Genome-wide analysis identifies novel loci associated with ovarian cancer outcomes: findings from the Ovarian Cancer Association Consortium

Sharon E. Johnatty[1*], Jonathan P. Tyrer[2*], Siddhartha Kar[2], Jonathan Beesley[1], Yi Lu[1], Bo Gao[3,4], Peter A. Fasching[5,6], Alexander Hein[7], Arif B. Ekici[6], Matthias W. Beckmann[7], Diether Lambrechts[8,9], Els Van Nieuwenhuysen[10], Ignace Vergote[10], Sandrina Lambrechts[10], Mary Anne Rossing[11,12], Jennifer A. Doherty[13], Jenny Chang-Claude[14], Francesmary Modugno[15-17], Roberta B. Ness[18], Kirsten B. Moysich[19], Douglas A. Levine[20], Lambertus A. Kiemeney[21], Leon F.A.G. Massuger[22], Jacek Gronwald[23], Jan Lubiński[23], Anna Jakubowska[23], Cezary Cybulski[23], Louise Brinton[24], Jolanta Lissowska[25], Nicolas Wentzensen[24], Honglin Song[26], Valerie Rhenius[26], Ian Campbell[27,28], Diana Eccles[29],Weiva Sieh[30], Alice S.Whittemore[30], Valerie McGuire[30], Joseph H. Rothstein[30], Rebecca Sutphen[31], Hoda Anton-Culver[32], Argyrios Ziogas[33], Simon A.Gayther[34], Aleksandra Gentry-Maharaj[35], Usha Menon[35], Susan J.Ramus[34], Celeste L Pearce[34,36], Malcolm C Pike[34,37], Daniel O. Stram[34], Anna H. Wu[34], Jolanta Kupryjanczyk[38], Agnieszka Dansonka-Mieszkowska[38], Iwona K. Rzepecka[38], Beata Spiewankiewicz[39], Marc T. Goodman[40,41], Lynne R. Wilkens[42], Michael E. Carney[42], Pamela J Thompson[40,41], Florian Heitz[43,44], Andreas du Bois[43,44], Ira Schwaab[45], Philipp Harter[43,44], Jacobus Pisterer[46], Peter Hillemanns[47] on behalf of the AGO Study Group, Beth Y. Karlan[48], Christine Walsh[48], Jenny Lester[48], Sandra Orsulic[48], Stacey J Winham[49], Madalene Earp[49], Melissa C. Larson[49], Zachary C. Fogarty[49], Estrid Høgdall[50,51], Allan Jensen[50], Susanne Kruger Kjaer[50,52], Brooke L. Fridley[53], Julie M. Cunningham[54], Robert A. Vierkant[49], Joellen M. Schildkraut[55,56], Edwin S. Iversen[57], Kathryn L. Terry[58,59], Daniel W. Cramer[58,59], Elisa V. Bandera[60], Irene Orlow[61], Tanja Pejovic[62,63], Yukie Bean[62,63], Claus Høgdall[64], Lene Lundvall[64], Ian McNeish[65], James Paul[66], Karen Carty[66], Nadeem Siddiqui[67], Rosalind Glasspool[66], Thomas Sellers[68], Catherine Kennedy[3,4], Yoke-Eng

Chiew[3,4], Andrew Berchuck[69], Stuart MacGregor[1], Anna deFazio[3,4], Paul D.P. Pharoah[2], Ellen L. Goode[49*], and Anna deFazio[3,4], Penelope M. Webb[70], Georgia Chenevix-Trench[1*] on behalf of the Australian Ovarian Cancer Study Group.

[*]These authors contributed equally

1. Department of Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia.
2. Department of Oncology, Department of Public Health and Primary Care, University of Cambridge, Strangeways Research Laboratory, Cambridge, UK.
3. Department of Gynaecological Oncology, Westmead Hospital, Sydney, NSW, Australia.
4. Center for Cancer Research, University of Sydney at Westmead Millennium Institute, Sydney, NSW, Australia.
5. University of California at Los Angeles, David Geffen School of Medicine, Department of Medicine, Division of Hematology and Oncology, Los Angeles, CA, USA.
6. University Hospital Erlangen, Institute of Human Genetics, Friedrich-Alexander-University Erlangen-Nuremberg, Erlangen, Germany.
7. University Hospital Erlangen, Department of Gynecology and Obstetrics, Friedrich-Alexander-University Erlangen-Nuremberg, Comprehensive Cancer Center Erlangen-EMN, Universitaetsstrasse 21-23, 91054 Erlangen, Germany.
8. Vesalius Research Center, VIB, Leuven, Belgium.
9. Laboratory for Translational Genetics, Department of Oncology, University of Leuven, Leuven, Belgium.
10. Department of Gynecologic Oncology, Leuven Cancer Institute, University of Leuven, Leuven, European Union.
11. Program in Epidemiology, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA.
12. Department of Epidemiology, University of Washington, Seattle, WA, USA.
13. Department of Community and Family Medicine, Section of Biostatistics & Epidemiology, The Geisel School of Medicine at Dartmouth, Lebanon, NH, USA.
14. German Cancer Research Center (DKFZ), Division of Cancer Epidemiology, Heidelberg, Germany.
15. Department of Obstetrics, Gynecology and Reproductive Sciences, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA.
16. Department of Epidemiology, University of Pittsburgh Graduate School of Public Health, Pittsburgh, PA, USA.
17. Women's Cancer Research Program, Magee-Women's Research Institute and University of Pittsburgh Cancer Institute, Pittsburgh, PA, USA.
18. The University of Texas School of Public Health, Houston, TX, USA.
19. Department of Cancer Prevention and Control, Roswell Park Cancer Institute, Buffalo, NY, USA.
20. Gynecology Service, Department of Surgery, Memorial Sloan-Kettering Cancer Center, New York, NY, USA.

21. Radboud University Medical Centre, Radboud Institute for Health Sciences, Nijmegen, The Netherlands
22. Radboud University Medical Centre, Radboud Institute for Molecular Sciences, Nijmegen, The Netherlands.
23. International Hereditary Cancer Center, Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland.
24. Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA.
25. Department of Cancer Epidemiology and Prevention, Maria Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland.
26. Department of Oncology, University of Cambridge, Strangeways Research Laboratory, Cambridge, UK.
27. Cancer Genetics Laboratory, Research Division, Peter MacCallum Cancer Centre, Melbourne, VIC, Australia.
28. Department of Oncology, The University of Melbourne, Melbourne, Victoria, Australia.
29. Faculty of Medicine, University of Southampton, University Hospital Southampton, Southampton, Hampshire, UK.
30. Department of Health Research and Policy - Epidemiology, Stanford University School of Medicine, Stanford, CA, USA.
31. Epidemiology Center, College of Medicine, University of South Florida, Tampa, Florida, USA.
32. Department of Epidemiology, Center for Cancer Genetics Research and Prevention, School of Medicine, University of California Irvine, Irvine, CA, USA.
33. Department of Epidemiology, University of California Irvine, Irvine, CA, USA.
34. Department of Preventive Medicine, Keck School of Medicine, University of Southern California Norris Comprehensive Cancer Center, Los Angeles, CA, USA.
35. Women's Cancer, UCL EGA Institute for Women's Health, London, UK.
36. Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, Michigan, USA
37. Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, NY, USA.
38. Department of Pathology and Laboratory Diagnostics, The Maria Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland.
39. Department of Gynecologic Oncology, The Maria Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland.
40. Cancer Prevention and Control, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA.
41. Community and Population Health Research Institute, Department of Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, CA, USA.
42. Cancer Epidemiology Program, University of Hawaii Cancer Center, Honolulu, HI, USA.
43. Department of Gynecology and Gynecologic Oncology, Kliniken Essen-Mitte, Essen, Germany.
44. Department of Gynecology and Gynecologic Oncology, Dr. Horst Schmidt Kliniken Wiesbaden, Wiesbaden, Germany.
45. Institut für Humangenetik Wiesbaden, Wiesbaden, Germany.
46. Zentrum für Gynäkologische Onkologie, Kiel, Germany.

47. Department of Obstetrics and Gynaecology, Hannover Medical School, Hannover, Germany.
48. Women's Cancer Program at the Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA.
49. Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA.
50. Department of Virus, Lifestyle and Genes, Danish Cancer Society Research Center, Copenhagen, Denmark.
51. Molecular Unit, Department of Pathology, Herlev Hospital, University of Copenhagen, Copenhagen, Denmark.
52. Department of Gynecology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark.
53. Biostatistics and Informatics Shared Resource, University of Kansas Medical Center, Kansas City, KS, USA.
54. Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA.
55. Department of Community and Family Medicine, Duke University Medical Center, Durham, NC, USA.
56. Cancer Control and Population Sciences, Duke Cancer Institute, Durham, NC, USA.
57. Department of Statistical Science, Duke University, Durham, NC, USA.
58. Obstetrics and Gynecology Epidemiology Center, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA.
59. Harvard School of Public Health, Boston, MA, USA.
60. Cancer Prevention and Control Program, Rutgers Cancer Institute of New Jersey, The State University of New Jersey, New Brunswick, NJ, USA.
61. Memorial Sloan Kettering Cancer Center, Department of Epidemiology and Biostatistics, Epidemiology Service, New York, NY, USA.
62. Department of Obstetrics and Gynecology, Oregon Health and Science University, Portland, OR, USA.
63. Knight Cancer Institute, Portland, OR, USA.
64. Department of Gynaecology, The Juliane Marie Centre, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark.
65. Institute of Cancer Sciences, University of Glasgow, Wolfson Wohl Cancer Research Centre, Beatson Institute for Cancer Research, Glasgow, UK.
66. Cancer Research UK Clinical Trials Unit, Glasgow, The Beatson West of Scotland Cancer Centre, Glasgow, UK.
67. Department of Gynaecological Oncology, Glasgow Royal Infirmary, Glasgow, UK.
68. Department of Cancer Epidemiology, Moffitt Cancer Center, Tampa, FL, USA.
69. Department of Obstetrics and Gynecology, Duke University Medical Center, Durham, NC, USA.
70. Department of Population Health, QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia.

Financial Disclosure

AUS studies (Australian Ovarian Cancer Study and the Australian Cancer Study) were funded by Army Medical Research and Materiel Command (DAMD17-01-1-0729), National Health & Medical Research Council of Australia, Cancer Councils of New South Wales, Victoria, Queensland, South Australia and Tasmania, Cancer Foundation of Western Australia; National Health and Medical Research Council of Australia (199600 and 400281). The grant numbers for AOCS Cancer Council funding are as follows- Multi-State Application Numbers 191, 211 and 182. The Bavarian study (BAV) was supported by ELAN Funds of the University of Erlangen-Nuremberg. The Belgian study (BEL) was funded by Nationaal Kankerplan. The Diseases of the Ovary and their Evaluation (DOV) study was funded by National Institutes of Health R01-CA112523 and R01-CA87538. The German Ovarian Cancer Study (GER) was supported by the German Federal Ministry of Education and Research of Germany, Programme of Clinical Biomedical Research (01 GB 9401) and the German Cancer Research Center (DKFZ). The Hawaii Ovarian Cancer Study (HAW) was supported by R01 CA 058598. The Hormones and Ovarian Cancer Prediction study (HOP) was supported by US National Cancer Institute: K07-CA80668; R01CA095023; P50-CA159981; R01-CA126841; US Army Medical Research and Materiel Command: DAMD17-02-1-0669; NIH/National Center for Research Resources/General Clinical Research Center grant MO1- RR000056. The Women's Cancer Program (LAX) was supported by the American Cancer Society Early Detection Professorship (120950-SIOP-06-258-06-COUN) and the National Center for Advancing Translational Sciences (NCATS), Grant UL1TR000124. The Mayo Clinic Case-Only Ovarian Cancer Study (MAC) was funded by the National Institutes of Health (R01-CA122443, P30-CA15083, P50-CA136393). The Mayo Clinic Ovarian Cancer Case-Control Study (MAY) was supported by National Institutes of Health (R01-CA122443, P30-CA15083, P50-CA136393); Mayo Foundation; Minnesota Ovarian Cancer Alliance; Fred C. and Katherine B. Andersen Foundation. The MALOVA study (MAL) was funded by The National Cancer Institute (grant RO1-CA 61107), The Danish Cancer Society (grant (94-222-52) and The Mermaid I project. The North Carolina Ovarian Cancer Study (NCO) National Institutes of Health (R01-CA76016) and the Department of Defense (DAMD17-02-1-0666). The New England-based Case-Control Study of Ovarian Cancer (NEC) was supported by NIH grants R01 CA 054419-10 and P50 CA105009, and Department of Defense CDMRP grant W81XWH-10-1-0280. The New Jersey Ovarian Cancer Study (NJO) was funded by the National Cancer Institute (NIH-K07 CA095666, R01-CA83918, NIH-K22-CA138563, and P30-CA072720) and the Cancer Institute of New Jersey. The Oregon study (ORE) was funded by the Sherie Hildreth Ovarian Cancer Research Fund and the OHSU Foundation. The Polish Ovarian Cancer Case Control Study (POL) was funded by Intramural Research Program of the National Cancer Institute. The SEARCH study (SEA) was supported by Cancer Research UK (C490/A8339, C490/A10119, C490/A10124 and C490/A16561) and UK National Institute for Health Research Biomedical Research Centre at the University of Cambridge. The Scottish Randomised Trial in Ovarian Cancer (SRO) was funded by Cancer Research UK (C536/A13086, C536/A6689) and Imperial Experimental Cancer Research Centre (C1312/A15589). The Gynaecological Oncology Biobank at Westmead (WMH) is a member of the Australasian Biospecimen Network-Oncology group, funded by the Australian National Health and Medical Research Council Enabling Grants ID 310670 & ID 628903 and

the Cancer Institute NSW Grant ID 12/RIG/1-17. The United Kingdom Ovarian cancer Population Study (UKO) 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. The UK Familial Ovarian Cancer Registry (UKR) Cancer Research UK (C490/A6187); UK National Institute for Health Research Biomedical Research Centres at the University of Cambridge. The Los Angeles County Case-Control Studies of Ovarian Cancer-3 (USC) P01CA17054, P30CA14089, R01CA61132, N01PC67010, R03CA113148, R03CA115195, N01CN025403, and California Cancer Research Program (00-01389V-20170, 2II0200). The Warsaw Ovarian Cancer Study (WOC) Polish Ministry of Science and Higher Education (4 PO5C 028 14, 2 PO5A 068 27), The Maria Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland.

Anna deFazio was funded by the University of Sydney Cancer Research Fund and the Cancer Institute NSW through the Sydney West-Translational Cancer Research Centre. Dr. Beth Y. Karlan is supported by American Cancer Society Early Detection Professorship (SIOP-06-258-01-COUN) and the National Center for Advancing Translational Sciences (NCATS), Grant UL1TR000124. Irene Orlow was supported by NCI CCSG award (P30-CA008748).

Funding for the iCOGS infrastructure came from: the European Community's Seventh Framework Programme under grant agreement n° 223175 (HEALTH-F2-2009-223175) (COGS), Cancer Research UK (C1287/A10118, C1287/A 10710, C12292/A11174, C1281/A12014, C5047/A8384, C5047/A15007, C5047/A10692, C8197/A16565), the National Institutes of Health (CA128978) and Post-Cancer GWAS initiative (1U19 CA148537, 1U19 CA148065 and 1U19 CA148112 - the GAME-ON initiative), the Department of Defence (W81XWH-10-1-0341), the Canadian Institutes of Health Research (CIHR) for the CIHR Team in Familial Risks of Breast Cancer, Komen Foundation for the Cure, the Breast Cancer Research Foundation, and the Ovarian Cancer Research Fund.

Running title: Genome-wide analysis of ovarian cancer outcomes

Keywords: progression-free survival, overall survival, epithelial ovarian cancer, lncRNA, chemotherapy

Correspondence to:

Georgia Chenevix-Trench
QIMR Berghofer Medical Research Institute,
300 Herston Road,
Herston 4006, Australia.
Ph: +61 7 3362 0390.
Email: georgia.trench@qimrberghofer.edu.au

There are no conflicts of interest to disclose

Abstract word count: 246
Main manuscript word count excluding references: 4,423
Tables & Figures: 6
References: 50

Translational Relevance

Although several genetic loci have been identified for ovarian cancer risk, finding loci associated with outcome remains a challenge primarily because of treatment heterogeneity and small sample sizes. We comprehensively analyzed ~2.8 million variants in the largest collection to date of epithelial ovarian cancer cases with detailed chemotherapy and clinical follow-up data, and identified SNPs in three long non-coding RNAs (lncRNAs) that were associated with progression-free survival, one of which lies within a super-enhancer recently shown to be associated with poor prognosis in another solid tumor. There is a growing body of evidence that lncRNAs are cancer-specific regulators in signalling pathways underlying metastasis and disease progression. While additional work is needed to delineate the role of associated SNPs on lncRNA expression and validate their role in a larger sample, our findings have important implications for the development of diagnostic markers of progression and novel therapeutic targets for epithelial ovarian cancer.

Abstract

Purpose: Chemotherapy resistance remains a major challenge in the treatment of ovarian cancer. We hypothesize that germline polymorphisms might be associated with clinical outcome.

Experimental Design: We analyzed ~2.8 million genotyped and imputed SNPs from the iCOGS experiment for progression-free survival (PFS) and overall survival (OS) in 2,901 European epithelial ovarian cancer (EOC) patients who underwent firstline treatment of cytoreductive surgery and chemotherapy regardless of regimen, and in a subset of 1,098 patients treated with ≥4 cycles of paclitaxel and carboplatin at standard doses. We evaluated the top SNPs in 4,434 EOC patients including patients from The Cancer Genome Atlas. Additionally we conducted pathway analysis of all intragenic SNPs and tested their association with PFS and OS using gene set enrichment analysis.

Results: Five SNPs were significantly associated ($p \leq 1.0 \times 10^{-5}$) with poorer outcomes in at least one of the four analyses, three of which, rs4910232 (11p15.3), rs2549714 (16q23) and rs6674079 (1q22) were located in long non-coding RNAs (lncRNAs) RP11-179A10.1, RP11-314O13.1 and RP11-284F21.8 respectively ($p \leq 7.1 \times 10^{-6}$). ENCODE ChIP-seq data at 1q22 for normal ovary shows evidence of histone modification around RP11-284F21.8, and rs6674079 is perfectly correlated with another SNP within the super-enhancer MEF2D, expression levels of which were reportedly associated with prognosis in another solid tumor. YAP1- and WWTR1 (TAZ)-stimulated gene expression, and HDL-mediated lipid transport pathways were associated with PFS and OS, respectively, in the cohort who had standard chemotherapy ($p_{GSEA} \leq 6 \times 10^{-3}$).

Conclusion: We have identified SNPs in three lncRNAs that might be important targets for novel EOC therapies.

Introduction

Approximately 238,000 women are diagnosed with ovarian cancer each year. It is the leading cause of death from gynecological cancers and globally approximately 152,000 women will die annually from the disease (1). Over the past three decades, significant advances have been made in chemotherapy for epithelial ovarian cancer (EOC), and the combination of cytoreductive surgery followed by the doublet of a taxane (paclitaxel 135 – 175 mg/m^2) and platinum (carboplatin AUC > 5) repeated every three weeks has been the most common regimen for primary treatment of this disease, with initial tumor response rates ranging from 70-80% (2, 3). Although survival rates have improved in the past decade, resistance to chemotherapy remains a major challenge, and the majority of patients with advanced disease succumb to the disease despite initial response to first line treatment (4). The identification of genes relevant to response to chemotherapy and survival of ovarian cancer may contribute to a better understanding of prognosis, and potentially guide the selection of treatment options to help circumvent this obstacle.

It is well recognized that genetic variation can have a direct effect on inter-individual variation in drug responses, although patient response to medication is dependent on multiple factors ranging from patient age, disease type, organ functions, concomitant therapy and drug interactions (5). Comparisons of intra-patient and inter-patient variability in both population-based and twin studies have demonstrated that the smallest differences in drug metabolism and their effects are between monozygotic twins, which is consistent with the hypothesis that genetics may play a significant role in drug responses (6, 7). While many cancer treatments have been successful in shrinking or eradicating tumor cells, studies of genetic factors related to drug responses are particularly challenging because tumor cell and the non-cancerous host tissue from which they arise share the same genetic background, and failure of treatment may

be due to the presence of *de novo* or acquired somatic alterations in tumors rather than germline variation (8).

To date several candidate gene studies have explored germline polymorphisms for an association with response to chemotherapy for ovarian cancer (9). Some obvious candidates are genes that encode drug-metabolizing enzymes and drug transporters that can influence toxicity or treatment response. The most clinically relevant drug metabolising enzymes are member of the cytochrome P450 (CYP) superfamily, of which CYP1, CYP2, and CYP3 contribute to the metabolism of more than 90% of clinically used drugs. There is considerable evidence that polymorphisms in the CYP genes have a significant impact on drug disposition and response, and >60% of Food and Drug Administration (FDA)-approved drug labels regarding genomic biomarkers pertain to polymorphisms in the CYP enzymes (10). Similarly the *ABCB1* gene, the most extensively studied ATP-binding cassette (ABC) transporter involved in transport of a wide range of anti-cancer drugs including paclitaxel (11), was previously shown to be associated with response to first-line paclitaxel-based chemotherapy regimens for ovarian cancer (12, 13). A systematic review of the most commonly evaluated genes in gynecologic cancers, including *ABCB1*, showed inconsistent findings across studies (14). Other studies including a comprehensive study of *ABCB1* SNPs putatively associated with progression-free survival (PFS) undertaken by the Ovarian Cancer Association Consortium (OCAC) did not replicate the association with PFS, although the possibility of subtle effects from one SNP on overall survival (OS) could not be discounted (13). Recently several ABCA transporters were explored in expression studies using cell-based models and shown to be associated with outcome in serous EOC patients (15), although this finding would need to be replicated in a larger independent study.

However, inter-individual variation in response to chemotherapy and post-treatment outcomes cannot be fully explained by genetic variations in the genes encoding drug metabolizing enzymes, transporters, or drug targets. Recent studies by the OCAC and the Australian Ovarian Cancer Study (AOCS) found that EOC patients carrying *BRCA1* or *BRCA2* germline mutations had better response to treatment and better short-term survival (5 years) than non-carriers (16, 17). This survival advantage is supported by *in vitro* studies of *BRCA1/2* mutated ovarian cancer cell lines that were shown to be more sensitive to platinum-based chemotherapy (18, 19). Genome-wide approaches that integrate SNP genotypes, drug-induced cytotoxicity in cell lines and gene expression data have been proposed as models for identifying predictors of treatment outcome (20), although their utility when applied to patient data proved inconclusive (21).

While *in vitro* studies have suggested functional relevance for genes and associated SNPs, the clinical utility of these findings remains in question mainly due to inconsistent results from under-powered and heterogeneous patient studies. In this report we present the findings from a comprehensive large-scale analysis of ~2.8 million genotyped and imputed SNPs from the Collaborative Oncological Gene-environment Study (COGS) project in relation to progression-free and overall survival as surrogate markers of response to chemotherapy in ~3,000 EOC patients with detailed first-line chemotherapy and follow-up data from the OCAC. In a secondary analysis, we also evaluated the association between OS and ~2.8 million SNPs in ~11,000 EOC patients irrespective of treatment regimen.

Materials and methods

Study Populations

The main analysis was restricted to invasive EOC patients with detailed chemotherapy and clinical follow-up for disease progression and survival following first-line treatment from

thirteen OCAC studies in the initial phase, with an additional four OCAC studies and patients from The Cancer Genome Atlas (TCGA) included in the validation phase (Supplementary Tables 1). Patients were included if they received a minimum of cytoreductive surgery as part of primary treatment, and were of European ancestry, determined using the program LAMP (22) to assign intercontinental ancestry based upon a set of unlinked markers also used to perform principal component (PC) analysis within each major population subgroup (23). A total of 2,901 patients were eligible for the main analysis, a subset of whom (n=1,098) were treated with ≥4 cycles of standard doses of paclitaxel and carboplatin intravenously (IV) at 3-weekly intervals. Clinical definitions and criteria for progression across studies have been previously described (13). Data from TCGA (http://cancergenome.nih.gov/) was downloaded through the TCGA data portal and assessed for ancestral outliers to determine those of European descent. A secondary analysis of OS in ~11,000 European EOC patients was also done using patients from 30 OCAC studies (Supplementary Table 2). All studies received approval from their respective human research ethics committees, and all OCAC participants provided written informed consent.

Genotyping and imputation

The Collaborative Oncological Gene-environment Study (COGS) and two ovarian cancer GWAS have been described in detail elsewhere (24). Briefly, 211,155 SNPs were genotyped in germline DNA from cases and controls from 43 studies participating in OCAC using a custom Illumina Infinium iSelect array (iCOGS) designed to evaluate genetic variants for association with risk of breast, ovarian and prostate cancers. In addition, two new ovarian cancer GWAS were included which used Illumina 2.5M and Illumina OmniExpress arrays. Genotypes were imputed to the European subset of the phased chromosomes from the 1000 Genome project (version 3). Approximately 8 million SNPs with a minor allele frequency

(MAF) of at least 0.02 and an imputation r^2>0.3 were available for analysis, ~2.8 million of which were well imputed (imputation $r^2 \geq 0.9$) and were retained in survival analyses. DNA extraction, iPLEX genotyping methods and quality assurance for additional samples genotyped for the validation analysis have also been previously described (25).

Statistical Analysis

The main analyses were the association between ~2.8 million SNPs and progression-free survival (PFS) and overall survival (OS). Analyses of PFS and OS were conducted separately for all patients known to have had a minimum of cytoreductive surgery for first-line treatment regardless of chemotherapy, hereafter referred to as the 'all chemo' analysis, and in a subset of patients known to have received standard of care first-line treatment of cytoreductive surgery and ≥4 cycles of paclitaxel and carboplatin IV at 3-weekly intervals, hereafter referred to as the 'standard chemo' subgroup (Supplementary Table 1). The majority of patients in the 'standard chemo' cohort were known to have had paclitaxel at 175 or 135 mg/m^2 and carboplatin AUC 5 or 6; for the remainder, standard dose was assumed based on treatment schedules. PFS was defined as the interval between the date of histological diagnosis and the first confirmed sign of disease progression or death, as previously described (13); OS was the interval between the date of histological diagnosis and death from any cause. Patients who had an interval of >12 months between the date of histological diagnosis and DNA collection were excluded from the analysis to avoid survival bias. A secondary analysis was OS in the largest available dataset of European invasive EOC patients regardless of treatment (n=11,311), hereafter referred to 'all OCAC'.

For the main analysis of PFS and OS in 'all chemo' and 'standard chemo', we obtained the per-allele hazard ratio [log(HR)] and standard error for each SNP using Cox regression models including study, the first two PCs, residual disease (nil vs. any), tumor stage (FIGO

stages I-IV), histology (5 subtypes), tumor grade (low vs. high), and age at diagnosis (OS analysis only) as covariates. To avoid inflation for rare SNPs, the likelihood ratio test was used to estimate the standard error for iCOGS SNPs and meta-analyzed with samples included in the US GWAS and U19 studies based on expected imputation accuracy for imputed SNPs. For secondary analysis of OS in the 'all OCAC' dataset, Cox regression models included study, age, and the first two PCs and histology as covariates. For the US GWAS and U19 studies, the principal components were estimated separately and the top two and top principal components used respectively. All tests for association were two-tailed and performed using in-house software programmed in C++ and STATA SE v. 11 (Stata Corp., USA). Manhattan and QQ plots were generated using the R project for Statistical Computing version 3.0.1 (http://www.r-project.org/), and meta-analysis was done using the program Metal (26), and between-study heterogeneity was assessed using the likelihood ratio test to compare regression models with and without a genotype-by-study interaction term.

SNP selection for validation

Preliminary analyses suggested that dosage scores from imputed SNPs with imputation r^2 <0.9 were not representative of actual genotypes in this sample (Supplementary Methods & Supplementary Table 3). We therefore selected SNPs with imputation $r^2 \geq 0.9$ and adjusted $p \leq 1.0 \times 10^{-5}$ in at least one of the four main analyses (PFS and OS in 'all chemo' and 'standard chemo') for genotype validation. SNPs were binned into LD blocks defined by pairwise correlation (r^2) > 0.8. We used Sequenom Assay Designer 4.0 to design two multiplexes in order to capture at least one SNP representing each block, although some blocks contained SNPs for which an iPLEX assay could not be designed (n=10). All patients for whom we had DNA, clinical follow-up and chemotherapy data were genotyped. We then meta-analyzed estimates from the genotyped samples with non-overlapping iCOGS samples and

TCGA data to obtain effect estimates from the largest possible dataset. SNPs that were significant at $p \leq 1.0 \times 10^{-5}$ in at least one outcome in the final analysis were queried for association with expression of protein-coding genes within 1Mb of the lead SNP using GEO, EGA and TCGA expression array data analyzed in KM-plotter (27).

Pathway analysis

All intragenic SNPs of the ~8 million (MAF ≥ 0.02 and imputation $r^2 > 0.3$) with p-values for association with PFS and OS in the 'standard chemo' cohort were mapped to 25,004 genes annotated with hg19 start and end positions. The boundaries of each gene were extended by 50 kb on both sides for SNP-to-gene mapping to include *cis*-regulatory variation. A total of 23,490 genes were captured by at least one SNP. The negative logarithm (base 10) of the p-value of the most significant SNP in each gene, adjusted for the number of SNPs in the gene (±50 kb) by a modification of the Sidak correction (28, 29) was used to rank genes based on their association with PFS and OS ('standard chemo'). A total of 837 known biological pathways (containing between 15 to 500 genes each) from the Kyoto Encyclopedia of Genes and Genomes (KEGG), BioCarta, and Reactome, three standard expert-curated pathway repositories, were accessed via the Molecular Signatures Database (version 4.0; http://www.broadinstitute.org/gsea/msigdb). The pathways were tested for their association with PFS and OS using gene set enrichment analysis (GSEA) run to 1,000 permutations (30). Specifically, we applied the "preranked" GSEA algorithm with default settings and the original GSEA implementation of correction for testing multiple pathways using false discovery (FDR) and familywise error rates (FWER). The genes in each pathway driving the GSEA signal (core genes) were defined as described previously (30).

Results

SNP associations

An overview of the analytic approaches in this study is provided in Supplementary Figure 1. There were 158 and 236 SNPs in analysis of OS in 'all chemo' and 'standard chemo' respectively, and 107 and 252 SNPs in analysis of PFS in 'all chemo' and 'standard chemo' that were above the minimal p-value threshold for suggestive significance ($p=1.0x10^{-5}$) but none reached the nominal level of genome-wide significance ($p= 5\times10^{-8}$; Figure 1). QQ plots and estimates of inflation of the test statistic (λ) revealed some inflation ($\lambda \leq 1.15$; Supp. Figure 2) which could not be accounted for by SNPs with low MAF (<0.1). Manhattan and QQ plots for the 'all OCAC' OS analysis showed similar effects (Supplementary Figure 3). We selected 130 iCOGS SNPs with imputation $r^2 \geq 0.9$ and adjusted $p \leq 1.0x10^{-5}$ in at least one of the four analyses (Supplementary Table 4), and genotyped 48 SNPs at 22 loci in all patients with chemotherapy and outcome data. To obtain effect estimates from the largest possible sample for PFS and OS in 'all chemo' and 'standard chemo' for these 48 SNPs, we meta-analyzed estimates from iPLEX genotyped samples (n=3,303), iCOGS imputed data on non-overlapping samples (n=821), and TCGA data (n=310; Supplementary Table 5).

Estimates for the most promising SNPs from meta-analysis ($p \leq 1.0x10^{-5}$ in at least one of the four analyses) are summarized in Table 1. The strongest association was for rs4910232 at 11p15.3 and PFS in the 'all chemo' analysis (HR=1.17, 95% CI 1.10-1.24; $p=4.7x10^{-7}$). The Kaplan Meier (KM) plot of genotyped samples for rs4910232 showed a significant trend in worse PFS associated with each additional minor allele (Figure 2A) and there was no evidence of between-study heterogeneity ($p= 0.7$, Figure 2B). This SNP lies within the long non-coding RNA (lncRNA) RP11-179A10.1. Two other SNPs, rs2549714 at 16q23 and rs6674079 at 1q22 were associated with worse OS in 'standard chemo' ($p=5.0x10^{-6}$) and 'all chemo' analyses ($p=7.1x10^{-6}$) respectively, and are also located in lncRNAs (Table 1). We further explored SNPs within a 1Mb region of rs6674079 at the 1q22 locus using ENCODE ChiP-Seq data and found that rs6674079 is perfectly correlated with rs11264489 which lies

within the super-enhancer MEF2D. Histone modification tracks from ENCODE for normal ovarian cancer cell lines suggest a strong regulatory potential for this SNP (Figure 3). The KM plot for rs6674079 clearly showed a significant per-allele trend in worse OS (Figure 4A) and study-specific estimates and heterogeneity tests showed no evidence of between-study heterogeneity (p=0.4, Figure 4B). Forest plots for other significant SNPs (rs7950311, rs2549714 and rs3795247) showed an overall trend in worse survival probabilities per minor allele (Supplementary Figure 4A-C) and there was no evidence of between-study heterogeneity for any of these SNPs ($p \geq 0.14$).

We further queried protein-coding genes within a 1Mb region of each of these lead SNPs at 1q22, 11p15.4, 11p15.3, 16q23 and 19p12 (Table 1) using KM-plotter to identify gene expressions that might be associated with PFS and OS using all available data (1,170 and 1,435 patients respectively), and in a subset of cases restricted to optimally debulked serous cases treated with Taxol and platin chemotherapy (330 and 387 patients respectively). Of a total of 55 expression probes for 174 genes queried across the five loci, significant associations that met our Bonferroni-corrected significance threshold of $p \leq 2.3 \times 10^{-4}$ were observed for 11 probes in at least one analysis (Supplementary Table 6). The strongest association with outcome was observed for PFS and high (defined as above the median) expression of *SLC25A44* (probe 32091_at) in the unrestricted dataset of 1,170 ovarian cancer patients (HR=1.56, 95% CI 1.33-1.82, log-rank $p=1.9 \times 10^{-8}$; Supplementary Figure 5A). This association was upheld, although more weakly, in the subset restricted to optimally debulked serous cases treated with Taxol and platin chemotherapy (n=330, HR=1.66, 95% CI 1.24-2.23, log-rank p-value=6.8×10^{-4}). High expression of *SEMA4A* (probe 219259_at) was significantly associated with better PFS in the unrestricted dataset (HR 0.71, 95% CI 0.61 - 0.82, log-rank $p=4.2 \times 10^{-6}$; Supplementary Figure 5B) and marginally with OS (unrestricted dataset log-rank $p=3.3 \times 10^{-4}$ and restricted dataset log-rank $p=5.7 \times 10^{-4}$). Significantly better

PFS was also observed for high expression of *SH2D2A* (probe 207351_s_at) in the unrestricted datasets (HR=0.67, 95% CI 0.57 - 0.77, log-rank $p=8.4x10^{-8}$; Supplementary Figure 5C) with a marginal association for OS in the unrestricted dataset (log-rank $p=8.7x10^{-4}$).

We also evaluated associations between OS and SNPs in the larger 'all OCAC' dataset with minimal adjustment. A total of 70 SNPs with imputation $r^2 \geq 0.9$ at 4 loci achieved a $p \leq 1.0x10^{-5}$ (Supplementary Table 7). The top SNP was rs2013459 (HR=1.14, 95% CI 1.08-1.20, $p= 9.7x10^{-7}$ at *PARK2* located at 6q26. Significant SNPs were also identified at *FAR1* (11p15), *ANKLE1*, *BABAM1* and *ABHD8* (all at 19p13) and *SYNE2* (6q25).

Pathway Analysis

We also explored the polygenic signal in our data using pathway-based analysis. This enrichment analysis of genome-wide single-variant summary statistics from the 'standard chemo' subgroup in the context of known biological pathways suggested heterogeneity in the pathways that may be associated with PFS and OS. Eight of the 837 pathways tested were associated with PFS in the 'standard chemo' dataset at nominal significance ($p_{GSEA}<0.05$ and FWERGSEA<1), with the "YAP1- and WWTR1 (TAZ)- stimulated gene expression" pathway from the Reactome pathway database emerging as the most significant ($p_{GSEA}=1x10^{-3}$, $FDR_{GSEA}=0.868$, $FWER_{GSEA}=0.575$, Table 2). Nine of the 837 pathways were associated with OS in the 'standard chemo' data set at the same threshold for nominal significance and the Reactome pathway "HDL-mediated lipid transport" was the top pathway ($p_{GSEA}=6x10^{-3}$, $FDR_{GSEA}=0.303$, $FWER_{GSEA}=0.268$, Table 2). Interestingly, the other nominally significant pathways suggested possible involvement of cell cycle genes in determining PFS and of xenobiotic and insulin metabolism genes in determining OS in the 'standard chemo' cohort (Table 2).

Discussion

We have evaluated ~2.8 million SNPs across the genome for an association with outcome following first-line chemotherapy in a large cohort of EOC patients and identified SNPs at five loci with p-values that ranged from 1.05×10^{-5} to 4.7×10^{-7}. Three SNPs, rs6674079, rs4910232 and rs2549714, were located in long non-coding RNAs (lncRNA) RP11-284F21.8, RP11-179A10.1 and RP11-314O13.1 respectively (Table 1). LncRNAs are RNA transcripts that have been implicated in a wide range of regulatory functions including epigenetic control and regulation of chromatin structure at the cellular level to tumor suppressors and regulators of angiogenesis and metastasis (31). It has been shown that alterations in the function of some lncRNAs, particularly those involved in transcriptional regulation, can play a critical role in cancer progression and exert its effect on genes located on other chromosomes. A well characterized example of this is the lncRNA *HOTAIR* which has been linked to invasiveness and poor prognosis of breast cancer (32). *HOTAIR* is expressed from the *HOXC* gene cluster on chromosome 12, and has been shown to mediate repression of transcription of *HOXD* genes on chromosome 2 via PRC2 (33). While little is known about the specific lncRNAs that we have identified or their target genes, it is likely that associated SNPs in these lncRNAs might exert their effects on chromatin modifying proteins that regulate genes involved in ovarian cancer progression. ENCODE ChIP-seq data for normal ovarian cell lines at the 1q22 locus shows evidence of histone modification in the region of RP11-284F21.8, and rs6674079 at this locus is perfectly correlated with rs11264489 which lies within the super-enhancer MEF2D (Figure 4). Expression studies of MEF2D in hepatocellular carcinoma showed that elevated expression promoted cancer cell growth and was correlated with poor prognosis in patients (34). Further analysis of rs6674079 and other SNPs identified in this study in lncRNAs would be necessary to

determine their putative regulatory effects and potential impact on ovarian cancer metastasis and progression.

Several protein-coding genes within 1Mb of rs6674079 at 1q22 were also found to be significantly associated with ovarian cancer progression in unrestricted analyses of KM-plotter data (Supplementary Table 6). Above-median expression of *SLC25A44* (probe 32091_at), a recently identified member of the SLC25 family of mitochondrial carrier proteins, was significantly associated with worse PFS in analysis in the larger unrestricted dataset of epithelial ovarian cancer (log-rank $p \leq 1.9 \times 10^{-8}$; Supplementary Figure 4A). While relatively little is known about specific functions or disease-gene associations with *SLC25A44*, changes in expression of some members of the SLC25 family of transporters have been implicated in resistance to cell death in other cancers (35). Similarly high expression of the signalling protein *SEMA4A* (probe 219259_at; Supplementary Figure 4B) was significantly associated with better PFS (log-rank $p = 4.2 \times 10^{-6}$). SEMA4A is a member of the semaphorin family of soluble and transmembrane proteins which mediate their signal transduction effects through plexins, both of which have been shown to have tumorigenic properties and are aberrantly expressed in human cancers, (36, 37). Also high expression of *SH2D2A* (probeset 21925_at) which encodes a T-cell-specific adaptor protein (TSAd), was associated with significantly better PFS (log-rank $p = 8.4 \times 10^{-8}$; Supplementary Figure 4C). Chromosmal imbalance at 1q22 was previously identified as a candidate region for response to chemotherapy in human glioma cell lines (38) and it has been shown that alterations on the long arm of chromosome 1, particularly gain of function, are among the most commonly reported chromosomal abnormalities in human cancers (39). Further studies would be necessary to delineate the relevance of these novel findings in EOC outcome.

We found that that PFS-associated SNPs in the 'standard chemo' dataset were most significantly enriched in a pathway containing target genes of the transcriptional co-activators YAP1 and WWTR1 and the antisense RNA gene TAZ (40, 41). *YAP1*, an established ovarian cancer oncogene (42), is known to regulate the cell cycle and epithelial-mesenchymal transition, promoting tumor survival even in the absence of oncogenic KRAS signaling (43, 44). A gene expression signature representing YAP1 activation in ovarian tumors has also recently been found to be predictive of response to taxane-based adjuvant chemotherapy regimens and is associated with overall survival in ovarian cancer (45). The HDL-mediated lipid transport pathway driven by genes that included *APOA1* was associated with OS in the setting of standard chemotherapy. Higher *APOA1* expression in serous ovarian cancer effusions has previously been associated with improved overall survival in a small cohort (46). Apolipoprotein A-I activity has been shown to reduce viability of platinum-resistant human ovarian cancer cells *in vitro* and inhibit tumor development in a mouse model of ovarian cancer (47).

In our exploratory histology-adjusted analysis of OS in 'all OCAC' we observed significant associations with SNPs in *PARK2* and decreased survival. *PARK2*, a component of E3 ubiquitin ligase complexes that drive cyclin D and E degradation, is frequently lost in human cancers, and knock-down in a range of cancer cell lines has been shown to correlate with increased cell proliferation and transcription of genes related to cell cycle control, suggesting a role in disease progression and prognosis (48). *ANKLE1* and *BABAM1* at 19p13.11 ($p \leq 9.5 \times 10^{-6}$; Supplementary Table 8) were also identified and SNPs at this locus were previously implicated in ovarian cancer risk and survival (49). However in our fully adjusted analysis of ~2900 patients for which we had all covariates, we observed no significant association for any SNP at this locus ($p \geq 0.002$). This may be accounted for by the lower power to detect the effects seen in the larger 'all OCAC' analysis, or the fact that the lower p-

value in the 'all OCAC' analysis is an artefact resulting from partial adjustment for confounders of outcome. Further analyses including FIGO stage, grade and residual disease would be necessary to evaluate this locus. We also observed no significant association for candidate SNPs previously identified to be associated with response to chemotherapy using the NHGRI GWAS catalog (http://www.genome.gov/gwastudies/) with any of our four analyses (Supplementary Table 9).

Our validation analysis of genotyped data also highlighted the potential for spurious associations using imputed data in smaller samples sets. Although current strategies of 'pre-phasing' has improved imputation accuracy for SNPs with MAF 1-3% and prior imputation r^2 as low as 0.6 in Europeans (50), we observed a high degree of discordance in estimates from imputed data compared to actual genotypes, even for SNPs with reasonable imputation quality (r^2=0.6-0.9) and particularly for SNPs with MAF<3% (Supplementary Methods and Supplementary Table 3). We therefore selected SNPs for validation from ~2.8 million SNPs with good imputation quality ($r^2 \geq 0.9$) to reduce the risk of false positives.

In conclusion we have identified three SNPs in lncRNAs that have not been previously reported on that were associated with PFS in ovarian cancer regardless of chemotherapy regimens. We also identified two other SNPs, rs7950311 at 11p15.4 associated with OS in the 'standard chemo' analysis and rs3795247 at 19p12 associated with PFS in the 'all chemo' analysis, both of which reside in genes that have not been previously implicated in solid tumors. To our knowledge this is the largest study that comprehensively analyzes genetic variation across the genome for an association with ovarian cancer outcomes, both with regard to first-line standard-of-care chemotherapy and regardless of treatment. Since residual disease is a strong predictor of overall and progression-free survival, patients were included in our main analyses if they received a minimum of cytoreductive surgery and had available

information on level of residual disease. SNPs were prioritized on the basis of good imputation quality ($r^2 \geq 0.9$) and final estimates were derived from meta-analysis of all available data imputed and genotyped samples from OCAC and publicly available TCGA data. To circumvent methodological flaws we restricted the analysis to European invasive EOC patients participating in the OCAC with standardized definitions of clinical and pathological characteristics. Despite our rigorous analysis approach, there are inherent limitations in the observational design of our study that a randomized clinical trial would circumvent, in that standardized treatment and outcome measurements would be available, and the presence of a control group receiving an alternative treatment would allow assessment of a likely causal relationship between the putative associations and treatment modalities.

Pharmacogenomic studies hold the promise of improving treatment approaches by the identification of genetic markers which may enhance the clinical approaches and cost-effectiveness of these treatment approaches. However, large clinical trials or well-designed prospective cohort studies that take into account differential responses according to EOC tumor types, as well as functional studies that shed light on putative associations are required to succeed in defining the role of genetics in ovarian cancer progression and survival.

Acknowledgements

This study would not have been possible without the contributions of the following: Per Hall (COGS); Douglas F. Easton, Paul Pharoah, Kyriaki Michailidou, Manjeet K. Bolla, Qin Wang (BCAC), Andrew Berchuck (OCAC), Rosalind A. Eeles, Douglas F. Easton, Ali Amin Al Olama, Zsofia Kote-Jarai, Sara Benlloch (PRACTICAL), Georgia Chenevix-Trench, Antonis Antoniou, Lesley McGuffog, Fergus Couch and Ken Offit (CIMBA), Joe Dennis,

Alison M. Dunning, Andrew Lee, and Ed Dicks, Craig Luccarini and the staff of the Centre for Genetic Epidemiology Laboratory, Javier Benitez, Anna Gonzalez-Neira and the staff of the CNIO genotyping unit, Jacques Simard and 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 are grateful to the family and friends of Kathryn Sladek Smith for their generous support for the Ovarian Cancer Association Consortium through their donations to the Ovarian Cancer Research Fund. The authors wish to thank Margie Riggan for her tireless dedication to the Ovarian Cancer Association Consortium through her excellent project and data management. The Australian Ovarian Cancer Study Management Group (D. Bowtell, G. Chenevix-Trench, A. deFazio, D. Gertig, A. Green, P. Webb) and ACS Investigators (A. Green, P. Parsons, N. Hayward, P. Webb, D. Whiteman) thank all the clinical and scientific collaborators (see http://www.aocstudy.org/) and the women for their contribution. GCT & PW are supported by Fellowships from NHMRC. The Belgian study (BEL) would like to thank Gilian Peuteman, Thomas Van Brussel and Dominiek Smeets for technical assistance. The German Ovarian Cancer Study (GER) thank Ursula Eilber and Tanja Koehler for competent technical assistance. The International Collaborative Ovarian Neoplasm study (ICON)7 trial team would like to thank the Medical Research Council (MRC) Clinical Trial Unit (CTU) at the University of London (UCL), the ICON7 Translational Research Sub-group, and the University of Leeds for their work on the coordination of samples and data from the ICON7 trial. The Mayo Clinic Ovarian Cancer Case-Control Study (MAY) thank C. Hilker, S. Windebank, and J. Vollenweider for iSelect genotyping. The Study of

Epidemiology and Risk Factors in Cancer Heredity (SEA) would like to acknowledge Craig Luccarini, Caroline Baynes, Don Conroy. The Scottish Randomised Trial in Ovarian Cancer (SRO) thank all members of Scottish Gynaecological Clinical Trails group and SCOTROC1 investigators. 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. Information about TCGA can be found at http://cancergenome.nih.gov/. The United Kingdom Ovarian Cancer Population Study (UKO) particularly thank I. Jacobs, M. Widschwendter, E. Wozniak, A. Ryan, J. Ford and N. Balogun for their contribution to the study. The Westmead Hospital Molecular Biology of Gynaecologic Disease (WMH) 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.

References

1. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, et al. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11. http://globocaniarcfr, accessed on 13/11/2014. 2013.
2. Marchetti C, Pisano C, Facchini G, Bruni GS, Magazzino FP, Losito S, et al. First-line treatment of advanced ovarian cancer: current research and perspectives. Expert Rev Anticancer Ther. 2010;10:47-60.
3. Ozols RF, Bundy BN, Greer BE, Fowler JM, Clarke-Pearson D, Burger RA, et al. Phase III trial of carboplatin and paclitaxel compared with cisplatin and paclitaxel in patients with optimally resected stage III ovarian cancer: a Gynecologic Oncology Group study. J Clin Oncol. 2003;21:3194-200.
4. Jayson GC, Kohn EC, Kitchener HC, Ledermann JA. Ovarian cancer. Lancet. 2014;384:1376-88.
5. Evans WE, McLeod HL. Pharmacogenomics--drug disposition, drug targets, and side effects. N Engl J Med. 2003;348:538-49.
6. Kalow W, Tang BK, Endrenyi L. Hypothesis: comparisons of inter- and intra-individual variations can substitute for twin studies in drug research. Pharmacogenetics. 1998;8:283-9.
7. Vesell ES. Pharmacogenetic perspectives gained from twin and family studies. Pharmacology & therapeutics. 1989;41:535-52.
8. Relling MV, Dervieux T. Pharmacogenetics and cancer therapy. Nat Rev Cancer. 2001;1:99-108.
9. Vella N, Aiello M, Russo AE, Scalisi A, Spandidos DA, Toffoli G, et al. 'Genetic profiling' and ovarian cancer therapy (review). Molecular medicine reports. 2011;4:771-7.
10. Li J, Bluth MH. Pharmacogenomics of drug metabolizing enzymes and transporters: implications for cancer therapy. Pharmacogenomics and personalized medicine. 2011;4:11-33.
11. Leschziner GD, Andrew T, Pirmohamed M, Johnson MR. ABCB1 genotype and PGP expression, function and therapeutic drug response: a critical review and recommendations for future research. Pharmacogenomics J. 2007;7:154-79.
12. Green H, Soderkvist P, Rosenberg P, Horvath G, Peterson C. mdr-1 single nucleotide polymorphisms in ovarian cancer tissue: G2677T/A correlates with response to paclitaxel chemotherapy. Clin Cancer Res. 2006;12:854-9.
13. Johnatty SE, Beesley J, Gao B, Chen X, Lu Y, Law MH, et al. ABCB1 (MDR1) polymorphisms and ovarian cancer progression and survival: a comprehensive analysis from the Ovarian Cancer Association Consortium and The Cancer Genome Atlas. Gynecol Oncol. 2013;131:8-14.
14. Diaz-Padilla I, Amir E, Marsh S, Liu G, Mackay H. Genetic polymorphisms as predictive and prognostic biomarkers in gynecological cancers: a systematic review. Gynecol Oncol. 2012;124:354-65.
15. Hedditch EL, Gao B, Russell AJ, Lu Y, Emmanuel C, Beesley J, et al. ABCA transporter gene expression and poor outcome in epithelial ovarian cancer. J Natl Cancer Inst. 2014;106.
16. Candido Dos Reis FJ, Song H, Goode EL, Cunningham JM, Fridley BL, Larson MC, et al. Germline mutation in BRCA1 or BRCA2 and ten-year survival for women diagnosed with epithelial ovarian cancer. Clin Cancer Res. 2014.
17. Alsop K, Fereday S, Meldrum C, deFazio A, Emmanuel C, George J, et al. BRCA mutation frequency and patterns of treatment response in BRCA mutation-positive women

with ovarian cancer: a report from the Australian Ovarian Cancer Study Group. J Clin Oncol. 2012;30:2654-63.
18. Samouelian V, Maugard CM, Jolicoeur M, Bertrand R, Arcand SL, Tonin PN, et al. Chemosensitivity and radiosensitivity profiles of four new human epithelial ovarian cancer cell lines exhibiting genetic alterations in BRCA2, TGFbeta-RII, KRAS2, TP53 and/or CDNK2A. Cancer Chemother Pharmacol. 2004;54:497-504.
19. Husain A, He G, Venkatraman ES, Spriggs DR. BRCA1 up-regulation is associated with repair-mediated resistance to cis-diamminedichloroplatinum(II). Cancer Res. 1998;58:1120-3.
20. Huang RS, Duan S, Shukla SJ, Kistner EO, Clark TA, Chen TX, et al. Identification of genetic variants contributing to cisplatin-induced cytotoxicity by use of a genomewide approach. Am J Hum Genet. 2007;81:427-37.
21. Huang RS, Johnatty SE, Gamazon ER, Im HK, Ziliak D, Duan S, et al. Platinum sensitivity-related germline polymorphism discovered via a cell-based approach and analysis of its association with outcome in ovarian cancer patients. Clin Cancer Res. 2011;17:5490-500.
22. Sankararaman S, Sridhar S, Kimmel G, Halperin E. Estimating local ancestry in admixed populations. Am J Hum Genet. 2008;82:290-303.
23. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006;38:904-9.
24. 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;45:362-70.
25. Johnatty SE, Beesley J, Paul J, Fereday S, Spurdle AB, Webb PM, et al. ABCB1 (MDR 1) polymorphisms and progression-free survival among women with ovarian cancer following paclitaxel/carboplatin chemotherapy. Clin Cancer Res. 2008;14:5594-601.
26. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010;26:2190-1.
27. Gyorffy B, Lanczky A, Szallasi Z. Implementing an online tool for genome-wide validation of survival-associated biomarkers in ovarian-cancer using microarray data from 1287 patients. Endocrine-related cancer. 2012;19:197-208.
28. Christoforou A, Dondrup M, Mattingsdal M, Mattheisen M, Giddaluru S, Nothen MM, et al. Linkage-disequilibrium-based binning affects the interpretation of GWASs. Am J Hum Genet. 2012;90:727-33.
29. Segre AV, Consortium D, investigators M, Groop L, Mootha VK, Daly MJ, et al. Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. PLoS genetics. 2010;6.
30. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102:15545-50.
31. Fritah S, Niclou SP, Azuaje F. Databases for lncRNAs: a comparative evaluation of emerging tools. Rna. 2014;20:1655-65.
32. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature. 2010;464:1071-6.
33. Mercer TR, Mattick JS. Structure and function of long noncoding RNAs in epigenetic regulation. Nature structural & molecular biology. 2013;20:300-7.

34. Ma L, Liu J, Liu L, Duan G, Wang Q, Xu Y, et al. Overexpression of the transcription factor MEF2D in hepatocellular carcinoma sustains malignant character by suppressing G2-M transition genes. Cancer Res. 2014;74:1452-62.
35. Gutierrez-Aguilar M, Baines CP. Physiological and pathological roles of mitochondrial SLC25 carriers. The Biochemical journal. 2013;454:371-86.
36. Rehman M, Tamagnone L. Semaphorins in cancer: biological mechanisms and therapeutic approaches. Seminars in cell & developmental biology. 2013;24:179-89.
37. Malik MF, Ye L, Jiang WG. The Plexin-B family and its role in cancer progression. Histology and histopathology. 2014;29:151-65.
38. Weber RG, Rieger J, Naumann U, Lichter P, Weller M. Chromosomal imbalances associated with response to chemotherapy and cytotoxic cytokines in human malignant glioma cell lines. Int J Cancer. 2001;91:213-8.
39. Gregory SG, Barlow KF, McLay KE, Kaul R, Swarbreck D, Dunham A, et al. The DNA sequence and biological annotation of human chromosome 1. Nature. 2006;441:315-21.
40. Murakami M, Nakagawa M, Olson EN, Nakagawa O. A WW domain protein TAZ is a critical coactivator for TBX5, a transcription factor implicated in Holt-Oram syndrome. Proc Natl Acad Sci U S A. 2005;102:18034-9.
41. Oh H, Irvine KD. Yorkie: the final destination of Hippo signaling. Trends in cell biology. 2010;20:410-7.
42. Hall CA, Wang R, Miao J, Oliva E, Shen X, Wheeler T, et al. Hippo pathway effector Yap is an ovarian cancer oncogene. Cancer Res. 2010;70:8517-25.
43. Kapoor A, Yao W, Ying H, Hua S, Liewen A, Wang Q, et al. Yap1 activation enables bypass of oncogenic Kras addiction in pancreatic cancer. Cell. 2014;158:185-97.
44. Shao DD, Xue W, Krall EB, Bhutkar A, Piccioni F, Wang X, et al. KRAS and YAP1 converge to regulate EMT and tumor survival. Cell. 2014;158:171-84.
45. Jeong W, Kim SB, Sohn BH, Park YY, Park ES, Kim SC, et al. Activation of YAP1 is associated with poor prognosis and response to taxanes in ovarian cancer. Anticancer Res. 2014;34:811-7.
46. Tuft Stavnes H, Nymoen DA, Hetland Falkenthal TE, Kaern J, Trope CG, Davidson B. APOA1 mRNA expression in ovarian serous carcinoma effusions is a marker of longer survival. American journal of clinical pathology. 2014;142:51-7.
47. Su F, Kozak KR, Imaizumi S, Gao F, Amneus MW, Grijalva V, et al. Apolipoprotein A-I (apoA-I) and apoA-I mimetic peptides inhibit tumor development in a mouse model of ovarian cancer. Proc Natl Acad Sci U S A. 2010;107:19997-20002.
48. Bartek J, Hodny Z. PARK2 orchestrates cyclins to avoid cancer. Nat Genet. 2014;46:527-8.
49. 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;42:880-4.
50. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. Nat Genet. 2012;44:955-9.

Table 1: Results of meta-analysis of estimates from iPLEX genotyped, non-overlapping iCOGS and TCGA datasets for selected promising SNPs

						OVERALL SURVIVAL				**PROGRESSION-FREE SURVIVAL**			
						All Chemo (N=4,426)		**Standard Chemo (N=1,799)**		**All Chemo (N=4,095)**		**Standard Chemo (N=1,598)**	
SNP	Chr	Position	Nearest Gene	Effect/ Ref Allele	[a]Effect Allele Freq.	[b]HR (95% CI)	P	[b]HR (95% CI)	P	[b]HR (95% CI)	P	[b]HR (95% CI)	P
rs6674079	1q22	156486061	RP11-284F21.8	G/A	0.28	1.15 (1.08 -1.23)	**$7.1x10^{-6}$**	1.07 (0.97-1.18)	$1.9x10^{-1}$	1.07 (1.01-1.13)	$2.8x10^{-2}$	0.98 (0.90-1.07)	$6.8x10^{-1}$
rs7950311	11p15.4	5672354	HBG2	C/T	0.48	1.10 (1.04-1.17)	$1.7x10^{-3}$	1.28 (1.16-1.42)	**$6.8x10^{-7}$**	1.03 (0.98-1.09)	$2.5\ x10^{-1}$	1.08 (0.99-1.18)	$7.8x10^{-2}$
rs4910232	11p15.3	11120369	RP11-179A10.1	G/T	0.32	1.12 (1.05-1.19)	$9.4x10^{-4}$	1.20 (1.08-1.33)	$5.3x10^{-4}$	1.17 (1.10-1.24)	**$4.7x10^{-7}$**	1.24 (1.12-1.56)	$1.2x10^{-5}$
rs2549714	16q23	80875263	RP11-314O13.1	C/A	0.06	1.20 (1.06-1.36)	$3.4x10^{-3}$	1.53 (1.28-1.84)	**$5.0x10^{-6}$**	1.14 (1.01-1.28)	$2.8\ x10^{-2}$	1.29 (1.08-1.55)	$5.6x10^{-3}$
rs3795247	19p12	21906428	ZNF100	C/T	0.08	1.16 (1.04-1.30)	$8.8x10^{-3}$	1.34 (1.13-1.60)	$9.7x10^{-4}$	1.26 (1.14-1.40)	**$1.05x10^{-5}$**	1.39 (1.18-1.65)	$9.2x10^{-5}$

[a] Effect allele frequency from genotyped samples

[b] Estimates are adjusted for residual disease (nil vs. any), FIGO stage (I-IV), tumor histology (serous, mucinous, endometrioid, clear cell, other epithelial), grade (low vs. high), site, age at diagnosis (OS only) and the first 3 principal components (imputed data only). BAV & NCO included only in OS analysis.

Table 2: Gene set enrichment (pathway-level) analysis results for PFS and OS associations in the 'standard chemo' data set

Pathway	[a]Genes	p-value	[b]FDR	[c]FWER	Core genes
Pathways associated with PFS in 'standard chemo' at p<0.05 and FWER<1					
REACTOME_YAP1_AND_WWTR1_TAZ_STIMULATED_GENE_EXPRESSION	23	0.001	0.868	0.575	*CTGF,TBL1X,NCOA6,TEAD3,MED1,PPARA,TEAD1,NCOA3, KAT2B*
REACTOME_G0_AND_EARLY_G1	23	0.012	1	0.991	*RBL2,CDC25A,MYBL2,LIN9,HDAC1,CCNA1,LIN52*
REACTOME_AMINE_DERIVED_HORMONES	15	0.025	1	0.993	*CGA,TPO,SLC5A5,TH*
REACTOME_FORMATION_OF_INCISION_COMPLEX_IN_GG_NER	21	0.010	1	0.994	*ERCC2,RAD23B,GTF2H1,GTF2H2,RPA1,ERCC1,DDB2,XPA, DDB1*
REACTOME_G_PROTEIN_ACTIVATION	27	0.007	1	0.999	*GNB2,GNAT1,GNAI2,GNAI1,POMC,GNB3,GNG4,GNGT2,GNAO1,GNG8,GNG3*
REACTOME_LYSOSOME_VESICLE_BIOGENESIS	22	0.013	1	0.999	*CLTA,AP1B1,AP1S1,DNAJC6,AP1G1,GNS,M6PR,VAMP8,BLOC1S1*
REACTOME_INHIBITION_OF_INSULIN_SECRETION_BY_ADRENALINE_NORADRENALINE	25	0.014	1	0.999	*GNB2,GNAI2,CACNB2,GNAI1,ADRA2A,GNB3,GNG4,GNGT2 ,GNAO1,GNG8,GNG3*
REACTOME_CYCLIN_A_B1_ASSOCIATED_EVENTS_DURING_G2_M_TRANSITION	15	0.025	1	0.999	*CDC25A,PLK1,CCNA1,WEE1,CDC25B,PKMYT1,XPO1*
Pathways associated with OS in 'standard chemo' at p<0.05 and FWER<1					
REACTOME_HDL_MEDIATED_LIPID_TRANSPORT	15	0.006	0.303	0.268	*BMP1,CETP,APOA1,APOC3,ABCG1*
REACTOME_XENOBIOTICS	15	0.009	1	0.891	*CYP2A13,CYP2B6,CYP2F1*
REACTOME_LIPOPROTEIN_METABOLISM	28	0.005	1	0.979	*BMP1,CETP,APOA1,APOC3,APOA5,ABCG1*
REACTOME_INSULIN_SYNTHESIS_AND_PROCESSING	20	0.005	0.915	0.980	*SNAP25,INS,EXOC5,ERO1L,PCSK1,EXOC4,PCSK2*
BIOCARTA_MTA3_PATHWAY	19	0.013	0.772	0.982	*TUBA1A,TUBA1C,HDAC1,MBD3,ALDOA,CDH1,MTA1,SNAI 2,TUBA3C*
REACTOME_ACETYLCHOLINE_BINDING_AND_DOWNSTREAM_EVENTS	15	0.022	0.781	0.994	*CHRNG,CHRND*
REACTOME_SYNTHESIS_OF_BILE_ACIDS_AND_BILE_SALTS_VIA_7ALPHA_HYDROXYCHOLESTEROL	15	0.032	0.716	0.996	*SLC27A5, HSD17B4, AKR1D1, SLC27A2, CYP27A1, ACOX2, HSD3B7, ABCB11*
KEGG_MATURITY_ONSET_DIABETES_OF_THE_YOUNG	23	0.006	0.658	0.997	*ONECUT1, INS, HNF1A, BHLHA15, NR5A2, FOXA3*
REACTOME_IMMUNOREGULATORY_INTERACTIONS_BETWEEN_A_LYMPHOID_AND_A_NON_LYMPHOID_CELL	56	0.001	0.621	0.998	*CD96, CD8A, CD8B, IFITM1, KIR3DL2, CRTAM, ICAM2, KIR3DL1, FCGR3A, LILRB2, CD19, LILRB5, LILRB3, CD200R1, RAET1E, FCGR2B, SELL,ULBP2, ULBP1, KIR2DL4, B2M, CDH1, CD81*

[a]Number of genes; [b]False Discovery Rate; [c]Familywise Error Rates

Figure Legends

Figure 1: Manhattan plots of ~2.8 million SNPs in four analyses of the cohort selected for first-line chemotherapy. SNPs with MAF ≥0.02 and imputation r^2 ≥0.9 associated with Overall Survival in A. 'All Chemo' and B. 'Standard chemo', and Progression-free survival in C. 'All chemo' and D. 'Standard chemo'; the blue line represents suggestive significance ($p=1x10^{-5}$) and the red line represents genome wide significance ($p=5x10^{-8}$).

Figure 2: Progression-free survival in 'all chemo' analysis for rs4910232. A. Kaplan Meier curve for PFS in 'all chemo' dataset (n=3,177); P-values derived from adjusted Cox PH models of genotyped samples; 0=common homozygotes AA, 1=heterozygotes AG, 2=rare homozygotes GG. B**.** Forest plot showing site-specific estimates for PFS and rs4910232 in 'all chemo' dataset.

Figure 3: ENCODE ChIP-seq data at 1q22 locus. Manhattan plot of all iCOGS imputed/genotyped SNPs at 1q22, black enclosed circles represent genotyped SNPs while open red circles are imputed SNPs. Hash marks indicate location of highly correlated SNPs (r^2 >0.9). Colored histograms denote histone modification for H3K4me1 and H3K27ac in normal ovary ChIP-seq data from UCSD and ENCODE.

Figure 4: Overall survival in 'all chemo' for rs6674079. A. Kaplan Meier curve for OS in the 'all chemo' dataset. P-value derived from adjusted Cox PH models of genotyped samples (n=4,399): 0=common homozygotes AA, 1=heterozygotes AG, 2=rare homozygotes GG. B. Forest plot showing site-specific estimates for OS and rs6674079 in 'all chemo' dataset

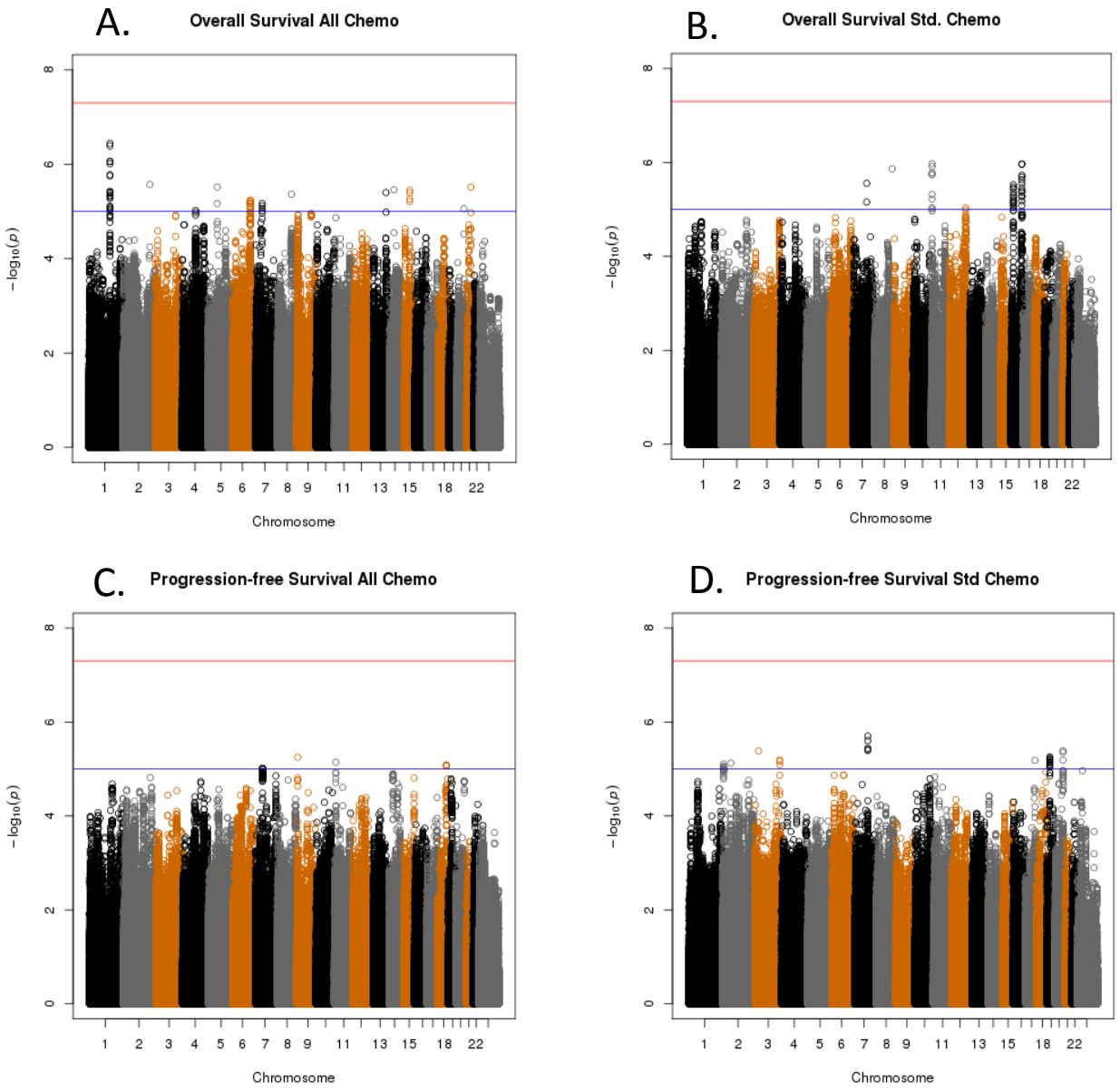

Figure 1: Manhattan plots of ~2.8 million SNPs (imputation $r^2 \geq 0.9$) in four analyses of cohorts selected according to first-line chemotherapy.

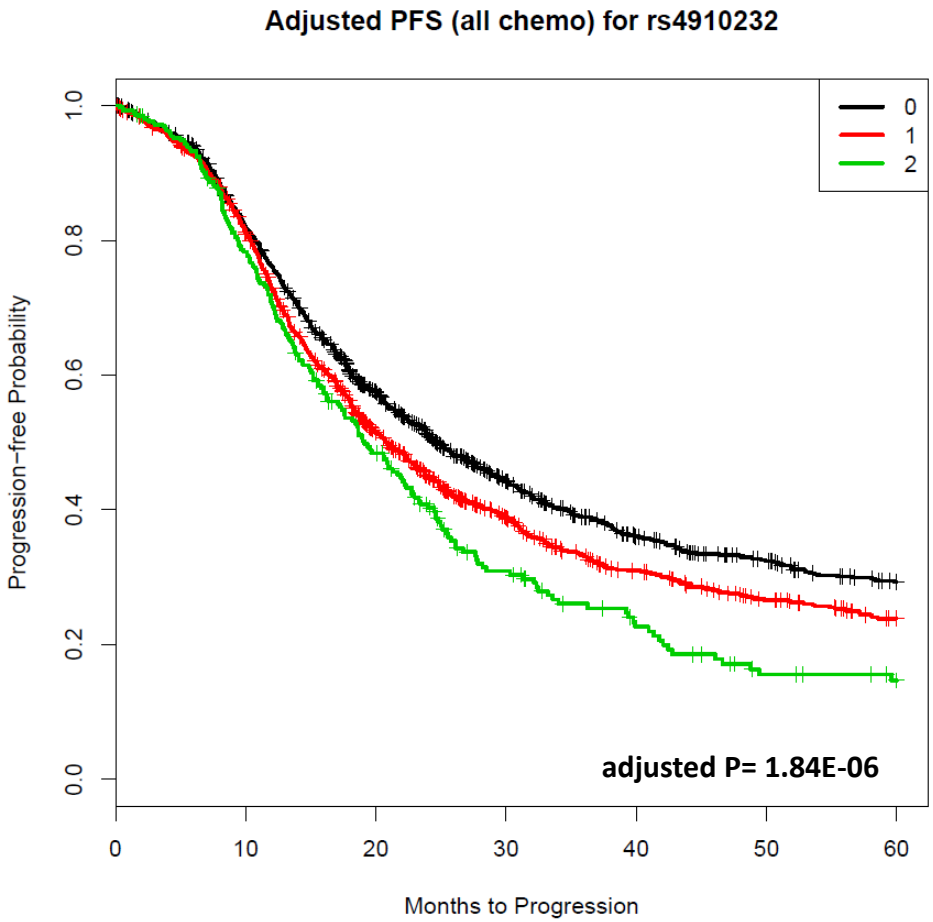

Figure 2A: PFS in 'all chemo' analysis for rs4910232. Kaplan Meier curve for PFS in 'all chemo' dataset (n=3,177); P-values derived from adjusted Cox PH models of genotyped samples; 0=common homozygotes AA, 1=heterozygotes AG, 2=rare homozygotes GG

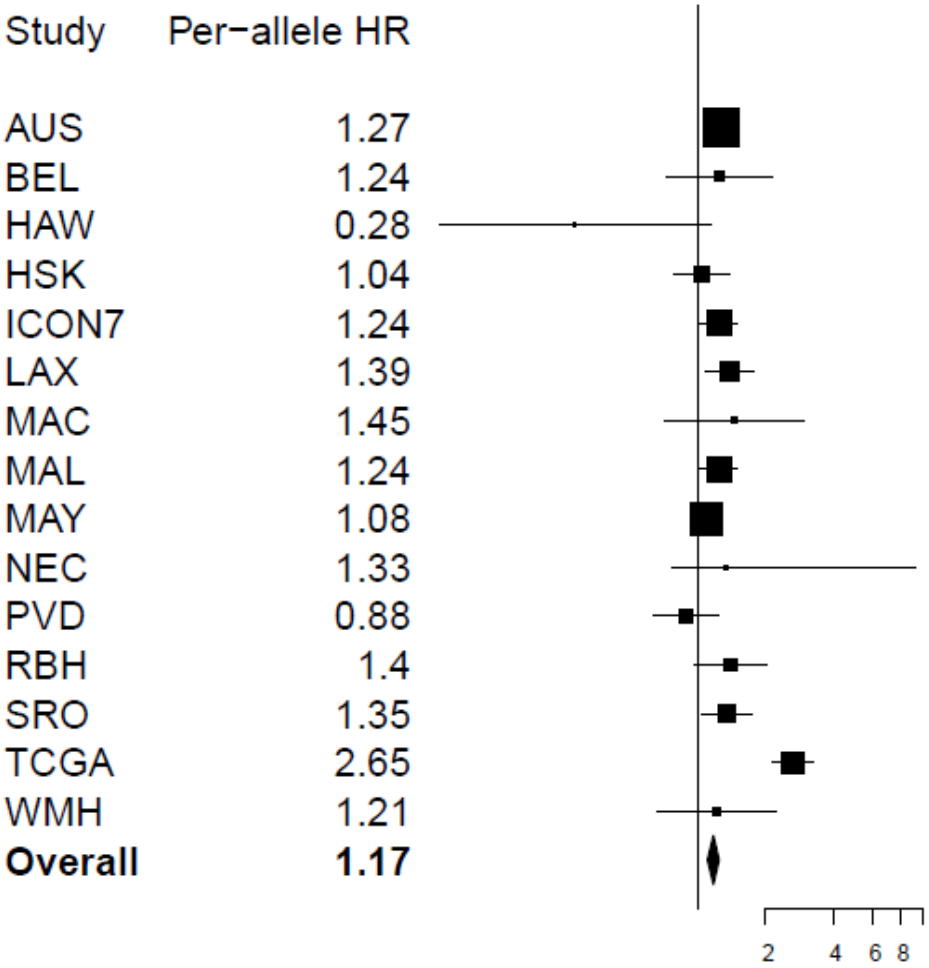

Figure 2B: Forest plot showing site-specific estimates for PFS and rs4910232 in 'all chemo' dataset

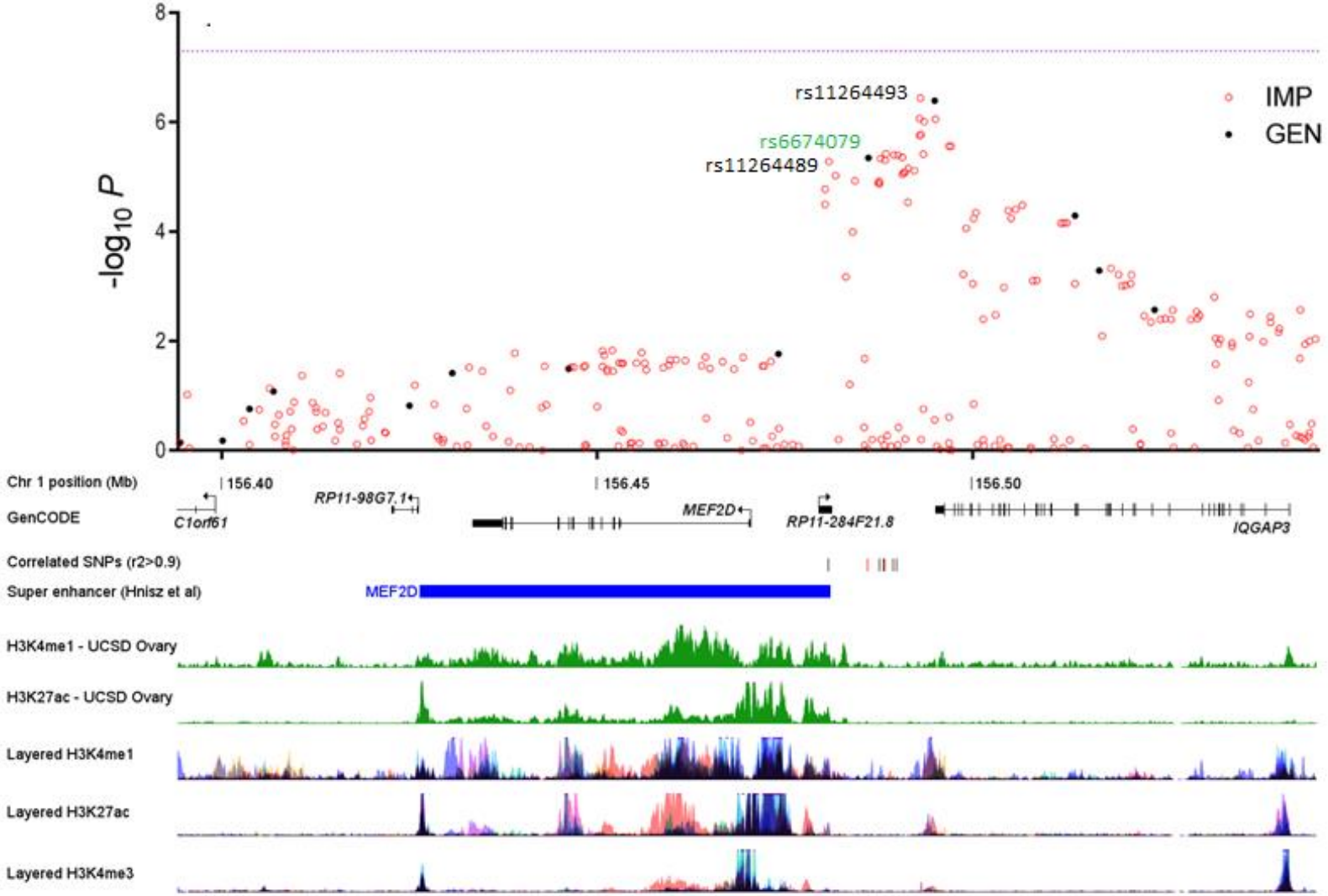

Figure 3: ENCODE ChIP-seq data at 1q22 locus. Black enclosed circles represent genotyped SNPs while open red circles are imputed SNPs. Hash marks indicate location of highly correlated SNPs (r^2 >0.9). Colored histograms denote histone modification for H3K4me1 and H3K27ac in normal ovary ChIP-seq data from UCSD and ENCODE.

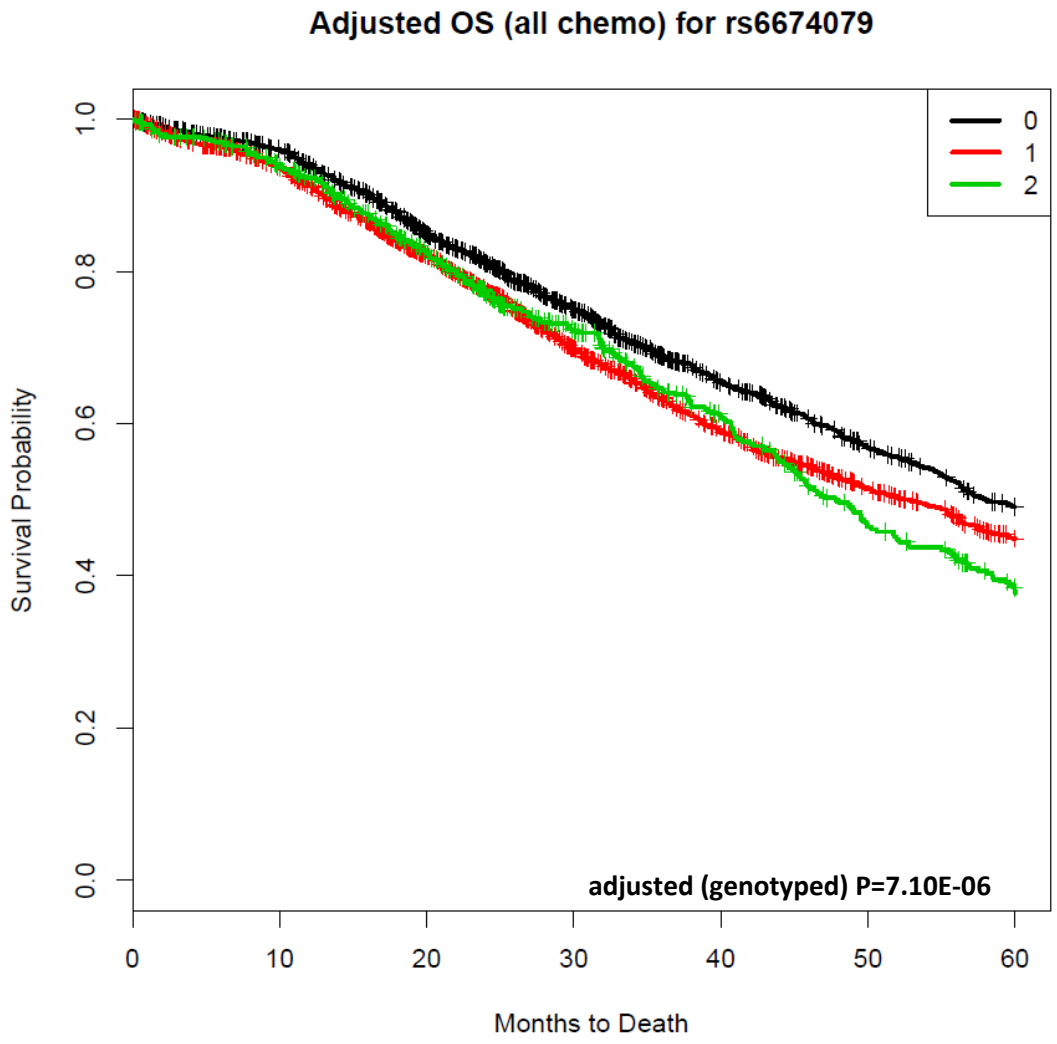

Figure 4A: OS in 'all chemo' analysis for rs6674079. P-value derived from adjusted Cox PH models of genotyped samples (n=4,399); 0=common homozygotes AA, 1=heterozygotes AG, 2=rare homozygotes GG.

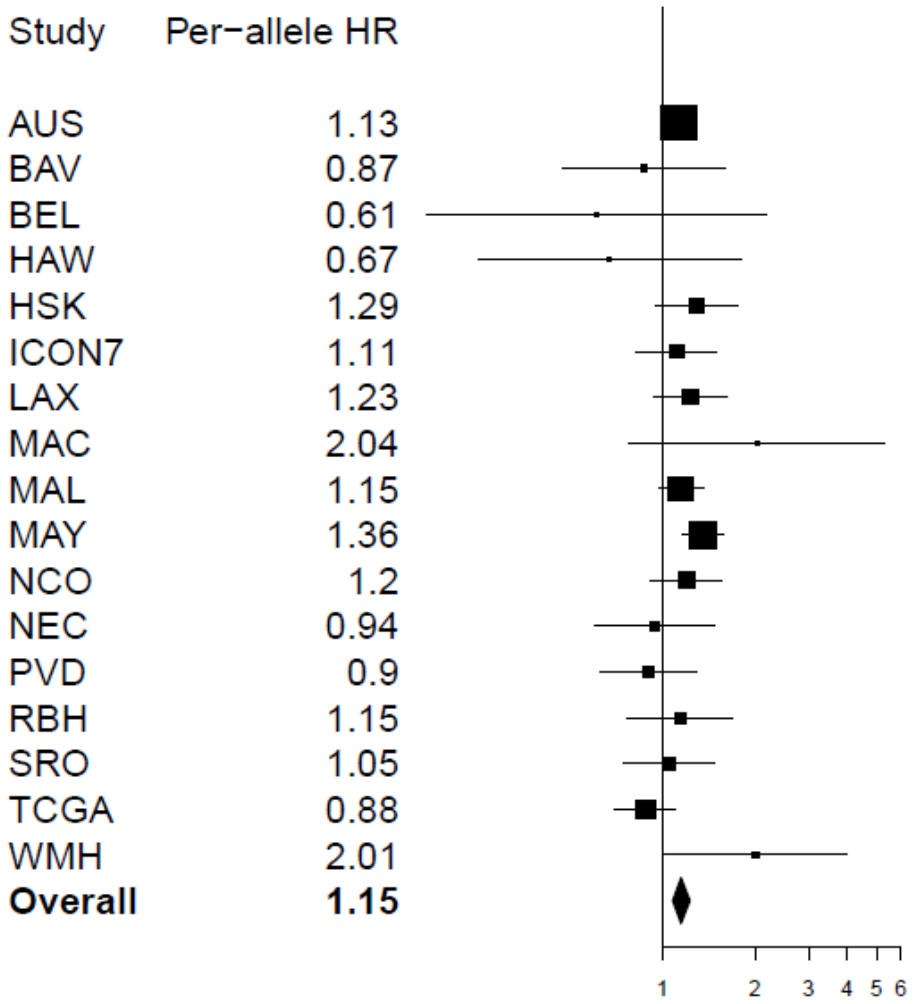

Figure 4B: Forest plot showing site-specific estimates for OS and rs6674079 in the 'all chemo' dataset.

AACR American Association for Cancer Research

Clinical Cancer Research

Genome-wide analysis identifies novel loci associated with ovarian cancer outcomes: findings from the Ovarian Cancer Association Consortium

Sharon Johnatty, Jonathan P Tyrer, Siddhartha P Kar, et al.

Clin Cancer Res Published OnlineFirst July 7, 2015.