Genetic overlap between endometriosis and endometrial cancer: evidence from cross-disease genetic correlation and GWAS meta-analyses


1Department of Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia
2Wellcome Trust for Human Genetics, University of Oxford, Oxford, UK
3Department of Biostatistics, University of Liverpool, Liverpool, UK
4Wellcome Trust Centre for Human Genetics and Oxford NIHR Biomedical Research Centre, University of Oxford, Oxford, UK
5Department of Clinical Genetics, St George’s, University of London, London, UK
6Institute for Molecular Bioscience, University of Queensland, Brisbane, Queensland, Australia
7Hunter Medical Research Institute, John Hunter Hospital, Newcastle, New South Wales, Australia
8Centre for Clinical Epidemiology and Biostatistics, School of Medicine and Public Health, University of Newcastle, Callaghan, New South Wales, Australia
9Division of Molecular Medicine, Pathology North, John Hunter Hospital, Newcastle, New South Wales, Australia
10Discipline of Medical Genetics, School of Biomedical Sciences and Pharmacy, Faculty of Health, University of Newcastle, Callaghan, New South Wales, Australia
11QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia
12Department of Gynaecology and Obstetrics, University Hospital Erlangen, Comprehensive Cancer Center Erlangen-EMN, Friedrich-Alexander University Erlangen-Nuremberg, Erlangen, Germany
13David Geffen School of Medicine, Department of Medicine Division of Hematology and Oncology, University of California at Los Angeles, Los Angeles, California
14Institute of Human Genetics, University Hospital Erlangen, Comprehensive Cancer Center Erlangen-EMN, Friedrich-Alexander University Erlangen-Nuremberg, Erlangen, Germany
15Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
16Gynaecology Research Unit, Hannover Medical School, Hannover, Germany
17Department of Gynaecology, Jena University Hospital - Friedrich Schiller University, Jena, Germany
18VIB Center for Cancer Biology, VIB, Leuven, Belgium
19Laboratory for Translational Genetics, Department of Human Genetics, University of Leuven, Leuven, Belgium
20Department of Obstetrics and Gynecology, Division of Gynecologic Oncology, University Hospitals KU Leuven, University of Leuven, Leuven, Belgium
21Vesalius Research Center, VIB, Leuven, Belgium
22Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota
23Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota
24Department of Obstetrics and Gynecology, Division of Gynecologic Oncology, Mayo Clinic, Rochester, Minnesota
25Centre for Cancer biomarkers, Department of Clinical Science, University of Bergen, Bergen, Norway
26Department of Obstetrics and Gynecology, Haukeland University Hospital, Bergen, Norway
27Centre for Information Based Medicine, University of Newcastle, Callaghan, New South Wales, Australia
28School of Medicine and Public Health, University of Newcastle, Callaghan, New South Wales, Australia
29Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden
30Clinical Genetics, Karolinska Institutet, Stockholm, Sweden
31Department of Women’s and Children’s Health, Karolinska Institutet, Stockholm, Sweden

© 2018 The Authors. Cancer Medicine published by John Wiley & Sons Ltd.
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
Endometriosis and Endometrial Cancer

Introduction

Endometriosis (defined as tissue resembling endometrium in extrauterine sites) and endometrial cancer (cancer of the uterine corpus) are serious gynecological diseases with major impacts on the quality of life of affected women. Endometriosis is a relatively common disease affecting 6–10% of women of reproductive age and 35–50% of infertile women [1, 2]. Affected women commonly
experience severe menstrual pain, pelvic pain, subfertility or infertility, and bowel-related symptoms. Endometrial cancer is the most common invasive gynecological cancer in Australia, ranking sixth for incident cancers in women [3]. This disease is associated with significant morbidity due to surgery and radiotherapy [4], and treatment is further complicated by the fact that most patients present at relatively older age and with major comorbidities, notably obesity and diabetes. Finding the genes and pathways underlying these complex diseases is an essential step toward developing better diagnostic and therapeutic tools for both diseases. Both diseases are known to have a genetic component, with twin studies showing heritability for both diseases. Both diseases are known to have a genetic component, with twin studies showing heritability for both diseases. Both diseases are known to have a genetic component, with twin studies showing heritability for both diseases. Both diseases are known to have a genetic component, with twin studies showing heritability for both diseases.

Genetic overlap between endometriosis and endometrial cancer: datasets and analyses

This study utilized data from four previously published genetic datasets for endometriosis and endometrial cancer (outlined below and in the following section; Table 1) [26–28]. Three of the datasets were GWAS datasets, genotyped using Illumina 610Quad and 670Quad BeadChips (Illumina Inc, San Diego, CA) and containing data for 462,430 SNPs in common between them. Of these, the
endometriosis GWAS dataset included 3194 Australian (QIMR Berghofer Medical Research Institute (QIMR)) and UK (Oxford) women with surgically confirmed endometriosis as cases [26]. The first endometrial cancer GWAS dataset included 1262 Australian (ANECS) and UK (SEARCH) endometrioid subtype endometrial cancer patients [27], and the second (NSECG) included 795 UK endometrial cancer cases and 895 nonoverlapping controls [28]. All endometrial cancer cases were histologically confirmed to be invasive cancer of the endometrium lining [27]. As previously published, the endometriosis and ANECS-SEARCH endometrial cancer GWAS datasets included the same sets of controls—1870 Australian controls and 5190 UK Wellcome Trust Case Control Consortium (WTCCC) controls. Hence to avoid overlapping control samples in this study, the controls were redistributed as follows: The 1870 Australian controls and two-third of the WTCCC controls (n = 3460, randomly assigned) were included in the endometriosis GWAS dataset, while an additional set of 1241 Australian controls [28] and the remaining one-third of the WTCCC controls (n = 1730) were included in the ANECS-SEARCH endometrial cancer GWAS dataset.

Following quality control [26–28], association analyses were performed for each GWAS dataset using PLINK [29]. Australian and UK cases and controls were analyzed as separate strata within the same GWAS for the endometriosis and ANECS-SEARCH endometrial cancer datasets, adjusting for the first two (endometriosis, ANECS, NSECG) or three (SEARCH) principal components of the genomic kinship matrix [26–28]. The summary results for the ANECS-SEARCH and NSECG datasets were then included in an inverse variance, fixed effects meta-analysis performed using METAL [30], to produce one set of endometrial cancer GWAS results. A fixed effect model was considered more appropriate than a random effect model as our hypothesis is that a proportion of SNPs will be associated with both diseases with the same direction of effect, and a fixed effect model is conservative given no expectation that the effect size is similar. The degree of genetic overlap between endometriosis and endometrial cancer was then examined using two programs that test the degree of genetic correlation/concordance between diseases using GWAS summary results (individual SNP effect sizes and P-values), SNP effect concordance analysis (SECA) [31] and LD Score regression [32].

To account for linkage disequilibrium (LD) between SNPs, SECA employs a “P-value informed” SNP clumping procedure to extract a subset of independent SNPs [31] (23,817 SNPs for the current analysis). These SNPs are then partitioned into 12 P-value “bins” (e.g., P ≤ 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1.0) for each disease. Using the default settings, a number of binomial and Fisher exact tests were performed on SNPs across all bins (12 × 12 bins = 144 SNP subset combinations), and on SNP subsets within bins (see Results), to determine the degree to which individual SNPs are concordant in their P-value level and direction of effect across two diseases, which can indicate the presence of genetic concordance and SNP pleiotropy [31]. For these analyses, the endometriosis dataset was designated as Dataset 1 and endometrial cancer as Dataset 2.

Taking a different approach, cross-trait LD Score regression utilizes the presence of LD, calculating an LD score between SNPs within a 1 cM window and then regressing the product of the SNP association results (z scores) from the two diseases against the LD score [32]. Following the recommendations at https://github.com/bulik/ldsc/wiki/Heritability-and-Genetic-Correlation, the “–no-intercept” option was used to constrain the LD Score regression intercept to 0 as there was no sample overlap between the two disease datasets.

### Cross-disease meta-analysis between endometriosis and endometrial cancer

The cross-disease meta-analysis was performed using an inverse variance, fixed effects model in METAL [30] to search for genetic loci potentially contributing to the increased risk of both endometriosis and endometrial cancer.
cancer. Heterogeneity was assessed using Cochran’s Q-test. The results for the top SNPs \((P \leq 10^{-3})\) from the endometriosis–endometrial cancer meta-analysis were then compared with results for the same SNPs from the fourth dataset included in this study, a separate, independent sample of 4402 endometrial cancer cases and 28,758 controls genotyped at 211,155 SNPs using a custom Illumina Infinium iSelect array by the Collaborative Oncological Gene-environment Study (“iCOGS”) [33, 34]. SNPs not included on the iCOGS array were imputed (including all SNPs within 1 Mb of the target SNP) using IMPUTE(v2) software [35] and the 1000 Genomes Project (2012 release) as the reference panel [9]. Imputation quality scores ranged from 0.34 to 1.00. Association testing on the iCOGS SNPs was performed using SNPTEST (v2) [36] employing frequentist tests with a logistic regression model adjusting for eight separate strata and the first 10 principal components \([9, 28]\). These results were then included in the replication meta-analysis, which included all four datasets and was conducted as described for the cross-disease meta-analysis above.

**Results**

**Genetic overlap between endometriosis and endometrial cancer**

Genetic correlation analyses of GWAS datasets for endometriosis and endometrial cancer revealed the presence of weak to moderate, but significant, genetic overlap between the two diseases. The LD Score regression analysis indicated moderate but significant genetic correlation \((r_g = 0.23, P = 9.3 \times 10^{-3})\) between the two diseases \((r_g = 0.23, P = 9.3 \times 10^{-3})\). The SECA primary test for the overlap of associated effects, including all 144 SNP subsets, revealed more subsets than expected by chance showing at least nominally significant pleiotropy between endometriosis and endometrial cancer \((P = 6.0 \times 10^{-3})\): The pair of SNP subsets producing the minimum exact binomial test \(P\)-value for pleiotropy (endometriosis SNP subset with \(P \leq 0.002\) and endometrial cancer SNP subset with \(P \leq 0.86\)) had \(P = 3.3 \times 10^{-4}\).

The primary test for concordant effects between endometriosis and endometrial cancer also revealed that the number of SNP subsets with nominally significant concordant effects \((P \leq 0.05)\) was significantly more than expected by chance \((P = 2.0 \times 10^{-3})\): The pair of SNP subsets producing the minimum Fisher’s exact test \(P\)-value for effect correlation (endometriosis SNP subset with \(P \leq 0.37\) and endometrial cancer SNP subset with \(P \leq 1\)) had \(P = 2.1 \times 10^{-4}\). The primary results indicate that SNP effects are correlated, with the presence of allelic effects that increase the risk of both traits. Including only specific (default) SNP subsets in the analyses \([31]\), SNP effects were positively, although not significantly, correlated for SNPs at \(P \leq 0.05\) in both datasets \((P = 8.4 \times 10^{-2})\) and for SNPs with \(P \leq 1.0 \times 10^{-3}\) in the larger endometriosis dataset and with \(P \leq 0.05\) in the endometrial cancer dataset \((P = 6.8 \times 10^{-2})\). Together, these results indicate that overall more SNPs than expected by chance were associated with the same direction of effect for both diseases, particularly amongst nominally or marginally associated SNPs.

**Cross-disease genomewide association analyses**

For the cross-disease meta-analysis including the endometriosis and endometrial cancer (ANECS-SEARCH and NSECG) GWAS datasets, two SNPs had \(P\)-values reaching genomewide significance \((rs6782972, P = 3.3 \times 10^{-4} \text{ and } rs2218868, P = 4.1 \times 10^{-4}; \text{Table S1})\). A further 92 SNPs were suggestively \((P \leq 10^{-3})\) associated in the combined analysis of both diseases (Table S1). Following inclusion of the iCOGs association results in the meta-analysis, the \(P\)-values for both rs6782972 and rs2218868 dropped below the threshold for suggestive significance \((rs6782972, P = 1.2 \times 10^{-2}, \text{OR} = 0.95, 95\% \text{ CI} = 0.90–0.99; \text{rs2218868, } P = 1.4 \times 10^{-3}, \text{OR} = 1.06, 95\% \text{ CI} = 1.02–1.09); \text{hence, neither locus was validated by the endometrial cancer replication dataset. Including all four datasets, 13 loci showed evidence for replication (P-values} \leq 10^{-5}; \text{Table 2; Fig. S1}), with a genomewide significant signal detected for SNP rs2475335 \((P = 4.9 \times 10^{-8}, \text{OR} = 1.11, 95\% \text{ CI} = 1.07–1.14). After adjusting for the multiple testing of 13 SNPs, two SNPs showed evidence of significant heterogeneity between studies (Table 2, rs9865110 \((P_{\text{het}} = 2.2 \times 10^{-3})\) and rs7515106\((P_{\text{het}} = 3.6 \times 10^{-3})\).
Table 2. Replicated associations (\(P \leq 10^{-5}\)) from the cross-disease meta-analysis for endometriosis and endometrial cancer.

<table>
<thead>
<tr>
<th>Chr</th>
<th>BP</th>
<th>Lead SNP</th>
<th>Flanking gene/allele genes(^1)</th>
<th>Effect allele</th>
<th>Noneffect allele</th>
<th>GWAS meta-analysis OR (95% CI)</th>
<th>P-value</th>
<th>Replication meta-analysis OR (95% CI)</th>
<th>P-value</th>
<th>(P_{\text{int}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>10260263</td>
<td>rs2475335</td>
<td>PTPRD</td>
<td>T</td>
<td>C</td>
<td>1.15 (1.09–1.21)</td>
<td>2.1 \times 10^{-6}</td>
<td>1.11 (1.07–1.15)</td>
<td>4.9 \times 10^{-8}</td>
<td>0.46</td>
</tr>
<tr>
<td>3</td>
<td>73981773</td>
<td>rs9865110</td>
<td>PCDNA3.1-NF-κB</td>
<td>C</td>
<td>A</td>
<td>1.11 (1.06–1.16)</td>
<td>9.4 \times 10^{-5}</td>
<td>1.10 (1.06–1.14)</td>
<td>2.6 \times 10^{-6}</td>
<td>0.002</td>
</tr>
<tr>
<td>17</td>
<td>4626217</td>
<td>rs2278868</td>
<td>SKAP1</td>
<td>C</td>
<td>T</td>
<td>0.87 (0.82–0.93)</td>
<td>4.7 \times 10^{-7}</td>
<td>0.92 (0.88–0.96)</td>
<td>5.5 \times 10^{-6}</td>
<td>0.05</td>
</tr>
<tr>
<td>12</td>
<td>89299016</td>
<td>rs9349553</td>
<td>TFAP2D</td>
<td>T</td>
<td>C</td>
<td>1.14 (1.08–1.19)</td>
<td>4.5 \times 10^{-6}</td>
<td>1.09 (1.05–1.13)</td>
<td>9.0 \times 10^{-6}</td>
<td>0.02</td>
</tr>
<tr>
<td>6</td>
<td>50715407</td>
<td>rs9349553</td>
<td>PDZRN3-CNTN3</td>
<td>T</td>
<td>C</td>
<td>1.14 (1.08–1.19)</td>
<td>4.5 \times 10^{-6}</td>
<td>1.09 (1.05–1.13)</td>
<td>9.0 \times 10^{-6}</td>
<td>0.02</td>
</tr>
<tr>
<td>17</td>
<td>4626217</td>
<td>rs2278868</td>
<td>SKAP1</td>
<td>C</td>
<td>T</td>
<td>1.12 (1.07–1.17)</td>
<td>3.3 \times 10^{-5}</td>
<td>1.08 (1.05–1.12)</td>
<td>1.7 \times 10^{-5}</td>
<td>0.054</td>
</tr>
<tr>
<td>6</td>
<td>57186294</td>
<td>rs10459129</td>
<td>PARP1-CCND2</td>
<td>A</td>
<td>G</td>
<td>0.89 (0.83–0.95)</td>
<td>6.8 \times 10^{-5}</td>
<td>0.90 (0.86–0.95)</td>
<td>2.1 \times 10^{-5}</td>
<td>0.44</td>
</tr>
<tr>
<td>19</td>
<td>3207608</td>
<td>rs17693745</td>
<td>ZNF536-TSHZ3</td>
<td>T</td>
<td>C</td>
<td>1.1 (1.06–1.16)</td>
<td>8.4 \times 10^{-5}</td>
<td>1.09 (1.05–1.13)</td>
<td>2.4 \times 10^{-5}</td>
<td>0.49</td>
</tr>
<tr>
<td>9</td>
<td>32907608</td>
<td>rs2278868</td>
<td>SKAP1</td>
<td>C</td>
<td>T</td>
<td>1.12 (1.07–1.17)</td>
<td>3.3 \times 10^{-5}</td>
<td>1.08 (1.05–1.12)</td>
<td>1.7 \times 10^{-5}</td>
<td>0.054</td>
</tr>
<tr>
<td>12</td>
<td>63665556</td>
<td>rs10459129</td>
<td>PARP1-CCND2</td>
<td>A</td>
<td>G</td>
<td>0.86 (0.79–0.93)</td>
<td>1.2 \times 10^{-5}</td>
<td>0.90 (0.86–0.95)</td>
<td>4.0 \times 10^{-5}</td>
<td>0.07</td>
</tr>
<tr>
<td>17</td>
<td>22473410</td>
<td>rs7515106</td>
<td>WNT4-ZBTB40</td>
<td>C</td>
<td>T</td>
<td>1.15 (1.08–1.21)</td>
<td>3.9 \times 10^{-5}</td>
<td>0.93 (0.89–0.97)</td>
<td>7.9 \times 10^{-5}</td>
<td>0.0036</td>
</tr>
</tbody>
</table>

\(^1\)Only one gene is listed for SNPs located within genes (introns or exons).

Endometriosis and Endometrial Cancer

PTPRD is an attractive candidate gene for regulation of cell growth and migration with potential for treatment of both diseases [42, 43]. While PTPRD is an attractive candidate gene for regulation of cell growth and migration with potential for treatment of both diseases [42, 43], this variant (rs2475335) has not yet been validated as a significant target for future studies. The SNP most implied by the epidemiological evidence is the receptor tyrosine protein tyrosine phosphatase receptor type D (PTPRD) as-endothelial cancer. The SNP most implied by the epidemiological evidence is the receptor tyrosine protein tyrosine phosphatase receptor type D (PTPRD).
additional neighboring genes (e.g., SNX11, HOXB2, and HOXB3) in various tissues.

Other SNPs of interest include rs12303900 on chromosome 12q21, located between the KITLG and DUSP6 genes. DUSP6 is a critical regulator of ERK signaling, a pathway dysregulated in both endometriosis and endometrial cancer and a potential target for treatment for both diseases [47–49]. SNP rs10008492, located on chromosome 4p14, is an eQTL for nearby toll-like receptors TLR1 and TLR6 (http://www.gtexportal.org). Both TLR1 and TLR6 are upregulated in endometriotic mesenchymal stem cells [50] and are expressed in endometrial cancer cell lines [51]. However, as for the PTPRD locus, all of these association results need to be further validated in additional replication datasets for both diseases, and relevant functional studies undertaken, before more is hypothesized about their genetic and biological effects on the risk of both endometriosis and endometrial cancer.

In this cross-disease genetic correlation and genomewide association study, we have provided evidence for overlap in genetic risk factors for endometriosis and endometrial cancer. Our genetic correlation analysis supports recent large epidemiological studies indicating an increased risk of endometrial cancer in women previously diagnosed with endometriosis, while the cross-disease meta-analysis has revealed plausible loci that could increase the risk of both diseases and which should be pursued further in functional studies. This work on endometriosis and endometrial cancer also adds further evidence to the utility of cross-disease genetic correlation and GWAS analyses as tractable and attractive methodologies to identify susceptibility loci that predispose to multiple diseases, which could lead to new diagnostic or treatment options for affected individuals.

Acknowledgments

We acknowledge with appreciation all women who participated in the QIMR Berghofer Medical Research Institute and OXGENE endometriosis studies and the ANECS-SEARCH, NSECG, and iCOGS endometrial cancer studies. For the endometriosis study, we thank Endometriosis Associations for supporting study recruitment, and the many hospital directors and staff, gynecologists, general practitioners, and pathology services in Australia and the UK who provided assistance with blood collection and confirming diagnoses. We are grateful to the many research assistants and interviewers for assistance with the studies contributing to the QIMR and OXGENE collections. The QIMR Study was supported by grants from the National Health and Medical Research Council (NHMRC) of Australia (241944, 339462, 389927, 389875, 389891, 389892, 389938, 443036, 442915, 442981, 496610, 496739, 552485, and 552498), the Cooperative Research Centre for Discovery of Genes for Common Human Diseases (CRC), Cerylid Biosciences (Melbourne), and donations from Neville and Shirley Hawkins. Genotyping of the OXGENE and QIMR cases was supported by a grant from the Wellcome Trust (WT084766/Z/08/Z) and makes use of WTCCC2 control data generated by the Wellcome Trust Case Control Consortium (full list of investigators at www.wtccc.org.uk). Funding for the WTCCC project was provided by the Wellcome Trust under award 076113 and 085475. For the endometrial cancer studies, ANECS recruitment was supported by project grants from the NHMRC (339435), Cancer Council Queensland (4196615), and Cancer Council Tasmania (403031 and 457636). The University of Newcastle, the Gladys M. Brawn Senior Research Fellowship scheme, the Vincent Fairfax Family Foundation, the Hunter Medical Research Institute, and the Hunter Area Pathology Service all contributed toward the costs of establishing the control dataset (HCS). SEARCH recruitment was funded by a program grant from Cancer Research UK (C490/A10124). Stage 1 and stage 2 case genotyping was supported by the NHMRC (552402 and 1031333). Control data were generated by the WTCCC, and we acknowledge use of DNA from the British 1958 Birth Cohort collection, funded by UK Medical Research Council grant G0009934 and Wellcome Trust grant 068545/Z/02; funding for this project was provided by the Wellcome Trust under award 085475. NSECG case recruitment was supported by the European Union’s Framework Programme 7 CHIBCHA grant and Wellcome Trust Centre for Human Genetics Core Grant 090532/Z/09/Z, and control (CORGI) recruitment by Cancer Research UK. The iCOGS endometrial cancer analysis was supported by an NHMRC project grant (1031333). Funding for iCOGS infrastructure came from the European Community’s Seventh Framework Programme under grant agreement 223175 (HEALTH-F2-2009-223175) (COGS), Cancer Research UK (C1287/A10118, C1287/A10710, C122292/A11174, C1281/A12014, C5047/A8384, C5047/A15007, C5047/A10692, and C8197/A16565), the US National Institutes of Health (R01 CA128978, U19 CA148537, U19 CA148065, and U19 CA148112), the US Department of Defense (W81XWH-10-1-0341), and the Canadian Institutes of Health Research (CIHR) for the CHIR Team in Familial Risks of Breast Cancer, the Susan G. Komen Foundation for the Cure, the Breast Cancer Research Foundation, and the Ovarian Cancer Research Fund. The Bavarian Endometrial Cancer Study (BECs) was partly funded by the ELAN fund of the University of Erlangen. The Hannover–Jena Endometrial Cancer Study was partly supported by the Rudolf Bartling Foundation. The Leuven Endometrium Study (LES) was supported by the Verelst Foundation for Endometrial Cancer. The Mayo
Endometriosis and Endometrial Cancer

References


endometriosis highlighting key genes involved in hormone metabolism. Nat. Commun. 8:15539.

**Supporting Information**

Additional supporting information may be found in the online version of this article:

**Figure S1.** Forest plots of association between the top 13 SNPs in the endometriosis-endometrial cancer meta-analysis and each of the datasets included in the analysis.

**Table S1.** Results for the top SNPs from the endometriosis and endometrial cancer (ANECS-SEARCH and NSECGR) GWAS meta-analysis, with iCOGS as the endometrial cancer replication dataset.