Identification of six new susceptibility loci for invasive epithelial ovarian cancer

Permalink
https://escholarship.org/uc/item/0807m7ph

Journal
Nature Genetics, 47(2)

ISSN
1061-4036

Authors
Kuchenbaecker, KB
Ramus, SJ
Tyrer, J
et al.

Publication Date
2015

DOI
10.1038/ng.3185

Peer reviewed
Identification of six new susceptibility loci for invasive epithelial ovarian cancer

A full list of authors and affiliations appears at the end of the article.

Genome-wide association studies (GWAS) have identified 12 epithelial ovarian cancer (EOC) susceptibility alleles. The pattern of association at these loci is consistent in \textit{BRCA1} and \textit{BRCA2} mutation carriers who are at high EOC risk. After imputation to the 1000 Genomes Project data, we assessed associations of 11 million genetic variants with EOC risk from 15,397 cases unselected for family history and 30,816 controls, 15,252 \textit{BRCA1} mutation carriers and 8,211 \textit{BRCA2} mutation carriers (3,096 with ovarian cancer), and combined the results in a meta-analysis. This new study design yielded increased statistical power, leading to the discovery of six new EOC susceptibility loci. Variants at 1p36 (nearest gene \textit{WNT4}), 4q26 (\textit{SYNPO2}), 9q34.2 (\textit{ABO}) and 17q11.2 (\textit{ATAD5}) were associated with...
EOC risk, and at 1p34.3 (RSPO1) and 6p22.1 (GPX6) specifically with the serous EOC subtype, at p<5×10^{-8}. Incorporating these variants into risk assessment tools will improve clinical risk predictions for BRCA1/2 mutation carriers.

The risk of developing invasive EOC is higher than the population average for relatives of women diagnosed with the disease\(^1,2\), indicating the importance of genetic factors in disease susceptibility. Approximately 25% of the familial aggregation of EOC is explained by rare, high-penetrance alleles of BRCA1 and BRCA2\(^3\). Furthermore, population-based GWAS have identified common variants associated with invasive EOC at 11 loci\(^4-9\) but only six have also been evaluated in BRCA1 and/or BRCA2 mutation carriers. All loci displayed associations in mutation carriers that were consistent with the associations observed in the general population\(^10-12\). In addition, the 4q32.3 locus is associated with EOC risk for BRCA1 mutation carriers only\(^13\). However, the common genetic variants explain less than 3.1% of the excess familial risk of EOC so additional susceptibility loci are likely to exist.

Women diagnosed with EOC and unaffected women from the general population ascertained through the Ovarian Cancer Association Consortium (OCAC)\(^14\) and BRCA1 and BRCA2 mutation carriers from the Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA)\(^15\) were genotyped as part of the Collaborative Oncological Gene-environment Study (COGS) using the iCOGS custom array. In addition, data were available for cases and controls from three EOC GWAS. We first evaluated whether the EOC susceptibility loci at 8q21.13, 10p12.31, 17q12, 5p15.33, and 17q21.31 recently identified by OCAC\(^7-9\) also show evidence of association in BRCA1 and BRCA2 mutation carriers. Using data from >200,000 genotyped SNPs\(^7,13,16\), we performed imputation of common variants from the 1000 Genomes Project data\(^17\) and evaluated the associations of these SNPs with invasive EOC risk in OCAC and in BRCA1 and BRCA2 mutation carriers from CIMBA. Given the strong evidence for a significant overlap in loci predisposing to EOC in the general population and those associated with risk in BRCA1 and BRCA2 mutation carriers, we carried out a meta-analysis of the EOC risk associations in order to identify novel EOC susceptibility loci.

Genotype data were available for imputation on 15,252 BRCA1 mutation carriers and 8,211 BRCA2 mutation carriers, of whom 2,462 and 631, respectively, were affected with EOC\(^13,16\). From OCAC, genotyping data were available from 15,437 women with invasive EOC (including 9,627 with serous EOC) and 30,845 controls from the general population\(^7\). Imputation was performed separately for BRCA1 carriers, BRCA2 carriers, OCAC-COGS samples and the three OCAC GWAS (Supplementary Tables 1–2; Supplementary Fig. 1; Supplementary Fig. 2). The meta-analysis was based on 11,403,952 SNPs (Supplementary Fig. 3).

Of five EOC susceptibility loci that have not yet been evaluated in mutation carriers, two were associated with EOC risk for both BRCA1 and BRCA2 mutation carriers at p<0.05 (10p12.31 and 17q21.31) (Supplementary Table 3). Overall, seven of the twelve known EOC susceptibility loci provided evidence of association in BRCA1 mutation carriers and six were associated in BRCA2 mutation carriers. However, with the exception of 5p15.33 (TERT), all loci had hazard ratio (HR) estimates in BRCA1 and BRCA2 carriers that were in
the same direction as the odds ratio (OR) estimates for serous subtype EOC from OCAC (Fig. 1). Analysing the associations jointly in BRCA1 and BRCA2 carriers and serous EOC in OCAC provided stronger evidence of association, with smaller p-values for eight of the susceptibility variants compared to the analysis in OCAC alone.

Using the imputed genotypes, we observed no novel associations at $p<5\times10^{-8}$ in the analysis of associations in BRCA1 or BRCA2 mutation carriers separately. However, we identified seven previously unreported associations ($p-values<5\times10^{-8}$) in either OCAC alone, the meta-analysis of EOC associations in BRCA1, BRCA2 carriers and OCAC, or in the meta-analysis in BRCA1 and BRCA2 carriers and serous EOC in OCAC (Supplementary Fig. 4; Supplementary Tables 4–5). SNPs in six of these loci remained genome-wide statistically significant after re-imputing genotypes with imputation parameters set to maximise accuracy (Table 1; Fig. 1). SNPs at 17q11.2 (near ATAD5) were found to be associated with invasive EOC in OCAC ($p<5\times10^{-8}$) (Table 1). For the lead SNP, chr17:29181220:I, the estimated HR estimate for BRCA1 mutation carriers was significantly different from the estimate in OCAC ($p=0.005$); the association for BRCA2 carriers was consistent with the OCAC OR estimate (BRCA2-OCAC meta-analysis $p=2.6\times10^{-9}$). SNPs at four loci were associated at $p<5\times10^{-8}$ with risk of all invasive EOC in the meta-analysis (Supplementary Fig. 5): 1p36, 1p34.3, 4q26, and 9q34.2. At 1p34.3, the most strongly associated SNP, rs58722170, displayed stronger associations in the meta-analysis of serous EOC for OCAC ($p=2.7\times10^{-12}$). In addition, SNPs at 6p22.1 were associated at genome-wide significance level in the meta-analysis of associations with serous EOC ($p=3.0\times10^{-8}$), but not in the meta-analysis of all invasive EOC associations ($p=6.8\times10^{-6}$).

The most significantly associated SNP at each of the six novel loci had high imputation accuracy ($r^2 \geq 0.83$). At the 1p34.3, 1p36, and 6p22.1 loci, there was at least one genome-wide significant genotyped SNP correlated with the lead SNP (pairwise $r^2 \geq 0.73$) (Supplementary Table 6; Supplementary Fig. 5; Supplementary Note). We genotyped the leading (imputed) SNPs of the three other loci in a subset of the samples using iPLEX (Supplementary Note). The correlations between the expected allele dosages from the imputation and the observed genotypes for the variants at 4q26 and 9q34.2, ($r^2=0.90$ and $r^2=0.84$, respectively) were consistent with the estimated imputation accuracy (0.93 and 0.83 for CIMBA samples). The lead SNP at 17q11.2 failed iPLEX design. However, the risk allele is highly correlated with the AA haplotype of two genotyped variants on the iCOGS array (rs9910051 and rs3764419). This haplotype is strongly associated with ovarian cancer risk in the subset of samples genotyped using iCOGS (BRCA2-OCAC meta-analysis $p=8.6\times10^{-8}$ for haplotype, and $p=1.8\times10^{-8}$ for chr17:29181220:I) (Supplementary Table 7).

None of the regions contained additional SNPs that displayed EOC associations at $p<10^{-4}$ in OCAC, BRCA1 carriers or BRCA2 carriers in multi-variable analyses adjusted for the lead SNP in each region, indicating that they each contain only one independent set of correlated highly associated variants (iCHAV). Relative to the 1000 Genomes Project data, we had genotyped or imputed data covering 91% of the genetic variation at 1p36, 84% at 1p34.3 and 83% at 4q26. The other three novel loci had coverage of less than 80% (Supplementary Note). There was evidence for heterogeneity at $p<0.05$ in the associations with histological
subtype in OCAC for the lead SNPs at 1p34.4 and 6p22.1, but not for at 1p36, 4q26, 9q34.2 and 17q11.2 (Table 2).

We carried out a competing risks association analysis in BRCA1 and BRCA2 mutation carriers in order to investigate whether these loci are also associated with breast cancer risk for mutation carriers (Supplementary Note). We used the most strongly associated genotyped SNPs for this purpose because the statistical method requires actual genotypes18. The EOC HR estimates were consistent with the estimates from the main analysis for all SNPs (Supplementary Table 8). None of the SNPs displayed associations with breast cancer risk at p<0.05.

At each of the six loci, we identified a set of SNPs with odds of less than 100 to 1 against being the causal variant; most are in non-coding DNA regions (Supplementary Table 9). None were predicted to have likely deleterious functional effects although some lie in or near chromatin biofeatures in fallopian tube and ovarian epithelial cells which may represent the functional regulatory targets of the risk SNPs (Table 3; Supplementary Table 10). We also evaluated the protein coding genes in each region for their role in EOC development, and as candidate susceptibility gene targets. Molecular profiling data from 496 HGSOCs performed by The Cancer Genome Atlas (TCGA) indicated frequent loss/deletion at four risk loci (1p36, 4q26, 9q34.2 and 17q11.2) (Supplementary Table 11). Consistent with this, WNT4 and ABO were significantly down-regulated in ovarian tumours while ATAD5 was up-regulated. Somatic coding sequence mutations in the six genes nearest the index SNPs were rare. We performed expression quantitative trait locus (eQTL) analysis in a series of 59 normal ovarian tissues (Supplementary Table 12) to evaluate the gene nearest the top ranked SNP at each locus. For the five genes expressed in normal cells, we found no statistically significant eQTL associations for any of the putative causal SNPs at each locus; neither did we find any significant tumour-eQTL associations for these genes based on data from TCGA (Supplementary Table 12). At the 1p36 locus, the most strongly associated variant, rs56318008, is located in the promoter region of WNT4 which encodes a ligand in the WNT signal transduction pathway, critical for cell proliferation and differentiation. Using a luciferase reporter assay we found no effect of these putatively causal SNPs on WNT4 transcription in iOSE4 normal ovarian cells (Fig. 2). Some of the putative causal SNPs at 1p36 are located in CDC42 and LINC00339, and several are in putative regulatory domains in ovarian tissues (Supplementary Table 10; Fig. 2). CDC42 is known to play a role in migration and signalling in ovarian and breast cancer19,20. SNPs at 1p36 are also associated with increased risk of endometriosis and WNT4, CDC42 and LINC00339 have all been implicated in endometriosis21, a known risk factor for endometrioid and clear cell EOC22.

The strongest associated variant at 1q34, rs58722170, is located in RSPO1, which encodes R-spondin 1, a protein involved in cell proliferation (Supplementary Fig. 6). RSPO1 is important in tumorigenesis and early ovarian development23,24, and regulates WNT4 expression in the ovaries25. SYNPO2 at 4q26 encodes myopodin which is involved in cell motility and growth26 and has a reported tumour suppressor role27–30. rs635634 is located upstream of the ABO gene (Supplementary Fig. 7). A moderately correlated variant (rs505922, r²=0.52) determines ABO blood group and is associated with increased risk of pancreatic cancer31,32. Previous studies in OCAC also showed a modestly increased risk of
EOC for individuals with the A blood group\textsuperscript{33}. The moderate correlation between rs635634 and rs505922 and considerably weaker EOC association of rs505922 (p=1.2×10\textsuperscript{-5}) suggests that the association with blood group is probably not driving the association with risk. The indel, 17:29181220:I, at 17q11.2 is located in \textit{ATAD5} which acts as a tumour suppressor gene\textsuperscript{34-36} (Supplementary Fig. 8). \textit{ATAD5} modulates the interaction between RAD9A and BCL2 in order to induce DNA damage related apoptosis. Finally, rs116133110, at 6p22.1, lies in \textit{GPX6} which has no known role in cancer.

The six novel loci reported in this study increase the number of genome-wide significant common variant loci so far identified for EOC to 18. Taken together, these explain approximately 3.9\% of the excess familial relative risk of EOC in the general population, and account for approximately 5.2\% of the EOC polygenic modifying variance in \textit{BRCA1} mutation carriers and 9.3\% in \textit{BRCA2} mutation carriers. The similarity in the magnitude of associations between \textit{BRCA1} and \textit{BRCA2} carriers and population-based studies suggests a general model of susceptibility whereby \textit{BRCA1} and \textit{BRCA2} mutations and common alleles interact multiplicatively on the relative risk scale for EOC\textsuperscript{37}. This model predicts large differences in absolute EOC risk between individuals carrying many alleles and individuals carrying few risk alleles of EOC susceptibility loci for \textit{BRCA1} and \textit{BRCA2} mutation carriers\textsuperscript{13,16}. Incorporating EOC susceptibility variants into risk assessment tools will improve risk prediction and may be particularly useful for \textit{BRCA1} and \textit{BRCA2} mutation carriers.

\section*{METHODS}

\subsection*{Study populations}

We obtained data on \textit{BRCA1} and \textit{BRCA2} mutation carriers through CIMBA. Eligibility in CIMBA is restricted to females 18 years or older with pathogenic mutations in \textit{BRCA1} or \textit{BRCA2}. The majority of the participants were sampled through cancer genetics clinics\textsuperscript{15}, including some related participants. Fifty-four studies from 27 countries contributed data. After quality control, data were available on 15,252 \textit{BRCA1} mutation carriers and 8,211 \textit{BRCA2} mutation carriers, of whom 2,462 and 631, respectively, were affected with EOC (Supplementary Table 1).

Data were available for the stage 1 of three population-based EOC GWAS. These included 2,165 cases and 2,564 controls from a GWAS from North America ("US GWAS")\textsuperscript{39}, 1,762 cases and 6,118 controls from a UK-based GWAS ("UK GWAS")\textsuperscript{6}, and 441 cases and 441 controls from the Mayo GWAS. Furthermore, 11,069 cases and 21,722 controls were genotyped using the iCOGS array ("OCAC-iCOGS" stage data). Overall, 43 studies from 11 countries provided data on 15,347 women diagnosed with invasive epithelial EOC, 9,627 of whom were diagnosed with serous EOC, and 30,845 controls from the general population.

All subjects included in this analysis were of European descent and provided written informed consent as well as data and blood samples under ethically approved protocols. Further details of the OCAC and CIMBA study populations as well as the genotyping, quality control and statistical analyses have been described elsewhere\textsuperscript{7,13,16}.

\textit{Nat Genet.} Author manuscript; available in PMC 2015 August 01.
Genotype data

Genotyping and imputation details for each study are shown in Supplementary Table 1.

Confirmatory genotyping of imputed SNPs

To evaluate the accuracy of the imputation of the SNPs we found to be associated with EOC risk, we genotyped rs17329882 (4q26) and rs635634 (9q34.2) in a subset of 3,541 subjects from CIMBA using Sequenon’s iPLEX technology. The lead SNP at 17q11.2, chr17:29181220:I failed iPLEX design. We performed quality control of the iPLEX data according to the CIMBA guidelines. After quality control, we used the imputation results to generate the expected allele dosage for each genotyped sample and computed the Pearson product-moment correlation coefficient between the expected allele dosage and the observed genotype. The squared correlation coefficient was compared to the imputation accuracy as estimated from the imputation.

Quality control of GWAS and iCOGS genotyping data

We carried out quality control separately for BRCA1 carriers, BRCA2 carriers, the three OCAC GWAS, and OCAC-iCOGS samples, but quality criteria were mostly consistent across studies. We excluded samples if they were not of European ancestry, if they had a genotyping call rate < 95%, low or high heterozygosity, if they were not female or had ambiguous sex, or were duplicates (cryptic or intended). In OCAC studies, one individual was excluded from each pair of samples found to be first-degree relatives and duplicate samples between the iCOGS stage and any of the GWAS were excluded from the iCOGS data. SNPs were excluded if they were monomorphic, had call rate<95%, showed evidence of deviation from Hardy-Weinberg equilibrium or had low concordance between duplicate pairs. For the Mayo GWAS and the UK GWAS, we also excluded rare SNPs (MAF<1% or allele count <5, respectively). We visually inspected genotype cluster plots for all SNPs with P<10^{-5} from each of the newly identified loci. We used the R GenABEL library version 1.6.7 for quality control.

Genotype data were available for analysis from iCOGS for 199,526 SNPs in OCAC-iCOGS, 200,720 SNPs in BRCA1 mutation carriers, and 200,908 SNPs in BRCA2 mutation carriers. After QC, for the GWAS, data were available on 492,956 SNPs for the US GWAS, 543,529 SNPs for the UK GWAS and 1,587,051 SNPs for the Mayo GWAS (Supplementary Table 2).

Imputation

We performed imputation separately for BRCA1 carriers, BRCA2 carriers, OCAC-iCOGS samples and each of the OCAC GWAS. We imputed variants from the 1000 Genomes Project data using the v3 April 2012 release as the reference panel. For OCAC-iCOGS, the UK GWAS and the Mayo GWAS, imputation was based on the 1000 Genomes Project data with singleton sites removed. To improve computation efficiency we initially used a two-step procedure, which involved pre-phasing in the first step and imputation of the phased data in the second. We carried out pre-phasing using the SHAPEIT software. We used the IMPUTE version 2 software for the subsequent imputation for all studies with the
exception of the US GWAS for which the MACH algorithm implemented in the minimac software version 2012.8.15, mach version 1.0.18 was used. To perform the imputation we divided the data into segments of approximately 5Mb each. We excluded SNPs from the association analysis if their imputation accuracy was $r^2 < 0.3$ or their minor allele frequency (MAF) was $< 0.005$ in BRCA1 or BRCA2 carriers or if their accuracy was $r^2 < 0.25$ in OCAC-iCOGS, the UK GWAS, UK GWAS or Mayo GWAS.

We performed more accurate imputation for the regions around the novel EOC loci from the joint analysis of the data from BRCA1 and BRCA2 carriers and the general population (any SNP with $P < 5 \times 10^{-8}$). The boundaries of these regions were set $+/- 500$kb from any significantly associated SNP in the region. As in the first run, the 1000 Genomes Project data v3 were used as the reference panel and the software IMPUTE2 was applied. However, for the second round of imputation, we imputed genotypes without pre-phasing in order to improve accuracy. To further increase the imputation accuracy we changed some of the default parameters in the imputation procedure. These included an increase of the MCMC iterations to 90 (out of which the first 15 were used as burn-in), an increase of the buffer region to 500kb and an increase of the number of haplotypes used as templates when phasing observed genotypes to 100. These changes were applied consistently for all data sets.

**Statistical analyses**

**Association analyses in the unselected ovarian cancer cases and controls from OCAC**—We evaluated the association between genotype and disease using logistic regression by estimating the associations with each additional copy of the minor allele (log-additive models). The analysis was adjusted for study and for population substructure by including the eigenvectors of the first five ancestry specific principal components as covariates in the model. We used the same approach to evaluate the SNP associations with serous ovarian cancer after excluding all cases with any other or with unknown tumour subtype. For imputed SNPs we used expected dosages in the logistic regression model to estimate SNP effect sizes and p-values. We carried out analyses separately for OCAC-iCOGS and the three GWAS and pooled thereafter using a fixed effects meta-analysis. We carried out the analysis of re-imputed genotypes of putative novel susceptibility loci jointly for the OCAC-iCOGS and GWAS samples. All results are based on the combined data from iCOGS and the three GWAS. We used custom written software for the analysis.

**Associations in BRCA1 and BRCA2 mutation carriers from CIMBA**—We carried out the ovarian cancer association analyses separately for BRCA1 and BRCA2 mutation carriers. The primary analysis was carried out within a survival analysis framework with time to ovarian cancer diagnosis as the endpoint. Mutation carriers were followed until the age of ovarian cancer diagnosis, or risk-reducing salpingo-oophorectomy (RRSO) or age at last observation. Breast cancer diagnosis was not considered as a censoring event. In order to account for the non-random sampling of BRCA1 and BRCA2 mutation carriers with respect to their disease status we conducted the analyses by modelling the retrospective likelihood of the observed genotypes conditional on the disease phenotype. We assessed the associations between genotype and risk of ovarian cancer using the 1 degree of freedom
score test statistic based on the retrospective likelihood\textsuperscript{18,43}. To account for the non-independence among related individuals in the sample, we used an adjusted version of the score test statistic, which uses a kinship adjusted variance of the score\textsuperscript{44}. We evaluated associations between imputed genotypes and ovarian cancer risk using a version of the score test as described above but with the posterior genotype probabilities replacing the genotypes. All analyses were stratified by the country of origin of the samples.

We carried out the retrospective likelihood analyses in CIMBA using custom written functions in Fortran and Python. The score test statistic was implemented in R version 3.0.1\textsuperscript{45}.

We evaluated whether there is evidence for multiple independent association signals in the region around each newly identified locus by evaluating the associations of genetic variants in the region while adjusting for the SNP with the smallest meta-analysis p-value in the respective region. This was done separately for \textit{BRCA1} carriers, \textit{BRCA2} carriers and OCAC.

For one of the novel associations, it was not possible to confirm the imputation accuracy of the lead SNP chr17:29181220:1 at 17q11.2 through genotyping. Therefore, we inferred two-allele haplotypes for rs9910051 and rs3764419, highly correlated with the lead SNP ($r^2=0.95$), using an in-house program. These variants were genotyped on the iCOGS array and therefore this analysis was restricted to 14,733 ovarian cancer cases and 9,165 controls from OCAC-COGS, and 8,185 \textit{BRCA2} mutation carriers that had available genotypes for both variants based on iCOGS. The association between the AA haplotype and risk was tested using logistic regression in OCAC and using Cox regression in \textit{BRCA2} mutation carriers.

**Meta-analysis**—We conducted a meta-analysis of the EOC associations in \textit{BRCA1}, \textit{BRCA2} carriers and the general population for genotyped and imputed SNPs using an inverse variance approach assuming fixed effects. We combined the logarithm of the per-allele hazard ratio estimate for the association with EOC risk in \textit{BRCA1} and \textit{BRCA2} mutation carriers and the logarithm of the per-allele odds ratio estimate for the association with disease status in OCAC. For the associations in \textit{BRCA1} and \textit{BRCA2} carriers, we used the kinship adjusted variance estimator\textsuperscript{44} which allows for inclusion of related individuals in the analysis. We only used SNPs with results in OCAC and in at least one of the \textit{BRCA1} or the \textit{BRCA2} analyses. We carried out two separate meta-analyses, one for the associations with EOC in \textit{BRCA1} carriers, \textit{BRCA2} carriers and EOC in OCAC, irrespective of tumour histological subtype, and a second using only the associations with serous EOC in OCAC. The number of \textit{BRCA1} and \textit{BRCA2} samples with tumour histology information was too small to allow for subgroup analyses. However, previous studies have demonstrated that the majority of EOCs in \textit{BRCA1} and \textit{BRCA2} mutation carriers are high-grade serous\textsuperscript{49–53}. Meta-analyses were carried out using the software “metal”, 2011-03-25 release\textsuperscript{54}.

**Candidate causal SNPs in each susceptibility region**—In order to identify a set of potentially causal variants we excluded SNPs with a likelihood of being causal of less than 1:100, by comparing the likelihood of each SNP from the association analysis with the one
of the most strongly associated SNP\textsuperscript{46}. The remaining variants were then analysed using pupasuite 3.1 to identify potentially functional variants (Supplementary Table 9).

Functional analysis

Expression quantitative trait locus (eQTL) analysis in normal OSE and FTSE cells—Early-passage primary normal ovarian surface epithelial cells (OSECs) and fallopian tube epithelial cells were harvested from disease-free ovaries and fallopian tubes. Normal ovarian epithelial cells were collected by brushing the surface of the ovary with a sterile cytobrush, and were cultured in NOSE-CM\textsuperscript{55}. Fallopian tube epithelial cells were harvested by Pronase digestion as previously described\textsuperscript{56}, plated onto collagen-coated plastics (Sigma) and cultured in DMEM/F12 (Sigma-Aldrich) supplemented with 2% Ultroser G (BioSepra) and 1x penicillin/streptomycin (Lonza). By the time of RNA harvesting, fallopian tube cultures tested consisted of PAX8 positive fallopian tube secretory epithelial cells (FTSECs), consistent with previous observations that ciliated epithelial cells from the fallopian tube do not proliferate \textit{in vitro}.

For gene expression analysis, RNA was harvested from 59 early passage samples: 54 OSECs and 5 FTSECs from cell cultures harvested at \~80\% confluence using the QIAGen miRNAeasy kit with on-column DNase 1 digestion. 500ng RNA was reverse transcribed using the TaqMan\textsuperscript{®} Preamp Mastermix; the resulting product was diluted 1:60 and used to quantify gene expression using the following TaqMan\textsuperscript{®} gene expression probes: WNT4, Hs01573504_m1; RSPO1, Hs00543475_m1; SYNPO2, Hs00326493_m1; ATAD5, Hs00227495_m1 and GPX6, Hs00699698_m1. Four control genes were also included: ACTB, Hs00357333_g1; GAPDH, Hs02758991_g1; HMBS, Hs00609293_g1 and HPRT1 Hs02800695_m1 (all Life Technologies). Assays were run on an ABI 7900HT Fast Real-Time PCR system (Life Technologies).

Data Analysis: Expression levels for each gene were normalized to the average of all four control genes. Relative expression levels were calculated using the \(\delta\delta\)Ct method. Genotyping was performed on the iCOGs chips, as described above. Where genotyping data were not available for the most risk-associated SNP, the next most significant SNP was used: rs3820282 at 1p36, rs12023270 at 1p34.3, rs752097 at 4q26, rs445870 at 6p22.1, rs505922 at 9q34.2 and rs3764419 at 17q11.2. Correlations between genotype and gene expression were calculated in ‘R’. Genotype specific gene expression in the normal tissue cell lines (eQTL analysis) was compared using the Jonckheere-Terpstra test. IData were normalized to the four control genes and we tested for eQTL associations, grouping OSECs and FTSECs together. Secondly, OSECs were analysed alone. eQTL analyses were performed using 3 genotype groups, or two groups (with the rare homozygote samples grouped together with the heterozygote samples).

eQTL analysis in primary ovarian tumours—eQTL analysis in primary tumours was based on the publicly available data available from The Cancer Genome Atlas (TCGA) project, which includes 489 primary high grade serous ovarian cancers. The methods have been described elsewhere\textsuperscript{57}. Briefly, we determined the ancestry for each case based on the
germ line genotype data using EIGENSTRAT software with 415 HapMap genotype profiles as a control set. Only populations of Northern and Western European ancestries were included. We first performed a \textit{cis}-eQTL analyses using a method we described previously, in which the association between 906,600 germ line genotypes and the expression levels of mRNA or miRNA (located within 500Kb on either side of the variant) were evaluated using linear regression model with the effects of somatic copy number and CpG methylation being deducted (For miRNA expression, the effect of CpG methylation is not adjusted for since the data are not available). To adjust for multiple tests, we adjusted the test P values using Benjamini-Hochberg method. A significant association was defined by a false discovery rate (FDR) of less than 0.1.

Having established a genome-wide \textit{cis}-eQTL associations in this series of tumours, we then evaluated \textit{cis}-eQTL associations for the top risk associations between each of the six new loci and the gene in closest proximity to the risk SNP. For each risk locus, we retrieved the genotype of all SNPs in ovarian cancer cases based on the Affymetrix 6.0 array. Using these genotypes and the impute2 March 2012 1000 Genomes Phase I integrated variant cosmopolitan reference panel of 1,092 individuals (Haplotypes were phased via SHAPEIT), we imputed the genotypes of SNPs in the 1000 Genomes Project in the target regions for TCGA samples\textsuperscript{58}. For each risk locus where data for the most risk-associated variant were not available, we retrieved the imputed variants tightly correlated with the most risk-associated variant. We then tested for association between imputed SNPs and gene expression using the linear regression algorithm described above, where each imputed SNP was coded as an expected allele count. Again, significant associations are defined by a false discovery rate (FDR) of less than 0.1.

**Regulatory profiling of normal ovarian cancer precursor tissues**—We performed genome-wide formaldehyde assisted regulatory element (FAIRE) and ChIP seq with histone 3 lysine 27 acetylation (H3K27ac) and histone 3 lysine 4 monomethylation (H3K4me) for two normal OSECs, two normal FTSECs and two HGSOC cell lines (UWB1.289 and CAOV3) [Shen et al. in preparation]. These datasets annotate epigenetic signatures of open chromatin, and collectively indicate transcriptional enhancer regions. We analysed the FAIRE-seq and ChIP-seq datasets and publically available genomic data on promoter and UTR domains, intron/exon boundaries, and positions of non-coding RNA transcripts to identify SNPs from the 100:1 likely causal set that align with biofeatures that may provide evidence of SNP functionality.

**Candidate Gene Analysis Using Genome Wide Profiling of Primary Ovarian Cancers**

**Data Sets:** The Cancer Genome Atlas (TCGA) Project and COSMIC Datasets

TCGA has performed extensive genomic analysis of tumours from a large number of tissue types including almost 500 high-grade serous ovarian tumours. These data include somatic mutations, DNA copy number, mRNA and miRNA expression and DNA methylation. COSMIC is the catalogue of somatic mutations in cancer that collates information on mutations in tumours from the published literature\textsuperscript{59}. They have also identified The Cancer
Gene Census, which is a list of genes known to be involved in cancer. Data are available on a large number of tissue types, including 2,809 epithelial ovarian tumours.

**Somatic coding sequence mutations:** We analysed all genes for coding somatic sequence mutations generated from either whole exome or whole genome sequencing. In TCGA, whole exome sequencing data were available for 316 high-grade serous EOC cases. In addition, we determined whether mutations had been reported in COSMIC and whether the gene was a known cancer gene in the Sanger Cancer Gene Census.

**mRNA expression in tumour and normal tissue:** Normalized and gene expression values (Level 3) gene expression profiling data were obtained from the TCGA data portal for three different platforms (Agilent, Affymetrix HuEx and Affymetrix U133A). We analysed only the 489 primary serous ovarian tumour samples included in the final clustering analysis and eight normal fallopian tube samples. The boxplot function in R was used to compare ovarian tumour samples to the fallopian tube for 91 coding genes with expression data on any platform within a 1MB region around the most significant SNP at the six loci. A difference in relative expression between EOC and normal tissue was carried out using the Wilcoxon rank-sum test.

**DNA copy number analysis:** Serous EOC samples for 481 tumours with log2 copy number data were analysed using the cBio portal for analysis of TCGA data. For each gene in a region the classes of copy number: homozygous deletion, heterozygous loss, diploid, gain, and amplification were queried individually using the advanced onco query language (OQL) option. The frequency of gain and amplification were combined as “gain”, and homozygous deletion and heterozygous loss were combined as “loss”.

**Analysis of copy number vs mRNA expression:** Serous EOC samples for 316 complete tumours (those with CNA, mRNA and sequencing data) were analysed. Graphs were generated using the cBio portal for analysis of TCGA data and the setting were mRNA expression data Z-score (all genes) with the Z-score threshold of 2 (default setting) and putative copy number alterations (GISTIC). The Z-score is the number of standard deviations away from the mean of expression in the reference population. GISTIC is an algorithm that attempts to identify significantly altered regions of amplification or deletion across sets of patients.

**Luciferase Reporter Assay**—The putative causal SNPs at the 1p36 locus lie in the WNT4 promoter and so we tested their effect on transcription in a luciferase reporter assay (Fig. 2D). Wild-type and risk haplotype (comprising five correlated variants) sequences corresponding to the region bound by hg19 co-ordinates chr1:22469416-22470869 were generated by Custom Gene Synthesis (GenScript Corporation), and then sub-cloned into pGL3-basic (Promega). Equimolar amounts of luciferase constructs (800 ng) and pRL-TK Renilla (50 ng) were co-transfected into ~8 × 10^4 iOSE4 normal ovarian cells in triplicate wells of 24 well plates using LipoFectamine 2000 (Life Technologies). Independent transfections were repeated three times. The Dual-Glo Luciferase Assay kit (Promega) was used to assay luciferase activity 24 hours post transfection using a BioTek Synergy H4 plate reader.

*Nat Genet. Author manuscript; available in PMC 2015 August 01.*
reader. The iOSE-4 cell line (derived by K. Lawrenson) was maintained under standard conditions and routinely tested for *Mycoplasma* and short tandem repeat profiled.

**Supplementary Material**

Refer to Web version on PubMed Central for supplementary material.

**Authors**

Karoline B. Kuchenbaecker\(^1\), Susan J. Ramus\(^2\), Jonathan Tyrer\(^3\), Andrew Lee\(^1\), Howard C. Shen\(^2\), Jonathan Beesley\(^4\), Kate Lawrenson\(^2\), Lesley McGuffog\(^1\), Sue Healey\(^4\), Janet M. Lee\(^2\), Tassa J. Spindler\(^2\), Yvonne G. Lin\(^5\), Tanja Pejovic\(^6,7\), Yukie Bean\(^6,7\), Qiyuan Li\(^8\), Simon Coetzee\(^9,10,11\), Dennis Hazelett\(^12,13\), Alexander Miron\(^14\), Melissa Southey\(^15\), Mary Beth Terry\(^16\), David E. Goldgar\(^17\), Saundra S. Buys\(^18\), Ramunas Janavicius\(^19,20\), Cecilia M. Dorfling\(^21\), Elizabeth J. van Rensburg\(^22\), Susan L. Neuhausen\(^22\), Yuan Chun Ding\(^22\), Thomas V. O. Hansen\(^23\), Lars Jønson\(^23\), Anne-Marie Gerdes\(^24\), Bent Ejertsen\(^25\), Daniel Barrowdale\(^1\), Joe Dennis\(^1,3\), Javier Benitez\(^26,27,28\), Ana Osorio\(^26,28\), Maria Jose Garcia\(^26,28\), Ian Komenaka\(^29\), Jeffrey N. Weitzel\(^30\), Pamela Ganschow\(^31\), Paolo Peterlongo\(^32\), Loris Bernard\(^33,34\), Alessandra Viel\(^35\), Bernardo Bonanni\(^36\), Bernard Peisse\(^37\), Siranoush Manoukian\(^37\), Paolo Radice\(^38\), Laura Papi\(^39\), Laura Ottini\(^40\), Florentia Fostira\(^41\), Irene Konstantopoulou\(^41\), Judy Garber\(^42\), Debra Frost\(^1\), Jo Perkins\(^1\), Radka Platte\(^1\), Steve Ellis\(^1\), EBRACE\(^43\), Andrew K. Godwin\(^44\), Rita Katharina Schmutzler\(^45,46,47,48\), Alfons Meindl\(^49\), Christoph Engel\(^50\), Christian Sutter\(^51\), Olga M. Sinilikova\(^52,53\), GEMO Study Collaborators\(^43\), Francesca Damiola\(^52\), Sylvie Mazoyer\(^52\), Dominique Stoppa-Lyonnet\(^54,55,56\), Kathleen Claes\(^57\), Kim De Leeneer\(^57\), Judy Kirk\(^58\), Gustavo C. Rodriguez\(^59\), Marion Piedmonte\(^60\), David M. O’Malley\(^61\), Miguel de la Hoya\(^62\), Trinidad Caldes\(^62\), Kristiina Aittomäki\(^63\), Heli Nevanlinna\(^64\), J. Margriet Collée\(^65\), Matti A. Rookus\(^66\), Jan C. Oosterwijk\(^67\), Breast Cancer Family Registry\(^43\), Laima Tihomirova\(^68\), Nadine Tung\(^69\), Ute Hamann\(^70\), Claudine Isaacs\(^71\), Marc Tischkowitz\(^72\), Evgeny N. Imyanitov\(^73\), Maria A. Caligo\(^74\), Ian Campbell\(^75\), Frans B.L. Hogervorst\(^76\), HEBON\(^43\), Edith Olah\(^77\), Orland Diez\(^78\), Ignacio Blanco\(^79\), Joan Brunet\(^80\), Conxi Lazaro\(^81\), Miquel Angel Pujana\(^82\), Anna Jakubowska\(^83\), Jacek Gronwald\(^83\), Jan Lubinski\(^83\), Grzegorz Sukiennicki\(^83\), Rosa B. Barkardottir\(^84\), Marie Plante\(^85\), Jacques Simard\(^86\), Penny Soucy\(^86\), Marco Montagna\(^87\), Silvia Tognazzo\(^87\), Manuel R. Teixeira\(^88,89\), KConFab Investigators\(^43\), Vernon S. Pankratz\(^90\), Xianshu Wang\(^91\), Noralane Lindor\(^90\), Csilla I. Szabo\(^92\), Noah Kauf\(^93\), Joseph Vijai\(^93\), Carol A. Aghajanian\(^93\), Georg Pfeifer\(^94\), Andreas Berger\(^94\), Christian F. Singer\(^94\), Muy-Kheng Tea\(^94\), Catherine M. Phelan\(^95\), Mark H. Greene\(^96\), Phuong L. Mai\(^96\), Gad Rennert\(^97\), Anna Marie Mulligan\(^98,99\), Sandrine Tchatchou\(^100\), Irene L. Andrus\(^98,101\), Gord Glendon\(^100\), Amanda Ewart Toland\(^102\), Uffe Birk Jensen\(^103\), Torben A. Kruse\(^104\), Mads Thomassen\(^104\), Anders Bojesen\(^105\), Jamal Zidan\(^106\), Eitan Friedman\(^107\), Yael Laitman\(^107\), Maria Soller\(^108\), Annelie Liljegren\(^109\), Brita Arve\(^109\), Zakaria Einbeigi\(^110\), Marie Stenmark-Askma\(^111\), Olufunmilayo I. Olopade\(^112\), Robert L. Nussbaum\(^113\), Timothy R. Rebbeck\(^114\), Katherine L. Nathanson\(^114\), Susan M. Domchek\(^114\), Karen H. Lu\(^115\), Beth Y. Nat Genet. Author manuscript; available in PMC 2015 August 01.
Affiliations

1Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK 2Department of Preventive Medicine, Keck School of Medicine, University of Southern California Norris Comprehensive Cancer Center, Los Angeles, CA, USA 3Department of Oncology, University of Cambridge, Strangeways Research Laboratory, Cambridge, UK 4Cancer Division, QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia 5Department of Obstetrics and Gynecology, University of Southern California, Los Angeles, CA, USA 6Department of Obstetrics and Gynecology, Oregon Health and Science University, Portland, OR, USA 7Knight Cancer Institute, Portland, OR, USA 8Department of Medical Oncology, The Center for Functional Cancer Epigenetics, Dana-Farber Cancer Institute, Boston, MA, USA 9Department of Genetics, Ribeirão Preto Medical School, University of São Paulo, Brazil 10Center For Cell Based Therapy, Monte Alegre, Ribeirão Preto, SP, Brazil 11Center for Integrative Systems Biology, Monte Alegre, Ribeirão Preto, SP, Brazil 12Department of Preventive Medicine, University of Southern California, Los Angeles, CA, USA 13Department of Urology, University of Southern California, Los Angeles, CA, USA 14Department of Genomics and Genome Sciences, Case Western Reserve University Medical School, Cleveland, OH, USA 15Center for Integrative Systems Biology, Minneapolis, MN, USA 16Probability and Data Science, University of California, San Francisco, CA, USA 17Center for Functional Cancer Epigenetics, Dana-Farber Cancer Institute, Boston, MA, USA 18Department of Obstetrics and Gynecology, Portland, OR, USA 19Department of Preventive Medicine, University of Southern California, Los Angeles, CA, USA 20Department of Urology, University of Southern California, Los Angeles, CA, USA 21Department of Genetics, Ribeirão Preto Medical School, University of São Paulo, Brazil 22Department of Population Sciences, Beckman Research Institute of City of Hope, Duarte, CA, USA 23Center for Genomic Medicine, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark 24Department of Clinical Genetics, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark 25Department of Oncology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark 26Human Genetics Group, Human Cancer Genetics Program, Spanish National Cancer Centre (CNIO), Madrid, Spain 27Genotyping Unit (CeGen), Human Cancer Genetics Program, Spanish National Cancer Centre (CNIO), Madrid, Spain 28Biomedical Network on Rare Diseases (CIBERER), Madrid, Spain 29Maricopa Medical Center, care of City of Hope Clinical Cancer Genetics Community Research Network, Duarte, CA, USA 30Clinical Cancer Genetics, for the City of Hope Clinical Cancer Genetics Community Research Network, Duarte, CA, USA 31Cook County Health and Hospital System, care of City of Hope Clinical Cancer Genetics Community Research Network, Duarte, CA, USA 32IFOM, Fondazione Istituto FIRC di Oncologia Molecolare, Milan, Italy 33Department of Experimental Oncology, Istituto Europeo di Oncologia, Milan, Italy
34Cogentech Cancer Genetic Test Laboratory, Milan, Italy 35Division of Experimental Oncology, CRO Aviano National Cancer Institute, Aviano (PN), Italy 36Division of Cancer Prevention and Genetics, Istituto Europeo di Oncologia, Milan, Italy 37Unit of Medical Genetics, Department of Preventive and Predictive Medicine, Fondazione IRCCS Istituto Nazionale Tumori (INT), Milan, Italy 38Unit of Molecular Bases of Genetic Risk and Genetic Testing, Department of Preventive and Predictive Medicine, Fondazione IRCCS Istituto Nazionale Tumori (INT), Milan, Italy 39Unit of Medical Genetics, Department of Biomedical, Experimental and Clinical Sciences, University of Florence, Florence, Italy 40Department of Molecular Medicine, University La Sapienza, Rome, Italy 41Molecular Diagnostics Laboratory, INRASTES, National Centre for Scientific Research "Demokritos", Aghia Paraskevi Attikis, Athens, Greece 42Cancer Risk and Prevention Clinic, Dana Farber Cancer Institute, Boston, MA, USA 43Department of Pathology and Laboratory Medicine, University of Kansas Medical Center, Kansas City, KS, USA 44Center for Hereditary Breast and Ovarian Cancer, Medical Faculty, University Hospital Cologne, Germany 45Center for Integrated Oncology (CIO), Medical Faculty, University Hospital Cologne, Germany 46Center for Molecular Medicine Cologne (CMMC), University of Cologne, Germany 47Department of Gynaecology and Obstetrics, Division of Tumor Genetics, Klinikum rechts der Isar, Technical University of Munich, Munich, Germany 48Institute for Medical Informatics, Statistics and Epidemiology University of Leipzig, Leipzig, Germany 49University Heidelberg, Heidelberg, Germany 50INSERM U1052, CNRS UMR5286, Université Lyon, Centre de Recherche en Cancérologie de Lyon, Lyon, France 51Unité Mixte de Génétique Constitutionnelle des Cancers Fréquents, Hospices Civils de Lyon – Centre Léon Bérard, Lyon, France 52Institut Curie, Department of Tumour Biology, Paris, France 53Institut Curie, INSERM U830, Paris, France 54Université Paris Descartes, Sorbonne Paris Cité, France 55Center for Medical Genetics, Ghent University, Ghent, Belgium 56Australia New Zealand Gynecologic Oncology Group (ANZGOG) and Familial Cancer Service, Westmead Hospital, Sydney, Australia 57Division of Gynecologic Oncology, NorthShore University HealthSystem, Evanston, IL, USA 58Gynecologic Oncology Group Statistical and Data Center, Roswell Park Cancer Institute, Buffalo, NY, USA 59James Cancer Center, Ohio State University, Columbus, OH, USA 60Molecular Oncology Laboratory, Hospital Clinico San Carlos, IdISSC, Madrid, Spain 61Department of Clinical Genetics, Helsinki University Central Hospital, Helsinki, Finland 62Department of Obstetrics and Gynecology, University of Helsinki and Helsinki University Central Hospital, Helsinki, HUS, Finland 63Department of Clinical Genetics, Family Cancer Clinic, Erasmus University Medical Center, Rotterdam, The Netherlands 64Department of Epidemiology, Netherlands Cancer Institute, Amsterdam, The Netherlands 65Department of Genetics, University Medical Center, Groningen University, Groningen, The Netherlands 66Latvian Biomedical Research and Study Centre, Riga, Latvia 67Department of Medical Oncology, Beth Israel Deaconess Medical Center 68Molecular Genetics of Breast Cancer, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany 69Lombardi Comprehensive Cancer Center, Georgetown University, Washington DC
DC, USA 72
Program in Cancer Genetics, McGill University, Montreal, Quebec, Canada 73
N.N. Petrov Institute of Oncology, St.Petersburg, Russia 74
Section of Genetic Oncology, Department of Laboratory Medicine, University and University Hospital of Pisa, Pisa, Italy 75
VBCRC Cancer Genetics Laboratory, Peter MacCallum Cancer Centre, Melbourne, Australia 76
Family Cancer Clinic, Netherlands Cancer Institute, Amsterdam, The Netherlands 77
Department of Molecular Genetics, National Institute of Oncology, Budapest, Hungary 78
Oncogenetics Group, University Hospital Vall d'Hebron, Vall d'Hebron Institute of Oncology (VHIO) and Universitat Autònoma de Barcelona; Barcelona, Spain 79
Genetic Counseling Unit, Hereditary Cancer Program, IDIBELL-Catalan Institute of Oncology, Barcelona, Spain 80
Genetic Counseling Unit, Hereditary Cancer Program, IDIBGI-Catalan Institute of Oncology, Girona, Spain 81
Molecular Diagnostic Unit, Hereditary Cancer Program, IDIBELL-Catalan Institute of Oncology, Barcelona, Spain 82
Breast Cancer and Systems Biology Unit, IDIBELL-Catalan Institute of Oncology, Barcelona, Spain 83
Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland 84
Department of Pathology, Landspitali University Hospital and BMC, Faculty of Medicine, University of Iceland, Reykjavik, Iceland 85
Gynaecologic Oncology Service, Centre Hospitalier Universitaire de Québec (CHUQ), Québec, Canada 86
Centre Hospitalier Universitaire de Québec Research Center, Laval University, Quebec City, Canada 87
Immunology and Molecular Oncology Unit, Istituto Oncologico Veneto IOV-IRCCS, Padua, Italy 88
Biomedical Sciences Institute (ICBAS), Porto University, Porto, Portugal 89
Department of Genetics, Portuguese Oncology Institute, Porto, Portugal 90
Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA 91
Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA 92
National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA 93
Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, USA 94
Department of Obstetrics and Gynecology, Comprehensive Cancer Center, Medical University of Vienna, Vienna, Austria 95
Department of Cancer Epidemiology, Moffitt Cancer Center, Tampa, FL, USA 96
Clinical Genetics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, MD, USA 97
Department of Community Medicine and Epidemiology, Carmel Medical Center, Haifa, Israel 98
Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON, Canada 99
Laboratory Medicine Program, University Health Network, Toronto, ON, Canada 100
Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, ON, Canada 101
Department of Molecular Genetics, University of Toronto, ON, Canada 102
Department of Molecular Virology, Immunology and Medical Genetics, Columbus, OH, USA 103
Department of Clinical Genetics, Aarhus University Hospital, Aarhus N, Denmark 104
Department of Clinical Genetics, Odense University Hospital, Odense C, Denmark 105
Department of Clinical Genetics, Vejle Hospital, Vejle, Denmark 106
Institute of Oncology, Rivka Ziv Medical Center, Zefat, Israel 107
Susanne Levy Gertner Oncogenetics Unit, Sheba Medical Center, Tel Aviv, Israel 108
Department of Clinical Genetics, Lund University Hospital, Lund, Sweden
Department of Oncology, Karolinska University Hospital, Stockholm, Sweden
Department of Oncology, Sahlgrenska University Hospital, Gothenburg, Sweden
Division of Clinical Genetics, Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden
Center for Clinical Cancer Genetics and Global Health, University of Chicago Medical Center, Chicago, USA
Department of Medicine and Genetics, University of California, San Francisco, USA
Abramson Cancer Center, Perelman School of Medicine, University of Pennsylvania, PA, USA
Department of Gynecologic Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA
Women's Cancer Program at the Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA
Department of Obstetrics and Gynecology, Erlangen University Hospital, University of Erlangen-Nuremberg, Erlangen, Germany
Institute of Human Genetics, University of Erlangen-Nuremberg, Erlangen, Germany
University of California at Los Angeles, David Geffen School of Medicine, Department of Medicine, Division of Hematology and Oncology, Los Angeles, CA, USA
Vesalius Research Center, VIB, Leuven, Belgium
Laboratory for Translational Genetics, Department of Oncology, University of Leuven, Belgium
Division of Gynecological Oncology, Department of Oncology, University Hospitals Leuven, Belgium
Department of Community and Family Medicine, Section of Biostatistics & Epidemiology, The Geisel School of Medicine at Dartmouth, Lebanon, NH, USA
Program in Epidemiology, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA
Department of Epidemiology, University of Washington, Seattle, WA, USA
German Cancer Research Center (DKFZ), Division of Cancer Epidemiology, Heidelberg, Germany
Department of Obstetrics and Gynecology, University of Ulm, Ulm, Germany
Department of Cancer Prevention and Control, Roswell Park Cancer Institute, Buffalo, NY, USA
Department of Gynecologic Oncology, Roswell Park Cancer Institute, Buffalo, NY, USA
Cancer Epidemiology Program, University of Hawaii Cancer Center, Hawaii, USA
Cancer Prevention and Control, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA
Community and Population Health Research Institute, Department of Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, CA, USA
Department of Gynecology, Jena University Hospital - Friedrich Schiller University, Jena, Germany
Clinics of Obstetrics and Gynaecology, Hannover Medical School, Hannover, Germany
Gynaecology Research Unit, Hannover Medical School, Hannover, Germany
Byelorussian Institute for Oncology and Medical Radiology Aleksandrov N.N., Minsk, Belarus
Department of Pathology, Helsinki University Central Hospital, Helsinki, Finland
Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA
Department of Obstetrics, Gynecology and Reproductive Sciences, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA
Ovarian Cancer Center of Excellence, University of Pittsburgh, Pittsburgh, PA, USA
Women's Cancer Research Program, Magee-Women's Research Institute and University of Pittsburgh Cancer Institute, Pittsburgh, PA, USA
The University

Kuchenbaecker et al. Nat Genet. Author manuscript; available in PMC 2015 August 01.
of Texas School of Public Health, Houston, TX, USA 143Department of Gynecology and Gynecologic Oncology, Dr. Horst Schmidt Kliniken Wiesbaden, Wiesbaden, Germany 144Department of Gynecology and Gynecologic Oncology, Kliniken Essen-Mitte, Essen, Germany 145Institut für Humangenetik Wiesbaden, Wiesbaden, Germany 146Department of Preventive Medicine, Kyushu University Faculty of Medical Sciences, Fukuoka, Fukuoka, Japan 147Division of Epidemiology and Prevention, Aichi Cancer Center Research Institute, Nagoya, Aichi, Japan 148Department of Virus, Lifestyle and Genes, Danish Cancer Society Research Center, Copenhagen, Denmark 149Department of Gynecology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark 150Molecular Unit, Department of Pathology, Herlev Hospital, University of Copenhagen, Copenhagen, Denmark 151Cancer Research Initiatives Foundation, Sime Darby Medical Centre, Subang Jaya, Malaysia 152Department of Obstetrics and Gynaecology, University Malaya Medical Centre, University Malaya, Kuala Lumpur, Malaysia 153University Malaya Cancer Research Institute, Faculty of Medicine, University Malaya Medical Centre, University Malaya, Kuala Lumpur, Malaysia 154Biostatistics and Informatics Shared Resource, University of Kansas Medical Center, Kansas City, KS, USA 155Department of Health Science Research, Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN, USA 156Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, VIC, Australia 157College of Pharmacy and Health Sciences, Texas Southern University, Houston, Texas, USA 158Department of Epidemiology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA 159Gynecology Service, Department of Surgery, Memorial Sloan-Kettering Cancer Center, New York, NY, USA 160Department of Obstetrics and Gynecology, Duke University Medical Center, Durham, NC, USA 161Department of Statistical Science, Duke University, Durham, North Carolina, USA 162Cancer Control and Population Sciences, Duke Cancer Institute, Durham, North Carolina, USA 163Department of Community and Family Medicine, Duke University Medical Center, Durham, NC, USA 164Department of Pathology, Immunology and Laboratory Medicine, University of Florida, Gainesville, FL, USA 165Genetics Institute, University of Florida, Gainesville, FL, USA 166Harvard School of Public Health, Boston, MA, USA 167Obstetrics and Gynecology Epidemiology Center, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA, USA 168Channing Division of Network Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA, USA 169Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA 170Cancer Prevention and Control, Rutgers Cancer Institute of New Jersey, New Brunswick, New Jersey, USA 171Memorial Sloan Kettering Cancer Center, Department of Epidemiology and Biostatistics, Epidemiology Service, New York, NY, USA 172Centre for Cancer Biomarkers, Department of Clinical Science, University of Bergen, Bergen, Norway 173Department of Gynecology and Obstetrics, Haukeland University Hospital, Bergen, Norway 174Department of Gynaecology, Radboud University Medical Centre, Nijmegen, The Netherlands 175Comprehensive Cancer Center The Netherlands, Utrecht, The Netherlands 176Department for Health Evidence, Radboud University Medical Centre, Nijmegen, The Netherlands
Acknowledgments

We thank all the individuals who took part in this study, and all the researchers, clinicians and technical and administrative staff who made possible the many studies contributing to this work (a full list is provided in the Supplementary Note), including Xiao Qing Chen for iPLEX genotyping. The COGS project is funded through a European Commission’s Seventh Framework Programme grant (agreement number 223175 - HEALTH-F2-2009-223175). The CIMBA data management and data analysis were supported by Cancer Research–UK grants C12292/A11174 and C1287/A10118. The Ovarian Cancer Association Consortium is supported by a grant from the Department of Urology, Radboud University Medical Centre, Nijmegen, Netherlands. Canada’s Michael Smith Genome Sciences Centre, BC Cancer Agency, Vancouver, BC, Canada. Department of Biomedical Physiology and Kinesiology, Simon Fraser University, Burnaby, BC Canada. Department of Public Health Sciences, College of Medicine, Medical University of South Carolina, Charleston, USA. Division of Epidemiology and Biostatistics, Department of Internal Medicine, University of New Mexico, Albuquerque, NM, USA. Cancer Control Research, BC Cancer Agency, Vancouver, BC, Canada. International Hereditary Cancer Center, Department of Genetics and Pathology, Clinic of Ophthalmology, Pomeranian Medical University, Szczecin, Poland. Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA. Department of Cancer Epidemiology and Prevention, Maria Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland. Department of Pathology, Rigshospitalet, University of Copenhagen, Denmark. Wessex Clinical Genetics Service, Princess Anne Hospital, Southampton, UK. Institute of Cancer Sciences, University of Glasgow, Wolfson Wohl Cancer Research Centre, Beatson Institute for Cancer Research, Glasgow, UK. Cancer Research UK Clinical Trials Unit, Glasgow, The Beatson West of Scotland Cancer Centre, Glasgow, UK. Department of Gynaecological Oncology, Glasgow Royal Infirmary, Glasgow, UK. Department of Health Research and Policy - Epidemiology, Stanford University School of Medicine, Stanford, CA, USA. Vanderbilt University School of Medicine, Nashville, TN, USA. Shanghai Cancer Institute, Shanghai, China. Department of Obstetrics and Gynecology, Faculty of Medicine, University of Toronto, Toronto, ON, Canada. Department of Gynecologic-Oncology, Princess Margaret Hospital, Toronto, ON, Canada. Department of Chronic Disease Epidemiology, Yale School of Public Health, New Haven, CT, USA. Prosserman Centre for Health Research, Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, ON, Canada. Women’s College Research Institute, University of Toronto, Toronto, ON, Canada. Department of Biostatistics, Moffitt Cancer Center, Tampa, FL, USA. Department of Epidemiology, University of California Irvine, Irvine, CA, USA. Women’s Cancer, UCL EGA Institute for Women’s Health, London, UK. Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, NY, USA. Department of Pathology, The Maria Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland. Department of Obstetrics, Gynecology and Oncology, Illd Faculty of Medicine, Warsaw Medical University and Brodnowski Hospital, Warsaw, Poland.
Ovarian Cancer Research Fund thanks to donations by the family and friends of Kathryn Sladek Smith (PPD/RCPI. 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. Funding for the project was provided 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 cBio portal is developed and maintained by the Computational Biology Center at Memorial Sloan-Kettering Cancer Center. SH is supported by an NHMRC Program Grant to GCT. Details of the funding of individual investigators and studies are provided in the Supplementary Note. This study made use of data generated by the Wellcome Trust Case Control consortium, funding for which was provided 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). A full list of the investigators who contributed to the generation of the data is available on the website (see URL). The cBio portal was developed and is maintained by the Computational Biology Center at Memorial Sloan-Kettering Cancer Center.

REFERENCES


Nat Genet. Author manuscript; available in PMC 2015 August 01.
40. GenABEL: genome-wide SNP association analysis. 2013
Figure 1.
Hazard ratios for the association with EOC of 12 previously reported epithelial ovarian cancer susceptibility variants and the six novel susceptibility variants for OCAC, BRCA1 mutation carriers and BRCA2 mutation carriers
Figure 2.
The 1p36 epithelial ovarian cancer susceptibility locus
A) The Manhattan plot depicts the strength of association between all imputed and genotyped SNPs across the region bound by hg19 co-ordinates chr1:21922893-22991643. The dotted line represents the genome-wide significance level $5 \times 10^{-8}$. Additional tracks show genes and enhancers in ovary as described in Hnisz et al\textsuperscript{38}. Positions of SNPs for which imputation $r^2<0.3$ and/or minor allele frequency $<0.005$ are shown in the bottom track as ‘untyped’ SNPs.
B) The shaded iCHAIV from (A) is shown depicting the genes and the location of the $WNT4$ promoter construct as a red box. Red ticks show the positions of the putative causal variants following likelihood ratio testing. Signals from FAIRE-seq data derived from ovarian cells are represented by black marks, and the locations of predicted $CDC42$ enhancers\textsuperscript{38} as blue boxes. The positions of genotyped SNPs, and those that were neither genotyped nor well imputed (‘untyped’), are shown.
C) Normalised luciferase reporter activity following triplicate transfections of wildtype and risk haplotype \textit{WNT4} promoter constructs in iOSE4 cells. Error bars represent standard error from three independent experiments.
Table 1

Association test results for loci associated at $p<5 \times 10^{-8}$ in the second imputation stage. Results reported for ovarian cancer in BRCA1 and BRCA2 mutation carriers, ovarian cancer as well as serous subtype in OCAC, the meta-analysis for ovarian cancer, and for the meta-analysis for all tumour histologies in BRCA1 and BRCA2 and serous ovarian cancer in OCAC. SNP with smallest $p$-value reported for each locus.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Nearest gene</th>
<th>rs#</th>
<th>Ref</th>
<th>Eff</th>
<th>EAF</th>
<th>OR (95% CI)</th>
<th>P</th>
<th>OR (95% CI)</th>
<th>P</th>
<th>HR (95% CI)</th>
<th>P</th>
<th>OR (95% CI)</th>
<th>P</th>
<th>HR (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1p36</td>
<td>WNT4</td>
<td>rs56318008</td>
<td>C</td>
<td>T</td>
<td>0.15</td>
<td>0.98</td>
<td>1.11</td>
<td>0.98</td>
<td>1.12</td>
<td>0.98</td>
<td>1.12</td>
<td>3.1×10^{-6}</td>
<td>0.98</td>
<td>1.11</td>
<td>3.1×10^{-3}</td>
</tr>
<tr>
<td>1p34.3</td>
<td>RSP01</td>
<td>rs58722170</td>
<td>G</td>
<td>C</td>
<td>0.23</td>
<td>0.85</td>
<td>1.08</td>
<td>0.85</td>
<td>1.12</td>
<td>0.85</td>
<td>1.12</td>
<td>9.7×10^{-5}</td>
<td>0.85</td>
<td>1.14</td>
<td>1.5×10^{-3}</td>
</tr>
<tr>
<td>4q26</td>
<td>SYNPO2</td>
<td>rs17329882</td>
<td>A</td>
<td>C</td>
<td>0.24</td>
<td>0.95</td>
<td>1.09</td>
<td>0.95</td>
<td>1.11</td>
<td>0.95</td>
<td>1.11</td>
<td>5.9×10^{-7}</td>
<td>0.95</td>
<td>1.08</td>
<td>0.93</td>
</tr>
<tr>
<td>6p22.1</td>
<td>GPX6</td>
<td>rs116133110f</td>
<td>C</td>
<td>T</td>
<td>0.31</td>
<td>0.99</td>
<td>0.93</td>
<td>0.99</td>
<td>0.91</td>
<td>0.99</td>
<td>0.91</td>
<td>9.0×10^{-8}</td>
<td>0.99</td>
<td>0.92</td>
<td>0.23</td>
</tr>
<tr>
<td>9q34.2</td>
<td>ABO</td>
<td>rs636634</td>
<td>C</td>
<td>T</td>
<td>0.19</td>
<td>0.85</td>
<td>1.11</td>
<td>0.85</td>
<td>1.12</td>
<td>0.85</td>
<td>1.12</td>
<td>1.1×10^{-7}</td>
<td>1.12</td>
<td>1.11</td>
<td>1.0×10^{-6}</td>
</tr>
<tr>
<td>17q11.22</td>
<td>ATAD5</td>
<td>chr17:29181220</td>
<td>A</td>
<td>T</td>
<td>0.28</td>
<td>0.95</td>
<td>0.91</td>
<td>0.95</td>
<td>0.91</td>
<td>0.95</td>
<td>0.91</td>
<td>5.4×10^{-8}</td>
<td>0.91</td>
<td>1.01</td>
<td>0.88</td>
</tr>
</tbody>
</table>

*Imputation accuracy $r^2$ estimate

1. $p$-value from the meta-analysis association test for ovarian cancer in OCAC and BRCA1 and BRCA2 carriers

2. $p$-value from the meta-analysis association test for ovarian cancer in BRCA1 and BRCA2 carriers and serous ovarian cancer in OCAC

3. $p$-value from the meta-analysis of ovarian cancer associations in BRCA2 carriers and OCAC only

4. rs116133110 listed as rs6456822 on dbSNP

5. chr17:29181220 listed as rs19661266 on dbSNP

6. Reference and effect allele

7. Effect allele frequency
### Table 2
Associations with ovarian cancer subtypes in OCAC samples for loci associated with ovarian cancer at $p<5\times10^{-8}$ in the meta-analysis

<table>
<thead>
<tr>
<th>Locus</th>
<th>rs#</th>
<th>All histologies OR (95% CI) P</th>
<th>Serous OR (95% CI) P</th>
<th>Endometrioid OR (95% CI) P</th>
<th>Clear cell OR (95% CI) P</th>
<th>Mucinous OR (95% CI) P</th>
<th>p-het*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1p36</td>
<td>rs56318008</td>
<td>1.11 (1.06–1.15) 8×10⁻⁷ 1.12 (1.06–1.17) 6×10⁻⁶ 1.09 (1.00–1.19) 0.05 1.24 (1.10–1.39) 5×10⁻⁴ 1.03 (0.91–1.17) 0.65 0.22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1p34.3</td>
<td>rs58722170</td>
<td>1.07 (1.03–1.11) 2×10⁻⁴ 1.12 (1.07–1.17) 4×10⁻⁷ 0.94 (0.87–1.02) 0.16 1.00 (0.89–1.12) 0.98 1.08 (0.97–1.21) 0.17 0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4q26</td>
<td>rs17329882</td>
<td>1.09 (1.06–1.13) 3×10⁻⁷ 1.11 (1.07–1.16) 3×10⁻⁷ 1.09 (1.01–1.18) 0.020 1.06 (0.96–1.18) 0.26 1.11 (0.99–1.23) 0.06 0.88</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6p22.1</td>
<td>rs116133110</td>
<td>0.94 (0.91–0.97) 9×10⁻³ 0.91 (0.87–0.94) 3×10⁻⁷ 0.95 (0.89–1.02) 0.16 1.05 (0.95–1.15) 0.34 1.03 (0.94–1.14) 0.53 0.008</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9q34.2</td>
<td>rs635634</td>
<td>1.12 (1.08–1.16) 9×10⁻⁹ 1.13 (1.08–1.18) 2×10⁻⁷ 1.12 (1.03–1.21) 0.007 1.03 (0.92–1.16) 0.58 1.23 (1.10–1.38) 3×10⁻⁴ 0.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17q11.2</td>
<td>chr17:29181</td>
<td>0.90 (0.87–0.93) 1×10⁻⁹ 0.90 (0.87–0.94) 2×10⁻⁷ 0.88 (0.82–0.95) 5×10⁻⁴ 0.88 (0.80–0.98) 0.020 1.01 (0.91–1.12) 0.84 0.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p-value for the heterogeneity in associations with different tumour subtypes*
Table 3

Summary of data on SNPs, closest gene and all genes in 1 MB region for each locus

<table>
<thead>
<tr>
<th>Loci</th>
<th>Position of top SNP</th>
<th># putatively causal SNPs</th>
<th>Genes in window of putatively causal SNPs</th>
<th># SNPs aligned w/ biofeatures</th>
<th>Normal eQTL closest gene</th>
<th>Tumour DNA copy number</th>
<th>Significant expression difference in tumour vs normal</th>
<th>Known role of gene in cancer</th>
<th># Genes in 1MB region</th>
<th>Other known cancer genes in 1MB region</th>
</tr>
</thead>
<tbody>
<tr>
<td>1p36</td>
<td>promoter region of WNT4</td>
<td>39</td>
<td>WNT4, CDC42, LINCO00339</td>
<td>11</td>
<td>NS</td>
<td>loss</td>
<td>Yes</td>
<td>11 RAP1GAP, CDC42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1p34.3</td>
<td>intron 3 of RSPO1</td>
<td>15</td>
<td>RSPO1</td>
<td>0</td>
<td>NS</td>
<td>gain</td>
<td>Yes</td>
<td>22 C1orf109, FHL3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4q26</td>
<td>intron 3 of SYNPO2</td>
<td>4</td>
<td>SYNPO2</td>
<td>2</td>
<td>NS**</td>
<td>loss</td>
<td>down</td>
<td>Yes</td>
<td>12 none</td>
<td></td>
</tr>
<tr>
<td>6p22.1</td>
<td>intron 1 of GPX6</td>
<td>22</td>
<td>GPX6, GPX5</td>
<td>1</td>
<td>N/A</td>
<td>gain</td>
<td></td>
<td>23 ZKSCAN3, TRIM27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9q34.2</td>
<td>4.3kb upstream of ABO</td>
<td>18</td>
<td>ABO, SLC2A6*</td>
<td>1</td>
<td>NS</td>
<td>loss</td>
<td>down</td>
<td>Yes</td>
<td>32 TSC1, RALGDS, RPL7A, VAV2</td>
<td></td>
</tr>
<tr>
<td>17q11.2</td>
<td>intron 6 of ATAD5</td>
<td>16</td>
<td>ATAD5, TEFM, ADAP2, CRLF3, SUZ12P1</td>
<td>0</td>
<td>NS</td>
<td>loss</td>
<td>up</td>
<td>Yes</td>
<td>17 NF1</td>
<td></td>
</tr>
</tbody>
</table>

Proximal promoter regions were defined as 1kb upstream of the transcription start site
N/A indicated no expression of GPX6 in normal tissues. NS, not significant.

♭ Biofeatures defined as open chromatin, H3K4me3 or H3K27ac marks detected in normal ovarian and/or fallopian cells

* There are 16 genes in this region; ABO, SURF6, MED22, RPL7A, SNORD24, SNORD36B, SNORD36A, SNORD36C, SURF1, SURF2, SURF4, C9orf96, REXO4, ADAMTS13, CACFD1, SLC2A6, however all SNPs are within or upstream of ABO or upstream of SLC2A6.

** Trend p=0.067