nature genetics

Identification of 12 new susceptibility loci for different histotypes of epithelial ovarian cancer

To identify common alleles associated with different histotypes of epithelial ovarian cancer (EOC), we pooled data from multiple genome-wide genotyping projects totaling 25,509 EOC cases and 40,941 controls. We identified nine new susceptibility loci for different EOC histotypes: six for serous EOC histotypes (3q28, 4q32.3, 8q21.11, 10q24.33, 18q11.2 and 22q12.1), two for mucinous EOC (3q22.3 and 9q31.1) and one for endometrioid EOC (5q12.3). We then performed meta-analysis on the results for high-grade serous ovarian cancer with the results from analysis of 31,448 *BRCA1* and *BRCA2* mutation carriers, including 3,887 mutation carriers with EOC. This identified three additional susceptibility loci at 2q13, 8q24.1 and 12q24.31. Integrated analyses of genes and regulatory biofeatures at each locus predicted candidate susceptibility genes, including *OBFC1*, a new candidate susceptibility gene for low-grade and borderline serous EOC.

EOC is a heterogeneous disease commonly classified into five major histotypes of invasive disease[1] (high-grade serous (HGSOC), low-grade serous (LGSOC), mucinous (MOC), endometrioid (ENOC) and clear cell (CCOC) carcinoma) and two histotypes of borderline disease (serous and mucinous). The histotypes have differences in lifestyle and genetic risk factors, precursor lesions, patterns of spread, molecular events during oncogenesis, response to chemotherapy and prognosis. HGSOC is thought to be derived from fallopian tube secretory epithelial cells through foci of endosalpingiosis existing as inclusion cysts lined with tubal epithelium at the ovarian and peritoneal surface[2]. In contrast, CCOC, ENOC and sero-endometrioid carcinomas appear to develop from endometriosis[3,4]. MOC resembles adenocarcinoma of the gastric pylorus, intestine or endocervix, and the majority of these tumors show gastrointestinal differentiation[5].

Approximately 20% of the familial component of EOC risk is attributable to high-to-intermediate risk genes[6]. An unknown fraction is due to more common, lower-risk genetic variation[7]. In European populations, genome-wide association studies (GWAS) have identified 23 EOC susceptibility alleles, including 18 common variants associated with all histologies and/or serous EOC[8–15], 1 associated with borderline serous tumors[13], 3 associated with MOC[16] and 1 associated with CCOC[12]. The majority of these loci also showed associations ($P < 0.05$) with EOC risk for *BRCA1* or *BRCA2* mutation carriers[15]. Five additional loci associated with EOC and breast and/or prostate cancer have been identified[17]; three of these were associated with susceptibility to EOC, breast cancer and prostate cancer, and two were associated only with breast cancer and EOC risk. However, the common genetic variants explain only 3.9% of the inherited component of EOC risk[15], and additional susceptibility loci are likely to exist, particularly for the less common non-serous histotypes.

We designed a custom Illumina array, named the OncoArray, to identify new cancer susceptibility loci[18]. The OncoArray includes ~533,000 variants (of which 260,660 formed a GWAS backbone) and has been used to genotype over 500,000 samples, including EOC case–control studies of the Ovarian Cancer Association Consortium (OCAC) and *BRCA1* and *BRCA2* mutation carriers of the Consortium of Investigators of Modifiers of *BRCA1/2* (CIMBA). These data were combined with genotype data from the Collaborative Oncological Gene–environment Study (COGS) project[14,19] and three EOC GWAS[8,9]. We present the results of these association analyses together with functional annotation of the new genome-wide significant EOC susceptibility loci.

RESULTS

Association analyses

We performed genetic association analyses using genotype data from 25,509 population-based EOC cases and 40,941 controls from OCAC and meta-analysis of these data with 19,036 *BRCA1* and 12,412 *BRCA2* mutation carriers from CIMBA, of whom 2,933 and 954, respectively, were affected with EOC. The numbers of participants by study for OCAC and CIMBA are shown in **Supplementary Tables 1** and **2**, respectively.

We used data from the 1000 Genomes Project[20] reference panel to impute genotypes for 11,403,952 common variants (minor allele frequency (MAF) >1%) and evaluated the associations of these SNPs with EOC risk. In OCAC alone, nine histotypes were investigated (all invasive, serous invasive, HGSOC, LGSOC, serous borderline, LGSOC and serous borderline combined, ENOC, CCOC and MOC). Association analyses identified six new loci associated with serous EOC histotypes at genome-wide significance ($P < 5 \times 10^{-8}$): rs9870207 at 3q28, rs13113999 at 4q32.3, rs150293538 at 8q21.11, rs7902587 at 10q24.33, rs8098244 at 18q11.2 and rs6005807 at 22q12.1. Five of these loci were associated with borderline serous EOC (3q28, 4q32.3, 8q21.11, 10q24.33 and 18q11.2), and four were associated with LGSOC tumors (3q28, 8q21.11, 10q24.33 and 18q11.2) (**Table 1**). We also identified two loci associated with MOC (rs112071820 at 3q22.3 and rs320203 at 9q31.1) and one locus associated with ENOC (rs555025179 at 5q12.3). The meta-analysis of OCAC and CIMBA

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

Received 4 June 2016; accepted 3 March 2017; published online 27 March 2017; doi:10.1038/ng.3826

Table 1 New genome-wide significant epithelial ovarian cancer susceptibility loci

SNP	Histotype	Chr.	Position (bp)	Risk allele	RAF	$r^{2\,a}$	OR	LCL	UCL	*P* value[b]	P_{het}[c]	BFRP
OCAC analyses												
rs112071820	Mucinous: invasive + borderline	3q22.3	138,849,110	G:GCCAG	0.33	0.86	1.29	1.20	1.37	1.5×10^{-13}	0.88	<1%
rs9870207	Serous: LG + borderline	3q28	190,525,516	G:A	0.73	0.97	1.19	1.12	1.27	4.5×10^{-8}	0.61	6%
rs13113999	Serous: borderline	4q32.3	167,187,046	G:T	0.56	0.86	1.23	1.14	1.32	4.7×10^{-8}	0.52	7%
rs555025179	Endometrioid	5q12.3	66,121,089	GTGACAC	0.56	0.86	1.18	1.11	1.26	4.5×10^{-8}	0.79	5%
rs150293538	Serous: LG + borderline	8q21.11	77,320,354	C:T	0.99	0.77	2.19	1.65	2.90	2.0×10^{-9}	0.38	3%
rs320203	Mucinous: invasive + borderline	9q31.1	104,943,226	C:A	0.88	0.98	1.29	1.18	1.41	1.7×10^{-8}	0.56	11%
rs7902587	Serous: LG + borderline	10q24.33	105,694,301	C:T	0.12	0.94	1.29	1.18	1.41	4.0×10^{-8}	0.99	7%
rs8098244	Serous: LG + borderline	18q11.2	21,405,553	G:A	0.31	0.98	1.19	1.12	1.27	3.9×10^{-8}	0.087	3%
rs6005807	Serous: invasive	22q12.1	28,934,313	T:C	0.91	0.99	1.17	1.11	1.23	4.5×10^{-9}	0.15	<1%
OCAC and CIMBA meta-analysis												
rs2165109	Serous HG + *BRCA1/2*	2q13	111,818,658	A:C	0.25	1.00	1.09	1.05	1.12	4.2×10^{-8}	0.66	2%
rs9886651	Serous HG + *BRCA1/2*	8q24.21	128,817,883	A:G	0.46	0.99	1.08	1.05	1.11	3.5×10^{-9}	0.26	<1%
rs7953249	Serous HG + *BRCA1/2*	12q24.31	121,403,724	A:G	0.42	1.00	1.08	1.06	1.06	1.1×10^{-9}	0.67	<1%

RAF, risk allele frequency; LCL, lower 95% confidence limit; UCL, upper 95% confidence limit; LG, low grade; HG, high grade; BFRP, Bayes false positive reporting probability assuming a prior of 1:10,000. Position is given with respect to genome build 37.
[a]Average imputation r^2 across the six data sets. [b]*P* value from analysis of imputed genotypes derived from one-step imputation (Online Methods). [c]Test for heterogeneity of effect between study strata in OCAC.

identified three additional serous EOC risk loci (rs2165109 at 2q13, rs9886651 at 8q24.21 and rs7953249 at 12q24.31). The 8q24.21 SNP rs9886651 is close to two SNPs previously associated with serous EOC (ref. 9 and A. Gjyshi, G.M.-F., J.P.T., N.T. Woods and K.L. *et al.*, unpublished data). Multivariable analysis of OCAC data showed that this is a third independent associated variant in this region (unadjusted odds ratio (OR) = 1.07, OR adjusted for rs1400482 and rs13255292 = 1.07). Variant rs6005807 at 22q12.1 was previously reported to be associated with serous EOC below genome-wide significance[21].

The association of the top SNP in each region with the nine EOC histotypes studied with EOC risk in *BRCA1* and *BRCA2* mutation carriers is shown in **Figure 1**. Four SNPs (rs8098244 (18q11.2), rs2165109 (2q13), rs9886651 (8q24.21) and rs7953249 (12q24.31)) showed associations with EOC risk for *BRCA1* mutation carriers, and one SNP (rs9886651 (8q24.21)) showed an association with risk for *BRCA2* mutation carriers ($P < 0.05$).

Eighteen of the 23 previously published loci were associated with the same histotype at genome-wide significance (excluding the 5 pleiotropic loci published by Kar *et al.*[17]; **Supplementary Table 3**). Of these, 11 showed association with EOC risk for *BRCA1* mutation carriers, and 8 showed association with risk for *BRCA2* mutation carriers ($P < 0.05$). There was significant heterogeneity of risk between the five main, non-overlapping histotypes (high-grade serous, low-grade/borderline serous, endometrioid, clear cell and invasive/borderline mucinous) for 28 of the 40 new and previously published loci (**Supplementary Table 3**).

We carried out a competing-risks association analysis in *BRCA1* and *BRCA2* mutation carriers to investigate whether the observed associations with ovarian cancer in mutation carriers are influenced by associations with breast cancer risk. For this analysis, we used the most significantly associated genotyped SNPs[22]. The EOC hazard ratio (HR) estimates were consistent with the estimates from the main analysis for all SNPs (data not shown). Some evidence suggested that rs7953249 at 12q24.31 was associated with reduced breast cancer risk in *BRCA1* mutation carriers (HR = 0.95, 95% confidence interval (CI) = 0.91–0.99, $P = 0.034$) and that SNP rs2165109 at 2q13 was associated with increased breast cancer risk in *BRCA2* mutation carriers (HR = 1.08, 95% CI = 1.01–1.14, $P = 0.02$). When these associations were analyzed by tumor estrogen receptor (ER) status, the associations for the two SNPs were restricted to ER-negative breast cancer for *BRCA1* ($P = 0.026$ for rs7953249) and *BRCA2* ($P = 0.02$ for rs2165109) mutation carriers.

Association analyses adjusted for the most significant SNP in each region (including three independent SNPs at 8q24.21) did not identify any additional independent association signals in these regions. At the 12 new EOC risk regions, 571 SNPs were deemed potentially causal (**Supplementary Table 4**) and carried forward for functional annotation, expression quantitative trait locus (eQTL) and methylation quantitative trait locus (mQTL) analyses.

Functional and molecular analyses

Of the 571 candidate causal variants in the 12 newly identified loci, 562 variants are located in noncoding DNA sequences and may influence the expression of nearby target genes[23]. We used a variety of *in silico* approaches to identify putative, tissue-specific regulatory biofeatures and candidate susceptibility genes associated with risk SNPs at each locus. For the few risk-associated nonsynonymous variants in protein-coding genes, we also evaluated predicted effects on protein function.

Functional annotation of candidate causal alleles. We mapped the set of 562 non-protein-coding candidate causal SNPs at the 12 susceptibility loci to regulatory biofeatures, using a variety of epigenomic marks profiled in normal and cancer tissues relevant to the cellular origins of the different ovarian cancer histotypes (**Supplementary Table 5**). The cell types interrogated included (i) fallopian tube (FT33 and FT246) and ovarian surface (IOSE4 and IOSE11) epithelial cell lines for serous precursor tissues; (ii) serous-related cancer cell lines, including HGSOC cell lines (UWB1.289 and CaOV3) and a LGSOC cell line (OAW42); (iii) endometriosis epithelial cells (EEC16), as a likely precursor of ENOC; and (iv) cell types relevant to MOC, including MOC cell lines (GTFR230 and MCAS) and both colonic normal (colon crypt) and cancer (HCT116 and HeLa-S3) tissues. The epigenomic marks annotated were open chromatin, identified using formaldehyde-assisted isolation of regulatory element sequencing (FAIRE–seq) and DNase I hypersensitivity sequencing (DNase–seq), and chromatin immunoprecipitation sequencing (ChIP–seq) of histone modifications, specifically histone H3 lysine 27 acetylation (H3K27ac; denoting active chromatin) and histone H3 lysine 4 monomethylation (H3K4me1; marking active and poised enhancers). SNPs were also intersected with Encyclopedia of DNA Elements (ENCODE) transcription factor ChIP–seq data. All tissue types were evaluated for all risk loci. The SNP–biofeature intersections by tissue type are illustrated in **Figure 2** and **Supplementary Table 6**.

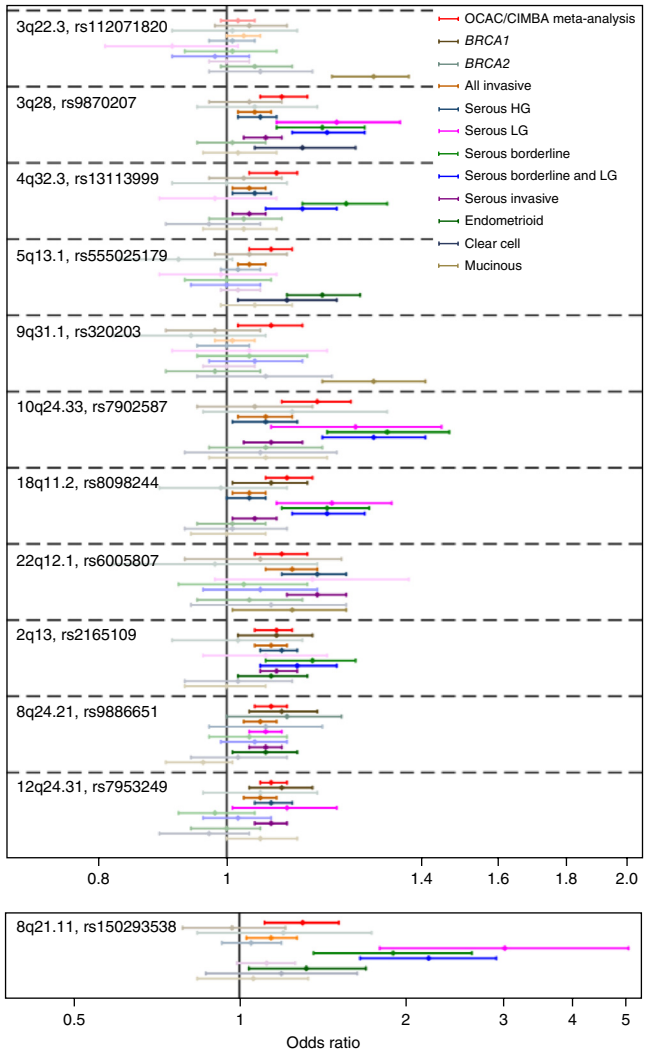

Figure 1 Histotype-specific associations (odds ratios) of the top SNP in 12 new EOC susceptibility regions. The forest plots show the point estimates of odds ratios with the 95% confidence interval around each estimate. Odds ratios and confidence intervals in bold correspond to histotypes significant at nominal $P < 0.05$. HG, high grade; LG, low grade.

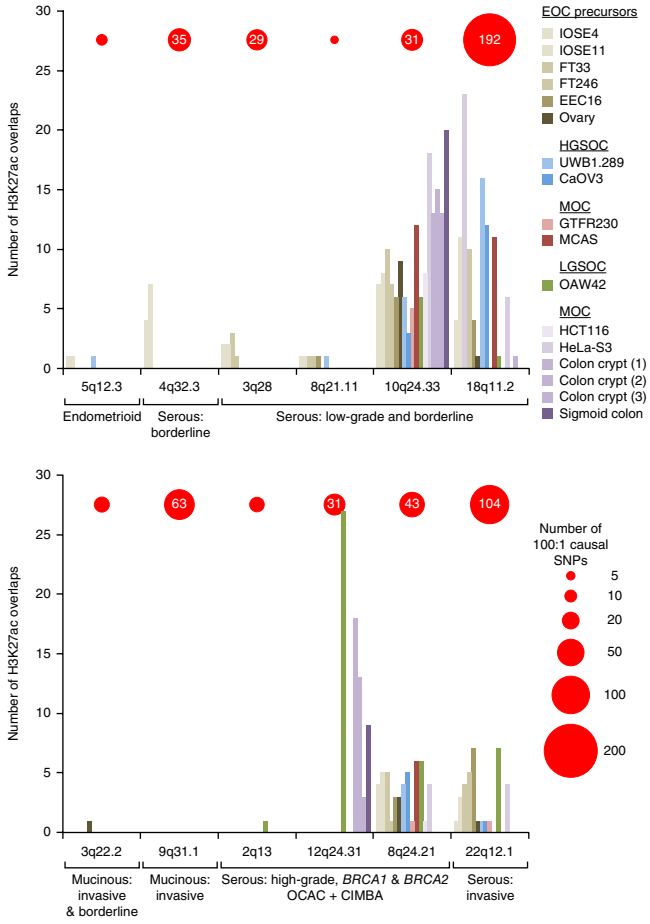

Figure 2 Number of overlaps between candidate causal SNPs and H3K27ac in relevant tissues and cell lines. The height of each bar in the histogram represents the number of candidate causal SNPs at each locus overlapping H3K27ac marks in a particular cell line. Loci are grouped according to their association with the different histotypes of ovarian cancer. The number of causal SNPs at 100:1 odds relative to the top regional SNP is indicated by the red circles. In the key, cell lines are grouped according to their likely relevance to the different histotypes. MOC, mucinous ovarian cancer; LGSOC, low-grade serous ovarian cancer; HGSOC, high-grade serous ovarian cancer; EOC, epithelial ovarian cancer.

Nine (1.6%) of the 571 candidate causal SNPs lie in protein-coding sequences. Five of these are synonymous and four are nonsynonymous but predicted to be benign by PolyPhen-2 (**Supplementary Table 6**). Four SNPs lie within the UTRs of protein-coding genes and so could affect mRNA stability: rs1051149 and rs1051150 in the 3′ UTR of *LAMA3* and rs12327412 in the 5′ UTR of *TTC39C*, all at the 18q11.2 locus, and rs1018128 in the 3′ UTR of *GMNC* at 3q28. The majority of SNP–biofeature intersections ($n = 166$; 29% of all candidate causal SNPs and 97% of candidate causal SNPs overlapping a biofeature) were for SNPs lying within active chromatin and/or open chromatin. Eleven SNPs lie in the promoters of four genes (*PVT1*, *HNF1A*, *TTC39C* and *TTC28*) (**Supplementary Table 6**).

At six serous risk loci (4q32.3, 3q28, 8q21, 18q11, 8q24 and 22q12), we observed extensive SNP–biofeature overlaps, particularly in serous-related tissue types. In contrast, the two MOC susceptibility loci (3q22.3 and 9q31.1) were biofeature-poor regions and showed little or no SNP–biofeature intersection in any of the tissue types under investigation, including MOC and ENCODE cell lines. At the endometrioid EOC risk locus (5q12.3), we observed enhancers in endometriosis, ovarian, fallopian and EOC cell types flanking the small number of risk-associated SNPs ($n = 8$), none of which coincided with regulatory elements.

Several studies have shown that common variant susceptibility alleles are significantly enriched for regulatory elements detected in disease-relevant tissue types. Therefore, we tested for enrichment of SNP–H3K27ac intersections at each locus because H3K27ac was the most comprehensively profiled regulatory feature across different tissue types (**Supplementary Table 7**). At the 12q24.31 locus, a large region of active chromatin spanning the *HNF1A* promoter drove a strong enrichment for risk SNP–H3K27ac intersection in the OAW42 LGSOC cell line ($P = 4.45 \times 10^{-22}$). At 10q24.33 (which is associated with LGSOC and borderline SOC), we identified a significant enrichment of H3K27ac in normal fallopian cells (FT33, $P = 1.09 \times 10^{-4}$; FT246, $P = 4.29 \times 10^{-3}$), HGSOC ovarian cancer cells (UWB1.289, $P = 6.23 \times 10^{-3}$) and MOC cells (GTFR230, $P = 5.16 \times 10^{-3}$) as well as, somewhat surprisingly, colorectal cancer cells (HCT116, $P = 2.64 \times 10^{-4}$) and cervical cancer cells (HeLa-S3, $P = 9.60 \times 10^{-12}$). This locus contains several clusters of H3K27ac activity and transcription factor

binding in ovarian and ENCODE data sets, and these highly active regions showed extensive overlap with candidate causal alleles (**Fig. 3**).

Identifying candidate susceptibility gene targets at risk loci. We used several approaches to identify candidate target genes at the 12 risk loci. First, we hypothesized that target genes underlying disease susceptibility are more likely to display prevalent copy number alterations in ovarian tumor tissues. Amplifications were the most frequent alteration at 6 of the 12 susceptibility loci (**Supplementary Fig. 1**). Contiguous genes were commonly amplified in the same sample, indicating segmental amplifications (data not shown). *HNF1A*, *ORAI1*, *CHEK2*, *XPB1*, *BUB1* and *FOXL2* are found inside the same topologically associating domain (TAD) as candidate causal SNPs and have been previously implicated in ovarian cancer development (**Supplementary Fig. 2**). Notably, *HNF1A*, *ORAI1* and *FOXL2* are amplified in >5% of EOC samples. No TAD was identified for 8q24.21, but *MYC* and *PVT1* appear to be the targets for multiple enhancer elements containing independent EOC risk associations for HGSOC at this locus (A. Gjyshi, G.M.-F., J.P.T., N.T. Woods and K.L. *et al.*, unpublished data).

We also performed eQTL and mQTL analyses in several data sets comprising methylation/expression profiling and germline genotyping in relevant tumor tissues (Online Methods). For eQTL analyses, we evaluated associations between the candidate causal SNPs and all genes profiled within 1 Mb of the index risk SNP at each locus, as this window will contain most *cis*-eQTL associations[23]. Results of the eQTL analyses in each data set were adjusted for tumor copy number and methylation status[24], and a meta-analysis of the two HGSOC data sets from The Cancer Genome Atlas (TCGA) and the Mayo Clinic are shown in **Supplementary Tables 8–11**. The most significant eQTL associations in both HGSOC data sets were identified between the candidate causal risk SNPs at the 10q24.33 risk locus and *OBFC1* expression (TCGA rs11597399, $P = 3.1 \times 10^{-10}$; Mayo rs7902587, $P = 4.0 \times 10^{-4}$; meta-analysis rs34379047, $P = 2.1 \times 10^{-11}$). The risk (T) allele was associated with reduced *OBFC1* expression in both data sets (**Fig. 3d**). We then evaluated all SNPs at this locus (not just the candidate causal SNPs) for eQTL associations; the SNPs with the most significant eQTL associations for *OBFC1* were also candidate causal SNPs for the risk association, reinforcing *OBFC1* as the likely target gene. No expression associations were identified at $P < 1 \times 10^{-4}$ for the candidate causal risk SNPs at any other locus in the eQTL meta-analysis. We used 32 ENOC samples to conduct an eQTL analysis focused on the 5q13.1 ENOC risk locus, but this did not identify any associations at $P < 0.05$ (**Supplementary Table 10**).

mQTL analyses were restricted to the set of 67 CpGs with the most significant association with decreased expression of the 74 genes of interest (within a 1-Mb region centered on the index SNP) in the 12 regions. Results are presented for the most significant mQTL associations for each SNP based on the reduced set of CpGs (**Supplementary Table 12**). We identified two regions with mQTL associations at $P < 0.005$. At 2q13, the risk allele (G) of rs56226558 was associated with reduced methylation of the CpG cg21469370 ($P = 1.4 \times 10^{-3}$), with methylation being associated with reduced expression of *BCL211* ($P = 1.1 \times 10^{-6}$), even though cg21469370 lies in the gene body of *ACOXL*. At 3q22.3, the risk allele (C) of rs68088905 was associated with reduced methylation of the CpG cg06726820 in the promoter of *RBP1* (mQTL $P = 4.9 \times 10^{-3}$). Methylation was

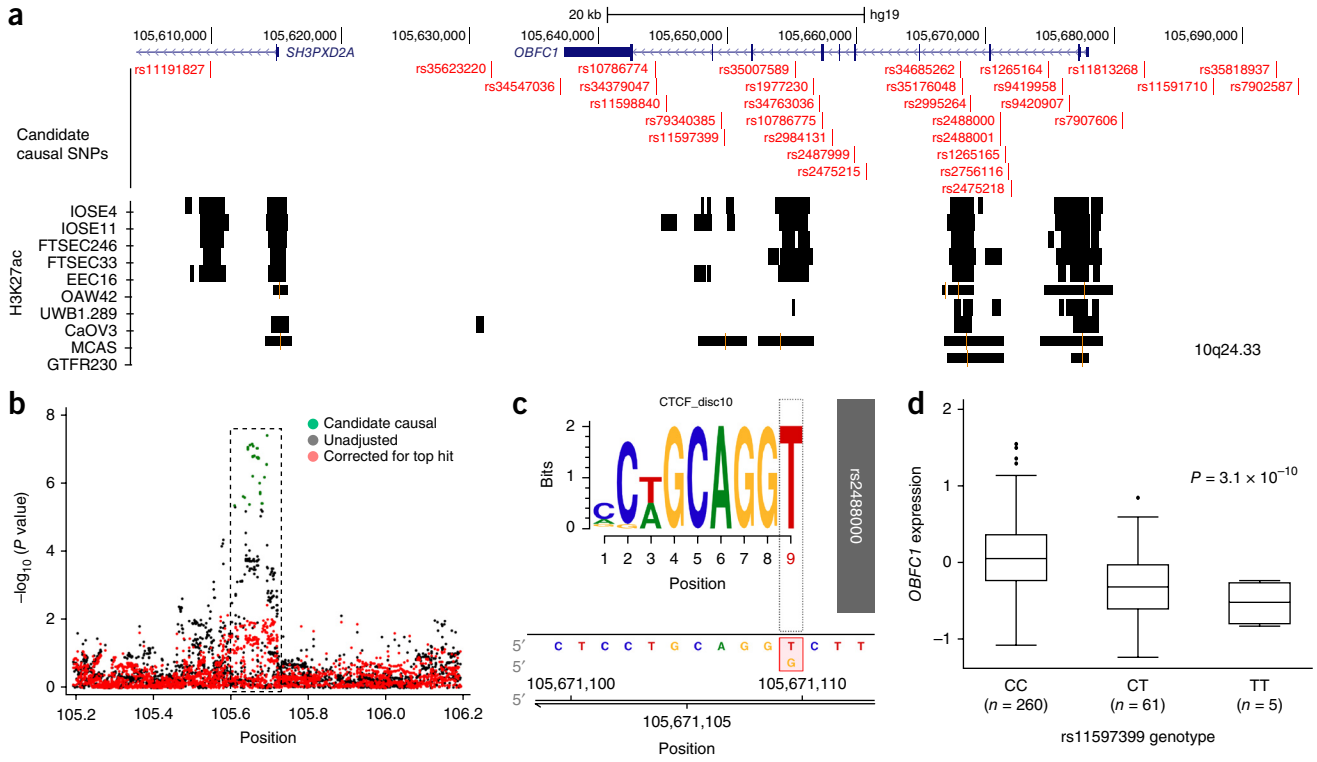

Figure 3 Functional analysis of the 10q24.33 locus. (**a**) Active chromatin, denoted by H3K27ac signal, in EOC-relevant cell types. (**b**) Regional association plot for genotyped and imputed SNPs. The dashed box highlights the region shown in **a**. (**c**) MotifbreakR analysis. A non-canonical CTCF motif is altered by SNP rs2488000. (**d**) eQTL analysis. *OBFC1* expression is associated with rs11597399 genotype in HGSOCs from TCGA. The box-and-whisker plot shows the median (horizontal line within the box), interquartile range (IQR; length of the box) and 1.5 times the IQR (whiskers) of *OBFC1* expression for each genotype.

strongly associated with reduced *RBP1* expression ($P = 1.7 \times 10^{-36}$). We found no highly significant mQTL associations for genes at any other locus, and the eQTL SNP at 10q24.33 was not associated with DNA methylation.

SNPs in the 10q24.33 locus showed the most significant eQTL effects on expression of the *OBFC1* gene. The most significant eQTL SNPs also showed the most epigenetic marks, including rs35007589 (eQTL $P = 2.3 \times 10^{-11}$), rs35176048 (eQTL $P = 2.6 \times 10^{-11}$) and rs34685262 (eQTL $P = 3.8 \times 10^{-11}$). These SNPs intersect regions of open chromatin, H3K27ac and H3K4me1 signal in normal ovarian and fallopian tube epithelial cells and, for rs35176048 and rs34685262, in HGSOC cell lines. These enhancers are not specific to ovarian cell types. At this locus, 11 candidate causal SNPs are predicted by motifbreakR[25] to alter transcription factor binding sites, of which 8 are predicted to have a strong effect on transcription factor binding (**Supplementary Table 13**). Of particular interest, rs2488000 (eQTL $P = 1.4 \times 10^{-10}$) is predicted *in silico* to strongly influence the binding of CTCF, a ubiquitously expressed transcriptional regulator that has a key role in insulator function and chromatin structure (**Fig. 3c**). Furthermore, in ENCODE, there is evidence from ChIP–seq experiments that CTCF does bind at this location in monocytes. Other SNPs predicted to have a strong effect on the binding of other cancer-relevant transcription factors are rs11813268 (*ETS1*), rs7907606 (*FOXP1*) and rs2995264 (*IRF8*) (**Supplementary Table 13**).

At 8q24.21, the candidate causal variants span a ~20-kb region that includes the promoter and first exon of *PVT1*, an oncogenic long noncoding RNA (lncRNA) with known roles in breast and ovarian cancers[26]. The 8q24 region is also a hotspot for association with other cancers[27], with *PVT1*, *MYC* and novel lncRNAs identified as candidate target genes. Five SNPs (rs10956390, rs10098831, rs6470578, rs6990534 and rs4410871) coincide with 11 or more biofeatures in normal ovarian and fallopian tube epithelial cells and in ovarian cancer cells.

DISCUSSION

We have identified 12 new loci associated with different histotypes of EOC at genome-wide significance. Despite the use of a stringent significance threshold, it is possible that some of these represent false positive associations. Wakefield has suggested the use of an approximate Bayes factor to calculate the Bayes false discovery probability (BFDP)[28]. We have estimated the BFDP on the basis of a plausible odds ratio of 1.2 and a prior probability of association of 0.0001. The BFDP was less than 10% for 11 of the 12 associations. We also calculated the BFDP values for the 22 previously reported loci, of which 17 were <1%, 1 was >1% but less than 10%, and 4 were greater than 10%. We did not calculate the BFDP for the five pleiotropic loci reported by Kar *et al.*[17]. These 5 loci, together with the 29 loci with BFDP <10%, bring the total number of susceptibility loci for different histotypes of EOC to 34 for women of European ancestry, of which 27 are associated with risk of invasive EOC at $P < 0.01$. Assuming a polygenic variance of 1.45 (ref. 29), the 27 loci account for approximately 6.4% of the polygenic risk in the population. Incorporating common EOC susceptibility variants into risk assessment tools will improve risk prediction and may be particularly useful for refining risk estimates in *BRCA1* and *BRCA2* mutation carriers.

Some strata in the OCAC analyses pooled data from several studies from the same country. This might increase the potential for bias because of population stratification, but we expect any inflation due to population stratification to be effectively removed by adjusting for the principal components. To evaluate the possible magnitude of such a bias, we compared the inflation of the median test statistic for the analysis of the OncoArray data stratified by study with that for an analysis in which all the cases and controls were combined into a single stratum. There was little difference ($\lambda = 1.054$ versus $\lambda = 1.078$). As these inflation factors are not adjusted for sample size, some of the difference is due to the higher effective sample size of the non-stratified analysis, suggesting that any bias due to pooling data from multiple studies will be minimal.

Consistent with previous studies in EOC and other cancer types, the vast majority of the risk-associated variants were located in non-protein-coding regions of the genome[30], suggesting that these variants influence target gene expression by altering the activity of functional element(s) that regulate the expression of one or more susceptibility genes. Because noncoding biofeatures, such as enhancers, show a high degree of tissue specificity, we intersected EOC risk SNPs with regions of active chromatin catalogued in cell lines representing the different EOC histotypes (HGSOC, ENOC, LGSOC and MOC) and in EOC precursor cells (OSEC/FTSEC for LGSOC/HGSOC, EEC for ENOC and colonic crypt for MOC). Enrichment analyses test for over-representation of cell-type-specific biofeatures intersecting risk variants at confirmed risk loci, in comparison to a lack of enrichment in non-disease-associated tissues. A major strength of our approach was the ability to interrogate histotype-specific epigenomic profiles, and so, in addition to identifying the putative functional targets of risk SNPs, these analyses can also indicate whether some cell types are more likely to be relevant to disease pathogenesis than other cell types. For example, we detected a significant enrichment of active chromatin marks coinciding with SNPs in fallopian tube epithelial cells at the 10q24.33 LGSOC/borderline serous locus, which could reflect recent pathological evidence that some of these tumors arise in the distal fallopian tube[31], as well as HGSOC[2]. At the same locus, we also identified enrichment for biofeatures in a primary MOC cell line, a cancer histotype that is often associated with deregulation of the MAPK pathway, which is also perturbed in LGSOC[32]. Given the growing evidence that regulatory mechanisms are highly tissue specific, it is perhaps to be expected that we find such enrichments in cell types associated with EOC development. However, the lack of enrichment at MOC and ENOC risk loci may indicate that alternative precursor cell types give rise to these histotypes rather than the cell types evaluated in the current study.

eQTL analysis identified associations between the most statistically significant risk-associated SNPs at 10q24.33 and *OBFC1*, many of which also coincide with epigenetic biofeatures. OBFC1 is a subunit of an alpha accessory factor that stimulates the activity of DNA polymerase α primase, the enzyme that initiates DNA replication. OBFC1 also appears to function in a telomere-associated complex that binds telomeric single-stranded DNA *in vitro* and localizes at telomeres *in vivo*[33]. Four SNPs in this region (rs2487999, rs4387287, rs9420907 and rs9419958) have previously been reported to be associated with telomere length (NHGRI-EBI GWAS catalog[27]; **Supplementary Table 14**). The r^2 values between these and rs7902587 are between 0.52 and 0.93 (1000 Genomes Project European populations). However, the associations of all four with LGSOC and borderline serous EOC are attenuated when adjusted for rs7902587, suggesting a single association peak. The alleles associated with an increase in leukocyte telomere length are associated with increased risk of LGSOC and borderline serous EOC. These findings are consistent with the association between borderline EOC and rs7705526 at 5p15 (adjacent to the telomerase reverse transcriptase gene *TERT*)[13]. Furthermore, the histotype specificity is consistent with the suggestion that telomere length differs between the different histotypes of EOC[34].

Candidate causal variants at 3 of the 12 new loci are associated with multiple traits in the NHGRI-EBI GWAS catalog ($P < 5 \times 10^{-8}$).

These traits converge on pathways involved in inflammation and immunity, including monocyte count, C-reactive protein (CRP) levels, γ-glutamyl transpeptidase levels, *N*-glycan levels, allergen sensitization and multiple sclerosis (**Supplementary Table 14**). For example, at the 12q24.31 HGSOC risk locus, the risk alleles of four candidate causal SNPs (rs7979473, rs1183910, rs2393791 and rs7310409) have previously been associated with elevated CRP levels in blood plasma, a marker of inflammation. This is consistent with the established link between chronic inflammation and increased cancer risk. In addition, SNPs within 500 kb of the top SNP at 2q13, 8q24.21, 10q24.33 and 22q12.1 are associated with several different cancers, although only one of these is a candidate causal EOC variant (rs2995264 at 10q24.33 associated with cutaneous malignant melanoma).

This study demonstrates the strength of large-scale collaborations in identifying common variant risk associations for complex traits such as EOC, which is rare, has a high mortality rate and exhibits heterogeneity by histotype. As the largest study thus far with over 90,000 EOC cases and controls, including an additional ~25,000 previously unstudied participants, we identify several new risk loci specific to rarer EOC histotypes: ENOC, MOC, LGSOC and borderline EOC. The histotype-specific nature of these associations adds to the somatic, epidemiological and clinical data indicating that EOC histotypes can be thought of as distinct diseases. The lack of heterogeneity between studies of varying designs, carried out in different populations, and the high levels of statistical significance, with confirmation of known EOC susceptibility loci, provide evidence that these are robust associations. Molecular analyses of genes and the tissue-specific regulatory architecture at these loci, which combined publicly available data sets with systematic, large-scale genome-wide profiling experiments, point to a small number of non-coding biofeatures and target genes that may have a histotype-specific role in EOC initiation and development. Detailed functional studies will be required to define the underlying biology of SNP–regulatory feature interactions to identify the causal SNP(s) at each locus, and to confirm which candidate susceptibility genes represent the targets of these risk SNPs. Evolving technologies, in particular CRISPR–Cas9 genome editing, now enable precise modification of risk SNPs to create isogenic models of different alleles[35], enabling the effects of each allele on disease pathogenesis to be studied, for example at 19p13 (ref. 36), 8q24 (ref. 14), 17q12 (ref. 12) and 5p15 (ref. 13). Finally, given that several previously identified EOC susceptibility alleles are associated with risk of other cancers[17] and that there are similarities in molecular phenotype and/or shared tissue of origin for endometrial cancer, endometriosis, and ENOC and CCOC[37] as well as basal-like breast cancer[38], we anticipate that the loci reported here may also be associated with risk of other cancer-related traits.

URLs. iCOGS, http://www.nature.com/icogs/; The Cancer Genome Atlas Project, http://cancergenome.nih.gov/; cBio Cancer Genomics Portal, http://www.cbioportal.org/; Pupasuite 3.1, http://pupasuite.bioinfo.cipf.es/; OCAC, http://apps.ccge.medschl.cam.ac.uk/consortia/ocac/; CIMBA quality control guidelines, http://ccge.medschl.cam.ac.uk/consortia/cimba/members/data%20management/CIMBA%20and%20BCAC%20Quality%20Control%20November%202008%20v2.doc.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

The OCAC OncoArray genotyping project was funded through grants from the US National Institutes of Health (CA1X01HG007491-01 (C.I.A.), U19-CA148112 (T.A.S.), R01-CA149429 (C.M.P.) and R01-CA058598 (M.T.G.)); Canadian Institutes of Health Research (MOP-86727 (L.E.K.)); and the Ovarian Cancer Research Fund (A.B.). Funding for the CIMBA OncoArray genotyping was provided by the Government of Canada through Genome Canada and the Canadian Institutes of Health Research, the Ministère de l'Économie, de la Science et de l'Innovation du Québec through Génome Québec, the Quebec Breast Cancer Foundation for the PERSPECTIVE project, the US National Institutes of Health (CA1X01HG007491-01 (C.I.A.)), the Odense University Hospital Research Foundation (M.T.), the National R&D Program for Cancer Control, Ministry of Health and Welfare, Republic of Korea (1420190 (S.K.P.)), the Italian Association for Cancer Research (IG16933 (L.O.)) and German Cancer Aid (110837 (R.K.S.). Funding sources for the contributing studies are provided in the **Supplementary Note**.

We pay special tribute to the contribution of Brian Henderson to the GAME-ON consortium and to Olga M. Sinilnikova for her contribution to CIMBA and for her part in the initiation and coordination of GEMO until she sadly passed away on 30 June 2014. We thank the study participants, doctors, nurses, clinical and scientific collaborators, healthcare providers and health information sources that have contributed to the many studies contributing to this manuscript. A full list is provided in the **Supplementary Note**.

AUTHOR CONTRIBUTIONS

Writing group: C.M.P., K.B.K., J.P.T., S.P.K., K.L., S.J.W., D.H., M.A.E., A.N.M., G.C.-T., E.L.G., S.J.R., T.A.S., S.A.G., A.C.A., P.D.P.P. **Coordination of OCAC OncoArray genotyping:** C.M.P., M.J.R., G.C. **Coordination of CIMBA OncoArray genotyping:** G.C.-T., L.M., J.S., P. Soucy. **OncoArray genotyping:** CIDR (M.A., T.S., K.F.D., J. Romm, E.P.), Mayo (J.M.C.), UCam (C. Luccarini). **OncoArray genotype calling** and **quality control:** K.B.K., D.F.E., J.D., D.B., E.D., A. Pirie, A. Lee, J.L., G.L. **Statistical analyses for OCAC:** J.P.T., P.D.P.P. **Statistical analyses for CIMBA:** K.B.K., A. Lee, A.C.A. **Meta-analyses:** K.B.K., A.C.A. **OCAC database management:** M.J.R., A. Berchuck. **CIMBA database management** and ***BRCA1* and *BRCA2* variant nomenclature and classification:** L.M., G.L., A.B.S. **Supervision of OCAC statistical analyses:** P.D.P.P. **Supervision of CIMBA statistical analyses:** A.C.A. **Conceiving and coordination of the synthesis of the OncoArray:** P.D.P.P., D.F.E., C.I.A., S. Chanock, S.G., D.J.H., A.C.A., J.S. **Functional analyses:** P.C.L., S. Coetzee, M.A.E., S.A.G., E.L.G., D.H., S.P.K., K.L., J.M.L, G.M.-F., A.N.M., S.J.W., G.C.-T., J. Beesley. **Provision of DNA samples and/or phenotypic data:** C.M.A., K.K.H.A., J. Adlard, I.L.A., H.A.-C., N. Antonenkova, G.A., N. Arnold, B.K.A., B.A., J. Azzollini, J. Balmaña, S.N.B., L. Barjhoux, R.B.B., Y.B., M.W.B., A.B.-F., J. Benitez, A. Berchuck, M. Bermisheva, M.Q. Bernardini, M.J. Birrer, L. Bjorge, A. Black, K. Blankstein, M.J. Blok, C. Bodelon, N.B., A. Bojesen, B. Bonanni, Å. Borg, A.R.B., J.D.B., C. Brewer, L. Brinton, P.B., A.B.-W., F.B., J. Brunet, B. Buecher, R.B., S.S.B., T. Caldes, M.A.C., I.C., R.C., M.E.C., T. Cescon, S.B.C., J.C.-C., X.Q.C., G.C.-T., Y.-E.C., J. Chiquette, W.K.C., K.B.M.C., T. Conner, J. Cook, L.S.C., F.J.C., D.W.C., A.A.D., M.B.D., F. Damiola, S.D.D., A.D., F. Dao, R.D., A.d.F., C.D., O.D., Y.C.D., J.A.D., S.M.D., C.M.D., T.D., L.D., M. Duran, M. Dürst, B.D., D.E., T.E., R.E., U.E., B.E., A.B.E., S.E., M.E., K.H.E., C.E., D.G.E., P.A.F., S.F., S.F.F., J.M.F., T.M.F., Z.C.F., R.T.F., F.F., W.D.F., G.F., B.L.F., E.F., D.F., P.A.G., J. Garber, M.J.G., V.G.-B., S.A.G., A.G., A.G.-M., A.-M.G., G.G.G. R.G., G. Glendon, A.K.G., D.E.G., E.L.G., M.T.G., T.G., M.G., M.H.G., J. Gronwald, E. Hahnen, C.A.H., N.H., U.H., T.V.O.H., P.A.H., H.R.H., J. Hauke, A. Hein, A. Henderson, M.A.T.H., P.H., S.H., C.K.H., E. Høgdall, F.B.L.H., H.H., M.J.H., K.H., R.-Y.H., P.J.H., J. Hung, D.G.H., T.H., E.N.I., C.I., E.S.I., L.I., A.I., A. Jakubowska, P.J., R.J., A. Jensen, M.J., U.B.J., E.M.J., S.J., M.E.J., P.K., B.Y.K., A. Karzenis, K.K., L.E.K., C.J.K., E.K., L.A.K., J.I.K., S.-W.K., S.K.K., M.K., R.K.K., T.A.K., J.K., A. Kwong, Y.L., D. Lambrechts, N.L., M.C.L., C. Lazaro, N.D.L., L.L.M., J.W.L., S.B.L., A. Leminen, D. Leroux, J. Lester, F.L., D.A.L., D. Liang, C. Liebrich, J. Lilyquist, L. Lipworth, J. Lissowska, K.H.L., J. Lubiński, L. Lundvall, P.L.M., S. Manoukian, L.F.A.G.M., T.M., S. Mazoyer, J.N.M., V.M., J.R.M., I.M., H.M.-H., A.M., U.M., A.R.M., M.A. Merritt, R.L.M., G.M., F.M., J.M.-S., M. Moffitt, M. Montagna, K.B.M., A.M.M., J.M., S.A.N., K.L.N., L.N., R.B.N., S.L.N., H.N., D.N., R.L.N., K. Odunsi, K. Offit, E.O., O.I.O., H.O., C.O., D.M.O'M., K.O., N.C.O.-M., N.O., S.O., A.O., L.O., D.P., L. Papi, S.K.P., T.-W.P.-S., J.P., C.L.P., I.S.P., P.H.M.P., B. Peissel, A. Peixoto, T. Pejovic, L.M.P., J.B.P., P. Peterlongo, L.P., G.P., P.D.P.P., C.M.P., K.-A.P., M.P., M.C.P., A.M.P., S.R.P., T. Pocza, E.M.P., B. Poppe, M.E.P., F.P., D. Palli, D. Prokofyeva, M.A.P., P. Pujol, P. Radice, S.J.R., J. Rantala, C.R.-F., G.R., K.R., P. Rice, A. Richardson, H.A.R., M.R., G.C.R., C.R.-A., M.A. Rookus, M.A. Rossing, J.H.R., A. Rudolph, I.B.R., H.B.S., D.P.S., J.M.S., R.K.S., M.J.S.,

T.A.S., L. Senter, V.W.S., G. Severi, P. Sharma, N.S., L.E. Side, W.S., J.S., C.F.S., H. Sobol, H. Song, P. Soucy, M.S., A.B.S., Z.S., D.S., D.S.-L., L.E.S.-C., G. Sukiennicki, R.S., C.S., A.J.S., C.I.S., L. Szafron, Y.Y.T., J.A.T., M.-K.T., M.R.T., S.-H.T., K.L.T., M. Thomassen, P.J.T., L.C.V.T., D.L.T., L.T., A.V.T., M. Tischkowitz, S.T., A.E.T., A. Tone, B.T., R.T., A. Trichopoulou, N.T., S.S.T., A.M.V.A., D.V.D.B., A.H.v.d.H., R.B.v.d.L., M.V.H., E.V.N., E.J.v.R., A. Vanderstichele, R.V.-M., A. Vega, D.V.E., I.V., J.V., R.A.V., A. Vratimos, L.W., C.W., D.W., S.W.-G., B.W., P.M.W., C.R.W., J.N.W., N.W., A.S.W., J.T.W., L.R.W., A.W., M.W., A.H.W., X.W., H.Y., D.Y., A.Z., K.K.Z.

All authors read and approved the final manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Prat, J. Ovarian carcinomas: five distinct diseases with different origins, genetic alterations, and clinicopathological features. *Virchows Arch.* **460**, 237–249 (2012).
2. Vang, R., Shih, Ie.-M. & Kurman, R.J. Fallopian tube precursors of ovarian low- and high-grade serous neoplasms. *Histopathology* **62**, 44–58 (2013).
3. Wang, Y. *et al.* Tubal origin of ovarian endometriosis and clear cell and endometrioid carcinoma. *Am. J. Cancer Res.* **5**, 869–879 (2015).
4. Gounaris, I. & Brenton, J.D. Molecular pathogenesis of ovarian clear cell carcinoma. *Future Oncol.* **11**, 1389–1405 (2015).
5. Brown, J. & Frumovitz, M. Mucinous tumors of the ovary: current thoughts on diagnosis and management. *Curr. Oncol. Rep.* **16**, 389 (2014).
6. Jervis, S. *et al.* Ovarian cancer familial relative risks by tumour subtypes and by known ovarian cancer genetic susceptibility variants. *J. Med. Genet.* **51**, 108–113 (2014).
7. Lu, Y. *et al.* Most common 'sporadic' cancers have a significant germline genetic component. *Hum. Mol. Genet.* **23**, 6112–6118 (2014).
8. Song, H. *et al.* A genome-wide association study identifies a new ovarian cancer susceptibility locus on 9p22.2. *Nat. Genet.* **41**, 996–1000 (2009).
9. Goode, E.L. *et al.* A genome-wide association study identifies susceptibility loci for ovarian cancer at 2q31 and 8q24. *Nat. Genet.* **42**, 874–879 (2010).
10. Bolton, K.L. *et al.* Common variants at 19p13 are associated with susceptibility to ovarian cancer. *Nat. Genet.* **42**, 880–884 (2010).
11. Permuth-Wey, J. *et al.* Identification and molecular characterization of a new ovarian cancer susceptibility locus at 17q21.31. *Nat. Commun.* **4**, 1627 (2013).
12. Shen, H. *et al.* Epigenetic analysis leads to identification of *HNF1B* as a subtype-specific susceptibility gene for ovarian cancer. *Nat. Commun.* **4**, 1628 (2013).
13. Bojesen, S.E. *et al.* Multiple independent variants at the *TERT* locus are associated with telomere length and risks of breast and ovarian cancer. *Nat. Genet.* **45**, 371–384 (2013).
14. Pharoah, P.D. *et al.* GWAS meta-analysis and replication identifies three new susceptibility loci for ovarian cancer. *Nat. Genet.* **45**, 362–370 (2013).
15. Kuchenbaecker, K.B. *et al.* Identification of six new susceptibility loci for invasive epithelial ovarian cancer. *Nat. Genet.* **47**, 164–171 (2015).
16. Kelemen, L.E. *et al.* Genome-wide significant risk associations for mucinous ovarian carcinoma. *Nat. Genet.* **47**, 888–897 (2015).
17. Kar, S.P. *et al.* Genome-wide meta-analyses of breast, ovarian, and prostate cancer association studies identify multiple new susceptibility loci shared by at least two cancer types. *Cancer Discov.* **6**, 1052–1067 (2016).
18. Amos, C.I. *et al.* The OncoArray Consortium: a network for understanding the genetic architecture of common cancers. *Cancer Epidemiol. Biomarkers Prev.* **26**, 126–135 (2017).
19. Couch, F.J. *et al.* Genome-wide association study in *BRCA1* mutation carriers identifies novel loci associated with breast and ovarian cancer risk. *PLoS Genet.* **9**, e1003212 (2013).
20. Abecasis, G.R. *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).
21. Lawrenson, K. *et al.* Common variants at the *CHEK2* gene locus and risk of epithelial ovarian cancer. *Carcinogenesis* **36**, 1341–1353 (2015).
22. Barnes, D.R., Lee, A., Easton, D.F. & Antoniou, A.C. Evaluation of association methods for analysing modifiers of disease risk in carriers of high-risk mutations. *Genet. Epidemiol.* **36**, 274–291 (2012).
23. Stranger, B.E. *et al.* Patterns of *cis* regulatory variation in diverse human populations. *PLoS Genet.* **8**, e1002639 (2012).
24. Li, Q. *et al.* Integrative eQTL-based analyses reveal the biology of breast cancer risk loci. *Cell* **152**, 633–641 (2013).
25. Coetzee, S.G., Coetzee, G.A. & Hazelett, D.J. motifbreakR: an R/Bioconductor package for predicting variant effects at transcription factor binding sites. *Bioinformatics* **31**, 3847–3849 (2015).
26. Guan, Y. *et al.* Amplification of *PVT1* contributes to the pathophysiology of ovarian and breast cancer. *Clin. Cancer Res.* **13**, 5745–5755 (2007).
27. Welter, D. *et al.* The NHGRI GWAS Catalog, a curated resource of SNP–trait associations. *Nucleic Acids Res.* **42**, D1001–D1006 (2014).
28. Wakefield, J. A Bayesian measure of the probability of false discovery in genetic epidemiology studies. *Am. J. Hum. Genet.* **81**, 208–227 (2007).
29. Jervis, S. *et al.* A risk prediction algorithm for ovarian cancer incorporating *BRCA1*, *BRCA2*, common alleles and other familial effects. *J. Med. Genet.* **52**, 465–475 (2015).
30. Coetzee, S.G. *et al.* Cell-type-specific enrichment of risk-associated regulatory elements at ovarian cancer susceptibility loci. *Hum. Mol. Genet.* **24**, 3595–3607 (2015).
31. Kurman, R.J. *et al.* Papillary tubal hyperplasia: the putative precursor of ovarian atypical proliferative (borderline) serous tumors, noninvasive implants, and endosalpingiosis. *Am. J. Surg. Pathol.* **35**, 1605–1614 (2011).
32. Diaz-Padilla, I. *et al.* Ovarian low-grade serous carcinoma: a comprehensive update. *Gynecol. Oncol.* **126**, 279–285 (2012).
33. Wan, M., Qin, J., Songyang, Z. & Liu, D. OB fold-containing protein 1 (OBFC1), a human homolog of yeast Stn1, associates with TPP1 and is implicated in telomere length regulation. *J. Biol. Chem.* **284**, 26725–26731 (2009).
34. Kuhn, E. *et al.* Telomere length in different histologic types of ovarian carcinoma with emphasis on clear cell carcinoma. *Mod. Pathol.* **24**, 1139–1145 (2011).
35. Spisák, S. *et al.* CAUSEL: an epigenome- and genome-editing pipeline for establishing function of noncoding GWAS variants. *Nat. Med.* **21**, 1357–1363 (2015).
36. Lawrenson, K. *et al.* Functional mechanisms underlying pleiotropic risk alleles at the 19p13.1 breast–ovarian cancer susceptibility locus. *Nat. Commun.* **7**, 12675 (2016).
37. Lu, Y. *et al.* Shared genetics underlying epidemiological association between endometriosis and ovarian cancer. *Hum. Mol. Genet.* **24**, 5955–5964 (2015).
38. Jönsson, J.M. *et al.* Molecular subtyping of serous ovarian tumors reveals multiple connections to intrinsic breast cancer subtypes. *PLoS One* **9**, e107643 (2014).

Catherine M Phelan[1,2,304], Karoline B Kuchenbaecker[3,4,304], Jonathan P Tyrer[5,304], Siddhartha P Kar[5,304], Kate Lawrenson[6,304], Stacey J Winham[7,304], Joe Dennis[5], Ailith Pirie[5], Marjorie J Riggan[8], Ganna Chornokur[1], Madalene A Earp[9], Paulo C Lyra Jr[1], Janet M Lee[10], Simon Coetzee[10], Jonathan Beesley[11], Lesley McGuffog[12], Penny Soucy[13], Ed Dicks[5], Andrew Lee[12], Daniel Barrowdale[12], Julie Lecarpentier[12], Goska Leslie[12], Cora M Aalfs[14], Katja K H Aben[15,16], Marcia Adams[17], Julian Adlard[18], Irene L Andrulis[19,20], Hoda Anton-Culver[21,22], Natalia Antonenkova[23]AOCS study group[24], Gerasimos Aravantinos[25], Norbert Arnold[26], Banu K Arun[27], Brita Arver[28], Jacopo Azzollini[29], Judith Balmaña[30], Susana N Banerjee[31], Laure Barjhoux[32], Rosa B Barkardottir[33,34], Yukie Bean[35,36], Matthias W Beckmann[37], Alicia Beeghly-Fadiel[38–41], Javier Benitez[42,43], Marina Bermisheva[44], Marcus Q Bernardini[45], Michael J Birrer[46], Line Bjorge[47,48], Amanda Black[49], Kenneth Blankstein[50], Marinus J Blok[51], Clara Bodelon[49], Natalia Bogdanova[52], Anders Bojesen[53], Bernardo Bonanni[54], Åke Borg[55], Angela R Bradbury[56], James D Brenton[57], Carole Brewer[58], Louise Brinton[49], Per Broberg[59], Angela Brooks-Wilson[60], Fiona Bruinsma[61], Joan Brunet[62], Bruno Buecher[63], Ralf Butzow[64], Saundra S Buys[65], Trinidad Caldes[66], Maria A Caligo[67], Ian Campbell[68,69], Rikki Cannioto[70], Michael E Carney[71], Terence Cescon[50], Salina B Chan[72], Jenny Chang-Claude[73,74], Stephen Chanock[49], Xiao Qing Chen[11], Yoke-Eng Chiew[75,76], Jocelyne Chiquette[77], Wendy K Chung[78], Kathleen B M Claes[79], Thomas Conner[65], Linda S Cook[80],

Jackie Cook[81], Daniel W Cramer[82], Julie M Cunningham[83], Aimee A D'Aloisio[84], Mary B Daly[85], Francesca Damiola[32], Sakaeva Dina Damirovna[86], Agnieszka Dansonka-Mieszkowska[87], Fanny Dao[88], Rosemarie Davidson[89], Anna DeFazio[75,76], Capucine Delnatte[90], Kimberly F Doheny[17], Orland Diez[91,92], Yuan Chun Ding[93], Jennifer Anne Doherty[94], Susan M Domchek[56], Cecilia M Dorfling[95], Thilo Dörk[96], Laure Dossus[97], Mercedes Duran[98], Matthias Dürst[99], Bernd Dworniczak[100], Diana Eccles[101], Todd Edwards[38–41], Ros Eeles[102], Ursula Eilber[73], Bent Ejlertsen[103], Arif B Ekici[104], Steve Ellis[12], Mingajeva Elvira[86], EMBRACE Study[24], Kevin H Eng[105], Christoph Engel[106], D Gareth Evans[107], Peter A Fasching[37,108], Sarah Ferguson[45], Sandra Fert Ferrer[109], James M Flanagan[110], Zachary C Fogarty[7], Renée T Fortner[73], Florentia Fostira[111], William D Foulkes[112], George Fountzilas[113], Brooke L Fridley[114], Tara M Friebel[115], Eitan Friedman[116,117], Debra Frost[12], Patricia A Ganz[118], Judy Garber[119], María J García[42,43], Vanesa Garcia-Barberan[66], Andrea Gehrig[120], GEMO Study Collaborators[24], Aleksandra Gentry-Maharaj[121], Anne-Marie Gerdes[122], Graham G Giles[61,123,124], Rosalind Glasspool[125], Gord Glendon[126], Andrew K Godwin[127], David E Goldgar[128], Teodora Goranova[57], Martin Gore[129], Mark H Greene[130], Jacek Gronwald[131], Stephen Gruber[132], Eric Hahnen[133], Christopher A Haiman[134], Niclas Håkansson[135], Ute Hamann[136], Thomas V O Hansen[137], Patricia A Harrington[5], Holly R Harris[135], Jan Hauke[133], HEBON Study[24], Alexander Hein[37], Alex Henderson[138], Michelle A T Hildebrandt[139], Peter Hillemanns[140], Shirley Hodgson[141], Claus K Høgdall[142], Estrid Høgdall[143,144], Frans B L Hogervorst[145], Helene Holland[11], Maartje J Hooning[146], Karen Hosking[5], Ruea-Yea Huang[147], Peter J Hulick[148], Jillian Hung[75,76], David J Hunter[149], David G Huntsman[150,151], Tomasz Huzarski[131], Evgeny N Imyanitov[152], Claudine Isaacs[153], Edwin S Iversen[154], Louise Izatt[155], Angel Izquierdo[62], Anna Jakubowska[131], Paul James[156,157], Ramunas Janavicius[158,159], Mats Jernetz[160], Allan Jensen[143], Uffe Birk Jensen[161], Esther M John[162], Sharon Johnatty[11], Michael E Jones[163], Päivi Kannisto[160], Beth Y Karlan[6], Anthony Karnezis[150,151], Karin Kast[164], KConFab Investigators[24], Catherine J Kennedy[75,76], Elza Khusnutdinova[44], Lambertus A Kiemeney[15], Johanna I Kiiski[165], Sung-Won Kim[166], Susanne K Kjaer[142,143], Martin Köbel[167], Reidun K Kopperud[47,48], Torben A Kruse[168], Jolanta Kupryjanczyk[87], Ava Kwong[169–171], Yael Laitman[116,117], Diether Lambrechts[172,173], Nerea Larrañaga[174,175], Melissa C Larson[7], Conxi Lazaro[176], Nhu D Le[177], Loic Le Marchand[178], Jong Won Lee[179], Shashikant B Lele[180], Arto Leminen[165], Dominique Leroux[181], Jenny Lester[6], Fabienne Lesueur[182,183], Douglas A Levine[88], Dong Liang[184], Clemens Liebrich[185], Jenna Lilyquist[186], Loren Lipworth[38,39,41], Jolanta Lissowska[187], Karen H Lu[188], Jan Lubiński[131], Craig Luccarini[5], Lene Lundvall[189], Phuong L Mai[130], Gustavo Mendoza-Fandiño[1], Siranoush Manoukian[29], Leon F A G Massuger[15], Taymaa May[45], Sylvie Mazoyer[190], Jessica N McAlpine[191], Valerie McGuire[192], John R McLaughlin[193], Iain McNeish[194], Hanne Meijers-Heijboer[195], Alfons Meindl[196], Usha Menon[121], Arjen R Mensenkamp[197], Melissa A Merritt[198], Roger L Milne[61,123], Gillian Mitchell[156,157], Francesmary Modugno[199,200,201], Joanna Moes-Sosnowska[87], Melissa Moffitt[36,202], Marco Montagna[203], Kirsten B Moysich[70], Anna Marie Mulligan[20,204], Jacob Musinsky[205], Katherine L Nathanson[56], Lotte Nedergaard[206], Roberta B Ness[207], Susan L Neuhausen[93], Heli Nevanlinna[165], Dieter Niederacher[208], Robert L Nussbaum[209,210], Kunle Odunsi[180], Edith Olah[211], Olufunmilayo I Olopade[212], Håkan Olsson[59,213], Curtis Olswold[186], David M O'Malley[214], Kai-ren Ong[215], N Charlotte Onland-Moret[216], OPAL study group[24], Nicholas Orr[217], Sandra Orsulic[6], Ana Osorio[42], Domenico Palli[218], Laura Papi[219], Tjoung-Won Park-Simon[140], James Paul[220], Celeste L Pearce[134,221], Inge Søkilde Pedersen[222], Petra H M Peeters[216], Bernard Peissel[29], Ana Peixoto[223], Tanja Pejovic[36,202], Liisa M Pelttari[165], Jennifer B Permuth[1], Paolo Peterlongo[224], Lidia Pezzani[29], Georg Pfeiler[225], Kelly-Anne Phillips[123,157,226], Marion Piedmonte[227], Malcolm C Pike[134,228], Anna M Piskorz[57], Samantha R Poblete[180], Timea Pocza[211], Elizabeth M Poole[229], Bruce Poppe[79], Mary E Porteous[230], Fabienne Prieur[231], Darya Prokofyeva[86], Elizabeth Pugh[17], Miquel Angel Pujana[232], Pascal Pujol[233], Paolo Radice[234], Johanna Rantala[235], Christine Rappaport-Fuerhauser[225], Gad Rennert[236,237], Kerstin Rhiem[133], Patricia Rice[50], Andrea Richardson[238], Mark Robson[239], Gustavo C Rodriguez[240], Cristina Rodríguez-Antona[43,241], Jane Romm[17], Matti A Rookus[242], Mary Anne Rossing[243], Joseph H Rothstein[192], Anja Rudolph[73], Ingo B Runnebaum[99], Helga B Salvesen[47,48], Dale P Sandler[244], Minouk J Schoemaker[163], Leigha Senter[245], V Wendy Setiawan[134], Gianluca Severi[246–249], Priyanka Sharma[250], Tameka Shelford[17], Nadeem Siddiqui[251], Lucy E Side[252], Weiva Sieh[192], Christian F Singer[225], Hagay Sobol[253], Honglin Song[5], Melissa C Southey[254], Amanda B Spurdle[11], Zsofia Stadler[239], Doris Steinemann[255], Dominique Stoppa-Lyonnet[63], Lara E Sucheston-Campbell[256], Grzegorz Sukiennicki[131], Rebecca Sutphen[257], Christian Sutter[258],

Anthony J Swerdlow[163,259], Csilla I Szabo[260], Lukasz Szafron[87], Yen Y Tan[225], Jack A Taylor[244], Muy-Kheng Tea[225], Manuel R Teixeira[223,261], Soo-Hwang Teo[262,263], Kathryn L Terry[82,264], Pamela J Thompson[265], Liv Cecilie Vestrheim Thomsen[47,48], Darcy L Thull[266], Laima Tihomirova[267], Anna V Tinker[268], Marc Tischkowitz[112,269], Silvia Tognazzo[203], Amanda Ewart Toland[270], Alicia Tone[45], Britton Trabert[49], Ruth C Travis[271], Antonia Trichopoulou[272,273], Nadine Tung[274], Shelley S Tworoger[229,264], Anne M van Altena[275], David Van Den Berg[134], Annemarie H van der Hout[276], Rob B van der Luijt[277], Mattias Van Heetvelde[79], Els Van Nieuwenhuysen[278], Elizabeth J van Rensburg[95], Adriaan Vanderstichele[278], Raymonda Varon-Mateeva[279], Ana Vega[280,281], Digna Velez Edwards[39,41,282], Ignace Vergote[278], Robert A Vierkant[7], Joseph Vijai[205], Athanassios Vratimos[111], Lisa Walker[283], Christine Walsh[6], Dorothea Wand[284], Shan Wang-Gohrke[285], Barbara Wappenschmidt[133], Penelope M Webb[286], Clarice R Weinberg[287], Jeffrey N Weitzel[288], Nicolas Wentzensen[49], Alice S Whittemore[192,289], Juul T Wijnen[290], Lynne R Wilkens[178], Alicja Wolk[135], Michelle Woo[150,151], Xifeng Wu[139], Anna H Wu[134], Hannah Yang[49], Drakoulis Yannoukakos[111], Argyrios Ziogas[22], Kristin K Zorn[266], Steven A Narod[291], Douglas F Easton[4,5], Christopher I Amos[292], Joellen M Schildkraut[293], Susan J Ramus[294,295], Laura Ottini[296], Marc T Goodman[265,297], Sue K Park[298–300], Linda E Kelemen[301], Harvey A Risch[302], Mads Thomassen[168], Kenneth Offit[205], Jacques Simard[13], Rita Katharina Schmutzler[133], Dennis Hazelett[303], Alvaro N Monteiro[1], Fergus J Couch[83], Andrew Berchuck[8], Georgia Chenevix-Trench[11,305], Ellen L Goode[9,305], Thomas A Sellers[1,305], Simon A Gayther[10,305], Antonis C Antoniou[4,5,305] & Paul D P Pharoah[4,5,305]

[1]Department of Cancer Epidemiology, Moffitt Cancer Center, Tampa, Florida, USA. [2]Department of Gynecologic Oncology, Moffitt Cancer Center, Tampa, Florida, USA. [3]Wellcome Trust Sanger Institute, Wellcome Genome Campus, Hinxton, UK. [4]Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Strangeways Research Laboratory, Cambridge, UK. [5]Department of Oncology, Centre for Cancer Genetic Epidemiology, University of Cambridge, Strangeways Research Laboratory, Cambridge, UK. [6]Women's Cancer Program at the Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA. [7]Division of Biomedical Statistics and Informatics, Department of Health Science Research, Mayo Clinic, Rochester, Minnesota, USA. [8]Department of Obstetrics and Gynecology, Duke University Medical Center, Durham, North Carolina, USA. [9]Department of Health Science Research, Division of Epidemiology, Mayo Clinic, Rochester, Minnesota, USA. [10]Center for Bioinformatics and Functional Genomics, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA. [11]Department of Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. [12]Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK. [13]Genomics Center, Centre Hospitalier Universitaire de Québec Research Center and Laval University, Quebec City, Quebec, Canada. [14]Department of Clinical Genetics, Academic Medical Center, Amsterdam, the Netherlands. [15]Radboud University Medical Center, Radboud Institute for Health Sciences, Department for Health Evidence, Nijmegen, the Netherlands. [16]Netherlands Comprehensive Cancer Organisation, Utrecht, the Netherlands. [17]Center for Inherited Disease Research, Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. [18]Yorkshire Regional Genetics Service, Chapel Allerton Hospital, Leeds, UK. [19]Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital and Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada. [20]Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada. [21]Genetic Epidemiology Research Institute, UCI Center for Cancer Genetics Research and Prevention, School of Medicine, University of California–Irvine, Irvine, California, USA. [22]Department of Epidemiology, University of California–Irvine, Irvine, California, USA. [23]N.N. Alexandrov National Cancer Centre of Belarus, Minsk, Belarus. [24]A full list of members appears in the **Supplementary Note**. [25]'Agii Anargiri' Cancer Hospital, Athens, Greece. [26]Department of Gynaecology and Obstetrics, University Hospital of Schleswig-Holstein, Campus Kiel, Christian Albrechts University Kiel, Kiel, Germany. [27]Department of Breast Medical Oncology and Clinical Cancer Genetics Program, University of Texas MD Anderson Cancer Center, Houston, Texas, USA. [28]Department of Oncology and Pathology, Karolinska University Hospital and Karolinska Institutet, Stockholm, Sweden. [29]Unit of Medical Genetics, Department of Preventive and Predictive Medicine, Fondazione IRCCS (Istituto di Ricovero e Cura a Carattere Scientifico) Istituto Nazionale Tumori (INT), Milan, Italy. [30]Department of Medical Oncology, University Hospital, Vall d'Hebron, Barcelona, Spain. [31]Gynaecology Unit, Royal Marsden Hospital, London, UK. [32]Bâtiment Cheney D, Centre Léon Bérard, Lyon, France. [33]Laboratory of Cell Biology, Department of Pathology, Landspitali, Reykjavik, Iceland. [34]BMC (Biomedical Centre), Faculty of Medicine, University of Iceland, Reykjavik, Iceland. [35]Department of Gynecologic Oncology, Oregon Health & Science University, Portland, Oregon, USA. [36]Knight Cancer Institute, Oregon Health & Science University, Portland, Oregon, USA. [37]Department of Gynecology and Obstetrics, Universtiy Hospital Erlangen, Friedrich Alexander University Erlangen-Nuremberg, Comprehensive Cancer Center Erlangen-Nuremberg, Erlangen, Germany. [38]Division of Epidemiology, Department of Medicine, Vanderbilt School of Medicine, Nashville, Tennessee, USA. [39]Vanderbilt Epidemiology Center, Department of Medicine, Vanderbilt Medical School, Nashville, Tennessee, USA. [40]Institute for Medicine and Public Health, Vanderbilt School of Medicine, Nashville, Tennessee, USA. [41]Vanderbilt-Ingram Cancer Center, Vanderbilt School of Medicine, Nashville, Tennessee, USA. [42]Human Genetics Group, Spanish National Cancer Centre (CNIO), Madrid, Spain. [43]Biomedical Network on Rare Diseases (CIBERER), Madrid, Spain. [44]Institute of Biochemistry and Genetics, Ufa Science Center, Russian Academy of Sciences, Ufa, Russian Federation. [45]Division of Gynecologic Oncology, Princess Margaret Hospital, University Health Network, Toronto, Ontario, Canada. [46]Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA. [47]Department of Gynecology and Obstetrics, Haukeland University Hospital, Bergen, Norway. [48]Centre for Cancer Biomarkers, Department of Clinical Science, University of Bergen, Bergen, Norway. [49]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland, USA. [50]Clinical Cancer Genetics for the City of Hope, Clinical Cancer Genetics Community Research Network, Duarte, California, USA. [51]Department of Clinical Genetics, Maastricht University Medical Center, Maastricht, the Netherlands. [52]Radiation Oncology Research Unit, Hannover Medical School, Hannover, Germany. [53]Department of Clinical Genetics, Vejle Hospital, Vejle, Denmark. [54]Division of Cancer Prevention and Genetics, Istituto Europeo di Oncologia (IEO), Milan, Italy. [55]Department of Oncology, Clinical Sciences, Lund University and Skåne University Hospital, Lund, Sweden. [56]Department of Medicine, Abramson Cancer Center, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania, USA. [57]Cancer Research UK (CRUK) Cambridge Institute, University of Cambridge, Cambridge, UK. [58]Department of Clinical Genetics, Royal Devon and Exeter Hospital, Exeter, UK. [59]Department of Cancer Epidemiology, University Hospital, Lund, Lund University, Lund, Sweden. [60]Canada's Michael Smith Genome Sciences Centre, BC Cancer Agency, Vancouver, British Columbia, Canada. [61]Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Victoria, Australia. [62]Genetic Counseling Unit, Hereditary Cancer Program, IDIBGI (Institut d'Investigació Biomèdica de Girona), Catalan Institute of Oncology, Girona, Spain. [63]Service de Génétique Oncologique, Institut Curie, Paris, France. [64]Department of Pathology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland. [65]Department of Medicine, Huntsman Cancer Institute, Salt Lake City, Utah, USA. [66]Molecular Oncology Laboratory, Hospital Clínico San Carlos, IdISSC (Instituto de Investigación Sanitaria del Hospital Clínico San Carlos), Madrid, Spain. [67]Section of Genetic Oncology, Department of Laboratory Medicine, University and University Hospital of Pisa, Pisa, Italy. [68]Cancer Genetics Laboratory, Research Division, Peter MacCallum Cancer Centre, Melbourne, Victoria, Australia. [69]Department of Pathology, University of Melbourne, Parkville, Victoria, Australia. [70]Cancer Pathology and Prevention, Division of Cancer Prevention and

Population Sciences, Roswell Park Cancer Institute, Buffalo, New York, USA. [71]Department of Obstetrics and Gynecology, John A. Burns School of Medicine, University of Hawaii, Honolulu, Hawaii, USA. [72]Cancer Risk Program, University of California, San Francisco, San Francisco, California, USA. [73]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany. [74]University Cancer Center Hamburg (UCCH), University Medical Center Hamburg-Eppendorf, Hamburg, Germany. [75]Centre for Cancer Research, Westmead Institute for Medical Research, University of Sydney, Sydney, New South Wales, Australia. [76]Department of Gynaecological Oncology, Westmead Hospital, Sydney, New South Wales, Australia. [77]Unité de Recherche en Santé des Populations, Centre des Maladies du Sein Deschênes-Fabia, Centre de Recherche FRSQ du Centre Hospitalier Affilié Universitaire de Québec, Quebec City, Quebec, Canada. [78]Departments of Pediatrics and Medicine, Columbia University, New York, New York, USA. [79]Center for Medical Genetics, Ghent University, Ghent, Belgium. [80]Division of Epidemiology, Biostatistics and Preventative Medicine, Department of Internal Medicine, University of New Mexico, Albuquerque, New Mexico, USA. [81]Sheffield Clinical Genetics Service, Sheffield Children's Hospital, Sheffield, UK. [82]Obstetrics and Gynecology Epidemiology Center, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA. [83]Division of Experimental Pathology, Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota, USA. [84]Social & Scientific Systems, Inc., Durham, North Carolina, USA. [85]Department of Clinical Genetics, Fox Chase Cancer Center, Philadelphia, Pennsylvania, USA. [86]Department of Genetics and Fundamental Medicine, Bashkir State University, Ufa, Russian Federation. [87]Department of Pathology and Laboratory Diagnostics, Maria Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland. [88]Gynecology Service, Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, New York, USA. [89]Department of Clinical Genetics, South Glasgow University Hospitals, Glasgow, UK. [90]Unité d'Oncogénétique, ICO–Centre René Gauducheau, Nantes, France. [91]Oncogenetics Group, Vall d'Hebron Institute of Oncology, Barcelona, Spain. [92]Clinical and Molecular Genetics Area, Vall d'Hebron University Hospital. Barcelona, Spain. [93]Department of Population Sciences, Beckman Research Institute of City of Hope, Duarte, California, USA. [94]Department of Epidemiology, Geisel School of Medicine at Dartmouth, Hanover, New Hampshire, USA. [95]Cancer Genetics Laboratory, Department of Genetics, University of Pretoria, Pretoria, South Africa. [96]Gynaecology Research Unit, Hannover Medical School, Hannover, Germany. [97]Nutrition and Metabolism Section, International Agency for Research on Cancer (IARC-WHO), Lyon, France. [98]Institute of Biology and Molecular Genetics, Universidad de Valladolid (IBGM-UVA), Valladolid, Spain. [99]Department of Gynecology, Jena University Hospital–Friedrich Schiller University, Jena, Germany. [100]Institute of Human Genetics, Münster, Germany. [101]University of Southampton Faculty of Medicine, Southampton University Hospitals NHS Trust, Southampton, UK. [102]Oncogenetics Team, Institute of Cancer Research and Royal Marsden NHS Foundation Trust, Sutton, UK. [103]Department of Oncology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark. [104]Institute of Human Genetics, Friedrich Alexander University Erlangen-Nuremberg, Erlangen, Germany. [105]Department of Biostatistics and Bioinformatics, Roswell Park Institute, Buffalo, New York, USA. [106]Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Leipzig, Germany. [107]Genomic Medicine, Manchester Academic Health Sciences Centre, University of Manchester, Central Manchester University Hospitals NHS Foundation Trust, Manchester, UK. [108]Division of Hematology and Oncology, Department of Medicine, UCLA David Geffen School of Medicine, Los Angeles, California, USA. [109]Laboratoire de Génétique Chromosomique, Hôtel Dieu Centre Hospitalier, Chambéry, France. [110]Department of Surgery and Cancer, Imperial College London, London, UK. [111]Molecular Diagnostics Laboratory, INRASTES, National Centre for Scientific Research 'Demokritos', Aghia Paraskevi Attikis, Athens, Greece. [112]Program in Cancer Genetics, Departments of Human Genetics and Oncology, McGill University, Montreal, Quebec, Canada. [113]Department of Medical Oncology, Papageorgiou, Hospital, Aristotle University of Thessaloniki School of Medicine, Thessaloniki, Greece. [114]Biostatistics and Informatics Shared Resource, University of Kansas Medical Center, Kansas City, Kansas, USA. [115]Harvard T.H. Chan School of Public Health and Dana-Farber Cancer Institute, Boston, Massachusetts, USA. [116]Susanne Levy Gertner Oncogenetics Unit, Institute of Human Genetics, Chaim Sheba Medical Center, Ramat Gan, Israel. [117]Sackler Faculty of Medicine, Tel Aviv University, Ramat Aviv, Israel. [118]Division of Cancer Prevention and Control Research, Jonsson Comprehensive Cancer Center, UCLA Schools of Medicine and Public Health, Los Angeles, California, USA. [119]Cancer Risk and Prevention Clinic, Dana-Farber Cancer Institute, Boston, Massachusetts, USA. [120]Centre of Familial Breast and Ovarian Cancer, Department of Medical Genetics, Institute of Human Genetics, University Würzburg, Würzburg, Germany. [121]Women's Cancer, Institute for Women's Health, University College London, London, UK. [122]Department of Clinical Genetics, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark. [123]Centre for Epidemiology and Biostatistics, School of Population and Global Health, University of Melbourne, Melbourne, Victoria, Australia. [124]Department of Epidemiology and Preventive Medicine, Monash University, Melbourne, Victoria, Australia. [125]Beatson West of Scotland Cancer Centre, Glasgow, UK. [126]Ontario Cancer Genetics Network: Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Ontario, Canada. [127]Department of Pathology and Laboratory Medicine, University of Kansas Medical Center, Kansas City, Kansas, USA. [128]Department of Dermatology, University of Utah School of Medicine, Salt Lake City, Utah, USA. [129]Department of Medicine, Royal Marsden Hospital, London, UK. [130]Clinical Genetics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, US National Institutes of Health, Bethesda, Maryland, USA. [131]Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland. [132]Keck School of Medicine and Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, California, USA. [133]Center for Familial Breast and Ovarian Cancer, Center for Integrated Oncology (CIO) and Center for Molecular Medicine Cologne (CMMC), University Hospital Cologne, Cologne, Germany. [134]Department of Preventive Medicine, Keck School of Medicine, University of Southern California Norris Comprehensive Cancer Center, Los Angeles, California, USA. [135]Division of Nutritional Epidemiology, Department of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. [136]Molecular Genetics of Breast Cancer, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany. [137]Center for Genomic Medicine, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark. [138]Institute of Genetic Medicine, Centre for Life, Newcastle-upon-Tyne Hospitals NHS Trust, Newcastle-upon-Tyne, UK. [139]Department of Epidemiology, University of Texas MD Anderson Cancer Center, Houston, Texas, USA. [140]Clinics of Obstetrics and Gynaecology, Hannover Medical School, Hannover, Germany. [141]Medical Genetics Unit, St George's University of London, London, UK. [142]Department of Gynecology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark. [143]Department of Virus, Lifestyle and Genes, Danish Cancer Society Research Center, Copenhagen, Denmark. [144]Molecular Unit, Department of Pathology, Herlev Hospital, University of Copenhagen, Copenhagen, Denmark. [145]Family Cancer Clinic, Netherlands Cancer Institute, Amsterdam, the Netherlands. [146]Department of Medical Oncology, Family Cancer Clinic, Erasmus University Medical Center, Rotterdam, the Netherlands. [147]Center for Immunotherapy, Roswell Park Cancer Institute, Buffalo, New York, USA. [148]Center for Medical Genetics, NorthShore University HealthSystem, University of Chicago Pritzker School of Medicine, Evanston, Illinois, USA. [149]Program in Genetic Epidemiology and Statistical Genetics, Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA. [150]Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, British Columbia, Canada. [151]British Columbia Cancer Agency, Vancouver, British Columbia, Canada. [152]N.N. Petrov Institute of Oncology, St. Petersburg, Russian Federation. [153]Lombardi Comprehensive Cancer Center, Georgetown University, Washington, DC, USA. [154]Department of Statistical Science, Duke University, Durham, North Carolina, USA. [155]Clinical Genetics, Guy's and St Thomas' NHS Foundation Trust, London, UK. [156]Familial Cancer Centre, Peter MacCallum Cancer Centre, Melbourne, Victoria, Australia. [157]Sir Peter MacCallum Department of Oncology, University of Melbourne, Parkville, Victoria, Australia. [158]Department of Molecular and Regenerative Medicine, Hematology, Oncology and Transfusion Medicine Center, Vilnius University Hospital Santariskiu Clinics, Vilnius, Lithuania. [159]State Research Institute Centre for Innovative Medicine, Vilnius, Lithuania. [160]Department of Obstetrics and Gynecology, Lund University Hospital, Lund, Sweden. [161]Department of Clinical Genetics, Aarhus University Hospital, Aarhus, Denmark. [162]Department of Epidemiology, Cancer Prevention Institute of California, Fremont, California, USA. [163]Division of Genetics and Epidemiology, Institute of Cancer Research, London, UK. [164]Department of Gynaecology and Obstetrics, University Hospital Carl Gustav Carus, Technical University Dresden, Dresden, Germany. [165]Department of Obstetrics and Gynecology, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland. [166]Department of Surgery, Breast Care Center, Daerim St. Mary's Hospital, Seoul, Republic of Korea. [167]Department of Pathology, University of Calgary, Calgary, Alberta, Canada. [168]Department of Clinical Genetics, Odense University Hospital, Odense, Denmark. [169]Hong Kong Hereditary Breast Cancer Family Registry, Hong Kong. [170]Department of Surgery, University of Hong Kong, Hong Kong. [171]Cancer Genetics Center and Department of Surgery, Hong Kong Sanatorium and Hospital, Hong Kong. [172]Vesalius Research Center, VIB, Leuven, Belgium. [173]Laboratory for Translational Genetics, Department of Oncology, KU Leuven, Leuven, Belgium. [174]Public Health Division of Gipuzkoa, Regional Government of the Basque Country, Donostia–San Sebastian, Spain. [175]CIBER of Epidemiology and Public Health, Madrid, Spain. [176]Molecular Diagnostic Unit, Hereditary Cancer Program, IDIBELL–Catalan Institute of Oncology, Barcelona, Spain. [177]Cancer Control Research, BC Cancer Agency, Vancouver, British Columbia, Canada. [178]Cancer Epidemiology Program, University of Hawaii Cancer Center, Honolulu, Hawaii, USA. [179]Department of Surgery, Ulsan College of Medicine and Asan Medical Center, Seoul, Republic of Korea. [180]Department of Gynecological Oncology, Roswell Park Cancer Institute, Buffalo, New York, USA. [181]Département de Génétique, Centre Hospitalier Universitaire de Grenoble, Grenoble, France. [182]Institut Curie, PSL Research Unviersity and INSERM U900, Paris, France. [183]Mines Paris Tech, Fontainebleau, France. [184]College of Pharmacy and Health Sciences, Texas Southern University, Houston, Texas, USA. [185]Cancer Center Wolfsburg, Clinics of Gynaecology, Wolfsburg, Germany. [186]Department of Health Sciences Research, Division of Epidemiology, Mayo Clinic, Rochester, Minnesota, USA. [187]Department of Cancer Epidemiology and Prevention, Maria Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland.

[188]Department of Gynecologic Oncology, University of Texas MD Anderson Cancer Center, Houston, Texas, USA. [189]Juliane Marie Centre, Department of Gynecology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark. [190]INSERM U1052, CNRS UMR 5286, Université Lyon 1, Centre de Recherche en Cancérologie de Lyon, Centre Léon Bérard, Lyon, France. [191]Ovarian Cancer Research (OVCARE) Program–Gynecologic Tissue Bank, Vancouver General Hospital and BC Cancer Agency, Vancouver, British Columbia, Canada. [192]Department of Health Research and Policy–Epidemiology, Stanford University School of Medicine, Stanford, California, USA. [193]Public Health Ontario, Toronto, Ontario, Canada. [194]Institute of Cancer Sciences, University of Glasgow, Wolfson Wohl Cancer Research Centre, Beatson Institute for Cancer Research, Glasgow, UK. [195]Department of Clinical Genetics, VU University Medical Centre, Amsterdam, the Netherlands. [196]Department of Gynaecology and Obstetrics, Division of Tumor Genetics, Klinikum Rechts der Isar, Technical University of Munich, Munich, Germany. [197]Department of Human Genetics, Radboud University Medical Centre, Nijmegen, the Netherlands. [198]Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK. [199]Division of Gynecologic Oncology, Department of Obstetrics, Gynecology and Reproductive Sciences, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA. [200]Department of Epidemiology, University of Pittsburgh Graduate School of Public Health, Pittsburgh, Pennsylvania, USA. [201]Ovarian Cancer Center of Excellence, Womens Cancer Research Program, Magee-Womens Research Institute and University of Pittsburgh Cancer Institute, Pittsburgh, Pennsylvania, USA. [202]Department of Obstetrics and Gynecology, Oregon Health & Science University, Portland, Oregon, USA. [203]Immunology and Molecular Oncology Unit, Veneto Institute of Oncology IOV–IRCCS, Padua, Italy. [204]Department of Laboratory Medicine and Keenan Research Centre of the Li Ka Shing Knowledge Institute, St Michael's Hospital, Toronto, Ontario, Canada. [205]Clinical Genetics Research Laboratory, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, New York, USA. [206]Department of Pathology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark. [207]University of Texas School of Public Health, Houston, Texas, USA. [208]Department of Gynaecology and Obstetrics, University Hospital Düsseldorf, Heinrich Heine University Düsseldorf, Düsseldorf, Germany. [209]Invitae Corporation, San Francisco, California, USA. [210]Department of Medicine, University of California, San Francisco, San Francisco, California, USA. [211]Department of Molecular Genetics, National Institute of Oncology, Budapest, Hungary. [212]Center for Clinical Cancer Genetics and Global Health, University of Chicago Medical Center, Chicago, Illinois, USA. [213]Department of Oncology, Lund University Hospital, Lund, Sweden. [214]Department of Obstetrics and Gynecology, Ohio State University and the James Cancer Center, Columbus, Ohio, USA. [215]West Midlands Regional Genetics Service, Birmingham Women's Hospital Healthcare NHS Trust, Edgbaston, Birmingham, UK. [216]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands. [217]Breast Cancer Now Toby Robins Research Centre, Institute of Cancer Research, London, UK. [218]Molecular and Nutritional Epidemiology Unit, Cancer Research and Prevention Institute–ISPO, Florence, Italy. [219]Unit of Medical Genetics, Department of Biomedical, Experimental and Clinical Sciences, University of Florence, Florence, Italy. [220]Cancer Research UK Clinical Trials Unit, Institute of Cancer Sciences, University of Glasgow, Glasgow, UK. [221]Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, Michigan, USA. [222]Section of Molecular Diagnostics, Clinical Biochemistry, Aalborg University Hospital, Aalborg, Denmark. [223]Department of Genetics, Portuguese Oncology Institute, Porto, Portugal. [224]IFOM, FIRC (Italian Foundation for Cancer Research) Institute of Molecular Oncology, Milan, Italy. [225]Department of Obstetrics and Gynecology, Medical University of Vienna and Comprehensive Cancer Center, Vienna, Austria. [226]Division of Cancer Medicine, Peter MacCallum Cancer Centre, East Melbourne, Victoria, Australia. [227]NRG Oncology, Statistics and Data Management Center, Roswell Park Cancer Institute, Buffalo, New York, USA. [228]Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, New York, USA. [229]Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA. [230]South East of Scotland Regional Genetics Service, Western General Hospital, Edinburgh, UK. [231]Service de Génétique Clinique Chromosomique et Moléculaire, Centre Hospitalier Universitaire de St Etienne, St Etienne, France. [232]Translational Research Laboratory, IDIBELL (Bellvitge Biomedical Research Institute), Catalan Institute of Oncology, Barcelona, Spain. [233]Unité d'Oncogénétique, Centre Hospitalier Universitaire Arnaud de Villeneuve, Montpellier, France. [234]Unit of Molecular Bases of Genetic Risk and Genetic Testing, Department of Preventive and Predictive Medicine, Fondazione IRCCS (Istituto di Ricovero e Cura a Carattere Scientifico) Istituto Nazionale Tumori (INT), Milan, Italy. [235]Department of Clinical Genetics, Karolinska University Hospital, Stockholm, Sweden. [236]Clalit National Israeli Cancer Control Center, Haifa, Israel. [237]Department of Community Medicine and Epidemiology, Carmel Medical Center and B. Rappaport Faculty of Medicine, Haifa, Israel. [238]Department of Pathology, Brigham and Women's Hospital, Dana-Farber Cancer Institute, Boston, Massachusetts, USA. [239]Clinical Genetics Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, New York, USA. [240]Division of Gynecologic Oncology, NorthShore University HealthSystem, University of Chicago, Evanston, Illinois, USA. [241]Hereditary Endocrine Cancer Group, Spanish National Cancer Research Center (CNIO), Madrid, Spain. [242]Department of Epidemiology, Netherlands Cancer Institute, Amsterdam, the Netherlands. [243]Program in Epidemiology, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA. [244]Epidemiology Branch, Division of Intramural Research, National Institute of Environmental Health Sciences, US National Institutes of Health, Research Triangle Park, North Carolina, USA. [245]Clinical Cancer Genetics Program, Division of Human Genetics, Department of Internal Medicine, Comprehensive Cancer Center, Ohio State University, Columbus, Ohio, USA. [246]Facultés de Médecin, Université Paris–Sud, Villejuif, France. [247]Institut Gustave Roussy, Villejuif, France. [248]Human Genetics Foundation (HuGeF), Turin, Italy. [249]Facultés de Médecin, Université Paris–Saclay, Paris, France. [250]Department of Hematology and Oncology, University of Kansas Medical Center, Kansas City, Kansas, USA. [251]Department of Gynaecological Oncology, Glasgow Royal Infirmary, Glasgow, UK. [252]North East Thames Regional Genetics Service, Great Ormond Street Hospital for Children NHS Trust, London, UK. [253]Département Oncologie Génétique, Institut Paoli-Calmettes, Marseille, France. [254]Genetic Epidemiology Laboratory, Department of Pathology, University of Melbourne, Parkville, Victoria, Australia. [255]Institute of Human Genetics, Hannover Medical School, Hannover, Germany. [256]Department of Cancer Prevention and Control, Roswell Park Cancer Institute, Buffalo, New York, USA. [257]Epidemiology Center, College of Medicine, University of South Florida, Tampa, Florida, USA. [258]Institute of Human Genetics, Department of Human Genetics, University Hospital Heidelberg, Heidelberg, Germany. [259]Division of Breast Cancer Research, Institute of Cancer Research, London, UK. [260]National Human Genome Research Institute, US National Institutes of Health, Bethesda, Maryland, USA. [261]Biomedical Sciences Institute (ICBAS), University of Porto, Porto, Portugal. [262]Cancer Research Initiatives Foundation, Sime Darby Medical Centre, Subang Jaya, Malaysia. [263]University Malaya Cancer Research Institute, Faculty of Medicine, University Malaya Medical Centre, University Malaya, Kuala Lumpur, Malaysia. [264]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA. [265]Cancer Prevention and Control, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA. [266]Department of Medicine Magee-Womens Hospital of UPMC, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA. [267]Latvian Biomedical Research and Study Centre, Riga, Latvia. [268]Ovarian Cancer Research (OVCARE) Program–Cheryl Brown Ovarian Cancer Outcomes Unit (CBOCOU), BC Cancer Agency, Vancouver, British Columbia, Canada. [269]Department of Medical Genetics, Addenbrooke's Treatment Centre, Addenbrooke's Hospital, Cambridge, UK. [270]Division of Human Cancer Genetics, Departments of Internal Medicine and Molecular Virology, Immunology and Medical Genetics, Comprehensive Cancer Center, Ohio State University, Columbus, Ohio, USA. [271]Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK. [272]Hellenic Health Foundation, Athens, Greece. [273]WHO Collaborating Center for Nutrition and Health, Unit of Nutritional Epidemiology and Nutrition in Public Health, Department of Hygiene, Epidemiology and Medical Statistics, University of Athens Medical School, Athens, Greece. [274]Department of Medical Oncology, Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA. [275]Division of Gynecologic Oncology, Department of Obstetrics and Gynaecology, Radboud University Medical Centre, Nijmegen, the Netherlands. [276]Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands. [277]Department of Medical Genetics, University Medical Center Utrecht, Utrecht, the Netherlands. [278]Division of Gynecologic Oncology, Department of Obstetrics and Gynaecology and Leuven Cancer Institute, University Hospitals Leuven, Leuven, Belgium. [279]Institute of Medical and Human Genetics, Charité Universitätsmedizin, Campus Virchov Klinikum, Berlin, Germany. [280]Fundación Pública Galega de Medicina Xenómica, Servizo Galego de Saúde (SERGAS), Instituto de Investigación Sanitaria (IDIS), Santiago de Compostela, Spain. [281]Grupo de Medicina Xenómica, Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Universidade de Santiago de Compostela (USC), Santiago de Compostela, Spain. [282]Vanderbilt Genetics Institute and Department of Obstetrics and Gynecology, Vanderbilt University Medical Center, Nashville, Tennessee, USA. [283]Oxford Regional Genetics Service, Churchill Hospital, Oxford, UK. [284]Institute of Human Genetics, University Hospital, Leipzig, Germany. [285]Department of Obstetrics and Gynecology, University of Ulm, Ulm, Germany. [286]Population Health Department, QIMR Berghofer Medical Research Institute, Herston, Queensland, Australia. [287]Biostatistics and Computational Biology Branch, Division of Intramural Research, National Institute of Environmental Health Sciences, US National Institutes of Health, Research Triangle Park, North Carolina, USA. [288]Clinical Cancer Genetics, City of Hope, Duarte, California, USA. [289]Department of Data Management Science, Stanford University School of Medicine, Stanford, California, USA. [290]Departments of Human Genetics and Clinical Genetics, Leiden University Medical Center, Leiden, the Netherlands. [291]Women's College Research Institute, University of Toronto, Toronto, Ontario, Canada. [292]Center for Genomic Medicine, Department of Biomedical Data Science, Geisel School of Medicine at Dartmouth, Lebanon, New Hampshire, USA. [293]Department of Public Health Sciences, University of Virginia, Charlottesville, Virginia, USA. [294]School of Women's and Children's Health, Lowy Cancer Research Centre, University of New South Wales, Sydney, New South Wales, Australia. [295]Kinghorn Cancer Centre, Garvan Institute of

Medical Research, Darlinghurst, New South Wales, Australia. [296]Department of Molecular Medicine, University La Sapienza, Rome, Italy. [297]Community and Population Health Research Institute, Department of Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, California, USA. [298]Department of Preventive Medicine, College of Medicine, Seoul National University, Seoul, Republic of Korea. [299]Seoul National University Cancer Research Institute, Seoul, Republic of Korea. [300]Department of Biomedical Science, Graduate School, Seoul National University, Seoul, Republic of Korea. [301]Department of Public Health Sciences, College of Medicine, Medical University of South Carolina, Charleston, South Carolina, USA. [302]Department of Chronic Disease Epidemiology, Yale School of Public Health, New Haven, Connecticut, USA. [303]Department of Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, California, USA. [304]These authors contributed equally to this work. [305]These authors jointly directed this work. Correspondence should be addressed to P.D.P.P. (pp10001@medschl.cam.ac.uk).

ONLINE METHODS

Study samples. Genotype data from six OCAC and two CIMBA genotyping projects were used for these analyses (**Supplementary Table 1**). All participating studies were approved by the relevant research ethics committee, and all participants provided written, informed consent.

OCAC. The OCAC OncoArray data comprised 63 genotyping project/case–control sets (**Supplementary Table 1**). Some studies (for example, SEARCH) contributed samples to more than one genotyping project, and some case–control sets are a combination of multiple individual studies. The following numbers are for the subjects of European ancestry that passed quality control. The analyses included 66,450 samples from seven genotyping projects: 40,941 controls, 22,406 invasive cases and 3,103 borderline cases. The number of cases by histotype was as follows: serous borderline (1,954), mucinous borderline (1,149), LGSOC (1,012), HGSOC (13,037), ENOC (2,810), CCOC (1,366) and other EOC (2,764).

CIMBA. Eligibility in CIMBA is restricted to females aged 18 years or older with pathogenic variants in *BRCA1* or *BRCA2*. The majority of the participants were sampled through cancer genetics clinics, including some related participants. Sixty-three studies contributed OncoArray and iCOGS genotype data on 31,448 mutation carriers. For the samples genotyped on the OncoArray, after quality control, data were available on 15,694 *BRCA1* mutation carriers and 10,988 *BRCA2* mutation carriers, of whom 2,372 and 849, respectively, were affected with EOC (**Supplementary Table 2**). We also obtained genotype data on 3,342 (561 affected) *BRCA1* and 1,424 (105 affected) *BRCA2* non-overlapping samples genotyped using the iCOGS array[19,39].

Genotype data and quality control. Data from all the genotyping projects apart from the OCAC and CIMBA OncoArray projects have been published previously[8,9,14,15,19]. Genotypes for OCAC samples were preferentially selected from the different projects in the following order: OncoArray, Mayo GWAS, COGS and other GWAS. Genotyping was performed at five centers: University of Cambridge, Center for Inherited Disease Research (CIDR), National Cancer Institute (NCI), Génome Québec and Mayo Clinic (**Supplementary Table 15**). OncoArray sample quality control was similar to that carried out for the other projects (as described in ref. 14). We excluded samples if they had a genotyping call rate <95%, if they had excessively low or high heterozygosity, if they were not female or if they were duplicates (cryptic or intended). Duplicates and close relatives were identified using in-house software that calculates a concordance matrix for all individuals. Samples with concordance >0.86 were flagged as duplicates, and samples with concordance between 0.74 and 0.86 were flagged as relatives. The comparison was performed among all the OncoArray samples and among all the previously genotyped samples. The concordance statistics were used to identify cryptic duplicates and expected duplicates whose genotypes did not match. We attempted to resolve these with the study investigators. If the discrepancy could not be resolved, both samples were excluded. In OCAC, for confirmed cryptic duplicates and relatives, we retained one sample in the analysis. For case–control pairs, we excluded the control, while for case–case and control–control pairs we excluded the sample with the lower call rate. In CIMBA, relatives were included in the analysis and the association tests were adjusted accordingly. For confirmed duplicates, the sample with the higher call rate was retained. SNP quality control was carried out according to the OncoArray QC Guidelines[18]. Only SNPs that passed quality control for all consortia were used for imputation. We excluded SNPs with a call rate <95%, SNPs deviating from Hardy–Weinberg equilibrium ($P < 1 \times 10^{-7}$ in controls or unrelated samples in CIMBA and $P < 1 \times 10^{-12}$ in cases) and SNPs with concordance <98% among 5,280 duplicate pairs. For imputation, we additionally excluded SNPs with MAF <1% and call rate <98%, SNPs that could not be linked to the 1000 Genomes Project reference or differed significantly in frequency from the 1000 Genomes Project (European frequency) and a further 1,128 SNPs where the cluster plot was judged to be inadequate. Of the 533,631 SNPs that were manufactured on the array, 494,813 SNPs passed the initial quality control and 470,825 SNPs were used for imputation. Samples with overall heterozygosity <5% or >40% were excluded.

Ancestry analysis. Intercontinental ancestry was calculated for the OCAC samples using the software package FastPop (http://sourceforge.net/projects/fastpop/)[40] developed specifically for the OncoArray. Only samples with >80% European ancestry were used for these analyses. For the CIMBA samples, 33,661 weakly correlated autosomal SNPs (pairwise $r^2 < 0.1$) were used to compute the genomic kinship between all pairs of individuals, along with 267 HapMap samples (CHB, JPT, YRI and CEU). These were converted to distances and subjected to multidimensional scaling. Using the first two components, we calculated the proportion of European ancestry for each individual and excluded samples with >27% non-European ancestry to ensure that samples of Ashkenazi Jewish ancestry were included in the final sample. Analysis using FastPop provided virtually identical results.

Principal-component analysis. Principal-component analysis for the OncoArray data was carried out using data from 33,661 uncorrelated SNPs (pairwise $r^2 < 0.1$) with MAF >0.05 using an in-house program (available at http://ccge.medschl.cam.ac.uk/software/pccalc/). Principal-component analysis for the other genotype data sets was carried out as previously described[14,19].

Imputation. We performed imputation separately for each genotyping project data set. We imputed genotypes into the reference panel from the 1000 Genomes Project (v3 October 2014)[20]. We initially used an efficient two-step procedure, which involved prephasing using SHAPEIT[41] followed by imputation of the phased data using IMPUTE2 (ref. 42). We then performed more accurate imputation for any region with a SNP with $P < 1 \times 10^{-6}$ in the OCAC analyses or the OCAC and CIMBA meta-analysis. The boundaries were set ±500 kb from the most significant SNP in each region. The single-step imputation used IMPUTE2 without prephasing with some of the default parameters modified. These included an increase of the MCMC iterations to 90 (of which the first 15 were used as burn-in), an increase of the buffer region to 500 kb and an increase to 100 of the number of haplotypes used as templates when phasing observed genotypes.

After imputation, 85% of common variants, including both single-nucleotide variants and small indels (MAF > 0.05), had imputation r^2 imputation accuracy >0.9, with 97% having imputation r^2 >0.7. Of the rare variants (0.001 < MAF < 0.05), 28% had imputation r^2 >0.9 and 58% had imputation r^2 >0.7.

Association analyses in the unselected ovarian cancer cases and controls from OCAC and CIMBA. We excluded SNPs from the association analysis if their r^2 imputation accuracy was <0.3 or MAF was <0.01. In total, genotypes for 11,595,112 million variants were available for analysis.

Association analyses in OCAC. We evaluated the association between genotype and disease using the imputed genotype dosage in a logistic regression model. We carried out initial genome-wide analyses separately for OncoArray, COGS and the five GWAS data sets and pooled the results using a fixed-effects meta-analysis. The analyses were adjusted for study and for population substructure by including the eigenvectors of project-specific principal components as covariates in the model (nine for OncoArray, five for COGS, two for UK GWAS, two for the US, BWH and POL GWAS, and a single principal component for the MAY GWAS). The number of eigenvectors chosen was based on the inflection point of a scree plot. After one-step imputation of the genotypes in the regions of interest, we used genotype dosages in a single logistic regression model with adjustment for each genotyping project/study combination and 19 principal components. Principal components were set to zero for samples not included in a given project. We used custom written software for the analysis.

EOC is a heterogeneous phenotype with five major histotypes for invasive disease (HGSOC, LGSOC, MOC, ENOC and CCOC) and two histotypes of borderline disease (serous and mucinous). The pattern of association across the different histotypes varies for the known EOC risk loci. We therefore carried out the association analysis on the following nine histotypes: all invasive disease; HGSOC; LGSOC; all invasive serous; serous borderline; LGSOC and borderline serous combined; ENOC; CCOC; and mucinous invasive/mucinous borderline combined.

Association analyses in CIMBA. We carried out the ovarian cancer association analyses separately for *BRCA1* and *BRCA2* mutation carriers and for OncoArray and COGS samples. The results were pooled using fixed-effects

meta-analysis. 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 study recruitment.

Breast cancer diagnosis was not considered as a censoring event. To account for the nonrandom sampling of *BRCA1* and *BRCA2* mutation carriers with respect to disease status, we conducted the analyses by modeling the retrospective likelihood of the observed genotypes conditional on the disease phenotype. We assessed the associations between genotype and risk of ovarian cancer using a score test statistic based on the retrospective likelihood[22]. 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[43]. 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 and for Ashkenazi Jewish origin.

We carried out the analyses using custom written functions in Fortran and Python. The score test statistic was implemented in R version 3.0.1.

OCAC and CIMBA meta-analysis. We conducted a meta-analysis of the EOC associations in *BRCA1* and *BRCA2* mutation carriers and in OCAC samples 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 *BRCA1* and *BRCA2* mutation carriers and the logarithm of the per-allele odds ratio estimate for the association with EOC (any subtype) and serous EOC in OCAC. The number of *BRCA1* and *BRCA2* mutation samples with tumor histology information was too small to allow for subgroup analyses. However, previous studies have demonstrated that the vast majority of EOCs in *BRCA1* and *BRCA2* mutation carriers are HGSOC. Meta-analyses were carried out using the software METAL (2011-03-25 release)[44].

We evaluated whether there was 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 *BRCA1* mutation carriers, *BRCA2* mutation carriers and OCAC.

Candidate causal SNPs in each susceptibility region. To identify a set of variants most likely to mediate the observed association (the credible causal variants), we excluded SNPs with causality odds of less than 1:100 by comparing the likelihood of each SNP from the association analysis with the likelihood of the most strongly associated SNP.

Functional annotation of risk-associated variants. *Expression and methylation quantitative trait locus analyses.* A TCGA data set[45] was available for 326 HGSOC tumors in women of European ancestry. Ancestry was estimated using the Local Ancestry in adMixed Populations (LAMP[46]) software package, and individuals with >95% European descent were retained for further analyses. Matched gene expression (measured on the Agilent 1M microarray), CpG methylation (measured on the Illumina Infinium HumanMethylation27 BeadChip), copy number alteration (called using the Affymetrix SNP 6.0 array) and germline genotype (called using the Affymetrix SNP 6.0 array) data were also available. A Mayo Clinic data set was available for 209 serous EOC tumors and 32 ENOC tumors in women of European ancestry. Matched gene expression (measured on the Agilent Technologies Whole Human Genome Microarray, 4x44K), CpG methylation (measured on the Illumina Infinium HumanMethylation450 BeadChip), copy number alteration (called using the OncoArray) and germline genotype (called using the OncoArray) data were available for all of these samples. Genotypes were imputed into the 1000 Genomes Project October 2014 (Phase 3, version 5)[20] European reference panel for both data sets. eQTL analyses were performed using linear regression as implemented in the R package Matrix eQTL[47]. Only variants with imputation accuracy r^2 >0.3 were analyzed. Prior to eQTL analyses, the effects of tumor copy number and methylation on gene expression were regressed out as previously described[24]. For the Mayo data set, we performed separate analyses on the HGSOC and ENOC samples. Results for the two HGSOC data sets were combined in a random-effects meta-analysis. We focused on local or *cis*-acting eQTLs between SNPs in the 1:100 list of potentially causal variants and all genes up to 1 Mb away on either side of these SNPs.

mQTL analyses for the 1:100 potentially causal SNPs in regions of interest (within 1 Mb on either side of the index SNP) were conducted using the Mayo data set only because methylation was assayed with the much denser Illumina Infinium HumanMethylation450 BeadChip and the Mayo sample included histologies other than HGSOC. Within each region, CpG probes were filtered on the basis of their association with gene expression. For each expression probe within the region, a linear model was fit by CpG probe adjusted for age and copy number variation (CNV) overlapping the expression probe. The CpG with the strongest negative test statistic for each gene (across multiple expression probes per gene) was retained for mQTL analysis to reduce the total number of tests. We performed VanderWaerden rank transformations of the beta values to account for skewed distributions in the values and conducted linear regression of the SNP genotypes on the transformed beta values, adjusted for age and CNVs overlapping the CpG probe; missing CNV values were imputed using the median for the non-missing samples within each region. As a sensitivity analysis, we also performed analyses adjusted only for age. Analyses were conducted for all histologies, as well as for the serous, HGSOC and ENOC subsets. Loci were eliminated from analyses where there were either no Agilent probes for the region on the array (9q31.1) or there were no negatively associated CpGs on the 450K array (8q21.11).

For eQTL and mQTL analyses, two-sided *P* values are reported.

Mapping risk SNPs to biofeatures. *Cell culture.* Cell lines were cultured in their respective media as follows: GTFR230, NOSE-CM (1:1 Medium 199: MCDB105 (both Sigma-Aldrich)) supplemented with 15% FBS (HyClone), 500 ng/ml hydrocortisone, 5 μg/ml insulin (both Sigma-Aldrich), 10 ng/ml epidermal growth factor and 34 μg protein/ml bovine pituitary extract (both Life Technologies); MCAS, EMEM supplemented with 15% FBS (Seradigm); RMG-II and JHOC5, RPMI plus 10% FBS and OAW42, DMEM containing 10% FBS, 20 μg/ml insulin and sodium pyruvate (Lonza). Cell lines were authenticated by profiling short tandem repeats using the Promega Powerplex 16HS Assay (performed at the University of Arizona Genetics Core facility), and all cultures tested negative for contaminating mycoplasma infections using mycoplasma-specific PCR. GTFR230 was established in house from a mucinous ovarian cancer; MCAS was a gift from M. Press (University of Southern California); RMG-II was a gift from H. Itamochi (Tottori University School of Medicine); OAW42 was a gift from G. Mills (MD Anderson); and JHOC5 was provided by the RIKEN repository (http://www2.brc.riken.jp/cache/cell/RCB1520).

Chromatin immunoprecipitation. Our ChIP protocol was based on the methods of Schmidt *et al.*[48]. Four 15-cm dishes of cells were fixed in formaldehyde for 10 min, before quenching the fixation with glycine. Cells were collected, lysed in a sarkosyl-containing lysis buffer and sonicated using the Covaris E220 evolution Focused Ultrasonicator to yield 100- to 300-bp genomic DNA fragments. Five micrograms of an antibody raised against H3K27ac (Diagenode, C15410196) was incubated with blocked magnetic Dynabeads (Life Technologies) at 4 °C for 4 h. Antibody–bead conjugates were incubated with 100 μg of chromatin at 4 °C overnight, with constant agitation. Beads were washed extensively with RIPA buffer and then treated with RNase and proteinase K (both Qiagen). DNA was then eluted from the beads in TE buffer and cleaned up using the QIAquick PCR Purification kit (Qiagen). Two independent immunoprecipitations and one input sample were sequenced for each cell line, and each sample was quality checked by site-specific qPCR before next-generation sequencing.

Next-generation sequencing. ChIP libraries were constructed using the Kapa Hyper Library Preparation kit, according to the manufacturer's instructions. Approximately two-thirds of the immunoprecipitated material was used as the starting amount. For undiluted input samples, 100–300 ng of starting material was used. Construction was carried out according to the manufacturer's instructions using Bio NextFLex adaptors diluted 1:50. Final PCR on a portion of the adaptor ligation was performed for 12 cycles. Products were evaluated on an Agilent Bioanalyzer, using high-sensitivity DNA chips. ChIP libraries were quantified using the Kapa Biosystems Illumina library quantification kit, and 12 were pooled for sequencing, which was carried out using single-end reads with 75 cycles on a NextSeq 500 instrument (with version 2 chemistry).

Analysis of ChIP–seq data. ChIP–seq data were processed using MACS2 with a *P*-value cutoff of 0.001. The smaller of input or signal was linearly scaled

doi:10.1038/ng.3826

to the same depth as the larger data set. To control the irreproducible discovery rate (IDR) in ChIP–seq analysis, we used the IDR version 2.0 pipeline[49]. A standard IDR threshold of $P < 0.05$ was applied.

Functional annotation of variants. We used shell scripts with BEDtools (http://bedtools.readthedocs.org/en/latest/) to generate overlap data between all variants in each associated region including likely causal SNPs and BED file versions of all the data represented in **Figure 2** and **Supplementary Table 6**. In addition, we included 3′ UTRs, 5′ UTRs, miRcode high-confidence conserved microRNA target sites, high-confidence microRNA target sites from http://microRNA.org/ and all coding exons. The overlap data thus obtained were converted to matrix form by means of Python scripts. MicroRNA target sites were only considered if they overlapped UTR gene regions. Exonic variants were further assessed for missense or nonsense mutations by MuTect software[50]. The NHGRI-EBI GWAS catalog was used to identify SNPs among the potentially causal set with other genome-wide significant associations (**Supplementary Table 14**).

Locus-specific tissue enrichment of variants. H3K27 acetylation peaks were collated from public sources (for HeLa-S3, HCT116, UCSD ovary, UCSD sigmoid colon and colon crypt) or from in-house data (IOSE4, IOSE11, FT33, FT246, EEC16, CaOV3, UWB1.289, OAW42, GFTR230 and MCAS) (**Supplementary Tables 5** and **16**). Overlaps were counted for all SNPs against which genotypes were imputed in the 1000 Genomes Project for each H3K27ac data set. The fraction of causal SNPs with overlaps was then tested for significance against this background for each cell type in the H3K27ac data sets using the hypergeometric distribution. Finally, *P* values were adjusted for multiple comparisons using Bonferroni's method.

Data availability. OncoArray germline genotype data for the OCAC studies have been deposited at the European Genome-phenome Archive (EGA; https://ega-archive.org/), which is hosted by the EBI and the CRG, under accession EGAS00001002305. Summary results are available from the Ovarian Cancer Association Consortium (http://ocac.ccge.medschl.cam.ac.uk/). A subset of the OncoArray germline genotype data for the CIMBA studies will be made publically available through the database of Genotypes and Phenotypes (dbGaP) under accession phs001321.v1.p1. The complete data set will not be made publically available because of restraints imposed by the ethics committees of individual studies; requests for further data can be made to the Data Access Coordination Committee (http://cimba.ccge.medschl.cam.ac.uk/). ChIP–seq data are available from the Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/) under accession GSE68104.

39. Gaudet, M.M. *et al.* Identification of a *BRCA2*-specific modifier locus at 6p24 related to breast cancer risk. *PLoS Genet.* **9**, e1003173 (2013).
40. Li, Y. *et al.* FastPop: a rapid principal component derived method to infer intercontinental ancestry using genetic data. *BMC Bioinformatics* **17**, 122 (2016).
41. Delaneau, O., Coulonges, C. & Zagury, J.F. Shape-IT: new rapid and accurate algorithm for haplotype inference. *BMC Bioinformatics* **9**, 540 (2008).
42. van Leeuwen, E.M. *et al.* Population-specific genotype imputations using minimac or IMPUTE2. *Nat. Protoc.* **10**, 1285–1296 (2015).
43. Antoniou, A.C. *et al.* A locus on 19p13 modifies risk of breast cancer in *BRCA1* mutation carriers and is associated with hormone receptor–negative breast cancer in the general population. *Nat. Genet.* **42**, 885–892 (2010).
44. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
45. Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature* **474**, 609–615 (2011).
46. Sankararaman, S., Sridhar, S., Kimmel, G. & Halperin, E. Estimating local ancestry in admixed populations. *Am. J. Hum. Genet.* **82**, 290–303 (2008).
47. Shabalin, A.A. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics* **28**, 1353–1358 (2012).
48. Schmidt, D. *et al.* ChIP–seq: using high-throughput sequencing to discover protein–DNA interactions. *Methods* **48**, 240–248 (2009).
49. Bickel, P.J., Brown, J.B., Huang, H. & Li, Q. An overview of recent developments in genomics and associated statistical methods. *Philos. Trans. A Math. Phys. Eng. Sci.* **367**, 4313–4337 (2009).
50. Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013).