HHS Public Access

Author manuscript

Nat Genet. Author manuscript; available in PMC 2011 June 29.

Published in final edited form as:

Nat Genet. 2010 October ; 42(10): 880–884. doi:10.1038/ng.666.

Common variants at 19p13 are associated with susceptibility to ovarian cancer

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

Abstract

Epithelial ovarian cancer (EOC) is the leading cause of death from gynecological malignancy in the developed world accounting for 4 percent of deaths from cancer in women1. We performed a three-phase genome-wide association study of EOC survival in 8,951 EOC cases with available survival time data, and a parallel association analysis of EOC susceptibility. Two SNPs at 19p13.11, rs8170 and rs2363956, showed evidence of association with survival (overall $P=5\times10^{-4}$ and 6×10^{-4}), but did not replicate in phase 3. However, the same two SNPs demonstrated genome-wide significance for risk of serous EOC ($P=3\times10^{-9}$ and 4×10^{-11} respectively). Expression analysis of candidate genes at this locus in ovarian tumors supported a role for the *BRCA1* interacting gene *C19orf62*, also known as *MERIT40*, which contains rs8170, in EOC development.

Factors related to tumor aggressiveness, response to therapy, and underlying patient health are major predictors of survival in EOC. Germline genetic variation could impact every step in the process from the likelihood of secondary mutational events to host tissue tolerance of a metastatic lesion and treatment response. Evidence for the role of germline genetics comes from the observations that rare EOC predisposition-alleles of *BRCA1* and *BRCA2* are associated with improved overall survival following a diagnosis of EOC2, 3. Many studies have investigated the association between common genetic variation in candidate genes and EOC survival, but no positive findings have been convincingly replicated. GWAS have successfully identified common genetic variants influencing a spectrum of phenotypes4; but, to date, there are no published reports of GWAS for cancer survival outcomes.

We conducted a three-phase GWAS to identify SNPs associated with variation in the time from invasive EOC diagnosis to death (Supplementary tables 1 and 2). Genotyping was carried out in parallel with a multi-phase GWAS of EOC susceptibility5. Phase 1 comprised 1,768 cases with invasive EOC from four UK studies. Survival time data, predominantly

Address correspondence to: Dr. Paul D.P. Pharoah, Strangeways Research Laboratory, Worts Causeway, Cambridge, UK CB18RN, paul1@srl.cam.ac.uk.

*K.L.B and J.T contributed equally to this work.

Author contributions

P.D.P.P., S.A.G., and D.F.E. designed the overall study and obtained financial support. P.D.P.P., S.A.G., S.J.R., and H.S. coordinated the studies used in phase 1 and phase 2. H.S., K.L.B. G.C.-T., and E.L.G. coordinated phase 3. J.T. and K.L.B. conducted primary phase 1 and phase 2 analysis and phase 3 SNP selection; K.L.B. conducted phase 3 and combined data statistical analyses; H.S., J.B., and J.M.C. conducted phase 3 genotyping. S.A.G., M.N., C.J. and T.S. designed and performed the functional analyses. The remaining authors coordinated contributing studies. K.L.B. and P.D.P.P. drafted the manuscript with substantial input from S.A.G, H.S., and S.J.R. All authors contributed to the final draft.

through routine notification of deaths through the Office of National Statistics, was available for 86 percent of cases. Controls were taken from two studies previously used as part of a GWAS for other phenotypes, the UK 1958 Birth Cohort and the UK Colorectal Control Cohort. Cases were genotyped using the Illumina Infinium 610K array and controls were genotyped using the similar 550k Illumina array5–7.

Association between SNP genotypes and survival were evaluated using a 1 degree of freedom trend test based on the Cox model (see methods). The 4,649 SNPs showing the strongest evidence for association with EOC survival were selected for genotyping in phase 2 together with 22,790 SNPs selected for the susceptibility study and 800 SNPs that reported on ancestry. Phase 2 comprised 4,238 cases and 4,810 controls from ten different studies across the USA, Europe and Australia; SNPs were genotyped using a custom Illumina iSelect array. The majority of cases (80 percent) had survival time data available through a variety of sources including death certificate flagging and medical records. Finally, we genotyped the three SNPs most strongly associated with survival - rs1125436, rs8170 and rs2363956 - in a phase 3 analysis that included 4,501 cases (of which 4,076 had survival time data) and 6,021 controls from twenty two additional studies that are part of the Ovarian Cancer Association Consortium (OCAC). The SNPs rs10426843 and rs8100241 that correlate perfectly with rs8170 and rs2363956, respectively, were included as proxies in the event of assay failure. We also genotyped thirty SNPs from the top nine loci from the analysis of susceptibility8. Genotyping of rs2363956 was poor for phase 3 studies genotyped by iPlex (see Methods and Supplementary note) and genotype data for the surrogate marker was used in analyses.

Characteristics of the cases by study phase are shown in Supplementary table 1. Cases from all three phases provided 21,127 person-years of follow-up; 3,358 deaths occurred within five years following diagnosis of EOC in the combined dataset. There was little evidence of any general inflation of the survival test statistics in either phase 1 or phase 2 (estimated inflation factor phase 1 λ_{1000} =0.99, phase 2 λ_{1000} =0.99) (Supplementary figure 1). In the analysis of the combined phase 1 and 2 data the SNP most strongly associated with risk of death was rs1125436 at 13q32 (HR=1.22, 95% CI 1.12–1.32, P=3×10^{-6}). There was no association of this SNP with EOC susceptibility (P=0.57). The next most strongly associated locus with survival was at 19p13, containing rs8170 (risk allele *t*) and rs2363956 (risk allele *t*) (HR = 1.18, 95% CI 1.09–1.27, P=2×10^{-5}, and HR = 1.13, 95% CI 1.06–1.21, P=2×10^{-4} respectively). Neither SNP reached the threshold of significance in phase 1 to be selected for phase 2 of the EOC susceptibility GWAS, but in the combined phase 1 and 2 data both showed some evidence for susceptibility to EOC (OR=1.15, 95% CI 1.08–1.23, P=7×10^{-6}, and OR=1.08, 95% CI 1.03–1.14, P=2×10^{-3} respectively). This association was stronger among ovarian cancers with serous histology (OR=1.22, 95% CI 1.13–1.31, P=1×10^{-7}, and OR=1.14 95% CI 1.07–1.21, P=2×10^{-5} respectively). These effects were similar in analyses unadjusted for population stratification by principal components (data not shown). Risk allele frequencies of these SNPs in cases and controls by study are shown in Supplementary table 3.

In the phase 3 data there was no evidence for the association of rs1125436, rs8170 or rs2363956 with survival time (P=0.12, 0.85 and 0.25 respectively) with the effect of

rs1125436 in the opposite direction to phases 1 and 2 (data not shown). The direction of the survival effect was the same for rs8170 and rs2363956, with the effect size being larger in phase 1 compared to phase 2 and 3 (Supplementary figure 2b). In the combined analysis of all three phases, rs8170 and rs2363956 showed similar levels of association with survival (HR 1.11, 95% CI 1.04–1.17, $P=5\times10^{-4}$ and HR 1.09, 95% CI 1.04–1.14, $P=6\times10^{-4}$; Table 1). The association with survival was not attenuated after adjusting for tumor grade, tumor stage, age at diagnosis and histology.

The phase 3 data, however, provided strong support for the association of rs8170 and rs2363956 with EOC susceptibility (Table 1). The association was considerably stronger when the analysis was restricted to serous cases and the association for both SNPs reached genome-wide significance in the combined data analysis of serous only cases ($P=3\times10^{-9}$ and 4×10^{-11} respectively). These remained highly significant ($P<10^{-9}$) after a conservative Bonferroni correction for three tests (all cases, serous cases, non-serous cases). There was little evidence of association with other histological subtypes (Table 2). No heterogeneity was seen in the OR of serous EOC risk or HR estimates for rs2363956 (Supplementary Figure 2a–b) or rs8170 (forest plots not shown) among studies for any phase. rs8170 and rs2363956 are separated by 4kB and are weakly correlated ($r^2 = 0.23$). In multivariate models, the associations with susceptibility to serous cancer and survival could not be fully explained by either SNP alone.

The SNP rs8170 localizes to *C19orf62*, also known as *MERIT40*, a gene with 5 distinct transcripts described to date. Depending on the alternative splice form, it is either synonymous (K279K) or non-synonymous (S281R). It may also act as an exonic splice enhancer (http://pupasuite.bioinfo.cipf.es/). rs2363956 is a non-synonymous SNP (W184L) in *ANKLE1*. Both amino acids are neutral and nonpolar suggesting this is a conservative change. Three recent reports have described interactions between MERIT40 and a complex including BRCA1, RAP80, BRCC45 and CCDC989–11. MERIT40 appears to regulate the retention of BRCA1 at double strand DNA breaks and maintain stability of this complex at the sites of DNA damage. Our data suggesting that common genetic variants in *MERIT40* may predispose women mainly to serous ovarian cancer are also consistent with a similar subtype specificity associated with inactivating germline *BRCA1* mutations12.

Common genetic variants can influence the expression of target genes through cis- and trans-regulation13. Because rs8170 and rs2363956 in *MERIT40* and *ANKLE1* respectively are located in the coding regions of these genes, we were able to evaluate cis-regulating expression by looking at both genotype associated expression and differential allelic expression, in 48 normal primary ovarian epithelial (POE) cell lines. We found no evidence of cis-regulated expression using either approach, although the power of these analyses was limited by the small sample size (Supplementary table 4 and Supplementary figure 3).

Array comparative genomic hybridization (aCGH) analysis was used to evaluate genomic alterations at the 19p13.11 locus in 105 high-grade serous ovarian cancers. Forty-six percent of tumors exhibit copy number gain/amplification of the p-arm of chromosome 19, with a peak of amplification in the region containing *MERIT40* and *ANKLE1* (Figure 1b and Figure 1c). This suggests that target genes in this region are functionally activated during tumor

development. We compared the expression of *MERIT40* and *ANKLE1* between 48 POE cell lines and 23 ovarian cancer (OC) cell lines. Consistent with aCGH data, *MERIT40* was significantly over expressed in OC cell lines compared to POE cell lines ($P=5\times10^{-9}$, Figure 1d), but there were no differences in expression of *ANKLE1* ($p = 0.54$) (Figure 1e). The data from The Cancer Genome Atlas (TCGA) Pilot Project analysis of 216 serous ovarian tumors also suggests that the expression of *MERIT40* (but not *ANKLE1*) is elevated in the majority of EOCs compared to normal tissues (Figure 1f).

The data suggesting a role for *MERIT40* in EOC development need to be treated with caution. The risk associated SNPs within *MERIT40* and *ANKLE1* may represent markers in linkage disequilibrium with the true functional variant(s) and target genes at this locus. Based on resequencing data from the 1000 genomes project (http://www.1000genomes.org/page.php) there are fifteen SNPs perfectly correlated with rs8170 and nine SNPs correlated with rs2363956. Thus, genotyping of additional SNPs will be required to fine map this region in order to nominate optimal variants to investigate function. The peak of DNA copy number gain identified by aCGH analysis in primary EOCs spans approximately 3.5Mb (nucl. 16390797–19830868; build v37) and contains 119 genes. Within this, a 330kb region defined by the haplotype block harboring rs8170 and rs2363956 contains 14 known genes (Supplementary table 5). Gene expression data from TCGA suggests other candidate genes that could be the targets of amplification at this locus, some of which some are plausible cancer associated genes. These include *NR2F6* (or *EAR*-2)14 which may be involved in regulation of disease progression in breast cancer, and *TMEM16H*, one of a family of trans-membrane proteins that may be over-expressed in several cancers15.

We can only speculate on the possible functional role of *MERIT40* in the initiation and development of serous subtype EOCs, if it is the target susceptibility gene at the 19p13 locus. Any hypotheses would need to consider the apparent paradox suggested by our data that *MERIT40* is over-expressed in EOCs, while BRCA1 is expected to show loss of function in its role in the double strand break (DSB) repair pathway. MERIT40 appears to act downstream of poly-ubiquitination of DNA (which occurs at all DSBs), and upstream of BRCA110. MERIT40 is necessary for BRCA1 assembly at γH2AX foci although BRCA1 is not usually a stable member of this complex9–11. Over-expression of *MERIT40* may ectopically stabilize mutant BRCA1 protein into the assembled complex. Since *MERIT40* knockdown makes cells more sensitive to ionizing radiation10, 11, *MERIT40* over-expression could have the opposite effect, protecting cells with dysfunctional BRCA1 and DSB repair activity and enabling them to tolerate more DNA damage.

The association with survival was only apparent in phases 1 and 2, and did not reach genome-wide significance overall. The clear evidence of association with serous EOC risk suggests that the survival association could still be of interest, but further study will be required to clarify the magnitude of the association. We would not have detected the association at 19p13 with risk of EOC if SNPs had not been selected for phase 2 as a result of its association with survival time. The failure to detect an association with susceptibility may simply be the play of chance – the power in phase 1 to detect an odds ratio of 1.12 (combined data estimate) at the P-value threshold required for a SNP to be taken into phase 2 was 50 percent. It may also have been the result of other factors such as disease

heterogeneity - the association was stronger for serous EOC and our initial analysis of phase 1 data (for selection of SNPs for Phase 2) was based on cases of all histological types. Furthermore, the majority of the phase 1 cases were prevalent and, if the association of this locus with survival time is real (but small), this would bias the susceptibility association towards the null.

These data add to a growing list of genetic loci with common susceptibility alleles for EOC. Our data suggesting that the *BRCA1* interacting gene *MERIT40* may be the gene underlying the genetic associations add weight to the significance of the 19p13 locus for susceptibility in EOC. This is further emphasized by the finding of Antoniou et al. in the accompanying article16 that genetic variants in this region appear to modify the risks of breast cancer in individuals carrying germline *BRCA1* mutations.

Methods

Study design

The ovarian cancer case-control studies that participated in phases 1, 2 and 3 are summarized in Supplementary table 2. Phase 1 comprised invasive epithelial ovarian cancer cases from UK and genotype data of UK controls from GWAS of other phenotypes. Phase 2 comprised ten case-control studies from the Ovarian Cancer Association Consortium. Phase 3 comprised 16 case-control studies from the OCAC and five case-only studies. All studies provided data on age at diagnosis and date of blood draw, self-reported ethnic group and histological subtype. Tumor histology was collected for all cases based on pathology reports or central pathological review and was categorized according to the World Health Organization classification system for ovarian cancer17.

Genotyping

Genotyping for phase 1 cases was conducted using the Illumina Infinium 610K array at Illumina Corporation. Existing data from two sets of controls, genotyped on the Infinium 550k array, were used in phase 1 analyses: the Welcome Trust Case-Control Consortium 1958 birth cohort and a national colorectal control study. All cases were from the UK and confirmed as invasive epithelial ovarian cancer. Genotyping the phase 2 studies was conducted using a custom Illumina iSelect array at Illumina Corporation.

For four phase 3 studies (TOR, NCO, MAY, MOF) genotype data were available from an independent, ongoing GWAS study that also used the Illumina Infinium 610K platform. Genotyping and QC were performed at the Mayo Clinic genotyping shared resource. deCODE ovarian cancer cases were assayed by single SNP genotyping on the Centaurus (Nanogen) platform and controls were from a GWAS using the Human Hap300 and HumanCNV370-duo Bead Arrays. The SNP rs2363956 was genotyped using ABI Taqman for five of the phase 3 case-only studies (LAX, PVD, SCO, YAL and additional cases from HOP). The remaining phase 3 studies were genotyped using Sequenom iPlex. Quality control procedures for all study phases are described in the supplementary materials.

Population stratification

We used the program LAMP18 to assign intercontinental ancestry to phase 1 samples based on the HapMap genotype frequency data for European, African and Asian populations (release no.22). LAMP was also used to assign ancestry to the Phase 2 samples using the HapMap data on European (CEU), African (ASW), East Asian (JPT-CHB-CHD), Mexican (MEX) and Indian (GIH). Subjects with less than 90 percent European ancestry were excluded. For both the phase 1 and 2 samples, we used AIMs to calculate principal components for the subjects of European ancestry. The first principle component explained 0.42 percent of the variability and was included as a covariate in subsequent association analyses. Subsequent principal components were not included as they explained less variability and there was little difference in their eigenvalues. In the phase 3 dataset, we excluded samples if their self-reported ethnicity was other than non-Hispanic white.

Imputation

We imputed missing genotype data for all the common variants in the HapMap for phase 1 samples in order to increase genome coverage. We used an in-house method that combines the features of fastPHASE19 and IMPUTE20 to impute the ungenotyped or missing SNPs, using the phase 2 HapMap data (CEU) which contains phased haplotypes for 60 individuals on 2.5 million SNPs. For each imputed genotype the expected number of minor alleles carried was estimated (weights). Genotyped SNPs were assigned weights of 0, 1 or 2 (actual number of minor alleles carried). We estimated the accuracy of imputation by calculating the estimated r2 between the imputed and actual SNP. SNPs with $r2 < 0.64$ were excluded ($n = 152,401$) leaving a total of 2,563,972 SNPs for phase 1 analysis.

Tests of association

In the analysis of the phase 1 and phase 2 data the effect of each SNP on time to all-cause mortality after EOC diagnosis was assessed using Cox regression stratified by study and modeling the per-allele effect as log-additive. The Cox proportional hazards assumption was evaluated by inspection of standard log-log plots. Individual level data for the deCODE study were not available and so for the analysis of the phase 3 data and for the combined analyses, each study was analyzed separately and the results pooled by estimating an average of the study specific $\log_e$ hazard ratios with each weighted by the inverse of its variance. Because the EOC cases showed a variable time from diagnosis to study entry, we allowed for left truncation with time at risk starting on date of diagnosis and time under observation beginning at the time of study entry. This generates an unbiased estimate of the hazard ratio provided the Cox proportional hazards assumption is correct21. The analysis of phase 1 data was right censored at 10 years after EOC diagnosis. In subsequent analyses, we right censored at 5 years after diagnosis in order to reduce the number of non-EOC related deaths. We used logistic regression to test for association between genotype and case-control status. For phase 1 and 2 data we adjusted for study phase and study by including phase and study specific indicators in the model. For phase 3 data we analyzed each study separately and then pooled the results using an inverse-variance weighted average of the study specific $\log_e$ odds ratios.

Array Comparative Genomic Hybridisation (aCGH) Analysis

aCGH analysis was performed using a whole genome tiling path microarray (http://www.instituteforwomenshealth.ucl.ac.uk/academic_research/gynaecologicalcancer/trl/arrayfacility) consisting of 32,450 BAC clones22. Regions containing >80 percent neoplastic cells were micro-dissected from formalin fixed paraffin embedded tumor tissue sections, and DNA extracted by proteinase K digestion. Tumor DNA and matching peripheral blood DNA were amplified using the GenomePlex whole genome amplification kit (Sigma) and fluorescently labelled using the BioPrime Total Kit (Invitrogen). Microarrays were co-hybridised with the labelled DNA as described previously23, scanned using a Scanarray Express laser scanner (Perking Elmer), and spot signal intensities extracted using BlueFuse (BlueGnome). Raw data were analysed using R and the Bioconductor packages MANOR, LIMMA, DNAcopy and CGHcall as described elsewhere. BAC clone locations were derived from NCBI Human Genome build 36 (HG18).

Gene expression analysis in POE and OC cell lines

Normal, primary ovarian epithelial (POE) cell lines were established from brushings of normal ovaries of patients undergoing total hysterectomies at University College London Hospital (UCLH), UK. All ovaries were histologically confirmed as free of disease. UCLH ethical committee approval was given for the collection and analysis of all patient samples. Short-term cultures of POE cells were established as previously described24. The non-neoplastic status and epithelial (non fibroblastic) nature of cells was confirmed by staining for the markers CA125, CK18, FVIII and FSP. RNA was extracted from POE and OC cell lines (Supplementary table 4) using RNAeasy Mini Kits (QIAgen). Reverse transcribed (RT) RNA was analyzed for candidate gene expression by semi-quantitative real-time PCR using the Applied Biosystems 7900HT genetic analyzer. Gene expression was normalized against 2 endogenous controls Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and β-actin. Real time expression data were analyzed using the comparative Delta-Delta Ct method. The expression values for genes in all cell lines that are given are relative to either the lowest or highest expression of a POE cell line, normalized against GAPDH and β-actin. Differences in the relative expression of each candidate gene between EOC and POE cell lines were assessed using the nonparametric two-sided Wilcoxon Rank sum test using R. For allele specific expression analysis, gene expression was calculated relative to the average expression of the common homozygotes for each candidate SNP normalized against the expression of the endogenous control genes. Wilcoxon Rank sum tests were used to assess the difference in expression between common homozygotes, heterozygotes and rare homozygotes.

Differential allelic expression analysis in POE cell lines

For each SNP, 8ng of cDNA from the heterozygous POE cell lines (10 for rs8170 and 15 for rs2363956) were analyzed by real time RT_PCR using Taqman custom genotyping assays (Applied Biosystems). Genomic DNA extracted from lymphocytes from two heterozygous individuals was used for a standard curve to adjust for dye bias as there would be equal copies of each allele. All samples were analyzed in triplicate. Differential allelic expression

was determined from the log2 ratio of the VIC allele / FAM allele with a cut-off of $\log_2(1.20)=0.263$ as described previously13.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Authors

Kelly L. Bolton[1,2,*], Jonathan Tyrer[1,*], Honglin Song[1], Susan J. Ramus[3], Maria Notaridou[3], Chris Jones[3], Tanya Sher[3], Aleksandra Gentry-Maharaj[3], Eva Wozniak[3], Ya-Yu Tsai[4], Joanne Weidhaas[5], Daniel Paik[6], David J. Van Den Berg[7], Daniel O. Stram[7], Celeste Leigh Pearce[7], Anna H. Wu[7], Wendy Brewster[8], Hoda Anton-Culver[8], Argyrios Ziogas[8], Steven A. Narod[9], Douglas A. Levine[10], Stanley B. Kaye[11], Robert Brown[12], Jim Paul[13], James Flanagan[12], Weiva Sieh[14], Valerie McGuire[14], Alice S. Whittemore[14], Ian Campbell[15], Martin E. Gore[16], Jolanta Lissowska[17], Hannah Yang[2], Krzysztof Medrek[18], Jacek Gronwald[18], Jan Lubinski[18], Anna Jakubowska[18], Nhu D. Le[19], Linda S. Cook[20,21], Linda E. Kelemen[21], Angela Brooks-Wilson[22,23], Leon F.A.G. Massuger[24], Lambertus A. Kiemeney[24], Katja K.H. Aben[25], Anne M. van Altena[24], Richard Houlston[26], Ian Tomlinson[27], Rachel T. Palmieri[28], Patricia G. Moorman[28], Joellen Schildkraut[28], Edwin S. Iversen[29], Catherine Phelan[4], Robert A. Vierkant[30], Julie M. Cunningham[31], Ellen L. Goode[30], Brooke L. Fridley[30], Susan Kruger-Kjaer[32], Jan Blaeker[33], Estrid Hogdall[32], Claus Hogdall[34], Jenny Gross[35], Beth Y. Karlan[35], Roberta B. Ness[36], Robert P. Edwards[37], Kunle Odunsi[38], Kirsten B. Moyisch[39], Julie A. Baker[40], Francesmary Modugno[41], Tuomas Heikkinenen[42], Ralf Butzow[42], Heli Nevanlinna[42], Arto Leminen[42], Natalia Bogdanova[43], Natalia Antonenkova[43], Thilo Doerk[44], Peter Hillemanns[44], Matthias Dürst[45], Ingo Runnebaum[45], Pamela J. Thompson[46], Michael E. Carney[46], Marc T. Goodman[46], Galina Lurie[46], Shan Wang-Gohrke[47], Rebecca Hein[48], Jenny Chang-Claude[48], Mary Anne Rossing[49], Kara L. Cushing-Haugen[49], Jennifer Doherty[49], Chu Chen[49], Thorunn Rafnar[50], Soren Besenbacher[50], Patrick Sulem[50], Kari Stefansson[50], Michael J. Birrer[51], Kathryn L. Terry[52], Dena Hernandez[53], Daniel W. Cramer[52], Ignace Vergote[54], Frederic Amant[54], Diether Lambrechts[55], Evelyn Despierre[54], Peter A. Fasching[56], Matthias W. Beckmann[57], Falk C. Thiel[58], Arif B. Ekici[59], Xiaoqing Chen[60], the Australian Ovarian Cancer Study Group[15], the Australian Cancer Study (Ovarian Cancer)[60], On behalf of the ovarian cancer association consortium, Sharon E. Johnatty[60], Penelope M. Webb[60], Jonathan Beesley[60], Stephen Chanock[2], Montserrat Garcia-Closas[2], Tom Sellers[4], Douglas F. Easton[1], Andrew Berchuck[28], Georgia Chenevix-Trench[60], Paul D.P. Pharoah[1], and Simon A. Gayther[3]

Affiliations

[1]Department of Oncology, University of Cambridge, Cambridge, UK [2]Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, Maryland, USA [3]Department of Gynaecological Oncology, University College London, EGA Institute for Women's Health, London, UK [4]H. Lee

Moffitt Cancer Center and Research Institute, Tampa, Florida, USA [5]Department of Therapeutic Radiology, Yale University, New Haven, Connecticut, USA [6]Department of Obstetrics and Gynecology, Yale University, New Haven, Connecticut, USA [7]Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California, USA [8]Department of Epidemiology, School of Medicine, University of California, Irvine, California, USA [9]Center for Research in Women's Health, Toronto, Canada [10]Gynecology Service, Department of Surgery, Memorial Sloan-Kettering Cancer Center, New York, New York, USA [11]Section of Medicine, Institute of Cancer Research, Sutton, UK [12]Department of Surgery and Cancer, Imperial College London, London, UK [13]Cancer Research UK Clinical Trials Unit, Glasgow University, Glasgow, UK [14]Department of Health Research and Policy, Stanford University School of Medicine, Standford, California, USA [15]Peter MacCallum Cancer Institute, Melbourne, Australia [16]The Royal Marsden Hospital, Gynecological Oncology Unit, London, UK [17]Department of Cancer Epidemiology and Prevention, M Sklodowska-Curie Cancer Center and Institute of Oncology, Warsaw, Poland [18]International Hereditary Cancer Center, Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland [19]Cancer Control Research, BC Cancer Agency, Vancouver, BC, Canada [20]Division of Epidemiology and Biostatistics, University of New Mexico, Albuquerque, New Mexico, USA [21]Alberta Health Services-Cancer Care, Calgary, AB, Canada [22]Genome Sciences Centre, BC Cancer Agency, Vancouver, BC, Canada [23]Department of Biomedical Physiology and Kinesiology, Simon Fraser University, Burnaby, BC, Canada [24]Department of Gynaecology, Radboud University, Nijmegen Medical Centre, Nijmegen, The Netherlands [25]Comprehensive Cancer Center East, Nijmegen, The Netherlands [26]Section of Cancer Genetics, Institute of Cancer Research, Sutton, UK [27]Population and Functional Genetics Lab, Wellcome Trust Centre for Human Genetics, Oxford, UK [28]Department of Community and Family Medicine, Duke University Medical Center, Durham, North Carolina, USA [29]Department of Statistics, Duke University, Durham, North Carolina, USA [30]Department of Health Sciences Research, Mayo Clinic College of Medicine, Rochester, Minnesota, USA [31]Department of Laboratory Medicine and Pathology, Mayo Clinic College of Medicine, Rochester, Minnesota, USA [32]Department of Virus, Hormones and Cancer Institute of Cancer Epidemiology, Danish Cancer Society, Copenhagen, Denmark [33]Aarhus University Hospital, Skejby, Aarhus, Denmark [34]The Gynaecologic Clinic, The Juliane Marie Centre, Rigshospitalet, Copenhagen, Denmark [35]Women's Cancer Research Institute at the Samuel Oschin Comprehensive Cancer Center, Cedars-Sinai Medical Center, Los Angeles, California, USA [36]University of Texas School of Public Health, Houston, Texas, USA [37]Magee-Womens Hospital, Pittsburgh, Pennsylvania, USA [38]Department of Gynecological Oncology, Roswell Park Cancer Institute, Buffalo, New York, USA [39]Department of Cancer Genetics, Roswell Park Cancer Institute, Buffalo, New York, USA [40]Department of Obstetrics and Gynecology, Brown University, Providence, Rhode Island, USA [41]Department of Epidemiology, University of

Pittsburgh, Pittsburgh, Pennsylvania, USA [42]Department of Obstetrics and Gynaecology, Helsinki University Central Hospital, Helsinki, Finland [43]Byelorussian Institute for Oncology and Medical Radiology Aleksandrov N.N., Minsk, Belarus [44]Clinics of Obstetrics and Gynaecology, Hannover Medical School, Hannover, Germany [45]Clinics of Obstetrics and Gynaecology, Friedrich Schiller University, Jena, Germany [46]Cancer Research Center of Hawaii, University of Hawaii, Honolulu, Hawaii, USA [47]Department of Obstetrics and Gynecology, University of Ulm, Ulm, Germany [48]Unit of Genetic Epidemiology, Division of Cancer Epidemiology, Deutsches Krebsforschungszentrum, Heidelberg, Germany [49]Program in Epidemiology, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA [50]deCODE Genetics, Reykjavik, Iceland [51]Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA [52]Obstetrics and Gynecology Epidemology Center, Brigham and Women's Hospital, Boston, Massachusetts, USA [53]National Institute of Aging, National Institutues of Health, Bethesda, Maryland, USA [54]Department of Gynaecologic Oncology, University Hospitals Leuven, Belgium [55]Vesalius Research Center, VIB and K.U.Leuven, Belgium [56]Division of Hematology and Oncology, Department of Medicine, David Geffen School of Medicine, Los Angeles, California, USA [57]Department of Cancer Epidemiology and Prevention, The M.Sklodowska-Curie Cancer Center and Institute of Oncology, Warsaw, Poland [58]Department of Gynecology and Obstetrics, University Hospital Erlangen, Erlangen, Germany [59]Institute of Human Genetics, Friedrich Alexander University Erlangen-Nuremberg, Erlangen, Germany [60]The Queensland Institute of Medical Research, Post Office Royal Brisbane Hospital, Australia

Acknowledgements

We thank all the individuals who took part in this study and all the researchers, clinicians and administrative staff who have enabled the many studies contributing to this work. In particular we thank A. Ryan and J. Ford (UKOPS); J. Morrison, P. Harrington and the SEARCH team (SEA), U. Eilber and T. Koehler (GER); D. Bowtell, A. deFazio, D. Gertig, A. Green (AOCS); A. Green, P. Parsons, N. Hayward, D. Whiteman (ACS); L. Gacucova (HMOCS); S. Haubold, P. Schürmann, F. Kramer, W. Zheng, T.-W. Park-Simon, K. Beer-Grondke and D. Schmidt (HJOCS); L. Brinton, M. Sherman, A. Hutchinson, N. Szeszenia- Dabrowska, B. Peplonska, W. Zatonski, A. Soni, P. Chao, M. Stagner (POL2).

The genotyping and data analysis for this study was supported by a project grant from Cancer Research UK. We acknowledge the computational resources provided by the University of Cambridge (CamGrid). This study makes use of data generated by the Wellcome Trust Case-Control consortium. A full list of the investigators who contributed to the generation of the data is available from www.wtccc.org.uk. Funding for the project was provided by the Wellcome Trust under award 076113. The Ovarian Cancer Association Consortium is supported by a grant from the Ovarian Cancer Research Fund thanks to donations by the family and friends of Kathryn Sladek Smith. The results published here are in part based upon data generated by The Cancer Genome Atlas Pilot Project established by the NCI and NHGRI. Information about TCGA and the investigators and institutions who constitute the TCGA research network can be found at http://cancergenome.nih.gov. S.J.R. is supported by the Mermaid/Eve Appeal, G.C.-T. and P.M.W. by the NHMRC, P.A.F. by the Deutsche Krebshilfe, MG acknowledges NHS funding to the NIHR Biomedical Research Centre at the Royal Marsden Hospital and D.F.E. is a Principal Research Fellow of Cancer Research UK. Funding of the constituent studies was provided by the Danish Cancer Society, the Ovarian Cancer Research Fund (PPD/RPCI.07), the Roswell Park Cancer Institute Alliance Foundation, the US National Cancer Institute (CA58860, CA92044, P50CA105009, R01CA122443, R01CA126841-01, R01CA16056, R01CA61107, R01CA71766, R01CA054419, R01CA114343, R01CA87538, R01CA112523, R01CA58598, N01CN55424, N01PC35137, and Intramural research funds), the US Army Medical Research and Material Command (DAMD17-01-1-0729), Cancer Council Victoria, Cancer Council Queensland, Cancer Council New

South Wales, Cancer Council South Australia, Cancer Council Tasmania and Cancer Foundation of Western Australia, the National Health and Medical Research Council of Australia (199600 and 400281), the German Federal Ministry of Education and Research of Germany Programme of Clinical Biomedical Research (01 GB 9401), the state of Baden-Wurttemberg through Medical Faculty of the University of Ulm (P.685), the Mayo Foundation, the Lon V. Smith Foundation (LVS-39420), the Oak Foundation, the University College Hospital National Institute for Health Research Biomedical Research Centre and the Royal Marsden Hospital Biomedical Research Centre.

Reference List

1. Ferlay, J.; Bray, F.; Pisani, P.; Parkin, DM. GLOBOCAN 2002: Cancer Incidence, Mortality and Prevalence Worldwide. Lyon: IARC Press; 2004.
2. Boyd J. Clinicopathologic features of BRCA-linked and sporadic ovarian cancer. J. Am. Med. Assoc. 2000; 283:2260–2265.
3. Majdak EJ, et al. Prognostic impact of BRCA1 pathogenic and BRCA1/BRCA2 unclassified variant mutations in patients with ovarian carcinoma. Cancer. 2005; 104:1004–1012. [PubMed: 16047333]
4. Hindorff LA, Junkins HA, Mehta JP, Manolio TA. A Catalog of Published Genome-Wide Association Studies. Internet. 2010 2-13-2010.
5. Song H, et al. A genome-wide association study identifies a new ovarian cancer susceptibility locus on 9p22.2. Nat. Genet. 2009; 41:996–1000. [PubMed: 19648919]
6. Tomlinson I, et al. A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. Nat. Genet. 2007; 39:984–988. [PubMed: 17618284]
7. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007; 447:661–678. [PubMed: 17554300]
8. Goode E. Identification of Four Novel Ovarian Cancer Susceptibility Loci Identified in a Genome-Wide Association Study. Nature Genetics. 2010
9. Feng L, Huang J, Chen J. MERIT40 facilitates BRCA1 localization and DNA damage repair. Genes Dev. 2009; 23:719–728. [PubMed: 19261748]
10. Shao G, et al. MERIT40 controls BRCA1-Rap80 complex integrity and recruitment to DNA double-strand breaks. Genes Dev. 2009; 23:740–754. [PubMed: 19261746]
11. Wang B, Hurov K, Hofmann K, Elledge SJ. NBA1, a new player in the Brca1 A complex, is required for DNA damage resistance and checkpoint control. Genes Dev. 2009; 23:729–739. [PubMed: 19261749]
12. Lakhani SR, et al. Pathology of ovarian cancers in BRCA1 and BRCA2 carriers. Clin Cancer Res. 2004; 10:2473–2481. [PubMed: 15073127]
13. Maia AT, et al. Extent of differential allelic expression of candidate breast cancer genes is similar in blood and breast. Breast Cancer Res. 2009; 11:R88. [PubMed: 20003265]
14. Yang C, Yu B, Zhou D, Chen S. Regulation of aromatase promoter activity in human breast tissue by nuclear receptors. Oncogene. 2002; 21:2854–2863. [PubMed: 11973645]
15. Galindo BE, Vacquier VD. Phylogeny of the TMEM16 protein family: some members are overexpressed in cancer. Int. J. Mol. Med. 2005; 16:919–924. [PubMed: 16211264]
16. Antoniou A. A locus on 19p13 locus modifies risk of breast cancer for BRCA1 mutation carriers and is associated with hormone receptor-negative breast cancer in the general population. Nature Genetics. 2010
17. Tavassoli, FA.; Devilee, P. World Health Organization Classification of Tumours. Pathology and Genetics of Tumours of the Breast and Female Genital Organs. Lyon: IARC Press; 2003.
18. Sankararaman S, Sridhar S, Kimmel G, Halperin E. Estimating local ancestry in admixed populations. Am. J. Hum. Genet. 2008; 82:290–303. [PubMed: 18252211]
19. Scheet P, Stephens M. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. Am. J. Hum. Genet. 2006; 78:629–644. [PubMed: 16532393]
20. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat. Genet. 2007; 39:906–913. [PubMed: 17572673]

21. Azzato EM, et al. Prevalent cases in observational studies of cancer survival: do they bias hazard ratio estimates? Br. J. Cancer. 2009; 100:1806–1811. [PubMed: 19401693]
22. Krzywinski M, et al. A set of BAC clones spanning the human genome. Nucleic Acids Res. 2004; 32:3651–3660. [PubMed: 15247347]
23. Dafou D, et al. Chromosomes 6 and 18 induce neoplastic suppression in epithelial ovarian cancer cells. Int. J. Cancer. 2009; 124:1037–1044. [PubMed: 19058220]
24. Li NF, et al. A modified medium that significantly improves the growth of human normal ovarian surface epithelial (OSE) cells in vitro. Lab Invest. 2004; 84:923–931. [PubMed: 15077121]

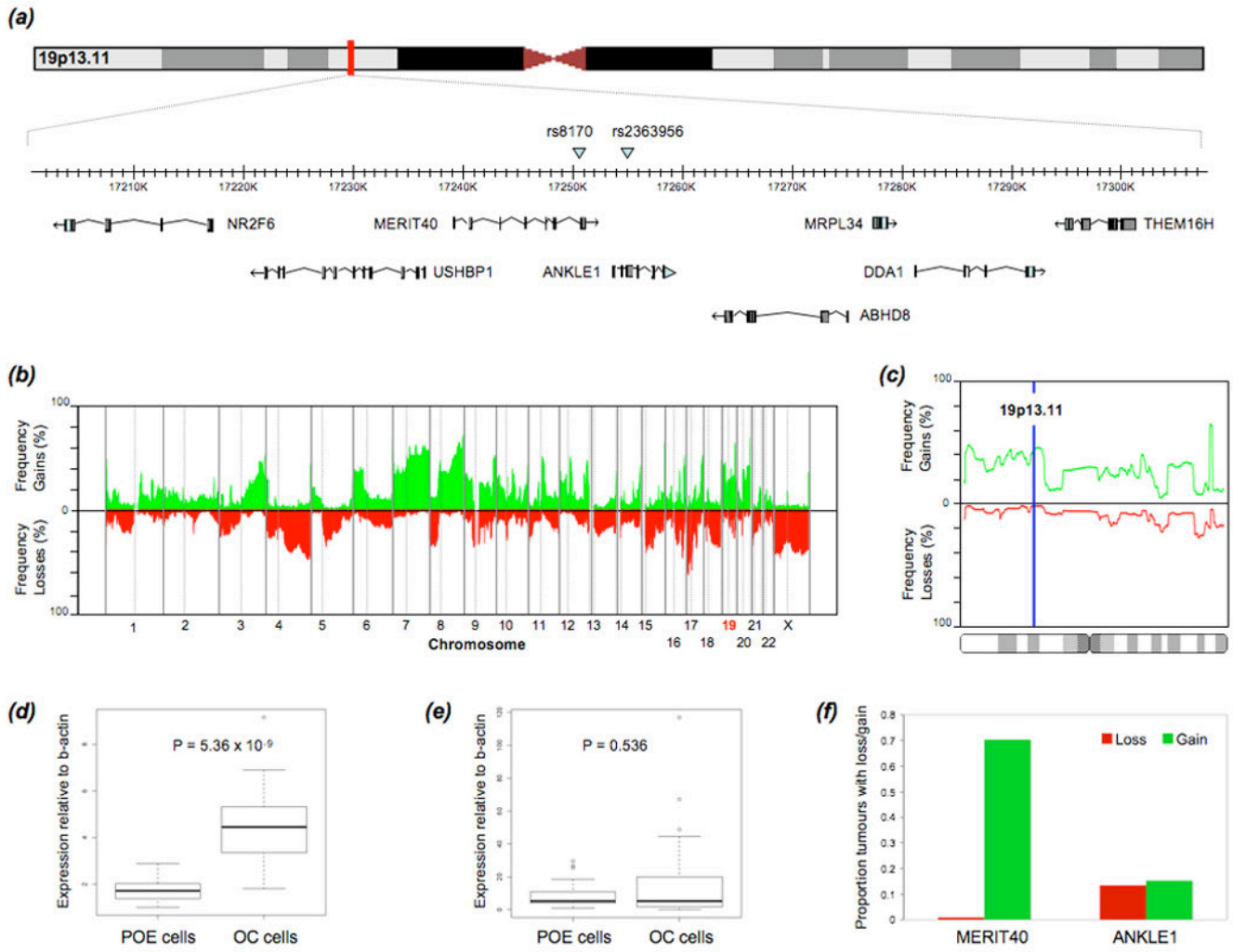

Figure 1. Genomic and transcript analysis of the *MERIT40* and *ANKLE1* genes in the 19p13 ovarian cancer susceptibility region

(a) Genomic architecture of the 19p13.11 region containing the two SNPs most significantly associated with EOC risk (rs8170 and rs2363956). SNPs are located with respect to genes within this region. rs8170 is located in *MERIT40* and rs2363956 is located in *ANKLE1*. **(b)** Whole genome array comparative genomic hybridization (aCGH) analysis of 105 serous, invasive ovarian cancers displays the range of copy number changes throughout the genome, along the length of each chromosome. Green = frequency of copy number gain; red = copy number loss. **(c)** Higher resolution aCGH map of chromosome 19 indicates that this chromosome is frequently amplified in EOCs with an amplification peak at the 19p13.11 susceptibility locus (blue line); 48/105 tumors (46%) showed copy number gain at 19p13.11 compared to 2/105 tumors (2%) that showed copy number loss. **(d & e)** Transcript expression of *MERIT40* and *ANKLE1* in 48 normal primary ovarian epithelial (POE) cell lines compared and 23 OC cell lines detected using real time RT-PCR. For each gene, transcript expression is normalized against β-actin; genes expression normalized against a second endogenous control, GAPDH, showed similar trends (Supplementary figure 4). *MERIT40* expression is significantly higher in OC cell lines compared to POE cells **(d)**, but

there was no difference in *ANKLE1* expression between OC and POE cells **(e)**. **(f)** Expression data from the Cancer Genome Atlas Project (http://cancergenome.nih.gov) for *MERIT40* and *ANKLE1* genes analyzed in 216 serous EOCs. The graph shows proportion of tumors that show loss or gain of expression with >0.5 fold change relative to pooled 'normal' samples.

Table 1

Association of rs8170 and rs2363956 with susceptibility and survival based on combined data for subjects of European ancestry.

		Susceptibility			Survival		
Tumor subtype	**Phase**	**No. of cases/controls**	**Per-allele OR (95% CI)**	P_{trend}	**No. of cases/deaths**	**Per-allele HR (95% CI)**	P_{trend}
rs8170‡							
All Cases	Phase 1	1768/2353	1.11(0.99–1.24)	0.08	1512/397	1.35(1.15–1.59)	2.4×10^{-4}
	Phase 2	4231/4806	1.17(1.09–1.26)	2.6×10^{-5}	3361/1470	1.13(1.04–1.24)	3.7×10^{-3}
	Phase 3	4497/6012	1.07(1.00–1.15)	0.05	4072/1487	1.01(0.92–1.10)	0.85
	Combined	10496/13172	1.12(1.07–1.17)	3.6×10^{-6}	8945/3354	1.11(1.04–1.17)	5.2×10^{-4}
Serous Cases	Phase 1	844/2354	1.22(1.07–1.41)	4.4×10^{-3}	767/266	1.32(1.09–1.61)	4×10^{-3}
	Phase 2	2509/4806	1.22(1.12–1.33)	7.0×10^{-6}	2034/1039	1.11(1.00–1.23)	0.04
	Phase 3	2550/6012	1.13(1.04–1.23)	2.7×10^{-3}	2383/959	0.97(0.87–1.08)	0.57
	Combined	5903/13172	1.18(1.12–1.25)	2.7×10^{-9}	5184/2264	1.07(1.00–1.15)	0.05
rs2363956†							
All Cases	Phase 1	1768/2354	1.06(0.97–1.16)	0.20	1512/397	1.22(1.05–1.40)	7.2×10^{-3}
	Phase 2	4236/4809	1.09(1.03–1.16)	3.1×10^{-3}	3363/1472	1.10(1.03–1.19)	6.4×10^{-3}
	Phase 3	4476/6013	1.13(1.06–1.20)	9.4×10^{-6}	4025/1473	1.04(0.97–1.12)	0.25
	Combined	10480/13176	1.10(1.06–1.15)	1.2×10^{-7}	8900/3342	1.09(1.04–1.14)	5.6×10^{-4}
Serous Cases	Phase 1	844/2354	1.15(1.03–1.29)	0.01	767/266	1.35(1.14–1.62)	7×10^{-4}
	Phase 2	2513/4809	1.13(1.06–1.21)	4.0×10^{-4}	2036/1041	1.08(0.99–1.18)	0.09
	Phase 3	2538/6013	1.19(1.11–1.28)	3.1×10^{-7}	2357/951	1.03(0.94–1.13)	0.57
	Combined	5895/13176	1.16(1.11–1.21)	3.8×10^{-11}	5160/2258	1.09(1.03–1.16)	5.2×10^{-3}

‡ *c* is the reference allele and *t* is the risk allele

† *g* is the reference allele and *t* is the risk allele

Table 2

Subtype specific odds ratios and hazard ratios based on combined data for subjects of European ancestry.

		Susceptibility		Survival		
Tumor subtype	**No. of cases/controls**	**Per-allele OR (95% CI)**	**P_{trend}**	**No. of cases/deaths**	**Per-allele HR (95% CI)**	**P_{trend}**
rs8170						
Mucinous	768/13172	1.02(0.90–1.17)	0.72	547/131	1.09(0.80–1.49)	0.58
Endometrioid	1584/13172	0.98(0.89–1.08)	0.74	1153/237	1.28(1.02–1.60)	0.03
Clear Cell	717/13172	0.98(0.86–1.13)	0.80	618/159	1.22(0.93–1.61)	0.16
Other Specified	590/13172	1.14(0.98–1.32)	0.09	663/275	1.19(0.98–1.46)	0.08
Epithelial, NOS	825/13172	1.12(0.99–1.27)	0.07	586/238	1.05(0.85–1.31)	0.64
rs2363956						
Mucinous	768/13176	1.00(0.90–1.11)	0.95	542/129	1.02(0.80–1.30)	0.85
Endometrioid	1582/13176	0.98(0.91–1.05)	0.57	1147/236	1.07(0.89–1.28)	0.48
Clear Cell	717/13176	1.10(0.99–1.22)	0.09	617/159	1.01(0.80–1.27)	0.93
Other Specified	588/13176	1.02(0.90–1.15)	0.77	656/272	1.16(0.98–1.39)	0.09
Epithelial, NOS	823/13176	1.10(1.00–1.22)	0.06	586/238	1.19(0.98–1.43)	0.08