This study (Ko-EVE) was supported by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), and the National R&D Program for Cancer Control, Ministry of Health & Welfare, Republic of Korea (HI16C1127; 0920010); **LAX:** American Cancer Society Early Detection Professorship (SIOP-06-258-01-COUN) and the National Center for Advancing Translational Sciences (NCATS), Grant UL1TR000124; **LUN:** ERC-2011-AdG 294576-risk factors cancer, Swedish Cancer Society, Swedish Research Council, Beta Kamprad Foundation; **MAC:** NIH (R01-CA122443, P30-CA15083, P50-CA136393); Mayo Foundation; Minnesota Ovarian Cancer Alliance; Fred C. and Katherine B. Andersen Foundation; Fraternal Order of Eagles; **MAL:** Funding for this study was provided by research grant R01- CA61107 from the National Cancer Institute, Bethesda, MD, research grant 94 222 52 from the Danish Cancer Society, Copenhagen, Denmark; and the Mermaid I project; **MAS:** Malaysian Ministry of Higher Education (UM.C/HIR/MOHE/06) and Cancer Research Initiatives Foundation; **MAY:** National

Institutes of Health (R01-CA122443, P30-CA15083, P50-CA136393); Mayo Foundation; Minnesota Ovarian Cancer Alliance; Fred C. and Katherine B. Andersen Foundation; **MCC:** Cancer Council Victoria, National Health and Medical Research Council of Australia (NHMRC) grants number 209057, 251533, 396414, and 504715; **MDA:** DOD Ovarian Cancer Research Program (W81XWH-07-0449); **MEC:** NIH (CA54281, CA164973, CA63464); **MOF:** Moffitt Cancer Center, Merck Pharmaceuticals, the state of Florida, Hillsborough County, and the city of Tampa; **NCO:** National Institutes of Health (R01-CA76016) and the Department of Defense (DAMD17-02-1-0666); **NEC:** National Institutes of Health R01-CA54419 and P50-CA105009 and Department of Defense W81XWH-10-1-02802; **NHS:** UM1 CA186107, P01 CA87969, R01 CA49449, R01-CA67262, UM1 CA176726; **NJO:** National Cancer Institute (NIH-K07 CA095666, R01-CA83918, NIH-K22-CA138563, P30-CA072720, and P30-CA008748) and the Cancer Institute of New Jersey; **NOR:** Helse Vest, The Norwegian Cancer Society, The Research Council of Norway; **NTH:** Radboud University Medical Centre; **OPL:** National Health and Medical Research Council (NHMRC) of Australia (APP1025142) and Brisbane Women's Club; **ORE:** OHSU Foundation; **OVA:** This work was supported by Canadian Institutes of Health Research grant (MOP-86727) and by NIH/NCI 1 R01CA160669-01A1; **PLC:** Intramural Research Program of the National Cancer Institute; **POC:** Pomeranian Medical University; **POL:** Intramural Research Program of the National Cancer Institute; **PVD:** Canadian Cancer Society and Cancer Research Society GRePEC Program; **RBH:** National Health and Medical Research Council of Australia; **RMH:** Cancer Research UK, Royal Marsden Hospital; **RPC:** National Institute of Health (P50 CA159981, R01CA126841); **SEA:** Cancer Research UK (C490/A10119 C490/A10124); UK National Institute for Health Research Biomedical Research Centres at the University of Cambridge; **SIS:** NIH, National Institute of Environmental Health Sciences, Z01 ES044005 and Z01-ES049033; **SMC:** The Swedish Research Council; **SON:** National Health Research and Development Program, Health Canada, grant 6613-1415-53; **SRO:** Cancer Research UK (C536/A13086, C536/A6689) and Imperial Experimental Cancer Research Centre (C1312/A15589); **STA:** NIH grants U01 CA71966 and U01 CA69417; **SWE:** Swedish Cancer foundation, WeCanCureCancer and årKampMotCancer foundation; **SWH:** NIH (NCI) grant R37-CA070867; **TBO:** National Institutes of Health (R01-CA106414-A2), American Cancer Society (CRTG-00-196-01-CCE), Department of Defense (DAMD17-98-1-8659), Celma Mastery Ovarian Cancer Foundation; **TOR:** NIH grants R01 CA063678 and R01 CA063682; **UCI:** NIH R01-CA058860 and the Lon V Smith Foundation grant LVS-39420; **UHN:** Princess Margaret Cancer Centre Foundation-Bridge for the Cure; **UKO:** The UKOPS study was funded by The Eve Appeal (The Oak Foundation) and supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre; **UKR:** Cancer Research UK (C490/A6187), UK National Institute for Health Research Biomedical Research Centres at the University of Cambridge; **USC:** P01CA17054, P30CA14089, R01CA61132, N01PC67010, R03CA113148, R03CA115195, N01CN025403, and California Cancer Research Program (00-01389V-20170, 2II0200); **VAN:** BC Cancer Foundation, VGH & UBC Hospital Foundation; **VTL:** NIH K05-CA154337; **WMH:** National Health and Medical Research Council of Australia, Enabling Grants ID 310670 & ID 628903. Cancer Institute NSW Grants 12/RIG/1-17 & 15/RIG/1-16; **WOC:** National Science Centren (N N301 5645 40). The Maria Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland. We are grateful to the family and friends of Kathryn Sladek Smith for their generous support of the Ovarian Cancer Association Consortium through their donations to the Ovarian Cancer Research Fund. The OncoArray and COGS genotyping projects would not have been possible without the

contributions of the following: Per Hall (COGS); Kyriaki Michailidou, Manjeet K. Bolla, Qin Wang (BCAC), Marjorie J. Riggan (OCAC), Rosalind A. Eeles, Ali Amin Al Olama, Zsofia Kote-Jarai, Sara Benlloch (PRACTICAL), Jonathan P. Tyrer, Siddhartha Kar, Alison M. Dunning, Andrew Lee, and Ed Dicks, Craig Luccarini and the staff of the Centre for Genetic Epidemiology Laboratory, Anna Gonzalez-Neira and the staff of the CNIO genotyping unit, Daniel C. Tessier, Francois Bacot, Daniel Vincent, Sylvie LaBoissière and Frederic Robidoux and the staff of the McGill University and Génome Québec Innovation Centre, Stig E. Bojesen, Sune F. Nielsen, Borge G. Nordestgaard and the staff of the Copenhagen DNA laboratory, and Julie M. Cunningham, Sharon A. Windebank, Christopher A. Hilker, Jeffrey Meyer and the staff of Mayo Clinic Genotyping Core Facility. We pay special tribute to the contribution of Professor Brian Henderson to the GAME-ON consortium and to Olga M. Sinilnikova for her contribution to CIMBA and for her part in the initiation and coordination of GEMO until she sadly passed away on the 30th June 2014. We thank the study participants, doctors, nurses, clinical and scientific collaborators, health care providers and health information sources who have contributed to the many studies contributing to this manuscript.

Acknowledgements for individual studies: **AOV:** We thank Jennifer Koziak, Mie Konno, Michelle Darago, Faye Chambers and the Tom Baker Cancer Centre Translational Laboratories; **AUS:** The AOCS also acknowledges the cooperation of the participating institutions in Australia and acknowledges the contribution of the study nurses, research assistants and all clinical and scientific collaborators to the study. The complete AOCS Study Group can be found at www.aocstudy.org. We would like to thank all of the women who participated in these research programs; **BEL:** We would like to thank Gilian Peuteman, Thomas Van Brussel, Annick Van den Broeck and Joke De Roover for technical assistance; **BGS:** The BGS is funded by Breast Cancer Now and the Institute of Cancer Research (ICR). ICR acknowledges NHS funding to the NIHR Biomedical Research Centre. We thank the Study staff, study participants, doctors, nurses, health care providers and health information sources who have contributed to the study; **BVU:** The dataset(s) used for the analyses described were obtained from Vanderbilt University Medical Center's BioVU, which is supported by institutional funding, the 1S10RR025141-01 instrumentation award, and by the Vanderbilt CTSA grant UL1TR000445 from NCATS/NIH; **CAM:** This work was supported by Cancer Research UK; the University of Cambridge; National Institute for Health Research Cambridge Biomedical Research Centre; **CHA:** Innovative Research Team in University (PCSIRT) in China (IRT1076); **CHN:** To thank all members of Department of Obstetrics and Gynaecology, Hebei Medical University, Fourth Hospital and Department of Molecular Biology, Hebei Medical University, Fourth Hospital; **COE:** Gynecologic Cancer Center of Excellence (W81XWH-11-2-0131); **CON:** The cooperation of the 32 Connecticut hospitals, including Stamford Hospital, in allowing patient access, is gratefully acknowledged. This study was approved by the State of Connecticut Department of Public Health Human Investigation Committee. Certain data used in this study were obtained from the Connecticut Tumor Registry in the Connecticut Department of Public Health. The authors assume full responsibility for analyses and interpretation of these data; **DKE:** OCRF; **EPC:** To thank all members and investigators of the Rotterdam Ovarian Cancer Study. Dutch Cancer Society (EMC 2014-6699); **GER:** The German Ovarian Cancer Study (GER) thank Ursula Eilber for competent technical assistance; **HOC:** The study was supported by the Helsinki University Research Fund; **JGO:** JSPS KAKENHI grant; **KRA:** This study (Ko-EVE) was

supported by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), and the National R&D Program for Cancer Control, Ministry of Health & Welfare, Republic of Korea (HI16C1127; 0920010); **LUN:** ERC - 2011-AdG, Swedish Cancer Society, Swedish Research Council; **MAS:** We would like to thank Famida Zulkifli and Ms Moey for assistance in patient recruitment, data collection and sample preparation. The Malaysian Ovarian Cancer Genetic Study is funded by research grants from the Malaysian Ministry of Higher Education (UM.C/HIR/MOHE/06) and charitable funding from Cancer Research Initiatives Foundation; **MCC:** MCCS cohort recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was further supported by Australian NHMRC grants 209057, 251553 and 504711 and by infrastructure provided by Cancer Council Victoria. Cases and their vital status were ascertained through the Victorian Cancer Registry (VCR) and the Australian Institute of Health and Welfare (AIHW), including the National Death Index and the Australian Cancer Database; **MOF:** 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 (P30-CA076292), Merck Pharmaceuticals and the state of Florida; **NHS:** 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; **NJO:** L. Paddock, M. King, U. Chandran, A. Samoila, and Y. Bensman; **OPL:** Members of the OPAL Study Group (http://opalstudy.qimrberghofer.edu.au/); **RPC:** National Institute of Health (P50 CA159981, R01CA126841); **SEA:** SEARCH team, Craig Luccarini, Caroline Baynes, Don Conroy; **SIS:** The Sister Study (SISTER) is supported by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences (Z01-ES044005 and Z01-ES049033); **SON:** National Health Research and Development Program, Health Canada, grant 6613-1415-53; **SRO:** To thank all members of Scottish Gynaecological Clinical Trails group and SCOTROC1 investigators; **SWE:** Swedish Cancer Foundation, WeCanCureCancer and årKampMotCancer Foundation; **SWH:** The SWHS is supported primarily by NIH grant R37-CA070867. We thank the participants and the research staff of the Shanghai Women's Health Study for making this study possible; **UCI:** The UCI Ovarian Cancer Study is supported by the National Institutes of Health, National Cancer Institute grants CA58860, and the Lon V Smith Foundation grant LVS-39420; **UHN:** Princess Margaret Cancer Centre Foundation-Bridge for the Cure; **UKO:** We particularly thank I. Jacobs, M.Widschwendter, E. Wozniak, A. Ryan, J. Ford and N. Balogun for their contribution to the study**; UKR:** Carole Pye; **VAN:** BC Cancer Foundation, VGH & UBC Hospital Foundation; **WMH:** We thank the Gynaecological Oncology Biobank at Westmead, a member of the Australasian Biospecimen Network-Oncology group, which is funded by the National Health and Medical Research Council Enabling Grants ID 310670 & ID 628903 and the Cancer Institute NSW Grants 12/RIG/1-17 & 15/RIG/1-16.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

References

1. Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, et al. Environmental and heritable factors in the causation of cancer–analyses of cohorts of twins from Sweden, Denmark, and Finland. N Engl J Med 2000; 343: 78–85.

2. Song H, Ramus SJ, Tyrer J, Bolton KL, Gentry-Maharaj A, Wozniak E, et al.A genome-wide association study identifies a new ovarian cancer susceptibility locus on 9p22.2. Nat Genet 2009; 41: 996–1000.

3. Bolton KL, Tyrer J, Song H, Ramus SJ, Notaridou M, Jones C, et al. Common variants at 19p13 are associated with susceptibility to ovarian cancer. Nat Genet 2010; 42: 880–884.

4. Goode EL, Chenevix-Trench G, Song H, Ramus SJ, Notaridou M, Lawrenson K, et al. A genome-wide association study identifies susceptibility loci for ovarian cancer at 2q31 and 8q24. Nat Genet 2010; 42: 874–879.

5. Pharoah PD, Tsai YY, Ramus SJ, Phelan CM, Goode EL, Lawrenson K, et al. GWAS meta-analysis and replication identifies three new susceptibility loci for ovarian cancer. Nat Genet 2013; 45: 362–370.

6. Bojesen SE,Pooley KA, Johnatty SE, Beesley J, Michailidou K, Tyrer JP, et al. Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer. Nat Genet 2013; 45: 371–384.

7. Kuchenbaecker KB, Ramus SJ, Tyrer J, Lee A, Shen HC, Beesley J, et al. Identification of six new susceptibility loci for invasive epithelial ovarian cancer. Nat Genet 2015 ;47: 164–171.

8. Kelemen LE, Lawrenson K, Tyrer J, Li Q, Lee JM, Seo JH, et al. Genome-wide significant risk associations for mucinous ovarian carcinoma. Nat Genet 2015; 47: 888–897.

9. Shen H, Fridley BL, Song H, Lawrenson K, Cunningham JM, Ramus SJ, et al. Epigenetic analysis leads to identification of HNF1B as a subtype-specific susceptibility gene for ovarian cancer. Nat Commun 2013; 4:1628.

10. Permuth-Wey J, Lawrenson K, Shen HC, Velkova A, Tyrer JP, Chen Z, et al. Identification and molecular characterization of a new ovarian cancer susceptibility locus at 17q21.31. Nat Commun 2013; 4:1627.

11. Kar SP, Beesley J, Amin Al Olama A, Michailidou K, Tyrer J, Kote-Jarai Z, et al. Genome-wide meta-analyses of breast, ovarian, and prostate cancer association studies identify multiple new susceptibility loci shared by at least two cancer types. Cancer Discov 2016; 6: 1052–1067.

12. Phelan CM, Kuchenbaecker KB, Tyrer JP, Kar SP, Lawrenson K, Winham SJ, et al. Identification of 12 new susceptibility loci for different histotypes of epithelial varian cancer. Nat Genet 2017; 49:680–691.

13. Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. PLos Genet 2010; 6: e1000888.

14. Consortium GT. Human genomics. The genotype-tissue expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science 2015; 348: 648–660.

15. Westra HJ, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat Genet 2013; 45: 1238–1243.

16. Li Q, Seo JH, Stranger B, McKenna A, Pe'er I, Laframboise T, et al. Integrative eQTL-based analyses reveal the biology of breast cancer risk loci. Cell 2013; 152: 633–641.

17. Das S, Forer L, Schonherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. Nat Genet 2016; 48: 1284–1287.

18. Wheeler HE, Shah KP, Brenner J, Garcia T, Aquino-Michaels K, Consortium GT, et al. Survey of the heritability and sparse architecture of gene expression traits across human tissues. PLos Genet 2016; 12 e1006423.

19. Stegle O, Parts L, Piipari M, Winn J, Durbin R. Using probabilistic estimation of expression residuals (PEER) to obtain increased power and interpretability of gene expression analyses. Nat Protoc 2012; 7: 500–507.

20. Genomes Project C, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, et al. An integrated map of genetic variation from 1,092 human genomes. Nature 2012; 491:56–65.

21. Gamazon ER, Wheeler HE, Shah KP, Mozaffari SV, Aquino-Michaels K, Carroll RJ, et al. A gene-based association method for mapping traits using reference transcriptome data. Nat Genet 2015; 47: 1091–1098.

22. Barbeira AN, Dickinson SP, Bonazzola R, Zheng J, Wheeler HE, Torres JM, et al. Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. Nat Commun 2018 May 8; 9:1825. doi: 10.1038/s41467-018-03621-1.

23. Guo X, Lin M, Rockowitz S, Lachman HM, Zheng D. Characterization of human pseudogene-derived non-coding RNAs for functional potential. PLoS One 2014; 9: e93972.

24. Yang J, Ferreira T, Morris AP, Medland SE, Genetic Investigation of ATCReplication DIG, et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat Genet 2012; 44: 369–375.

25. Prat J. New insights into ovarian cancer pathology. Ann Oncol 2012; 23: x111–117.

26. Lawrenson K, Kar S, McCue K, Kuchenbaeker K, Michailidou K, Tyrer J, et al. Functional mechanisms underlying pleiotropic risk alleles at the 19p13.1 breast-ovarian cancer susceptibility locus. Nat Commun 2016; 7: 12675.

27. Boyer A, Goff AK, Boerboom D. WNT signaling in ovarian follicle biology and tumorigenesis. Trends Endocrinol Metab 2010; 21: 25–32.

28. Hsieh M, Boerboom D, Shimada M, Lo Y, Parlow AF, Luhmann UF, et al. Mice null for Frizzled4 (Fzd4-/-) are infertile and exhibit impaired corpora lutea formation and function. Biol Reprod 2005; 73: 1135–1146.

29. Zhan T, Rindtorff N, Boutros M. Wnt signaling in cancer. Oncogene 2017; 36: 1461–1473.

30. Couch FJ, Wang X, McGuffog L, Lee A, Olswold C, Kuchenbaecker KB, et al. Genome-wide association study in BRCA1 mutation carriers identifies novel loci associated with breast and ovarian cancer risk. PLos Genet 2013; 9: e1003212.

31. Stefansson H, Helgason A, Thorleifsson G, Steinthorsdottir V, Masson G, Barnard J, et al. A common inversion under selection in Europeans. Nat Genet 2005; 37: 129–137.

32. Liang B, Wei DL, Cheng YN, Yuan HJ, Lin J, Cui XZ, et al. Restraint stress impairs oocyte developmental potential in mice: role of CRH-induced apoptosis of ovarian cells. Biol Reprod 2013; 89: 64.

33. De Jong S, Chepelev I, Janson E, Strengman E, Van den Berg LH, Veldink JH, et al. Common inversion polymorphism at 17q21.31 affects expression of multiple genes in tissue-specific manner. BMC Genomics 2012; 13: 458.

34. Calogero AE, Burrello N, Negri-Cesi P, Papale L, Palumbo MA, Cianci A, et al. Effects of corticotropin-releasing hormone on ovarian estrogen production in vitro. Endocrinology 1996; 137: 4161–4166.

35. Yu C, Li M, Wang Y, Liu Y, Yan C, Pan J, et al. MiR-375 Mediates CRH signaling pathway in inhibiting E2 synthesis in porcine ovary. Reproduction 2017; 153: 63–73.

36. Ghizzoni L, Mastorakos G, Vottero A, Barreca A, Furlini M, Cesarone A, et al. Corticotropin-releasing hormone (CRH) inhibits steroid biosynthesis by cultured human granulosa-lutein cells in a CRH and interleukin-1 receptor-mediated fashion. Endocrinology 1997; 138: 4806–4811.

37. Barrett ES, Parlett LE, Windham GC, Swan SH. Differences in ovarian hormones in relation to parity and time since last birth. Fertil Steril 2014; 101: 1773–1780.

38. Kar SP, Tyrer JP, Li Q, Lawrenson K, Aben KK, Anton-Culver H, et al. Network-based integration of GWAS and gene expression identifies a HOX-centric network associated with serous ovarian cancer risk. Cancer Epidemiol Biomark Prev 2015; 24: 1574–1584.

39. Lindblad O, Chougule RA, Moharram SA, Kabir NN, Sun J, Kazi JU, et al. The role of HOXB2 and HOXB3 in acute myeloid leukemia. Biochem Biophys Res Commun 2015; 467: 742–747.

40. Inamura K, Togashi Y, Okui M, Ninomiya H, Hiramatsu M, Satoh Y, et al. HOXB2 as a novel prognostic indicator for stage I lung adenocarcinomas. J Thorac Oncol 2007; 2: 802–807.

41. Lopez R, Garrido E, Pina P, Hidalgo A, Lazos M, Ochoa R, et al. HOXB homeobox gene expression in cervical carcinoma. Int J Gynecol Cancer 2006; 16: 329–335.

42. Segara D, Biankin AV, Kench JG, Langusch CC, Dawson AC, Skalicky DA, et al. Expression of HOXB2, a retinoic acid signaling target in pancreatic cancer and pancreatic intraepithelial neoplasia. Clin Cancer Res 2005; 11: 3587–3596.

43. Boimel PJ, Cruz C, Segall JE. A functional in vivo screen for regulators of tumor progression identifies HOXB2 as a regulator of tumor growth in breast cancer. Genomics 2011; 98: 64–72.

44. Loges NT, Olbrich H, Becker-Heck A, Haffner K, Heer A, Reinhard C, et al. Deletions and point mutations of LRRC50 cause primary ciliary dyskinesia due to dynein arm defects. Am J Hum Genet 2009; 85: 883–889.

45. McComb P, Langley L, Villalon M, Verdugo P. The oviductal cilia and Kartagener's syndrome. Fertil Steril 1986; 46: 412–416.

46. Luo J, Emanuele MJ, Li D, Creighton CJ, Schlabach MR, Westbrook TF, et al. A genome-wide RNAi screen identifies multiple synthetic lethal interactions with the Ras oncogene. Cell 2009; 137: 835–848.

47. Evans JR, Feng FY, Chinnaiyan AM. The bright side of dark matter: lncRNAs in cancer. J Clin Invest 2016; 126: 2775–2782.

48. Van Leeuwen FE, Klip H, Mooij TM, van de Swaluw AM, Lambalk CB, Kortman M, et al. Risk of borderline and invasive ovarian tumours after ovarian stimulation for in vitro fertilization in a large Dutch cohort. Hum Reprod 2011; 26: 3456–3465.

49. Kessous R, Davidson E, Meirovitz M, Sergienko R, Sheiner E. The risk of female malignancies after fertility treatments: a cohort study with 25-year follow-up. J Cancer Res Clin Oncol 2016; 142: 287–293.

50. Stewart LM, Holman CD, Finn JC, Preen DB, Hart R. In vitro fertilization is associated with an increased risk of borderline ovarian tumours. Gynecol Oncol 2013; 129: 372–376.