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.

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...
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-consorium 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
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 \(P<5.0 \times 10^{-8}\) and \(P>5.0 \times 10^{-8}\), 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 \(HFN1B\) 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.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Locus</th>
<th>Position$^a$</th>
<th>EA$^b$</th>
<th>EAF$^c$</th>
<th>OR</th>
<th>CI$_{0.5}$</th>
<th>P-value/BI$^d$</th>
<th>Gene$^e$</th>
<th>eQTL$^f$</th>
<th>Refs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>EnOC</td>
<td>rs55025179</td>
<td>05q12.3</td>
<td>del</td>
<td>0.56</td>
<td>1.18</td>
<td>(1.11-1.26)</td>
<td>4.5x10^{-8}</td>
<td>MAST4</td>
<td></td>
<td>(31)</td>
</tr>
<tr>
<td>EnOC</td>
<td>rs12951053</td>
<td>17p13.1</td>
<td>7674089</td>
<td>C</td>
<td>0.18</td>
<td>1.31</td>
<td>(1.05-1.62)</td>
<td>109.6</td>
<td>TP53</td>
<td>(93)</td>
</tr>
<tr>
<td>EOC</td>
<td>rs199661266</td>
<td>17q11.2</td>
<td>30854203</td>
<td>insT</td>
<td>0.28</td>
<td>0.91</td>
<td>(0.88-0.94)</td>
<td>5.4x10^{-9}</td>
<td>ATAD5, AC130324.2</td>
<td>(20)</td>
</tr>
<tr>
<td>EOC</td>
<td>rs587778134</td>
<td>17q25.2</td>
<td>61776459</td>
<td>insTT</td>
<td>0.02</td>
<td>8.13</td>
<td>(4.74-13.95)</td>
<td>2.8x10^{-14}</td>
<td>BRIP1</td>
<td>(86)</td>
</tr>
<tr>
<td>EOC-BRCA1</td>
<td>rs2165109</td>
<td>02q13</td>
<td>111061081</td>
<td>C</td>
<td>0.25</td>
<td>1.09</td>
<td>(1.05-1.12)</td>
<td>4.2x10^{-8}</td>
<td>ACOXL</td>
<td>(31)</td>
</tr>
<tr>
<td>EOC-BRCA1</td>
<td>rs4691139</td>
<td>04q32.3</td>
<td>164987569</td>
<td>G</td>
<td>0.47</td>
<td>1.20</td>
<td>(1.17-1.38)</td>
<td>3.4x10^{-8}</td>
<td>TRIM61, AC106872.4</td>
<td>(90)</td>
</tr>
<tr>
<td>EOC-BRCA1</td>
<td>rs1998651</td>
<td>17q24.21</td>
<td>127805637</td>
<td>G</td>
<td>0.46</td>
<td>1.08</td>
<td>(1.05-1.11)</td>
<td>3.5x10^{-9}</td>
<td>PV1</td>
<td>(31)</td>
</tr>
<tr>
<td>EOC-BRCA1</td>
<td>rs12951053</td>
<td>17p13.1</td>
<td>7674089</td>
<td>C</td>
<td>0.18</td>
<td>1.19</td>
<td>(1.01-1.38)</td>
<td>47.8</td>
<td>TP53</td>
<td>(93)</td>
</tr>
<tr>
<td>EOC-BRCA1</td>
<td>rs1361104</td>
<td>08q24</td>
<td>45349306</td>
<td>G</td>
<td>0.19</td>
<td>1.27</td>
<td>(1.17-1.38)</td>
<td>3.1x10^{-8}</td>
<td>PLEKHM1</td>
<td>(90)</td>
</tr>
<tr>
<td>EOC-BRCA1</td>
<td>rs988651</td>
<td>17q21</td>
<td>46710944</td>
<td>A</td>
<td>0.23</td>
<td>1.25</td>
<td>(1.16-1.35)</td>
<td>3.8x10^{-14}</td>
<td>HOXD9</td>
<td>(69)</td>
</tr>
<tr>
<td>HGSC</td>
<td>rs5755777</td>
<td>02q31.1</td>
<td>176178498</td>
<td>T</td>
<td>0.68</td>
<td>1.15</td>
<td>NR</td>
<td>9.0x10^{-14}</td>
<td>HAGLR, HAGLROS</td>
<td>(56)</td>
</tr>
<tr>
<td>HGSC</td>
<td>rs7651446</td>
<td>02q25</td>
<td>156689208</td>
<td>A</td>
<td>0.05</td>
<td>1.44</td>
<td>(1.35-1.53)</td>
<td>1.5x10^{-28}</td>
<td>TIPARP</td>
<td>(37)</td>
</tr>
<tr>
<td>HGSC</td>
<td>rs78561123</td>
<td>02q25.31</td>
<td>156747145</td>
<td>C</td>
<td>0.06</td>
<td>1.45</td>
<td>(1.32-1.59)</td>
<td>3.0x10^{-15}</td>
<td>LINCO0886</td>
<td>(77)</td>
</tr>
<tr>
<td>HGSC</td>
<td>rs62273959</td>
<td>02q25.31</td>
<td>156852891</td>
<td>A</td>
<td>0.06</td>
<td>1.45</td>
<td>(1.32-1.59)</td>
<td>3.5x10^{-14}</td>
<td>LEKR1</td>
<td>(77)</td>
</tr>
<tr>
<td>HGSC</td>
<td>rs10069690</td>
<td>05p15.33</td>
<td>1279675</td>
<td>T</td>
<td>0.26</td>
<td>1.15</td>
<td>(1.11-1.20)</td>
<td>3.1x10^{-11}</td>
<td>TERT</td>
<td>(48)</td>
</tr>
<tr>
<td>HGSC</td>
<td>rs7726159</td>
<td>05p15.33</td>
<td>1282204</td>
<td>A</td>
<td>0.34</td>
<td>1.12</td>
<td>(1.09-1.16)</td>
<td>3.4x10^{-9}</td>
<td>TERT</td>
<td>(47)</td>
</tr>
<tr>
<td>HGSC</td>
<td>rs11782652</td>
<td>08q24</td>
<td>81741409</td>
<td>G</td>
<td>0.07</td>
<td>1.24</td>
<td>(1.16-1.33)</td>
<td>7.0x10^{-10}</td>
<td>CHMP4C</td>
<td>(37)</td>
</tr>
<tr>
<td>HGSC</td>
<td>rs10088218</td>
<td>08q24</td>
<td>128531703</td>
<td>A</td>
<td>0.13</td>
<td>0.76</td>
<td>(0.70-0.81)</td>
<td>8.0x10^{-15}</td>
<td>MYC</td>
<td>(69)</td>
</tr>
<tr>
<td>HGSC</td>
<td>rs7814937</td>
<td>08q24</td>
<td>128529479</td>
<td>C</td>
<td>0.11</td>
<td>1.18</td>
<td>(1.13-1.24)</td>
<td>NR</td>
<td>LINCO0824</td>
<td>(55)</td>
</tr>
<tr>
<td>HGSC</td>
<td>rs3814113</td>
<td>09p22</td>
<td>16915023</td>
<td>C</td>
<td>0.37</td>
<td>0.77</td>
<td>(0.73-0.83)</td>
<td>4.1x10^{-21}</td>
<td>BNC2</td>
<td>(68)</td>
</tr>
<tr>
<td>HGSC</td>
<td>rs7084454</td>
<td>10p12.31</td>
<td>21532345</td>
<td>A</td>
<td>0.24</td>
<td>1.10</td>
<td>(1.06-1.14)</td>
<td>NR</td>
<td>MLLT10</td>
<td>(55)</td>
</tr>
<tr>
<td>HGSC</td>
<td>rs2287498</td>
<td>17p13.1</td>
<td>7689242</td>
<td>A</td>
<td>0.14</td>
<td>1.30</td>
<td>(1.07-1.57)</td>
<td>165.7</td>
<td>WRAP53, TP53</td>
<td>(93)</td>
</tr>
<tr>
<td>HGSC</td>
<td>rs12951053</td>
<td>17p13.1</td>
<td>7674089</td>
<td>C</td>
<td>0.18</td>
<td>1.19</td>
<td>(1.01-1.38)</td>
<td>47.8</td>
<td>TP53</td>
<td>(93)</td>
</tr>
<tr>
<td>HGSC</td>
<td>rs757210</td>
<td>17q12</td>
<td>37736525</td>
<td>A</td>
<td>0.37</td>
<td>1.12</td>
<td>(1.08-1.17)</td>
<td>8.1x10^{-10}</td>
<td>HNFI1B</td>
<td>(37)</td>
</tr>
<tr>
<td>HGSC</td>
<td>rs7405776</td>
<td>17q12</td>
<td>37733029</td>
<td>A</td>
<td>0.37</td>
<td>1.25</td>
<td>(1.17-1.34)</td>
<td>2.3x10^{-11}</td>
<td>HNFI1B</td>
<td>(36)</td>
</tr>
<tr>
<td>HGSC</td>
<td>rs2077606</td>
<td>17q21.31</td>
<td>45452177</td>
<td>A</td>
<td>0.08</td>
<td>1.15</td>
<td>(1.12-1.19)</td>
<td>NR</td>
<td>PLEKHM1</td>
<td>(55)</td>
</tr>
<tr>
<td>HGSC</td>
<td>rs2906000</td>
<td>17q21.31</td>
<td>4546987</td>
<td>C</td>
<td>0.18</td>
<td>1.16</td>
<td>(1.12-1.20)</td>
<td>3.3x10^{-10}</td>
<td>PLEKHM1, AC091132.2</td>
<td>(70)</td>
</tr>
<tr>
<td>HGSC</td>
<td>rs12942666</td>
<td>17q21.31</td>
<td>45422473</td>
<td>G</td>
<td>0.22</td>
<td>1.15</td>
<td>(1.11-1.20)</td>
<td>1.0x10^{-9}</td>
<td>PLEKHM1, ARHGAP27</td>
<td>(70)</td>
</tr>
<tr>
<td>HGSC</td>
<td>rs1052587</td>
<td>17q21.31</td>
<td>46025238</td>
<td>C</td>
<td>0.22</td>
<td>1.12</td>
<td>(1.08-1.17)</td>
<td>4.6x10^{-8}</td>
<td>MAPT</td>
<td>(70)</td>
</tr>
<tr>
<td>HGSC</td>
<td>rs7218345</td>
<td>17q21.32</td>
<td>48425805</td>
<td>T</td>
<td>0.31</td>
<td>1.12</td>
<td>(1.08-1.16)</td>
<td>NR</td>
<td>SKAP1</td>
<td>(55)</td>
</tr>
<tr>
<td>HGSC</td>
<td>rs2363956</td>
<td>19p13.11</td>
<td>17283315</td>
<td>T</td>
<td>0.49</td>
<td>1.16</td>
<td>(1.11-1.21)</td>
<td>3.8x10^{-11}</td>
<td>ANKLE1</td>
<td>(50)</td>
</tr>
<tr>
<td>SNP</td>
<td>Position</td>
<td>Gene</td>
<td>Refs</td>
<td>EAF</td>
<td>OR</td>
<td>95% CI</td>
<td>P-value/BF</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>----------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>--------</td>
<td>------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs8170</td>
<td>19p13.11</td>
<td>BABAM1, USHBP1, ABHD8, BABAM1</td>
<td>50</td>
<td>0.19</td>
<td>1.18</td>
<td>1.12-1.25</td>
<td>2.7x10^-9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs4808075</td>
<td>19p13.11</td>
<td>BABAM1, USHBP1, ABHD8, BABAM1</td>
<td>54</td>
<td>0.3</td>
<td>1.19</td>
<td>1.14-1.23</td>
<td>9.2x10^-20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs711530</td>
<td>19p13.11</td>
<td>PAX8, USHBP1, ABHD8</td>
<td>36</td>
<td>0.21</td>
<td>1.34</td>
<td>1.21-1.49</td>
<td>3.3x10^-8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs711530</td>
<td>19p13.11</td>
<td>PAX8</td>
<td>38</td>
<td>1.34</td>
<td>1.21-1.49</td>
<td>3.3x10^-8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs711530</td>
<td>19p13.11</td>
<td>ABHD8</td>
<td>42</td>
<td>1.30</td>
<td>1.20-1.40</td>
<td>7.5x10^-12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs711530</td>
<td>19p13.11</td>
<td>ABCD3, HOXD9</td>
<td>56</td>
<td>1.30</td>
<td>1.20-1.40</td>
<td>7.5x10^-12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs4808075</td>
<td>19p13.11</td>
<td>ABHD8, BABAM1</td>
<td>54</td>
<td>0.3</td>
<td>1.19</td>
<td>1.14-1.23</td>
<td>9.2x10^-20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs320203</td>
<td>09q31.1</td>
<td>GRIN3A</td>
<td>31</td>
<td>0.88</td>
<td>1.29</td>
<td>1.18-1.41</td>
<td>1.7x10^-8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs688187</td>
<td>19q13.2</td>
<td>IFNL3, IFNL4</td>
<td>38</td>
<td>0.32</td>
<td>0.67</td>
<td>0.60-0.75</td>
<td>6.8x10^-13</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table I. continued.**

<table>
<thead>
<tr>
<th>Gene(s) in closest proximity to the most significant ovarian cancer risk variant at each locus.</th>
</tr>
</thead>
</table>
| eQTL association with rs4808075, that has significant expression quantitative trait loci (eQTL) association with BABAM1, leads to stabilisation of the breast cancer susceptibility protein (BRCA1) 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).

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 $P \leq 5.0 \times 10^{-8}$ for MAF $\geq 0.1\%$ and $1 \times 10^{-8}$ for MAF $\leq 0.05\%$ (condi-tions: 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 $P \leq 0.05$. Variants meeting the threshold criteria discussed above ($P \leq 5.0 \times 10^{-8}$) 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 SNP will be within 1 Mb of at least one other SNP (54).

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 (r²=0.90) with rs7651446 in the aforementioned TIPARP gene (77). Finally, it is worth noting that rs58778134 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>
<thead>
<tr>
<th>SNP</th>
<th>Locus</th>
<th>Subtype</th>
<th>Amino acid change (Human SIFT PolyPhen-2 PolyPhen-2 clinVar clinVar Genome variation society)</th>
<th>SIFT scoreb</th>
<th>PolyPhen-2 prediction scoreb</th>
<th>prediction clinical significance</th>
<th>clinical significance phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs13831468</td>
<td>1p36.33</td>
<td>EOC</td>
<td>rs13831468</td>
<td>rs13831468</td>
<td>AGRN</td>
<td>NP_001223024:p.Ala237Ser</td>
<td>0.15</td>
</tr>
<tr>
<td>rs61757604</td>
<td>2p22.2</td>
<td>EOC</td>
<td>rs61757604</td>
<td>rs61757604</td>
<td>DHX57</td>
<td>NP_0013692:2:p.Gly49Ser</td>
<td>0.06</td>
</tr>
<tr>
<td>rs200337373</td>
<td>9p23.1</td>
<td>EOC</td>
<td>rs200337373</td>
<td>rs200337373</td>
<td>BTD</td>
<td>NP_001268649:p.Arg202Asn</td>
<td>0.09</td>
</tr>
<tr>
<td>rs7357391</td>
<td>0p11.2</td>
<td>EOC</td>
<td>rs7357391</td>
<td>rs7357391</td>
<td>ACTB1.2</td>
<td>NP_140962.2:p.Asp519Asp</td>
<td>0.04</td>
</tr>
<tr>
<td>rs139761238</td>
<td>0q25.2</td>
<td>EOC</td>
<td>rs139761238</td>
<td>rs139761238</td>
<td>SYNE1</td>
<td>NP_001268654:p.Asp202Asn</td>
<td>0.42</td>
</tr>
<tr>
<td>rs61757604</td>
<td>2p22.2</td>
<td>EOC</td>
<td>rs61757604</td>
<td>rs61757604</td>
<td>DHX57</td>
<td>NP_0013692:2:p.Gly49Ser</td>
<td>0.06</td>
</tr>
<tr>
<td>rs200337373</td>
<td>9p23.1</td>
<td>EOC</td>
<td>rs200337373</td>
<td>rs200337373</td>
<td>BTD</td>
<td>NP_001268649:p.Arg202Asn</td>
<td>0.09</td>
</tr>
<tr>
<td>rs7357391</td>
<td>0p11.2</td>
<td>EOC</td>
<td>rs7357391</td>
<td>rs7357391</td>
<td>ACTB1.2</td>
<td>NP_140962.2:p.Asp519Asp</td>
<td>0.04</td>
</tr>
<tr>
<td>rs139761238</td>
<td>0q25.2</td>
<td>EOC</td>
<td>rs139761238</td>
<td>rs139761238</td>
<td>SYNE1</td>
<td>NP_001268654:p.Asp202Asn</td>
<td>0.42</td>
</tr>
<tr>
<td>rs61757604</td>
<td>2p22.2</td>
<td>EOC</td>
<td>rs61757604</td>
<td>rs61757604</td>
<td>DHX57</td>
<td>NP_0013692:2:p.Gly49Ser</td>
<td>0.06</td>
</tr>
</tbody>
</table>

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

Subtype. Amino acid change (Human SIFT PolyPhen-2 PolyPhen-2 clinVar clinVar Genome variation society) SIFT scoreb PolyPhen-2 prediction scoreb PolyPhen-2 clinVar genome variation society. Prediction and clinical significance. **Pathogenic** = Biologically plausible deleterious variant predicted to cause disease. **Likely pathogenic** = Biologically plausible deleterious variant predicted to cause disease with an evidence score of not less than 0.05. **Benign** = Biologically plausible neutral variant predicted to cause disease. **Likely benign** = Biologically plausible neutral variant predicted to cause disease with an evidence score of not less than 0.05. **Uncertain** = No biological relevance is assigned.

- **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; EOC, 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.**

**Gene(s) in closest proximity to the most significant ovarian cancer risk variant at each locus.**
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 OC, 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


