The prospect of discovering new biomarkers for ovarian cancer based on current knowledge of susceptibility loci and genetic variation (Review)

MIKAEL KRONBORG CHRISTOPHERSEN[1], CLAUS HØGDALL[2] and ESTRID HØGDALL[1]

[1]Molecular Unit, Department of Pathology, Herlev and Gentofte Hospital, Copenhagen University Hospital, 2730 Herlev; [2]The Juliane Marie Centre, Department of Gynaecology, Rigshospitalet, Copenhagen University Hospital, 2100 Copenhagen, Denmark

Received May 31, 2019; Accepted July 30, 2019

DOI: 10.3892/ijmm.2019.4352

Abstract. Ovarian cancer is the most lethal gynaecological malignancy. The cancer initially presents with non-specific symptoms; thus, it is typically not discovered until the patient has reached the late, considerably more lethal, stages of the disease. Research focus is currently on finding novel biomarkers, especially for early detection and stratification of the disease. One promising approach has been to focus on mutations or variations in the genetic code that are associated with the risk of developing ovarian cancer. A certain heritable component is already known regarding genes such as *BRCA1/2*, *TP53*, *MSH6*, *BRIP1* and *RAD51C*, yet these are estimated to only account for ~3.1% of the total risk. Recent advances in sequencing technologies have enabled the investigation of hundreds of thousands of genetic variants in genome-wide association studies in tens of thousands of patients, which has led to the discovery of 108 (39 loci with $P<5.0x10^{-8}$) novel susceptibility loci for ovarian cancer, presented in this review. Using the published variants in a patient cohort screening, together with variants identified in our ongoing whole exome sequencing project, future aims are to ascertain whether certain of the novel variants could be used as biomarkers for early diagnosis and/or treatment decisions.

Contents

1. Introduction

Ovarian cancer (OC) is the 5th most common cancer and the most lethal gynaecological malignancy in European women (1). The International Federation of Gynaecology and Obstetrics characterises four major stages of OC, with stages I and II constituting tumours localised and mainly confined to the ovaries, which are associated with a good prognosis [5-year overall survival for stage I, 87.0-89.5% (2)], and the late stages III and IV, with confirmed spread to the peritoneum and/or distant metastasis, and poorer outcome [5-year overall survival for stage IV, 13.2-17.9% (2)]. Early-stage OC presents with non-specific symptoms [including pelvic or abdominal pain, loss of appetite, fatigue and unexplained weight loss (3,4)] commonly associated with other diseases or ailments. Additionally, OC is a relatively rare disease, meaning that general practitioners will encounter a small number of OC cases throughout their career (5). Combined, this means that

Correspondence to: Professor Estrid Høgdall, Molecular Unit, Department of Pathology, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev Ringvej 75, 2730 Herlev, Denmark
E-mail: estrid.hoegdall@regionh.dk

Abbreviations: BCAC, Breast Cancer Association Consortium; CIMBA, The Consortium of Investigators of Modifiers of BRCA1/2; COGS, Collaborative Oncological Gene-environment Study; eQTL, expression quantitative trait loci; GWAS, genome-wide association study; HGSC, high-grade serous (ovarian) carcinoma; LD, linkage disequilibrium; NGS, next-generation sequencing; OC, ovarian cancer; OCAC, Ovarian Cancer Association Consortium; OCCC, ovarian clear cell carcinoma; OR, odds ratio; SNP, single nucleotide polymorphism

Key words: ovarian cancer, susceptibility loci, next-generation sequencing, genetic variation, biomarkers, genome-wide association studies

most patients with malignant growth in the pelvic region are diagnosed in the late stages of OC.

2. OC subtypes and biomarkers

OC is commonly divided into two major groups, epithelial and non-epithelial. Epithelial OC (EOC) comprises four main subtypes, based on the tissue of origin: Serous adenocarcinoma [high-grade (HGSC) and low-grade]; endometrioid adenocarcinoma; ovarian clear cell adenocarcinoma (OCCC); and mucinous adenocarcinoma. Non-epithelial OC is subdivided into germ cell and sex chord/stromal OC. Overall, ~86% of OC cases are epithelial, and of these, 76% are serous histological subtype, with HGSC counting 83% (2). The characteristics of particularly the four main EOC types differ markedly in origin tissue, gene and microRNA (miRNA) expression, and morphology, and there is emerging consensus that they should be recognised as four distinct diseases (6-8).

Established biomarkers for OC include CA-125 (9) and human epididymis protein 4 (10), and various multivariate index assays measuring serum concentrations of these and other proteins, as well as taking ultrasound examination of the pelvic region, menopausal state, patient age and/or family history into account, have also been devised (11-15). While these schemes have increased the likelihood of differentiating malignant OC from a benign growth in the pelvis, they have not proved sufficient to decisively decrease mortality rates (16,17). Consequently, there is still a clear requirement for finding robust biomarkers, especially those capable of detecting OC at the early stages that can be used prognostically and to guide targeted treatment.

OC has a significant heritable component. Mutations in particularly *BRCA1* and *BRCA2*, but also in other genes, including *TP53*, *BRIP1*, *MSH6* and *RAD51C*, have been described as risk factors (18,19), yet the known and familial genetic factors are estimated to only account for 3.1% of the risk of developing EOC (20). Therefore, it is proposed that there are additional OC susceptibility loci yet to be discovered, and this has been a major focus area in the past decade, as genomic research and sequencing techniques have improved significantly.

This review will present the advances in applying next-generation sequencing (NGS) in large cohort studies in the search for genetic variants that act as susceptibility loci and/or driver mutations for OC.

3. Literature search and inclusion/exclusion criteria

This review was carried out according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines (21). Studies were selected based on the search criteria 'ovarian cancer' and 'susceptibility loci' in the biomedical databases Medline, EMBASE and Scopus. Studies reporting genome-wide association studies (GWAS) in large cohorts were preferred. The aim was to cover as much of the published literature as possible; however, studies reporting variants associated with low malignant potential (borderline) OC subtypes were omitted, and only studies reported in English were included. In total, 108 susceptibility loci from 28 studies published from 2008 to 2018 were included (Fig. 1).

4. NGS as a tool in searching for ovarian cancer biomarkers

Only a decade ago, sequencing the genome of a single individual took months, if not years. Subsequent advances in microarray and sequencing technologies prompted by large-scale sequencing efforts such as the Human Genome Project (22) and the 1,000 Genomes Project (23) have revolutionised the field of genomic research, and today this can be accomplished over ~1 week using high-throughput sequencing (24). Targeted sequencing of only parts of the genome, such as transcriptome or whole exome sequencing, or sequencing of a subset of genes known to be involved in tumorigenesis, have enabled scientists and clinicians to develop and tailor research and treatment to the individual patient, a fundamental premise for precision medicine initiatives, and the overall goal for the treatment of patients with OC (25).

As high-throughput sequencing evolved into NGS (also known by the more appropriate term, massively-parallel sequencing), several hundred thousand genetic variants in thousands of patients can now be investigated in only a fraction of the time (26). Naturally, this has spawned large cohort studies, often with participation of clinics across the world, as well as the invention of specific arrays or chips focused on variants or genes suspected to be the cause of specific diseases. The Cancer Genome Atlas (TCGA) Research Network investigated 33 different cancer forms using high-throughput single nucleotide polymorphism (SNP), exome and genome sequencing, as well as gene expression, copy number variation, DNA methylation and miRNA profiling; these findings were recently summarised (27). OC was one of the three cancer types selected for the pilot project, and a cohort of 489 patients with HGSC were selected for analysis. Among other findings, the researchers found *TP53* to be mutated in almost all cases, and were able to classify tumours into several subtypes depending on transcription, miRNA and methylation profiles (28).

NGS studies of OC have been reported in the last five years, mainly stemming from two large global initiatives with significant overlaps: The US-based OncoArray Network and its eponymous genotyping array chip (29); and the mega-consortium Collaborative Oncological Gene-environment Study (COGS) with the iCOGS array, and updated OncoChip (30). Established in 2005, the Ovarian Cancer Association Consortium (OCAC) is a major collaboration, with contributors from the United States, United Kingdom, Australia, The Netherlands, Denmark, Poland, Germany and numerous other countries, and consists of 25,509 population-based EOC cases and 40,941 controls (31). The consortium was included in COGS together with Breast Cancer Association Consortium (BCAC), Prostate Cancer Association Group to Investigate Cancer-Associated Alterations in the Genome and The Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA), with the aim of studying the genetics and risk factors of these three hormone-related cancers [summarised in (30)]. For this collaboration, a custom genotyping array chip called iCOGS capable of genotyping >211,000 SNPs was developed and used on >250,000 subjects (30). Like COGS, the OncoArray Network's research and the OncoArray chip capable of genotyping 570,000 SNPs have resulted in numerous articles on glioblastoma, breast, ovarian, prostate

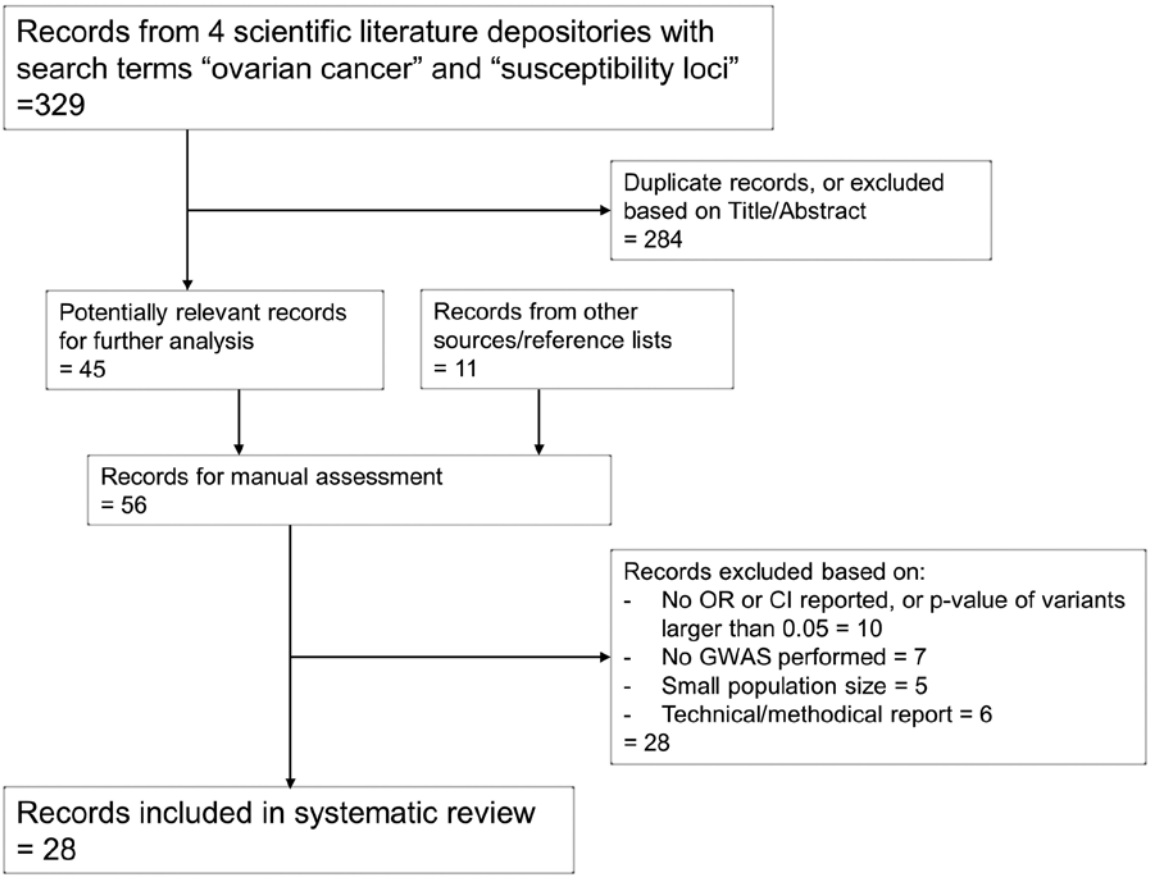

Figure 1. Systematic literature search according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines. Search criteria ‘ovarian cancer’ and ‘susceptibility loci’ were used in the databases PubMed/Medline, EMBASE, Web of Science and Scopus. Only studies in English were included. OR, odds ratio; CI, confidence interval; GWAS, genome-wide association study.

and lung cancers, using, among others, the OCAC and BCAC cohorts (31-35).

5. GWAS identifies numerous susceptibility loci for ovarian cancer

In total, 108 susceptibility loci for OC were identified following a systematic literature search (summarised in Table I and Table SI for variants with $P<5.0x10^{-8}$ and $P>5.0x10^{-8}$, respectively). These loci were mainly found via GWAS, in which genetic variations in a cohort of patients are compared to a cohort of healthy controls to isolate variants that may contribute to developing the disease. Variants are given an odds ratio (OR) score, depending on whether the variant is found predominantly in the patient cohorts (OR >1) or in the healthy controls (OR <1).

In total, >50% of the OC susceptibility loci were found to be involved in HGSC (59/108), which was perhaps expected, as this is by far the most prevalent subtype of OC and thus the one most frequently encountered. Certain variants have been reported in >1 subtype, most notably rs757210, which seems to be linked with poor prognosis in HGSC (odds ratio 1.12), but predicts superior outcomes in OCCC (OR 0.80), demonstrating the importance of stratifying GWAS findings by OC subtype (36-38). rs757210 sits in the promoter region of *HFN1B*, which is known to be overexpressed in OCCC (39) and downregulated in serous OC (36), as well as being a susceptibility gene for diabetes type II (40), prostate cancer (41,42), uterine corpus cancer (43) and endometrial cancer (44,45). Shen *et al* (36) hypothesised that the difference in expression levels could be due to promoter methylation of *HNF1B* in proximity to this variant, which was later confirmed (46).

Mutation hotspots are a common feature in cancer genomics, and some of the identified susceptibility loci were situated in or near genes that are frequently altered in cancer cells. As such, Pooley *et al* and Bojesen *et al* (47,48) investigated the telomerase gene *TERT*, which maintains chromosome telomeres. Somatic mutations, especially in the promoter region of *TERT*, have been found in cancers of the brain, thyroid gland, bladder and skin (49). Bojesen *et al* (48) reported a locus associated with HGSC (rs10069690) in intron 4 with the minor allele conferring increased risk of disease and creates an alternative splice site that results in a truncated protein and impaired telomerase function. This reinforces the hypothesis that shorter telomeres increase cancer risk.

Bolton *et al* (50) examined a known breast cancer locus on chromosome 19p13, and found susceptibility loci that were significantly associated with risk of serous OC. Presumably, the SNP rs8170 located in the *BABAM1* (previously *MERIT40*) gene may explain this risk, as *BABAM1* has been shown to interact with, and stabilise, *BRCA1* interactions with a complex including *RAP80*, *BRCC45* and *CCDC98* (51-53). Although counterintuitive, as breast cancer and OC frequently present

Table I. Susceptibility loci identified through 28 published genome-wide association studies in ovarian cancer cohorts. Subtype.

	SNP	Locus	Position[a]	EA[b]	EAF[c]	OR	CI_{95}	P-value/BF[d]	Gene[e]	eQTL[f]	Refs.
EnOC	rs555025179	05q12.3	66825262	delACA CACAC	0.56	1.18	(1.11-1.26)	$4.5x10^{-8}$	MAST4		(31)
EnOC	rs12951053	17p13.1	7674089	C	0.18	1.31	(1.05-1.62)	109.6[g]	TP53		(93)
EOC	rs199661266	17q11.2	30854203	insT	0.28	0.91	(0.88-0.94)	$5.4x10^{-9}$	ATAD5, AC130324.2		(20)
EOC	rs587778134	17q23.2	61776459	insTT	0.02	8.13	(4.74-13.95)	$2.8x10^{-14}$	BRIP1	BRIP1	(86)
EOC-BRCA1	rs2165109	02q13	111061081	C	0.25	1.09	(1.05-1.12)	$4.2x10^{-8}$	ACOXL		(31)
EOC-BRCA1	rs4691139	04q32.3	164987569	G	0.47	1.20	(1.17-1.38)	$3.4x10^{-8}$	TRIM61, AC106872.4		(90)
EOC-BRCA1	rs9886651	08q24.21	127805637	G	0.46	1.08	(1.05-1.11)	$3.5x10^{-9}$	PVT1		(31)
EOC-BRCA1	rs7953249	12q24.31	120965921	G	0.42	1.08	(1.06-1.10)	$1.1x10^{-9}$	HNF1A	HNF1A, OASL	(31)
EOC-BRCA1	rs17631303	17q21	45439036	G	0.19	1.27	(1.17-1.38)	$1.4x10^{-8}$	PLEKHM1		(90)
EOC-BRCA1	rs183211	17q21	46710944	A	0.23	1.25	(1.16-1.35)	$3.1x10^{-8}$	PLEKHM1	PLEKHM1	(90)
HGSC	rs2072590	02q31	176177905	T	0.32	1.20	(1.14-1.25)	$3.8x10^{-14}$	HOXD9	HOXD1, HOXD3, HOXD9	(69)
HGSC	rs6755777	02q31.1	176178498	T	0.68	1.15	NR	$9.0x10^{-14}$	HAGLR, HAGLROS		(56)
HGSC	rs7651446	03q25	156689208	A	0.05	1.44	(1.35-1.53)	$1.5x10^{-28}$	TIPARP	TIPARP	(37)
HGSC	rs78561123	03q25.31	156747145	C	0.06	1.45	(1.32-1.59)	$3.0x10^{-15}$	LINC00886		(77)
HGSC	rs62273959	03q25.31	156852891	A	0.06	1.45	(1.32-1.59)	$3.5x10^{-14}$	LEKR1		(77)
HGSC	rs10069690	05p15.33	1279675	T	0.26	1.15	(1.11-1.20)	$1.3x10^{-11}$	TERT	TERT	(48)
HGSC	rs7726159	05p15.33	1282204	A	0.34	1.12	(1.09-1.16)	$3.4x10^{-9}$	TERT	TERT	(47)
HGSC	rs11782652	08q21	81741409	G	0.07	1.24	(1.16-1.33)	$7.0x10^{-10}$	CHMP4C	CHMP4C	(37)
HGSC	rs10088218	08q24	128531703	A	0.13	0.76	(0.70-0.81)	$8.0x10^{-15}$	MYC	MYC, THEM75	(69)
HGSC	rs7814937	08q24	128529479	C	0.11	1.18	(1.13-1.24)	NR	LINC00824		(55)
HGSC	rs3814113	09p22	16915023	C	0.27	0.77	(0.73-0.81)	$4.1x10^{-21}$	BNC2	BNC2	(68) (73)
HGSC	rs7084454	10p12.31	21532345	A	0.24	1.10	(1.06-1.14)	NR	MLLT10		(55)
HGSC	rs2287498	17p13.1	7689242	A	0.14	1.30	(1,07-1,57)	165.7[g]	WRAP53, TP53		(93)
HGSC	rs12951053	17p13.1	7674089	C	0.18	1.19	(1,01-1,38)	47.8[g]	TP53		(93)
HGSC	rs757210	17q12	37736525	A	0.37	1.12	(1.08-1.17)	$8.1x10^{-10}$	HNF1B	HNF1B	(37) (36)
HGSC	rs7405776	17q12	37733029	A	0.37	1.25	(1.17-1.34)	$2.3x10^{-11}$	HNF1B	HFN1B	(36) (67)
HGSC	rs2077606	17q21.31	45452177	A	0.08	1.15	(1.12-1.19)	NR	PLEKHM1		(55)
HGSC	rs2960000	17q21.31	45456987	C	0.18	1.16	(1.12-1.20)	$3.3x10^{-10}$	PLEKHM1, AC091132.2		(70)
HGSC	rs12942666	17q21.31	45422473	G	0.22	1.15	(1.11-1.20)	$1.0x10^{-9}$	PLEKHM1, ARHGAP27	PLEKHM1	(70)
HGSC	rs1052587	17q21.31	46025238	C	0.22	1.12	(1.08-1.17)	$4.6x10^{-8}$	MAPT		(70)
HGSC	rs7218345	17q21.32	48425805	T	0.31	1.12	(1.08-1.16)	NR	SKAP1		(55)
HGSC	rs2363956	19p13.11	17283315	T	0.49	1.16	(1.11-1.21)	$3.8x10^{-11}$	ANKLE1	ANKLE1	(50)

Table I. Continued.

	SNP	Locus	Position[a]	EA[b]	EAF[c]	OR	CI_{95}	P-value/BF[d]	Gene[e]	eQTL[f]	Refs.
HGSC	rs8170	19p13.11	17278895	T	0.19	1.18	(1.12-1.25)	$2.7x10^{-9}$	BABAM1, USHBP1	ANKLE1, BABAM1, ABHD8	(50)
HGSC	rs4808075	19p13.11	17279482	C	0.3	1.19	(1.14-1.23)	$9.2x10^{-20}$	USHBP1, BABAM1	ABHD8, BABAM1	(54)
MOC	rs752590	02q13	113215368	G	0.21	1.34	(1.21-1.49)	$3.3x10^{-8}$	PAX8	PAX8	(38)
MOC	rs711830	02q31.1	176172583	A	0.32	1.30	(1.20-1.40)	$7.5x10^{-12}$	HOXD3, AC009336.2	HOXD3, HOXD9	(38)
MOC	rs112071820	03q22.3	139130269	insCCA GATTCA GAAT	0.33	1.29	(1.20-1.37)	$1.5x10^{-13}$	BPESC1, MRPS22		(31)
MOC	rs320203	09q31.1	102180944	A	0.88	1.29	(1.18-1.41)	$1.7x10^{-8}$	GRIN3A		(31)
MOC	rs688187	19q13.2	39242112	A	0.32	0.67	(0.60-0.75)	$6.8x10^{-13}$	IFNL3	IFNL3, IFNL4	(38)

Susceptibility loci with an OR P<$5.0x10^{-8}$ are included. [a]Human genome reference assembly GRCh38/hg38. [b]EA is not necessarily on the leading strand. Preferably quoted by original publication, otherwise the minor allele, or with frequency corresponding to EAF. [c]EAF, preferably in case cohort, otherwise in the European ancestry population in the 1,000 Genomes Project. [d]BF >10 is considered strong confidence that the result will withstand further investigation. [e]Gene(s) in closest proximity to the most significant ovarian cancer risk variant at each locus. [f]Gene with published functional association with phenotype and thus the likely gene responsible. Preferably quoted by reference; otherwise Ensembl's Biomart database ('Associated gene with phenotype'). [g]BF. BF, Bayes factor; CI, confidence interval; EA, effect allele; EAF, EA frequency; EnOC, endometrioid ovarian cancer; EOC, epithelial ovarian cancer (no subtype specified); EOC-BRCA1, EOC cases including BRCA1 carriers (The Consortium of Investigators of Modifiers of BRCA1/2); eQTL, expression quantitative trait loci; HGSC, high grade serous carcinoma; MOC, mucinous ovarian cancer; NR, not reported; OR, odds ratio; SNP, single nucleotide polymorphism.

with inactivating mutations in *BRCA1*, the authors speculated that an overexpressed *BABAM1* leads to stabilisation of the breast cancer susceptibility protein (BRCA) complex, and thus an increased tolerance to DNA damage. Lawrenson *et al* (54) fine-mapped the region near rs8170 in the BCAC, OCAC and CIMBA cohorts, and found rs4808075 to be the strongest candidate causal variant. Further investigation led to the discovery of rs4808616, in strong linkage disequilibrium (LD) with rs4808075, that has significant expression quantitative trait loci (eQTL) association with *ABHD8*, a gene neighbour of *BABAM1*, in serous OC. Functional analyses revealed an association between the *ABHD8* promoter and rs4808616 via chromatin conformation capture, whereas overexpression of *ABHD8*, but not *BABAM1*, affected ovarian epithelial cells *in vitro*. Therefore, *BABAM1* and rs8170 identified by Bolton *et al* may have been a proxy for the real driver variant, potentially either rs4808075 or rs4808616 in *ABHD8*. *ABHD8* is a notable gene that has recently been speculated to be involved in the migration and invasion of OC tumour cells through a homeobox-containing transcription factor network (55,56).

Ghoussaini *et al* (57) used cohorts that were later included in COGS (Studies of Epidemiology and Risk Factors in Cancer Heredity, MALignant OVArian cancer study, Family Registry for Ovarian Cancer Study, United Kingdom Ovarian Cancer Population Study and United Kingdom Genetic Prostate Cancer Study) to examine the 8q24 gene desert, where the two closest genes are well-known cancer susceptibility genes (*c-MYC* and *FAM84B*). *c-MYC* was functionally validated to be implicated in OC (Table I) and was also identified in the TCGA dataset (28).

6. GWAS statistical significance threshold

GWAS analyses have become the golden standard for finding disease susceptibility loci, but there are certain limitations as well. With hundreds of thousands of variants examined on a single chip, the risk of false positives increases dramatically, and stringent data processing must be employed. In general, GWAS studies favour common variants in the population, meaning that fine-mapping and additional filtering are required to discover variants with a minor allele frequency (MAF) of <5% (58). Genetic variation occurs semi-randomly and is widespread throughout the genome, and large cohort sizes in the thousands are required to obtain statistically significant and reliable results. There is an ongoing debate regarding which P-value threshold should be the standard for GWAS, or whether Bayesian approaches should be employed instead (59). The generally accepted P-value is P≤$5.0x10^{-8}$ for common variants, as first introduced by The International HapMap Project (60) and subsequently by Pe'er *et al* (61), and which has been recently evaluated and confirmed (62). This latest study concluded that this threshold is too relaxed for rare variants (MAF ≤0.5%), and cut-offs for these should be: $3x10^{-8}$ for MAF ≥1%, $2x10^{-8}$ for MAF ≥0.5% and $1x10^{-8}$ for MAF ≥0.1% (conditions: Whole-genome sequencing studies in European populations with all variants having an LD r^2>0.8). For the present review, and contrary to two recent reviews of GWAS susceptibility loci (63,64), it was determined that all variants reported by the original articles would be included, with a more relaxed cut-off of P≤0.05. Variants meeting the threshold criteria discussed above (P≤$5.0x10^{-8}$) are presented in Table I; the remaining variants are included in Table SI. For simplicity, for articles fine-mapping a

susceptibility loci region and finding additional SNPs with lower P-values, but in strong LD with the index SNP (48,54), only the novel variant with the strongest association was included.

7. Susceptibility loci region fine-mapping

Genetic variants are not randomly distributed in the genome, but often aggregate in the same populations (23). Genomic research has taken advantage of this, by examining only those variants or polymorphisms already reported in large population studies. Nevertheless, the number of genetic variants in the human genome amounts to tens of millions, which is not feasible to investigate in a research setting on a large number of patients. Instead, an array of representative or index SNPs are frequently used to cover all variants in a genomic region, utilising the fact that neighbouring SNPs are often in tight LD and thus inherited together (65,66). Data from The 1,000 Genomes Project estimates that any given trait-associated variant in the National Human Genome Research Institute GWAS database will have 56 neighbouring variants in LD with $r^2 \geq 0.5$ (28). It follows that fine-mapping of the region is required to determine if the index SNP is indeed the causal variant, or merely a proxy for other SNPs in the region. Three examples of this have been described in the subsection 'Genome-wide association studies identify numerous susceptibility loci for ovarian cancer': Bojesen *et al* fine-mapped the *TERT* locus and SNPs in LD with rs10069690 and rs7705526 (48); Lawrenson *et al* examined the *ABHD8*/*ANKLE1* locus and rs4808075 (54); and Shen *et al* investigated the *HNF1B* region and rs7405776 and rs11651755 (36). Following the initial findings of the COGS initiative, Earp *et al* (67) analysed 11 known susceptibility regions and found novel associated variants with more robust P-values and ORs than those previously reported (Table I) (20,37,48,50,68-71).

Several studies over the last few years have fine-mapped the 9p22.2 region by rs3814113 first reported by Song *et al* in 2009 (68). eQTL analyses concluded the nearby zinc finger protein basonuclin-2 (*BNC2*), which has been implicated in oocyte differentiation (72), to be the most likely causal candidate gene (69,73). Additional SNPs were found to be associated with abnormal ovarian ultrasound results (74) and to modify OC risk in *BRCA1/2* mutation carriers (75). *BNC2* was reported to contribute to a *HOX*-centric network of transcription factors associated with serous OC risk (55), and Carter *et al* (76) found a significant association between germline rs3814113 and tumour formation in OC. Finally, a recent study by Buckley *et al* (73) reported additional SNPs in LD with rs3814113, as well as SNPs located in the regulatory regions of *BNC2*, including some in this gene's scaffold/matrix attachment region, suggesting that they influence chromosomal three-dimensional organisational optimization for transcription in an allele-specific manner.

Finally, the cluster of OC-related variants in the *TIPARP*/*LEKR1* region on chromosome 3q25 is of note in an OC disease setting and has been studied thoroughly (37,67,69,77). *TIPARP* (also known as *PARP7*) codes for a poly ADP ribose polymerase (PARP), a group of proteins that have been the target of PARP inhibitor cancer treatments showing great promise in the targeted treatment of patients with breast, prostate and OC carrying *BRCA1/2* mutations (78).

8. Several variants cause potentially functionally relevant amino acid changes

The potential impact of a genetic variant is associated with its location in the gene. Only 1% of the human genome codes for proteins; the remaining regions are intra/intergenic, promoters, enhancers and long stretches of 'gene deserts', where genes are tens or hundreds of kilobases apart (79). It follows that a variant within a protein-coding region is potentially more detrimental to the cell than one located in a gene desert. In the present study, 21 of the 108 identified variants alter amino acid sequences (Table II). Two algorithms have been developed to evaluate the potential damage caused by these changes: SIFT (80) and PolyPhen-2 (81) scores. Both have values between 0 and 1, but the values have reciprocal interpretation. A variant with a SIFT score approaching 0 is considered deleterious, while one with a PolyPhen-2 score approaching 1 is considered damaging. Several variants in Table II are located in notable genes from an OC perspective: *ANKLE1* and *BRCA2*, as discussed earlier in this review; *BTD*, which has shown promise as a biomarker for breast (82) and cervical (83) cancers; *ZFHX3*, which is a tumour suppressor gene frequently mutated in prostate (84) and endometrial (85) cancer; and *LEKR*, which, although the variant rs62273959 is considered benign, was found to be in tight LD (r^2=0.90) with rs7651446 in the aforementioned *TIPARP* gene (77). Finally, it is worth noting that rs587778134 causes a frame shift mutation in the DNA repair gene *BRIP1* (also known as *FANCJ*), which is a well-known OC susceptibility gene that interacts with BRCA1 (19,86). Although it is not found in the SIFT/PolyPhen-2 databases, it does have an entry for OC susceptibility (RCV000409984.1) in the ClinVar database of potentially clinically relevant genetic variants and is designated as 'Likely pathogenic' (Table II) (87).

9. OC as a hormone-related disease

OC is suspected to be a hormonal disease and related to breast and prostate cancers (88,89). Three new susceptibility loci were found by investigating a cohort of patients with breast or ovarian cancer harbouring mutations in *BRCA1* (90). As part of the COGS initiative, the three cancers were examined in individual GWAS projects (20,32,91), whereas Kar *et al* (92) combined the results from the three projects in a single three-cancer meta-analysis, as well as one-by-one comparisons. The findings showed clear pleiotropy among the diseases, with three susceptibility loci identified in all three cancers (rs17041869, rs7937840 and rs1469713), and four loci shared between breast and ovarian cancers (rs635634, rs11571833, rs200182588 and rs8037137). No shared loci were found for prostate and ovarian cancer alone.

10. Most susceptibility loci are found outside classic OC causal genes

Surprisingly few of the susceptibility loci were found in genes commonly associated with OC, such as *TP53* (93), *BRCA2* (92) or *HNF1B* (36,37). This may be explained by the fact that a GWAS only detects susceptibility loci of a single or few nucleotides, often conferring subtle differences in gene expression, whereas some mutations in the classic causal genes are large deletions or inactivating mutations that are detrimental

Table II. Functional relevance of identified susceptibility loci.Subtype.

	SNP	Locus	Gene[a]	Amino acid change (Human Genome variation society)	SIFT score[b]	SIFT prediction	PolyPhen-2 score[b]	PolyPhen-2 prediction	ClinVar clinical significance	ClinVar phenotype
EnOC	rs138031468	01p36.33	AGRN	NP_001292204:p.Ala237Ser	0.15	Tolerated	0.805	Possibly damaging	Likely benign	Not specified
EnOC	rs61757604	02p22.2	DHX57	NP_001316892:p.Gly49Ser	0.06	Tolerated-low confidence	0.05	benign		
EOC	rs200337373	03p25.1	BTD	NP_001268654:p.Asp202Asn	0.42	Tolerated	0.895	Possibly damaging	Pathogenic	Biotinidase deficiency
EOC	rs73757391	05q11.2	AC025470.2 | ACTBL2	NP_001017992:p.Glu108Lys	0.04	Deleterious-low	0.989	Probably damaging		
EOC	rs199761238	06q25.2	SYNE1	NP_149062:p.Asn4519Asp		Confidence	0.517	Possibly damaging	Uncertain significance	not specified
EOC	rs147432497	15q12	ATP10A	NP_077816:p.Arg999Cys	0	deleterious	0.99	Probably damaging		
EOC	rs587778134	17q23.2	BRIP1	NP_114432.2:p.Leu680fs					Likely pathogenic	Neoplasm of ovary
EOC	rs13181	19q13.32	ERCC2	NP_000391:p.Lys751Gln	0.45	Tolerated	0	Benign	Benign	NSCLC
EOC	rs141200301	22q11.2	MMP11	NP_005931:p.Arg334Cys	0.01	Deleterious	0.984	Probably damaging		
HGSC	rs62273959	03q25.31	LEKR1	NP_001004316:p.Val58Ile	0.29	Tolerated	0.026	Benign		
HGSC	rs381852	05q11.2	CDC20B | GPX8	NP_001008398:p.Lys182Arg	0.62	Tolerated	0	Benign		
HGSC	rs2297980	06p12.1	TINAG	NP_055279:p.Gln22Arg	0.28	Tolerated	0.001	Benign		
HGSC	rs130072	06p21.33	CCHCR1	NP_061925:p.Arg627Gln	0.03	Deleterious	0.998	Probably damaging		
HGSC	rs2073724	06p21.33	TCF19	NP_001070979:p.Pro241Leu	0.1	Tolerated	0.521	Possibly damaging		
HGSC	rs2233976	06p21.33	C6orf15	NP_054789:p.Gly48Arg	0.11	Tolerated	0.324	Benign		
HGSC	rs145514333	11q13.1	PYGM	NP_005600:p.Arg61His	0	Deleterious	0.947	Probably damaging	Uncertain significance benign	Not specified Hereditary BC and OC
HGSC	rs11571833	13q13	BRCA2	NP_000050.2:p.Lys3326Ter						
HGSC	rs147445846	16q22.3	ZFHX3	NP_008816:p.Leu379Val	0.04	Deleterious	0.703	Possibly damaging		
HGSC	rs150321809	17q21.2	KRT13 | AC019349.1	NP_705694:p.Arg429His	0.08	Tolerated	0.043	Benign	Likely benign	White sponge nevus of cannon
HGSC	rs104894658	18p11.21	MC2R	NP_000520:p.Ser74Ile	0	Deleterious	0.995	Probably damaging	Pathogenic	ACTH resistance
HGSC	rs2363956	19p13.11	ANKLE1	NP_001265373:p.Leu185Trp	0.03	Deleterious	0.162	Benign		

[a]Gene(s) in closest proximity to the most significant ovarian cancer risk variant at each locus. [b]SIFT and PolyPhen-2 scores are reciprocal-both range from 0 to 1, with SIFT scores near 0 considered deleterious, and PolyPhen-2 scores near 0 considered benign. ACTH, adrenocorticotropic hormone; BC, breast cancer; EnOC, endometrioid ovarian cancer; EOC, Epithelial ovarian cancer (no subtype specified); HGSC, high grade serous carcinoma; NSCLC, non-small cell lung cancer; OC, ovarian cancer; SNP, single nucleotide polymorphism.

to normal protein function. Additionally, most of the variants have relatively low ORs (<2.0), meaning that they only have moderate effects on OC risk. Incidentally, the variant with the second-highest OR (rs587778134, OR=8.13) was located in *BRIP1*, and mutations in this gene have been established to confer a moderate to high risk of developing OC (19,86).

11. Future directions

Much attention has been focused on finding isolated susceptibility loci on a genome-wide scale over the past decade. A large number of the identified variants were situated in intergenic regions far from the genes they potentially affect, and while several GWAS have been performed and analysed, few studies have fine-mapped and functionally validated any of the findings. With most loci situated in genes not previously associated with OC, including hits in long noncoding RNAs (94), there is an urgent requirement and potential for examining these further, particularly those that are near or in genes implicated in oocyte and ovary development, or tumour progression.

The focus must be on finding candidate causal genes for OC. Promising studies have been released in recent years, including the transcriptome-wide association study by Lu *et al* (95). In this study, they performed a 'reverse GWAS' by cross-matching existing OC-specific gene expression profiles with all known susceptibility loci and candidate SNPs, and reported the Frizzled gene *FZD4* as a novel candidate causal gene. This is an area that complements and overlaps well with the search for novel susceptibility loci.

We are currently performing whole exome sequencing of patients with HGSC and OCCC, to identify variants that are subtype- and survival-specific. Combined with the published variants summarised in this review, a screen of a large number of patients with OC will be performed to identify potential biomarkers for the early detection of OC that may decrease the mortality rates for patients.

Acknowledgements

Not applicable.

Funding

This study was supported by the Danish Mermaid III project, who was not involved in the decision to write the paper.

Availability of data and materials

All data generated or analyzed during this study are included in this published article.

Authors' contributions

All authors designed the study. MKC drafted and edited the manuscript, EH supervised and edited, and CH supervised. All authors contributed to and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Stewart BW and Wild CP: World Cancer Report 2014, 2014.
2. Danish Gynecologic Cancer Group: Annual Report of the Danish Gynecologic Cancer Database 2016-17. Danish Gynecol Cancer Database, 2017.
3. Goff BA, Mandel LS, Drescher CW, Urban N, Gough S, Schurman KM, Patras J, Mahony BS and Andersen MR: Development of an ovarian cancer symptom index: Possibilities for earlier detection. Cancer 109: 221-227, 2007.
4. Hamilton W, Peters TJ, Bankhead C and Sharp D: Risk of ovarian cancer in women with symptoms in primary care: Population based case-control study. BMJ 339: b2998, 2009.
5. Sundar S, Neal RD and Kehoe S: Diagnosis of ovarian cancer. BMJ 351: h4443, 2015.
6. Köbel M, Kalloger SE, Boyd N, McKinney S, Mehl E, Palmer C, Leung S, Bowen NJ, Ionescu DN, Rajput A, *et al*: Ovarian carcinoma subtypes are different diseases: Implications for biomarker studies. PLoS Med 5: e232, 2008.
7. Prat J: Ovarian carcinomas: Five distinct diseases with different origins, genetic alterations, and clinicopathological features. Virchows Arch 460: 237-249, 2012.
8. Vaughan S, Coward JI, Bast RC Jr, Berchuck A, Berek JS, Brenton JD, Coukos G, Crum CC, Drapkin R, Etemadmoghadam D, *et al*: Rethinking ovarian cancer: Recommendations for improving outcomes. Nat Rev Cancer 11: 719-725, 2011.
9. Bast RC Jr, Klug TL, St John E, Jenison E, Niloff JM, Lazarus H, Berkowitz RS, Leavitt T, Griffiths CT, Parker L, *et al*: A radioimmunoassay using a monoclonal antibody to monitor the course of epithelial ovarian cancer. N Engl J Med 309: 883-887, 1983.
10. Hellström I, Raycraft J, Hayden-Ledbetter M, Ledbetter JA, Schummer M, McIntosh M, Drescher C, Urban N and Hellström KE: The HE4 (WFDC2) protein is a biomarker for ovarian carcinoma. Cancer Res 63: 3695-3700, 2003.
11. Karlsen MA, Fagö-Olsen C, Høgdall E, Schnack TH, Christensen IJ, Nedergaard L, Lundvall L, Lydolph MC, Engelholm SA and Høgdall C: A novel index for preoperative, non-invasive prediction of macro-radical primary surgery in patients with stage IIIC-IV ovarian cancer-a part of the Danish prospective pelvic mass study. Tumor Biol 37: 12619-12626, 2016.
12. Jacobs I, Oram D, Fairbanks J, Turner J, Frost C and Grudzinskas JG: A risk of malignancy index incorporating CA 125, ultrasound and menopausal status for the accurate preoperative diagnosis of ovarian cancer. Br J Obstet Gynaecol 97: 922-929, 1990.
13. Moore RG, McMeekin DS, Brown AK, DiSilvestro P, Miller MC, Allard WJ, Gajewski W, Kurman R, Bast RC Jr and Skates SJ: A novel multiple marker bioassay utilizing HE4 and CA125 for the prediction of ovarian cancer in patients with a pelvic mass. Gynecol Oncol 112: 40-46, 2009.
14. Skates SJ: Ovarian cancer screening: Development of the risk of ovarian cancer algorithm (ROCA) and ROCA screening trials. Int J Gynecol Cancer 22 (Suppl 1): S24-S26, 2012.
15. Ueland FR, Desimone CP, Seamon LG, Miller RA, Goodrich S, Podzielinski I, Sokoll L, Smith A, van Nagell JR Jr and Zhang Z: Effectiveness of a multivariate index assay in the preoperative assessment of ovarian tumors. Obstet Gynecol 117: 1289-1297, 2011.
16. Karlsen MA, Sandhu N, Høgdall C, Christensen IJ, Nedergaard L, Lundvall L, Engelholm SA, Pedersen AT, Hartwell D, Lydolph M, *et al*: Evaluation of HE4, CA125, risk of ovarian malignancy algorithm (ROMA) and risk of malignancy index (RMI) as diagnostic tools of epithelial ovarian cancer in patients with a pelvic mass. Gynecol Oncol 127: 379-383, 2012.
17. Jacobs IJ and Menon U: Progress and challenges in screening for early detection of ovarian cancer. Mol Cell Proteomics 3: 355-366, 2004.
18. Norquist BM, Harrell MI, Brady MF, Walsh T, Lee MK, Gulsuner S, Bernards SS, Casadei S, Yi Q, Burger RA, *et al*: Inherited mutations in women with ovarian carcinoma. JAMA Oncol 2: 482-490, 2016.

19. Ramus SJ, Song H, Dicks E, Tyrer JP, Rosenthal AN, Intermaggio MP, Fraser L, Gentry-Maharaj A, Hayward J, Philpott S, *et al*: Germline mutations in the BRIP1, BARD1, PALB2, and NBN genes in women with ovarian cancer. J Natl Cancer Inst 107: djv214, 2015.
20. Kuchenbaecker KB, Ramus SJ, Tyrer J, Lee A, Shen HC, Beesley J, Lawrenson K, McGuffog L, Healey S, Lee JM, *et al*: Identification of six new susceptibility loci for invasive epithelial ovarian cancer. Nat Genet 47: 164-171, 2015.
21. Moher D, Liberati A, Tetzlaff J, Altman DG and PRISMA Group: Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Med 6: e1000097, 2009.
22. International Human Genome Sequencing Consortium: Finishing the euchromatic sequence of the human genome. Nature 431: 931-945, 2004.
23. The 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA and Abecasis GR: A global reference for human genetic variation. Nature 526: 68-74, 2015.
24. Metzker ML: Sequencing technologies-the next generation. Nat Rev Genet 11: 31-46, 2010.
25. Ashley EA: Towards precision medicine. Nat Rev Genet 17: 507-522, 2016.
26. Goodwin S, McPherson JD and McCombie WR: Coming of age: Ten years of next-generation sequencing technologies. Nat Rev Genet 17: 333-351, 2016.
27. Blum A, Wang P and Zenklusen JC: SnapShot: TCGA-analyzed tumors. Cell 173: 530, 2018.
28. Cancer Genome Atlas Research Network: Integrated genomic analyses of ovarian carcinoma. Nature 474: 609-615, 2011.
29. Amos CI, Dennis J, Wang Z, Byun J, Schumacher FR, Gayther SA, Casey G, Hunter DJ, Sellers TA, Gruber SB, *et al*: The oncoarray consortium: A network for understanding the genetic architecture of common cancers. Cancer Epidemiol Biomarkers Prev 26: 126-135, 2017.
30. Sakoda LC, Jorgenson E and Witte JS: Turning of COGS moves forward findings for hormonally mediated cancers. Nat Genet 45: 345-348, 2013.
31. Phelan CM, Kuchenbaecker KB, Tyrer JP, Kar SP, Lawrenson K, Winham SJ, Dennis J, Pirie A, Riggan MJ, Chornokur G, *et al*: Identification of 12 new susceptibility loci for different histotypes of epithelial ovarian cancer. Nat Genet 49: 680-691, 2017.
32. Michailidou K, Beesley J, Lindstrom S, Canisius S, Dennis J, Lush MJ, Maranian MJ, Bolla MK, Wang Q, Shah M, *et al*: Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. Nat Genet 47: 373-380, 2015.
33. McKay JD, Hung RJ, Han Y, Zong X, Carreras-Torres R, Christiani DC, Caporaso NE, Johansson M, Xiao X, Li Y, *et al*: Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. Nat Genet 49: 1126-1132, 2017.
34. Schumacher FR, Al Olama AA, Berndt SI, Benlloch S, Ahmed M, Saunders EJ, Dadaev T, Leongamornlert D, Anokian E, Cieza-Borrella C, *et al*: Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. Nat Genet 50: 928-936, 2018.
35. Melin BS, Barnholtz-Sloan JS, Wrensch MR, Johansen C, Il'yasova D, Kinnersley B, Ostrom QT, Labreche K, Chen Y, Armstrong G, *et al*: Genome-wide association study of glioma subtypes identifies specific differences in genetic susceptibility to glioblastoma and non-glioblastoma tumors. Nat Genet 49: 789-794, 2017.
36. Shen H, Fridley BL, Song H, Lawrenson K, Cunningham JM, Ramus SJ, Cicek MS, Tyrer J, Stram D, Larson MC, *et al*: Epigenetic analysis leads to identification of HNF1B as a subtype-specific susceptibility gene for ovarian cancer. Nat Commun 4: 1628, 2013.
37. Pharoah PD, Tsai YY, Ramus SJ, Phelan CM, Goode EL, Lawrenson K, Buckley M, Fridley BL, Tyrer JP, Shen H, *et al*: GWAS meta-analysis and replication identifies three new susceptibility loci for ovarian cancer. Nat Genet 45: 362-370, 2013.
38. Kelemen LE, Lawrenson K, Tyrer J, Li Q, Lee JM, Seo JH, Phelan CM, Beesley J, Chen X, Spindler TJ, *et al*: Genome-wide significant risk associations for mucinous ovarian carcinoma. Nat Genet 47: 888-897, 2015.
39. Tsuchiya A, Sakamoto M, Yasuda J, Chuma M, Ohta T, Ohki M, Yasugi T, Taketani Y and Hirohashi S: Expression profiling in ovarian clear cell carcinoma: Identification of hepatocyte nuclear factor-1beta as a molecular marker and a possible molecular target for therapy of ovarian clear cell carcinoma. Am J Pathol 163: 2503-2512, 2003.
40. Gudmundsson J, Sulem P, Steinthorsdottir V, Bergthorsson JT, Thorleifsson G, Manolescu A, Rafnar T, Gudbjartsson D, Agnarsson BA, Baker A, *et al*: Two variants on chromosome 17 confer prostate cancer risk and the one in TCF2 protects against type 2 diabetes. Nat Genet 39: 977-983, 2007.
41. Sun J, Zheng SL, Wiklund F, Isaacs SD, Purcell LD, Gao Z, Hsu FC, Kim ST, Liu W, Zhu Y, *et al*: Evidence for two independent prostate cancer risk-associated loci in the HNF1B gene at 17q12. Nat Genet 40: 1153-1155, 2008.
42. Thomas G, Jacobs KB, Yeager M, Kraft P, Wacholder S, Orr N, Yu K, Chatterjee N, Welch R, Hutchinson A, *et al*: Multiple loci identified in a genome-wide association study of prostate cancer. Nat Genet 40: 310-315, 2008.
43. Spurdle AB, Thompson DJ, Ahmed S, Ferguson K, Healey CS, O'Mara T, Walker LC, Montgomery SB, Dermitzakis ET; Australian National Endometrial Cancer Study Group, *et al*: Genome-wide association study identifies a common variant associated with risk of endometrial cancer. Nat Genet 43: 451-455, 2011.
44. Painter JN, O'Mara TA, Batra J, Cheng T, Lose FA, Dennis J, Michailidou K, Tyrer JP, Ahmed S, Ferguson K, *et al*: Fine-mapping of the HNF1B multicancer locus identifies candidate variants that mediate endometrial cancer risk. Hum Mol Genet 24: 1478-1492, 2015.
45. Setiawan VW, Haessler J, Schumacher F, Cote ML, Deelman E, Fesinmeyer MD, Henderson BE, Jackson RD, Vöckler JS, Wilkens LR, *et al*: HNF1B and endometrial cancer risk: Results from the PAGE study. PLoS One 7: e30390, 2012.
46. Ross-Adams H, Ball S, Lawrenson K, Halim S, Russell R, Wells C, Strand SH, Ørntoft TF, Larson M, Armasu S, *et al*: HNF1B variants associate with promoter methylation and regulate gene networks activated in prostate and ovarian cancer. Oncotarget 7: 74734-74746, 2016.
47. Pooley KA, Bojesen SE, Weischer M, Nielsen SF, Thompson D, Amin Al Olama A, Michailidou K, Tyrer JP, Benlloch S, Brown J, *et al*: A genome-wide association scan (GWAS) for mean telomere length within the COGS project: Identified loci show little association with hormone-related cancer risk. Hum Mol Genet 22: 5056-5064, 2013.
48. Bojesen SE, Pooley KA, Johnatty SE, Beesley J, Michailidou K, Tyrer JP, Edwards SL, Pickett HA, Shen HC, Smart CE, *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.
49. Vinagre J, Almeida A, Pópulo H, Batista R, Lyra J, Pinto V, Coelho R, Celestino R, Prazeres H, Lima L, *et al*: Frequency of TERT promoter mutations in human cancers. Nat Commun 4: 2185, 2013.
50. Bolton KL, Tyrer J, Song H, Ramus SJ, Notaridou M, Jones C, Sher T, Gentry-Maharaj A, Wozniak E, Tsai YY, *et al*: Common variants at 19p13 are associated with susceptibility to ovarian cancer. Nat Genet 42: 880-884, 2010.
51. Shao G, Patterson-Fortin J, Messick TE, Feng D, Shanbhag N, Wang Y and Greenberg RA: MERIT40 controls BRCA1-Rap80 complex integrity and recruitment to DNA double-strand breaks. Genes Dev 23: 740-754, 2009.
52. Wang B, Hurov K, Hofmann K and Elledge SJ: NBA1, a new player in the Brca1 A complex, is required for DNA damage resistance and checkpoint control. Genes Dev 23: 729-739, 2009.
53. Feng L, Huang J and Chen J: MERIT40 facilitates BRCA1 localization and DNA damage repair. Genes Dev 23: 719-728, 2009.
54. Lawrenson K, Kar S, McCue K, Kuchenbaeker K, Michailidou K, Tyrer J, Beesley J, Ramus SJ, Li Q, Delgado MK, *et al*: Functional mechanisms underlying pleiotropic risk alleles at the 19p13.1 breast-ovarian cancer susceptibility locus. Nat Commun 7: 12675, 2016.
55. Kar SP, Tyrer JP, Li Q, Lawrenson K, Aben KK, Anton-Culver H, Antonenkova N, Chenevix-Trench G; Australian Cancer Study; Australian Ovarian Cancer Study Group, *et al*: Network-based integration of GWAS and gene expression identifies a HOX-Centric network associated with serous ovarian cancer risk. Cancer Epidemiol Biomarkers Prev 24: 1574-1584, 2015.
56. Lawrenson K, Li Q, Kar S, Seo JH, Tyrer J, Spindler TJ, Lee J, Chen Y, Karst A, Drapkin R, *et al*: Cis-eQTL analysis and functional validation of candidate susceptibility genes for high-grade serous ovarian cancer. Nat Commun 6: 8234, 2015.
57. Ghoussaini M, Song H, Koessler T, Al Olama AA, Kote-Jarai Z, Driver KE, Pooley KA, Ramus SJ, Kjaer SK, Hogdall E, *et al*: Multiple loci with different cancer specificities within the 8q24 gene desert. J Natl Cancer Inst 100: 962-966, 2008.

58. Panagiotou OA, Evangelou E and Ioannidis JP: Genome-wide significant associations for variants with minor allele frequency of 5% or less-an overview: A HuGE review. Am J Epidemiol 172: 869-889, 2010.
59. Stephens M and Balding DJ: Bayesian statistical methods for genetic association studies. Nat Rev Genet 10: 681-690, 2009.
60. International HapMap Consortium: A haplotype map of the human genome. Nature 437: 1299-1320, 2005.
61. Pe'er I, Yelensky R, Altshuler D and Daly MJ: Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. Genet Epidemiol 32: 381-385, 2008.
62. Fadista J, Manning AK, Florez JC and Groop L: The (in)famous GWAS P-value threshold revisited and updated for low-frequency variants. Eur J Hum Genet 24: 1202-1205, 2016.
63. Kar SP, Berchuck A, Gayther SA, Goode EL, Moysich KB, Pearce CL, Ramus SJ, Schildkraut JM, Sellers TA and Pharoah PDP: Common genetic variation and susceptibility to ovarian cancer: Current insights and future directions. Cancer Epidemiol Biomarkers Prev 27: 395-404, 2018.
64. Jones MR, Kamara D, Karlan BY, Pharoah PDP and Gayther SA: Genetic epidemiology of ovarian cancer and prospects for polygenic risk prediction. Gynecol Oncol 147: 705-713, 2017.
65. Elmas A, Ou Yang TH, Wang X and Anastassiou D: Discovering genome-wide tag SNPs based on the mutual information of the variants. PLoS One 11: e0167994, 2016.
66. Johnson GC, Esposito L, Barratt BJ, Smith AN, Heward J, Di Genova G, Ueda H, Cordell HJ, Eaves IA, Dudbridge F, *et al*: Haplotype tagging for the identification of common disease genes. Nat Genet 29: 233-237, 2001.
67. Earp M, Winham SJ, Larson N, Permuth JB, Sicotte H, Chien J, Anton-Culver H, Bandera EV, Berchuck A, Cook LS, *et al*: A targeted genetic association study of epithelial ovarian cancer susceptibility. Oncotarget 7: 7381-7389, 2016.
68. Song H, Ramus SJ, Tyrer J, Bolton KL, Gentry-Maharaj A, Wozniak E, Anton-Culver H, Chang-Claude J, Cramer DW, DiCioccio R, *et al*: A genome-wide association study identifies a new ovarian cancer susceptibility locus on 9p22.2. Nat Genet 41: 996-1000, 2009.
69. Goode EL, Chenevix-Trench G, Song H, Ramus SJ, Notaridou M, Lawrenson K, Widschwendter M, Vierkant RA, Larson MC, Kjaer SK, *et al*: A genome-wide association study identifies susceptibility loci for ovarian cancer at 2q31 and 8q24. Nat Genet 42: 874-879, 2010.
70. Permuth-Wey J, Lawrenson K, Shen HC, Velkova A, Tyrer JP, Chen Z, Lin HY, Chen YA, Tsai YY, Qu X, *et al*: Identification and molecular characterization of a new ovarian cancer susceptibility locus at 17q21.31. Nat Commun 4: 1627, 2013.
71. Chen K, Ma H, Li L, Zang R, Wang C, Song F, Shi T, Yu D, Yang M, Xue W, *et al*: Genome-wide association study identifies new susceptibility loci for epithelial ovarian cancer in Han Chinese women. Nat Commun 5: 4682, 2014.
72. Romano RA, Li H, Tummala R, Maul R and Sinha S: Identification of Basonuclin2, a DNA-binding zinc-finger protein expressed in germ tissues and skin keratinocytes. Genomics 83: 821-833, 2004.
73. Buckley MA, Woods NT, Tyrer JP, Mendoza-Fandiño G, Lawrenson K, Hazelett DJ, Najafabadi HS, Gjyshi A, Carvalho RS, Lyra PC Jr, *et al*: Functional analysis and fine mapping of the 9p22.2 ovarian cancer susceptibility locus. Cancer Res 79: 467-481, 2019.
74. Wentzensen N, Black A, Jacobs K, Yang HP, Berg CD, Caporaso N, Peters U, Ragard L, Buys SS, Chanock S and Hartge P: Genetic variation on 9p22 is associated with abnormal ovarian ultrasound results in the prostate, lung, colorectal, and ovarian cancer screening trial. PLoS One 6: e21731, 2011.
75. Vigorito E, Kuchenbaecker KB, Beesley J, Adlard J, Agnarsson BA, Andrulis IL, Arun BK, Barjhoux L, Belotti M, Benitez J, *et al*: Fine-scale mapping at 9p22.2 identifies candidate causal variants that modify ovarian cancer risk in BRCA1 and BRCA2 mutation carriers. PLoS One 11: e0158801, 2016.
76. Carter H, Marty R, Hofree M, Gross AM, Jensen J, Fisch KM, Wu X, DeBoever C, Van Nostrand EL, Song Y, *et al*: Interaction landscape of inherited polymorphisms with somatic events in cancer. Cancer Discov 7: 410-423, 2017.
77. Permuth JB, Pirie A, Ann Chen Y, Lin HY, Reid BM, Chen Z, Monteiro A, Dennis J, Mendoza-Fandino G; AOCS Study Group, *et al*: Exome genotyping arrays to identify rare and low frequency variants associated with epithelial ovarian cancer risk. Hum Mol Genet 25: 3600-3612, 2016.
78. Fong PC, Boss DS, Yap TA, Tutt A, Wu P, Mergui-Roelvink M, Mortimer P, Swaisland H, Lau A, O'Connor MJ, *et al*: Inhibition of poly(ADP-Ribose) polymerase in tumors from BRCA mutation carriers. N Engl J Med 361: 123-134, 2009.
79. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, *et al*: The sequence of the human genome. Science 291: 1304-1351, 2001.
80. Vaser R, Adusumalli S, Leng SN, Sikic M and Ng PC: SIFT missense predictions for genomes. Nat Protoc 11: 1-9, 2016.
81. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS and Sunyaev SR: A method and server for predicting damaging missense mutations. Nat Methods 7: 248-249, 2010.
82. Kang UB, Ahn Y, Lee JW, Kim YH, Kim J, Yu MH, Noh DY and Lee C: Differential profiling of breast cancer plasma proteome by isotope-coded affinity tagging method reveals biotinidase as a breast cancer biomarker. BMC Cancer 10: 114, 2010.
83. Huang L, Zheng M, Zhou QM, Zhang MY, Jia WH, Yun JP and Wang HY: Identification of a gene-expression signature for predicting lymph node metastasis in patients with early stage cervical carcinoma. Cancer 117: 3363-3373, 2011.
84. Sun X, Frierson HF, Chen C, Li C, Ran Q, Otto KB, Cantarel BL, Vessella RL, Gao AC, Petros J, *et al*: Frequent somatic mutations of the transcription factor ATBF1 in human prostate cancer. Nat Genet 37: 407-412, 2005.
85. Walker CJ, Miranda MA, O'Hern MJ, McElroy JP, Coombes KR, Bundschuh R, Cohn DE, Mutch DG and Goodfellow PJ: Patterns of CTCF and ZFHX3 mutation and associated outcomes in endometrial cancer. J Natl Cancer Inst 107: djv249, 2015.
86. Rafnar T, Gudbjartsson DF, Sulem P, Jonasdottir A, Sigurdsson A, Jonasdottir A, Besenbacher S, Lundin P, Stacey SN, Gudmundsson J, *et al*: Mutations in BRIP1 confer high risk of ovarian cancer. Nat Genet 43: 1104-1107, 2011.
87. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, Gu B, Hart J, Hoffman D, Jang W, *et al*: ClinVar: Improving access to variant interpretations and supporting evidence. Nucleic Acids Res 46: D1062-D1067, 2018.
88. Henderson BE and Feigelson HS: Hormonal carcinogenesis. Carcinogenesis 21: 427-433, 2000.
89. Henderson BE, Ross RK, Pike MC and Casagrande JT: Endogenous hormones as a major factor in human cancer. Cancer Res 42: 3232-3239, 1982.
90. Couch FJ, Wang X, McGuffog L, Lee A, Olswold C, Kuchenbaecker KB, Soucy P, Fredericksen Z, Barrowdale D, Dennis 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.
91. Al Olama AA, Kote-Jarai Z, Berndt SI, Conti DV, Schumacher F, Han Y, Benlloch S, Hazelett DJ, Wang Z, Saunders E, *et al*: A meta-analysis of 87,040 individuals identifies 23 new susceptibility loci for prostate cancer. Nat Genet 46: 1103-1109, 2014.
92. Kar SP, Beesley J, Amin Al Olama A, Michailidou K, Tyrer J, Kote-Jarai Z, Lawrenson K, Lindstrom S, Ramus SJ, Thompson DJ, *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.
93. Schildkraut JM, Goode EL, Clyde MA, Iversen ES, Moorman PG, Berchuck A, Marks JR, Lissowska J, Brinton L, Peplonska B, *et al*: Single nucleotide polymorphisms in the TP53 region and susceptibility to invasive epithelial ovarian cancer. Cancer Res 69: 2349-2357, 2009.
94. Johnatty SE, Tyrer JP, Kar S, Beesley J, Lu Y, Gao B, Fasching PA, Hein A, Ekici AB, Beckmann MW, *et al*: Genome-wide analysis identifies novel loci associated with ovarian cancer outcomes: Findings from the ovarian cancer association consortium. Clin Cancer Res 21: 5264-5276, 2015.
95. Lu Y, Beeghly-Fadiel A, Wu L, Guo X, Li B, Schildkraut JM, Im HK, Chen YA, Permuth JB, Reid BM, *et al*: A transcriptome-wide association study among 97,898 women to identify candidate susceptibility genes for epithelial ovarian cancer risk. Cancer Res 78: 5419-5430, 2018.