# 10 Years of GWAS discovery in endometrial cancer: Aetiology, function and translation

*Xuemin Wang,[1] Dylan M. Glubb,[1] and Tracy A. O'Mara **

Genetics and Computational Biology Department, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4006, Australia

## Summary

**Endometrial cancer is a common gynaecological cancer with increasing incidence and mortality. In the last decade, endometrial cancer genome-wide association studies (GWAS) have provided a resource to explore aetiology and for functional interpretation of heritable risk variation, informing endometrial cancer biology. Indeed, GWAS data have been used to assess relationships with other traits through correlation and Mendelian randomisation analyses, establishing genetic relationships and potential risk factors. Cross-trait GWAS analyses have increased statistical power and identified novel endometrial cancer risk variation related to other traits. Functional analysis of risk loci has helped prioritise candidate susceptibility genes, revealing molecular mechanisms and networks. Lastly, risk scores generated using endometrial cancer GWAS data may allow for clinical translation through identification of patients at high risk of disease. In the next decade, this knowledge base should enable substantial progress in our understanding of endometrial cancer and, potentially, new approaches for its screening and treatment.**

**eBioMedicine 2022;77: 103895**
Published online xxx
https://doi.org/10.1016/j.ebiom.2022.103895



**Keywords:** Endometrial cancer; Genome-wide association study (GWAS); Risk factors; Cross-trait analysis; Functional analysis; Polygenic risk scores

## Introduction

Endometrial cancer is the most common gynaecological cancer in countries with a high Human Development Index,[1] with increased incidence and mortality rates observed recently worldwide.[2,3] Over 380,000 new cases and nearly 90,000 deaths were estimated globally in 2018.[1] Historically, endometrial cancers have been classified into two subgroups using histology: endometrioid and non-endometrioid. The endometrioid subtype is oestrogen-responsive and accounts for about 80% of endometrial cases,[4] often with a favourable prognosis.[5] Conversely, the less common non-endometrioid subtypes (e.g. serous and clear cell) do not respond strongly to oestrogen and are often associated with poor prognosis.[5]

Genetics play a role in predisposition to endometrial cancer. Family history of endometrial cancer has been reported to increase a woman's risk of developing the cancer up to two-fold.[6] Lynch syndrome also increases endometrial cancer risk via pathogenic germline variants within mismatch-repair genes (i.e. *MLH1, MSH2, MSH6, PMS2* and *EPCAM*) and are estimated to account for 3% of endometrial cancer cases.[7] Although these mismatch-repair genes have large effects on endometrial cancer risk, the risk variants they contain are rare in the general population (reviewed by Dörk et al.[8]). Other rarer pathogenic germline variants in genes such as *PTEN, POLE* and *POLD1* also modestly increase endometrial cancer risk.[9]

The effects of common genetic variation (minor allele frequency > 1%) on endometrial cancer risk can be systematically assessed by genome-wide association studies (GWAS). GWAS typically involve the genotyping of millions of common germline genetic variants across the genome, using array-based technology and statistical imputation. In comparison with sequencing approaches, GWAS has enabled both coding and non-coding variation to be cheaply and efficiently evaluated for disease associations. GWAS have transformed the study of common genetic variation, identifying thousands of mostly non-coding variants that associate with complex traits or disease, revealing novel biological mechanisms that underlie these phenotypes and enabling identification of individuals with increased risk of diseases such as cancer (reviewed in).[10] The first endometrial cancer GWAS in 2011 analysed data from 1,265 cases and identified a single risk region that encompasses the *HNF1B* gene.[11] As endometrial cancer GWAS datasets have increased in size, so has the statistical power to detect associations, revealing further risk regions (reviewed in).[12] The largest endometrial cancer GWAS to date, conducted by the Endometrial Cancer

*Corresponding author.
*E-mail address:* tracy.omara@qimrberghofer.edu.au (T.A. O'Mara).

[1] These authors contributed equally to this work.

Association Consortium (ECAC), used data from nearly 13,000 cases and identified 16 endometrial cancer risk regions.[13] Genetic variation identified from this GWAS can explain one quarter of the genetic component for the two-fold familial relative endometrial cancer risk,[13] with future larger GWAS likely to detect more of this genetic component.

This review summarises the recent progress made by GWAS to: (i) explore endometrial cancer aetiology by assessing relationships between endometrial cancer and other traits; (ii) identify candidate susceptibility genes using functional analyses and inform endometrial cancer biology; and (iii) generate endometrial genetic risk scores and enable stratification of women for screening and risk reducing purposes.

## Endometrial cancer aetiology revealed using GWAS data

### Genetic correlations between endometrial cancer and other traits

Genetic correlation between traits indicates the presence of a shared genetic background or architecture that influences both traits. Importantly for epidemiology, germline genetic variants are not typically related to confounding factors that can lead to spurious associations in observational studies (e.g. ascertainment bias) nor are they affected by disease onset. Thus, genetic correlation analysis can help clarify associations from observational studies and prioritise potentially novel aetiological factors for follow-up studies.

Analyses reported from the largest endometrial cancer GWAS to date revealed positive genetic correlations between endometrial cancer risk and type 2 diabetes, body mass index (BMI) and other related anthropometric traits; while negative correlations were found with age at menarche and years of schooling.[13] All of these traits are either obesity-related or genetically correlated with BMI, one of the most established endometrial cancer risk factors from observational studies.[14] However, additional analyses are required to determine if any of these traits have a causal effect on endometrial cancer.

Genetic correlation analysis also allows for the disentangling of epidemiological associations. For example, observational studies have suggested that non-cancerous gynaecological diseases, such as polycystic ovarian syndrome (PCOS) and uterine fibroids, are associated with endometrial cancer.[15–18] These relationships are supported by genetic correlations[19,20] but after adjustment for BMI using GWAS data, the correlation between endometrial cancer and PCOS appears to be at least partly mediated by BMI.[20] Correlation analysis has helped clarify previous inconsistent reports of associations between endometrial cancer and another non-cancerous gynaecological disease, endometriosis,[15,16,21,22] with no evidence of a shared genetic background found using the largest available European GWAS datasets.[20] On the other hand, ovarian cancer, has been shown to share several epidemiological similarities with endometrial cancer[23] and correlation analysis further supports a relationship between these cancers by demonstrating shared genetics[3919].

There are now a large amount of publicly available GWAS summary statistics housed by the GWAS Catalog (https://www.ebi.ac.uk/gwas/summary-statistics), allowing for more comprehensive and systematic genetic correlation analyses to be performed for endometrial cancer. Additionally, the large number of association results provided by analysis of well-phenotyped and well-powered biobank studies, such as the UK Biobank analyses by the Neale Laboratory (http://www.nealelab.is/uk-biobank), provide further avenues for assessing genetic relationships between endometrial cancer and other traits.

### Use of Mendelian randomisation analysis to establish endometrial cancer risk factors

Although genetic correlation between endometrial cancer and another trait provides evidence of shared genetics, other analyses are needed to determine if a causal relationship exists. Mendelian randomisation analysis is a genetic approach that allows causality to be assessed by using trait-associated (typically GWAS-identified) variants as proxies for exposure to a potential risk factor. As for genetic correlation analysis, genetic variants are usually independent from environmental or lifestyle factors and less likely to be influenced by confounding and reverse causation. Importantly, the alleles of genetic variants are randomly assigned at conception and are thus comparable to a lifelong randomised trial. Mendelian randomisation analysis has been successfully used to assess a variety of traits for effects on endometrial cancer, identifying potential risk factors related to obesity, puberty, cytokines, hormones, circulating cholesterol, gynaecological disease and telomere length (Table 1).

To tease apart the effects of obesity on endometrial cancer, obesity-related traits have been assessed using Mendelian randomisation analysis. A number of such studies using European GWAS data have confirmed the causal effect of increasing BMI on elevated endometrial cancer risk.[13,24,25] These include associations with endometrioid endometrial cancer, and the novel association of BMI and risk of non-endometrioid histological subtypes.[13] The relationship between BMI and endometrial cancer has also been confirmed by analysis performed in a Japanese GWAS population.[26] Other obesity-related traits have been found to causally associate with endometrial cancer risk, including circulating cytokines related to obesity (i.e. plasminogen activator inhibitor-1 and tumour necrosis factor[27,28]) and type 2 diabetes.[29] Mendelian randomisation approaches can be used to

| Exposure assessed[a] | Endometrial cancer (all histologies) | | | | Endometrial cancer (endometrioid histology) | | | | Endometrial cancer (non-endometrioid histologies) | | | | Refs. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Nsnp | Ncases/ Ncontrols | OR (95% CI) | P-value | Nsnp | Ncases/ Ncontrols | OR (95% CI) | P-value | Nsnp | Ncases/ Ncontrols | OR (95% CI) | P-value | |
| Anthropometric Factors | | | | | | | | | | | | | |
| Adult height | 814 | 12906/108979 | 1.00 (0.95-1.06) | 0.90 | 814 | 8758/46126 | 0.99 (0.93-1.05) | 0.63 | 814 | 1230/35447 | 1.00 (0.88-1.15) | 0.95 | 13 |
| BMI | - | - | - | - | 32 | 1287/8273 | 3.86 (2.24-6.64) | $1.0 \times 10^{-6}$ | - | - | - | - | 24 |
| BMI | - | - | - | - | 97 | 2094/3867 | 1.11 (1.02-1.21) | 0.02 | - | - | - | - | 69 |
| BMI | 77 | 6609/37926 | 2.01 (1.94-2.28) | $3.4 \times 10^{-17}$ | - | - | - | - | - | - | - | - | 25 |
| BMI | 77 | 12906/108979 | 1.92 (1.63-2.25) | $1.7 \times 10^{-11}$ | 77 | 8758/46126 | 2.04 (1.69-2.46) | $8.6 \times 10^{-11}$ | 77 | 1230/35447 | 1.65 (1.13-2.14) | 0.01 | 13 |
| BMI[b] | 74 | 909/39556 | 2.08 (1.35-3.23) | 0.001 | - | - | - | - | - | - | - | - | 26 |
| Waist-hip ratio | 47 | 6609/37926 | 0.97 (0.63-1.31) | 0.86 | - | - | - | - | - | - | - | - | 25 |
| Waist-hip ratio | 47 | 12906/108979 | 0.95 (0.72-1.25) | 0.71 | 47 | 8758/46126 | 0.94 (0.71-1.24) | 0.66 | 47 | 1230/35447 | 1.27 (0.69-2.33) | 0.45 | 13 |
| Waist-hip ratio (sex-specific) | 34 | 6609/37926 | 1.02 (0.99-1.04) | 0.09 | - | - | - | - | - | - | - | - | 25 |
| Pubertal Factors | | | | | | | | | | | | | |
| Age at menarche | 237 | 6609/37926 | 0.78 (0.70-0.87) | $1.0 \times 10^{-5}$ | - | - | - | - | - | - | - | - | 31 |
| Age at menarche | 368 | 12906/108979 | 0.82 (0.77-0.87) | $2.2 \times 10^{-9}$ | 368 | 8758/46126 | 0.80 (0.74-0.86) | $1.9 \times 10^{-9}$ | 368 | 1230/35447 | 0.93 (0.79-1.08) | 0.33 | 13 |
| Age at natural menopause | 54 | 12906/108979 | 1.03 (1.00-1.06) | 0.06 | 54 | 8758/46126 | 1.02 (0.99-1.06) | 0.19 | 54 | 1230/35447 | 1.07 (0.99-1.14) | 0.08 | 13 |
| Age at natural menopause | 255 | 12270/ 46126 | 1.05 (1.03-1.07) | $2.2 \times 10^{-8}$ | | | | | | | | | 32 |
| Circulating Factors | | | | | | | | | | | | | |
| Adiponectin | 18 | 12906/108979 | 1.02 (0.89-1.17) | 0.75 | - | - | - | - | - | - | - | - | 27 |
| Bilirubin | 110 | 12906/108979 | 1.37 (0.99-1.89) | 0.06 | - | - | - | - | - | - | - | - | 70 |
| Leptin | 2 | 12906/108979 | 1.46 (0.69-3.06) | 0.32 | - | - | - | - | - | - | - | - | 27 |
| Soluble leptin receptor | 4 | 12906/108979 | 1.02 (1.00-1.05) | 0.09 | - | - | - | - | - | - | - | - | 27 |
| Plasminogen activator inhibitor-1 | 4 | 12906/108979 | 1.38 (1.04-1.82) | 0.03 | - | - | - | - | - | - | - | - | 27 |
| Selenium | 4 | 12906/108979 | 0.99 (0.87-1.14) | 0.93 | - | - | - | - | - | - | - | - | 71 |
| Insulin-like growth factor-1 | 265 | 12906/108979 | 0.98 (0.90-1.07) | 0.69 | | | | | | | | | 36 |
| Tumor necrosis factor | 3 | 1520/197318 | 0.25 (0.07-0.94) | 0.04 | - | - | - | - | - | - | - | - | 28 |
| Vitamin D | 75 | 12906/108979 | 0.95 (0.83−1.09) | 0.46 | 75 | 8758/46126 | 0.93 (0.81−1.08) | 0.36 | 75 | 1230/35447 | 1.02 (0.76−1.36) | 0.91 | 72 |
| Total testosterone | 248 | 12270/46126 | 1.39 (1.26-1.53) | $3.4 \times 10^{-11}$ | 248 | 8758/46126 | 1.39 (1.24-1.55) | $1.2 \times 10^{-8}$ | 224 | 1230/38224 | 1.26 (0.99-1.61) | 0.06 | 35 |
| Total testosterone | 229 | 12906/108979 | 1.37 (1.24-1.51) | $1.9 \times 10^{-10}$ | | | | | | | | | 36 |
| Bioavailable testosterone | 176 | 12270/46126 | 1.63 (1.43-1.87) | $1.5 \times 10^{-12}$ | 176 | 8758/46126 | 1.62 (1.39-1.88) | $3.9 \times 10^{-10}$ | 155 | 1230/38224 | 1.46 (1.05-2.02) | 0.02 | 35 |
| Bioavailable testosterone | 162 | 12906/108979 | 1.67 (1.45-1.92) | $6.1 \times 10^{-13}$ | | | | | | | | | 36 |
| Oestradiol (post-menopausal) | 1 | 6608/37925 | 1.15 (1.11-1.21) | $4.8 \times 10^{-11}$ | - | - | - | - | - | - | - | - | 34 |
| SHBG | 351 | 12270/46126 | 0.77 (0.67-0.89) | $2.6 \times 10^{-4}$ | 351 | 8758/46126 | 0.78 (0.67-0.91) | 0.002 | 315 | 1230/38224 | 0.78 (0.55-1.11) | 0.17 | 35 |
| SHBG | 315 | 12906/108979 | 0.78 (0.66-0.91) | 0.001 | | | | | | | | | 36 |
| HDL-cholesterol | 163 | 12906/108979 | 1.07 (1.00-1.14) | 0.04 | 161 | 8758/46126 | 1.03 (0.95-1.11) | 0.51 | 166 | 1230/35447 | 1.22 (1.01-1.47) | 0.04 | 30 |
| LDL-cholesterol | 140 | 12906/108979 | 0.88 (0.83-0.93) | $7.3 \times 10^{-6}$ | 142 | 8758/46126 | 0.89 (0.83-0.95) | $5.4 \times 10^{-4}$ | 145 | 1230/35447 | 0.76 (0.65-0.89) | $6.2 \times 10^{-4}$ | 30 |

***Table 1*** (Continued)

| Exposure assessed[a] | Endometrial cancer (all histologies) | | | | Endometrial cancer (endometrioid histology) | | | | Endometrial cancer (non-endometrioid histologies) | | | | Refs. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Nsnp | Ncases/ Ncontrols | OR (95% CI) | P-value | Nsnp | Ncases/ Ncontrols | OR (95% CI) | P-value | Nsnp | Ncases/ Ncontrols | OR (95% CI) | P-value | |
| Triglycerides | 104 | 12906/108979 | 0.96 (0.89-1.04) | 0.34 | 106 | 8758/46126 | 0.95 (0.86-1.05) | 0.31 | 104 | 1230/35447 | 1.10 (0.87-1.39) | 0.42 | 30 |
| Diabetic Factors | | | | | | | | | | | | | |
| Early insulin secretion | - | - | - | - | 17 | 1287/8273 | 1.40 (1.12-1.76) | $3.0 \times 10^{-3}$ | - | - | - | - | 24 |
| Increased fasting glucose | - | - | - | - | 36 | 1287/8273 | 1.00 (0.67-1.50) | 0.99 | - | - | - | - | 24 |
| Increased fasting glucose | 35 | 1931/196907 | 1.27 (0.95-1.70) | 0.11 | - | - | - | - | - | - | - | - | 29 |
| Increased fasting insulin | - | - | - | - | 18 | 1287/8273 | 2.34 (1.06-5.14) | 0.03 | - | - | - | - | 24 |
| Increased fasting insulin | 21 | 1931/196907 | 2.01 (1.07-3.78) | 0.03 | - | - | - | - | - | - | - | - | 29 |
| Type 2 diabetes | - | - | - | - | 49 | 1287/8273 | 0.91 (0.79-1.04) | 0.16 | - | - | - | - | 24 |
| Type 2 diabetes | 399 | 1931/196907 | 1.08 (1.01-1.15) | 0.03 | - | - | - | - | - | - | - | - | 29 |
| Diseases | | | | | | | | | | | | | |
| Parkinson's disease | 15 | 12906/108979 | 0.97 (0.80-1.18) | 0.78 | - | - | - | - | - | - | - | - | 73 |
| Endometriosis | 26 | 12270/46126 | 1.09 (1.00-1.20) | 0.34 | - | - | - | - | - | - | - | - | 20 |
| PCOS | 14 | 12270/46126 | 0.95 (0.91-0.99) | 0.26 | - | - | - | - | - | - | - | - | 20 |
| Uterine fibroids | 23 | 12270/46126 | 1.19 (1.11-1.27) | 0.01 | - | - | - | - | - | - | - | - | 20 |
| Other traits | | | | | | | | | | | | | |
| Alcohol consumption | 94 | 1931/196907 | 0.84 (0.45-1.58) | 0.59 | - | - | - | - | - | - | - | - | 74 |
| Smoking | 361 | 1931/196907 | 1.11 (0.91-1.37) | 0.31 | - | - | - | - | - | - | - | - | 74 |
| Telomere length | 12 | 6608/37925 | 1.31 (1.07-1.61) | 0.01 | - | - | - | - | - | - | - | - | 75 |

***Table 1*: Mendelian randomisation studies examining the causal relationship between genetically predicted putative risk factors and endometrial cancer susceptibility. Bolded estimations indicate significant associations determined by individual studies. Dash ("-") indicates analysis not performed.**

Abbreviations: Nsnp – number of variants used in Mendelian randomisation analyses; OR – odds ratio; CI – confidence interval; BMI - body mass index; PCOS - polycystic ovary syndrome; SHBG - sex hormone-binding protein

[a] Generalised summary-data based Mendelian Randomisation (GSMR) analyses were conducted for triglycerides, HDL- and LDL-cholesterol; whereas inverse variance weighted (IVW) method was used to infer the associations between all other exposures and endometrial cancer risk.

[b] This result was reported in a Japanese population, whereas results for all other exposures were reported in European populations.

disentangle these associations, allowing the effects of traits on endometrial cancer risk to be evaluated in the context of BMI. For example, a Mendelian randomisation study found genetically predicted early insulin secretion and increased fasting insulin levels were associated with endometrial cancer risk independent of BMI.[24] A further Mendelian randomisation study has shown that increased levels of LDL cholesterol and reduced levels of HDL cholesterol associate with lower risk of endometrial cancer but only the relationship between LDL cholesterol and endometrial cancer (non-endometrioid or all histologies) appears to be independent of BMI.[30]

A pubertal trait that is related to obesity and has been associated with endometrial cancer risk through Mendelian randomisation, independent of BMI, is earlier age at menarche.[13,31] Conversely, later age of natural menopause has also been found to associate with increased risk of endometrial cancer.[32] Both findings potentially indicate that increased lifelong exposure to oestrogen affects endometrial cancer susceptibility, consistent with the established relationship between endometrial cancer and oestrogen exposure unopposed by progesterone.[33] Mendelian randomisation analysis has been used to further explore the effects of sex hormones on endometrial cancer. Because of the cyclical expression of these hormones amongst pre-menopausal women and their low expression amongst postmenopausal women, genetic instruments for levels of these hormones are difficult to identify. Indeed, to date, only one GWAS variant has been associated with oestradiol levels amongst postmenopausal women.[34] This variant was used to demonstrate a relationship between increased postmenopausal oestradiol levels and endometrial cancer risk.[34] An important precursor of oestradiol is testosterone and, consistent with the finding for oestradiol, increased levels of genetically predicted testosterone have also been associated with increased endometrial cancer risk.[35,36] Both oestradiol and testosterone can be bound by sex hormone binding globulin (SHBG) and thus it is not surprising that Mendelian randomisation analysis has demonstrated that higher SHBG levels associate with reduced endometrial cancer risk.[35,36] Although the extent to which BMI affects these relationships remains to be investigated, it is interesting to note that bioavailable testosterone and SHBG are both genetically correlated with BMI (Figure 1).

While endometrial cancer has been found to be genetically correlated with both PCOS and uterine fibroids, Mendelian randomisation analysis only provides evidence for a causal relationship between uterine fibroids and endometrial cancer risk.[20] However, sensitivity analyses in this study suggested potential violation of the Mendelian randomisation assumption that a genetic variant only affects disease through its effect on the exposure. Further assessment of this relationship is recommended when larger GWAS datasets for uterine fibroids are available.

We have explored relationships between Mendelian randomisation identified endometrial cancer risk factors using genetic correlation to provide insights into the relationships between these traits (Figure 1). As expected, type 2 diabetes and fasting insulin positively correlated with BMI; whereas BMI was negatively correlated with age of menarche and HDL cholesterol. Another anticipated group of related traits included bioavailable testosterone and SHBG which had a strong negative correlation. Perhaps unexpectedly, HDL cholesterol was negatively correlated with bioavailable testosterone and positively correlated with SHBG. Uterine fibroids were positively correlated with oestradiol, age at natural menopause and telomere length, and may reflect relationships with oestrogen exposure and aging.[37] These correlations with uterine fibroids may explain the potential violation of the Mendelian randomisation assumptions observed in the analysis of uterine fibroids (i.e. genetic instruments for uterine fibroids may also associate with other risk factors and affect endometrial cancer risk independent of effects on uterine fibroids).

### Cross-trait GWAS meta-analysis to uncover novel endometrial cancer genetic risk loci

Genetic information from related traits can be exploited by joint analysis methods to increase effective sample sizes, boosting statistical power to help identify novel trait-associated loci. To this end, progress has been made to uncover additional endometrial cancer risk regions, highlighting loci that may be found in larger endometrial cancer GWAS and are related to the jointly analysed trait. For example, cross-trait meta-analyses have combined endometrial cancer GWAS data with those from related diseases such as endometriosis,[21] which shares a common tissue of origin, to identify risk loci that have subsequently been confirmed by larger endometrial cancer GWAS (reviewed by O'Mara et al.).[12] Subsequent functional analysis of such risk loci can reveal molecular mechanisms and pathways underlying biology shared between endometrial cancer and the jointly analysed trait, as will be discussed later.

Findings of genetic correlation between endometrial cancer and other gynaecological diseases (e.g. ovarian cancer, PCOS and uterine fibroids) have provided a rationale for cross-trait GWAS that have yielded further potential novel risk loci. For example, meta-analysis of endometrial and ovarian cancers and their subtypes, revealed three novel endometrial cancer risk regions (5p15.33, 9q34.2, and 10p12.31; Table 2); however, these loci have not yet been assessed for replication in an independent endometrial cancer GWAS dataset.[38] Meanwhile, meta-analysis of endometrial cancer with PCOS, and uterine fibroids GWAS identified a novel risk region (1p36.12; Table 2) that did replicate in an independent dataset.[20]

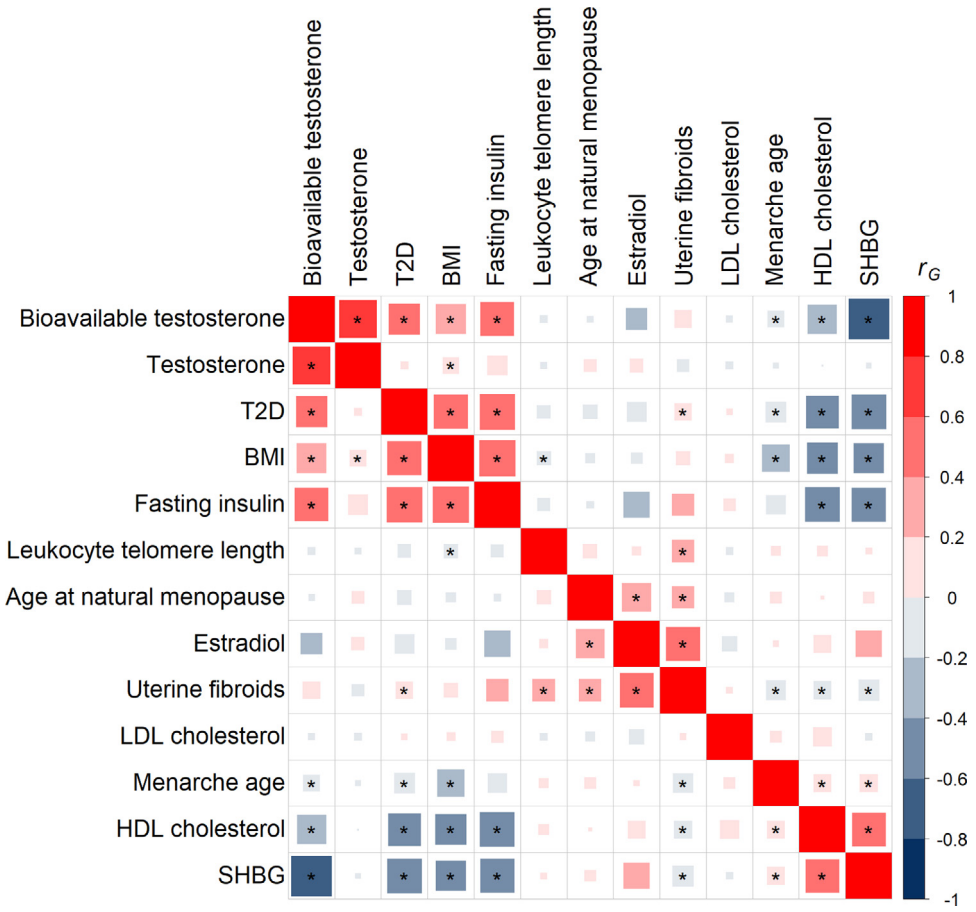


**Figure 1. Pairwise genetic correlations among risk factors of endometrial cancer.** Sex-combined GWAS summary statistics were used for type 2 diabetes (T2D), while female-stratified GWAS summary statistics were used for all other phenotypes (accessed from the GWAS Catalog https://www.ebi.ac.uk/gwas/summary-statistics). Genetic correlation with tumour necrosis factor could not be estimated due to its very low heritability; whereas genetic correlation with plasminogen activator inhibitor-1 was not included because GWAS summary statistics for this phenotype were not available. Abbreviations – T2D: type 2 diabetes; BMI: body mass index; LDL: low density lipoprotein; HDL: high density lipoprotein; SHBG: sex hormone binding globulin.

The four additional findings described above bring the number of potential risk loci for endometrial cancer to 20, although there are data from a preliminary report of a meta-analysis of endometrial cancer and three other hormone-related cancers (breast, ovarian and prostate) to support additional novel risk loci.[39] It has been estimated that the total number of independent risk loci for common susceptibility variants to be 1052 ± 772 for endometrial cancer.[40] Thus, to date, the identified risk loci for endometrial cancer only account for a small proportion of these variants and a large number remain to be detected by future well-powered studies.

## Functional analysis of endometrial cancer GWAS to inform endometrial cancer biology

At the 20 endometrial cancer risk regions described above, the vast majority of risk variation is located in non-coding regions, similar to most trait-associated GWAS loci, making it difficult to determine its function and identify genes that mediate its effects. Nevertheless, as the discovery of GWAS target genes, especially those with known function, can inform biology of the GWAS trait, post-GWAS functional analysis is a very area of active research (reviewed in).[41] Importantly, the proteins encoded by endometrial cancer GWAS genes may provide effective therapeutic targets as they are related to disease causality and this approach is supported by findings that GWAS of relevant traits or diseases have led to the rediscovery of many known drug targets.[9]

### Locus specific functional analysis

Functional genomic analyses of endometrial cell lines have indicated that non-coding regulatory elements (e.g. enhancers and promoters) are enriched for endometrial cancer GWAS risk variation[13] and, thus, the

| Risk locus | Independent index variant/s | Identification method of index variant/s | Studies reporting association |
|---|---|---|---|
| 1p34.3 | rs113998067 | Endometrial cancer GWAS | 13 |
| 1p36.12 | rs3820282 | Multi-trait analysis of endometrial cancer, PCOS and uterine fibroids GWAS | 20 |
| 2p16.1 | rs148261157 | Endometrial cancer GWAS | 13 |
| 5p15.33 | rs7725218 | Multi-trait analysis of endometrial cancer and ovarian cancer GWAS | 38 |
| 6q22.3 | rs1740828 | Endometrial cancer GWAS | 13,20,76 |
| 6q22.31 | rs2747716 | Endometrial cancer GWAS | 13,42,76 |
| 8q24.21 | rs35286446, rs4733613, rs139584729 | Endometrial cancer GWAS | 13,20,38,42 |
| 9p23 | rs2475339 | Multi-trait analysis of endometrial cancer and ovarian cancer GWAS | 38 |
| 9p21.3 | rs1679014 | Endometrial cancer GWAS | 13 |
| 9q34.2 | rs687289 | Multi-trait analysis of endometrial cancer and ovarian cancer GWAS | 38 |
| 10p12.31 | rs564819152 | Multi-trait analysis of endometrial cancer and ovarian cancer GWAS | 38 |
| 11p13 | rs10835920 | Endometrial cancer GWAS | 13,20 |
| 12p12.1 | rs9668337 | Endometrial cancer GWAS | 13 |
| 12q24.11 | rs3184504 | Endometrial cancer GWAS | 13,20,77 |
| 12q24.21 | rs10850382 | Endometrial cancer GWAS | 13 |
| 13q22.1 | rs7981863 | Endometrial cancer GWAS | 13,20,76 |
| 15q15.1 | rs937213 | Endometrial cancer GWAS | 13,42 |
| 15q21.2 | rs17601876 | Endometrial cancer GWAS | 13,20,34,42 |
| 17q11.2 | rs1129506 | Endometrial cancer GWAS | 13,20 |
| 17q12 | rs11263761 | Endometrial cancer GWAS | 11,13,38,42,43,76 |
| 17q21.32 | rs882380 | Endometrial cancer GWAS | 13,38 |

***Table 2*: Genetic risk regions identified for endometrial cancer by genome-wide association studies.**

targets of these elements could reveal candidate endometrial cancer susceptibility genes. Locus specific functional genetic studies have been performed at two endometrial cancer GWAS risk regions (13q22.1 and 17q12), providing evidence that two transcription factor genes (*KLF5* and *HNF1B*, respectively; **Table 3**) may be upregulated by risk variation located in regulatory regions.[42,43] Notably, *KLF5* and *HNF1B* have functions that are related to genetically established endometrial cancer risk factors: *KLF5* is involved in adipocyte differentiation[44]; and *HNF1B* is involved in insulin secretion.[45]

## Use of functional genomic data to interpret endometrial cancer risk loci

The functional analyses used to identify *KLF5* and *HNF1B* cannot easily be applied in a systematic fashion. Functional data generated at a genome-wide level are needed to efficiently identify candidate target genes at multiple risk loci. To address this issue, chromatin looping data have been generated in three tumoural and one normal immortalized endometrial cell lines, and identified chromatin loops linking enhancers with promoters.[46] Notably, endometrial cancer GWAS risk variation was enriched in the anchors of these enhancer/promoter-associated loops. Furthermore, by intersecting candidate causal risk variants with these loops, evidence was found for potential targeting of 103 genes at 13 of the 16 GWAS risk loci studied.[46] Importantly, these candidate target genes had biological relevance to endometrial cancer as evidenced by enrichment for: (i) genes that are differentially expressed in endometrial tumours; and (ii) somatically mutated cancer drivers among a protein-protein interaction network derived from the candidate gene set.[46] The enhancer/promoter chromatin looping data has also been used to assess the four endometrial cancer risk loci recently identified through cross-trait analyses, revealing further candidate target genes.[20,38] As these analyses have provided a relatively large number of candidate target genes, even allowing for multiple causal genes at each locus, these genes require prioritisation using additional data.

Expression quantitative trait loci (eQTL) data can identify candidate target genes at GWAS loci by revealing associations between gene expression and GWAS variation. This evidence is particularly compelling if the eQTL and GWAS variation colocalise, increasing the likelihood of causality.[47] Such an approach has been used to identify six candidate target genes at endometrial cancer GWAS loci, five of which had also been revealed through chromatin looping analyses (**Table 3**). eQTL data can be further integrated with GWAS data through transcriptome-wide association studies (TWAS) to assess associations between genetically predicted gene expression and GWAS traits.[48] Furthermore, data from multiple tissues can be combined to increase statistical power to detect associations. Using

| Locus | Gene | Protein Coding | Druggable[a] | Evidence[b] |
|---|---|---|---|---|
| 1p34.3 | *GNL2* | Yes | No | HiChIP (2), nearest gene |
| 1p36.12 | *CDC42* | Yes | Yes | HiChIP (4), eQTL (blood) |
| | *RAP1GAP* | Yes | No | HiChIP (2) |
| | *WNT4* | Yes | Yes | HiChIP (2), nearest gene |
| 2p16.1 | *BCL11A* | Yes | No | HiChIP (1), eQTL (endometrial tumor), nearest gene |
| 3q21.3 | *EEFSEC* | Yes | No | TWAS |
| 5p15.33 | *LPCAT1* | Yes | No | HiChIP (2) |
| | *MIR4635* | No | No | HiChIP (2) |
| | *NKD2* | Yes | No | HiChIP (2) |
| | *SLC6A18* | Yes | Yes | HiChIP (2) |
| | *SLC6A19* | Yes | Yes | HiChIP (3) |
| | *TERT* | Yes | Yes | HiChIP (3), nearest gene |
| 6q22.31 | *HEY2* | Yes | No | TWAS, nearest gene |
| 8q24.21 | *MYC* | Yes | No | HiChIP (4), pan-cancer driver gene |
| 9p21.3 | *CDKN2A* | Yes | No | HiChIP (2), pan-cancer driver gene |
| | *CDKN2B-AS1* | No | No | HiChIP (2), nearest gene |
| | *MIR31HG* | No | No | HiChIP (2) |
| 9q34.2 | *ABO* | Yes | No | eQTL (blood & endometrial tumor), circulating protein association, nearest gene |
| | *ADAMTS13* | Yes | Yes | HiChIP (2), nearest gene |
| | *ADAMTSL2* | Yes | No | HiChIP (2) |
| | *STKLD1* | Yes | Yes | HiChIP (2) |
| | *CACFD1* | Yes | No | HiChIP (3), nearest gene |
| | *CEL* | Yes | Yes | HiChIP (2) |
| | *CELP* | No | No | HiChIP (2) |
| | *DBH* | Yes | Yes | HiChIP (3) |
| | *DBH-AS1* | No | No | HiChIP (2) |
| | *FAM163B* | Yes | No | HiChIP (2) |
| | *LINC00094* | No | No | HiChIP (2) |
| | *MED22* | Yes | No | HiChIP (2) |
| | *OBP2B* | Yes | Yes | HiChIP (2) |
| | *RPL7A* | Yes | No | HiChIP (2) |
| | *RXRA* | Yes | Yes | HiChIP (2) |
| | *SLC2A6* | Yes | Yes | HiChIP (1), nearest gene |
| | *SNORD24* | No | No | HiChIP (2) |
| | *SNORD36A* | No | No | HiChIP (2) |
| | *SNORD36B* | No | No | HiChIP (2) |
| | *SNORD36C* | No | No | HiChIP (2) |
| | *SURF1* | Yes | No | HiChIP (2) |
| | *SURF2* | Yes | No | HiChIP (2) |
| | *SURF4* | Yes | No | HiChIP (2) |
| 10p12.31 | *MIR1915HG* | No | No | HiChIP (3), nearest gene |
| | *DNAJC1* | Yes | No | HiChIP (1), nearest gene |
| | *MIR1915* | No | No | HiChIP (3), nearest gene |
| | *MLLT10* | Yes | No | HiChIP (2), nearest gene |
| | *SKIDA1* | Yes | No | HiChIP (3), nearest gene |
| 11p13 | *WT1* | Yes | Yes | HiChIP (1), pan-cancer driver gene, nearest gene |
| | *WT1-AS* | No | No | HiChIP (1), nearest gene |
| 12p12.1 | *BHLHE41* | Yes | No | HiChIP (2) |
| | *SSPN* | Yes | No | HiChIP (1), nearest gene |
| 12q24.11 | *SH2B3* | Yes | No | HiChIP (3) |
| 12q24.21 | *TBX3* | Yes | No | HiChIP (1), pan-cancer driver gene |
| 13q22.1 | *KLF5* | Yes | Yes | Functional study,[42] pan-cancer driver gene |
| 15q.15.1 | *AC021755.3* | No | No | TWAS |

***Table 3*** (Continued)

| Locus | Gene | Protein Coding | Druggable[a] | Evidence[b] |
|---|---|---|---|---|
| | *C15orf56* | Yes | No | HiChIP (4) |
| | *ANKRD63* | Yes | No | HiChIP (2) |
| | *BAHD1* | Yes | No | HiChIP (3) |
| | *BMF* | Yes | No | HiChIP (3) |
| | *CCDC9B* | Yes | No | HiChIP (4) |
| | *EIF2AK4* | Yes | Yes | TWAS, HiChIP (1), nearest gene |
| | *GPR176* | Yes | Yes | HiChIP (2) |
| | *INAFM2* | Yes | No | HiChIP (2) |
| | *KNSTRN* | Yes | No | HiChIP (2) |
| | *PAK6* | Yes | Yes | HiChIP (4) |
| | *PLCB2* | Yes | Yes | HiChIP (3) |
| | *PLCB2-AS1* | No | No | HiChIP (2) |
| | *SRP14* | Yes | No | HiChIP (4), eQTL (blood), nearest gene |
| | *SRP14-AS1* | No | No | HiChIP (4), nearest gene |
| 15q21.2 | *CYP19A1* | Yes | Yes | TWAS, nearest gene |
| 17q11.2 | *EVI2A* | Yes | No | TWAS, nearest gene |
| | *MIR193A* | No | No | HiChIP (2) |
| | *RAB11FIP4* | Yes | No | HiChIP (3), nearest gene |
| | *RNU6ATAC7P* | No | No | HiChIP (2) |
| 17q12 | *HNF1B* | Yes | No | HiChIP (1), functional study,[43] nearest gene |
| 17q21.32 | *CBX1* | Yes | No | HiChIP (3), nearest gene |
| | *HOXB2* | Yes | No | HiChIP (1), eQTL (blood) |
| | *HOXB3* | Yes | No | HiChIP (3) |
| | *HOXB4* | Yes | No | HiChIP (3) |
| | *HOXB6* | Yes | No | HiChIP (2) |
| | *HOXB7* | Yes | No | HiChIP (3) |
| | *HOXB8* | Yes | No | HiChIP (4) |
| | *HOXB9* | Yes | No | HiChIP (3) |
| | *MIR10A* | No | No | HiChIP (3) |
| | *MIR1203* | No | No | HiChIP (3), nearest gene |
| | *MIR196A1* | No | No | HiChIP (2) |
| | *PRR15L* | Yes | No | HiChIP (2) |
| | *SKAP1* | Yes | No | TWAS, HiChIP (2), nearest gene |
| | *SKAP1-AS1* | No | No | HiChIP (2), nearest gene |
| | *SNX11* | Yes | No | TWAS, HiChIP (4), eQTL (blood), nearest gene |

***Table 3*: Candidate endometrial cancer GWAS target genes prioritised with supporting evidence.**

[a] Genes encoding druggable proteins were identified using The Drug Gene Interaction Database (https://www.dgidb.org/)

[b] Number of cell lines HiChIP chromatin looping observed provided in parentheses; references for functional studies are provided; pan-cancer driver genes are those listed in Bailey et al.[56]; TWAS genes, eQTLs and HiChIP targets are those reported using data from Kho et al.,[49] O'Mara et al[46] and Glubb et al.[38]

this multi-tissue TWAS approach, in conjunction with prioritisation by colocalisation or Mendelian randomisation analyses, eight candidate endometrial cancer susceptibility genes have been identified.[49] Seven of the eight candidate susceptibility genes were located at endometrial cancer GWAS susceptibility loci, with the novel candidate susceptibility gene *EEFSEC* located at a potentially novel risk locus (Table 3). Importantly, these approaches have allowed the identification of tissue specific associations. For example, adipose-specific expression of *CYP19A1*, encoding the aromatase enzyme, was associated with endometrial cancer risk, concordant with the production of oestrogen by aromatase in the adipose of postmenopausal women.[50] Using a similar approach to TWAS, the genetically predicted levels of nine circulating proteins have been found to associate with endometrial cancer risk,[51] one of which is encoded by *ABO*, a candidate target gene also identified through eQTL analyses (Table 3).

## Prioritisation of candidate target genes and further biological interpretation

It has been suggested that at ~30-70% of GWAS loci[52,53] the causal gene is the nearest gene to the GWAS variation and thus proximity of a gene to GWAS variation can be used in candidate target gene prioritisation. Candidate target genes that have been established

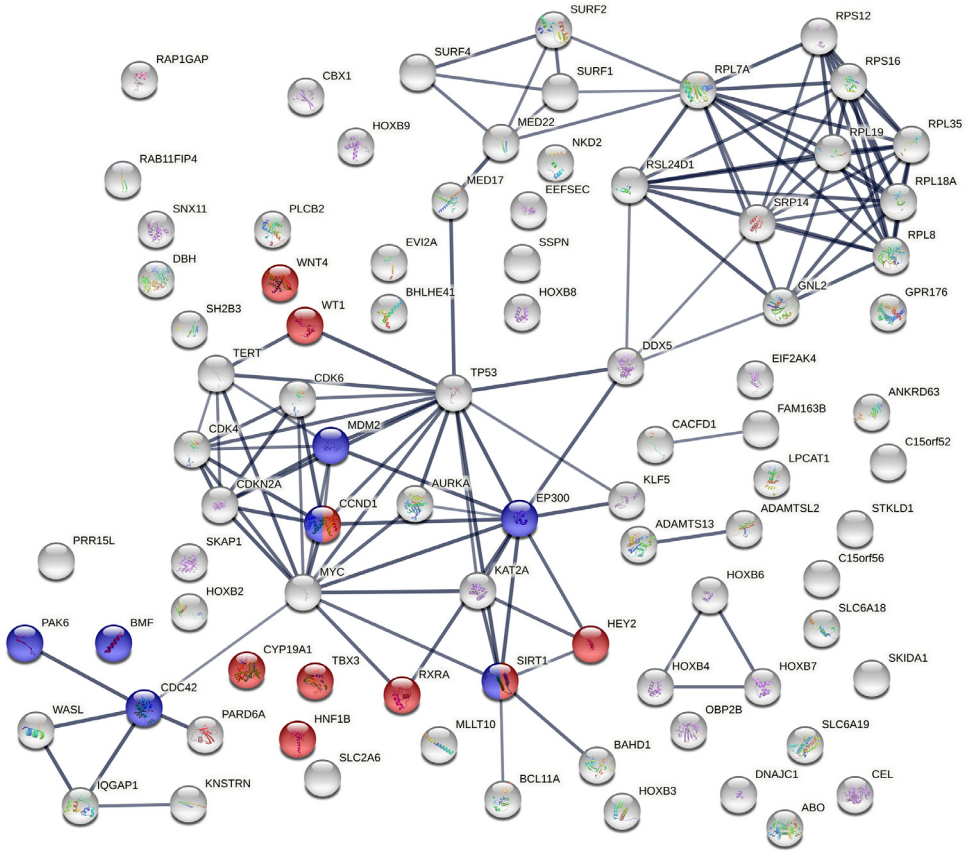


**Fig. 2. Network of candidate target genes found at endometrial cancer risk loci.** The network and gene enrichment analyses were performed using STRING version 11.5 (https://string-db.org/). Genes enriched in the Wikipathways "androgen receptor signaling pathway" are highlighted blue and genes belonging to the "reproductive structure development" ontology are highlighted red.

as pan-cancer driver genes (e.g. *CDKN2A* and *MYC*; Table 3) also support causality as a consequence of their somatic targeting in cancer. Indeed, cancer drivers have been found to be enriched among candidate susceptibility genes identified from GWAS of other cancers.[54,55] Integrating this information and the functional genomic data discussed above, we have prioritized a set of 88 candidate target genes that have at least two lines of evidence for causality at 19 of the 20 endometrial cancer GWAS risk loci (Table 3). This includes genes with evidence of potential targeting by endometrial GWAS variation through chromatin looping in at least two endometrial cell lines or those identified by TWAS with statistical prioritisation for causality. Importantly for drug development, proteins encoded by 19 of these genes are considered to be druggable (Table 3).

Network analysis of the prioritised candidate target gene set shows interactions and shared relationships (Figure 2), linking multiple candidate target genes and providing evidence that they interact with established endometrial cancer drivers such as *TP53*, *CCND1* and *EP300*.[56] Further analysis of this network demonstrates enrichment of genes in relevant gene ontologies (e.g. reproductive structure development) or pathways related to endometrial cancer risk factors such as testosterone (e.g. androgen receptor signalling) (Figure 2). These findings provide further insight into the biology underlying the endometrial cancer risk loci. For example, at the 1p36.12 locus, found through the cross-trait analysis with uterine fibroids and PCOS, one candidate target gene, *WNT4*, relevantly belongs to the reproductive structure development gene ontology; while another, *CDC42*, encodes a component of the androgen receptor signalling pathway, suggesting the involvement of testosterone in disease development related to this locus.

## Translation from endometrial cancer GWAS: population stratification based on genetic risk scores

Diagnosis of endometrial cancer is largely based on histological analysis of samples biopsied after presentation with postmenopausal vaginal bleeding, which occurs in

around 90% of postmenopausal women with endometrial cancer. However, only 9% of women have postmenopausal bleeding due to endometrial cancer.[57] Thus, it would be useful to identify women with high life-time risk of developing endometrial cancer for symptom education and their prioritisation for endometrial sampling upon postmenopausal bleeding. This would also allow for the development of targeted intervention strategies for high risk women, such as weight-loss or progestin treatment and could be potentially achieved by the incorporation of polygenic risk scores (PRS) that summarise the effect of endometrial cancer GWAS risk variation. Indeed, studies have reported comparable predictive performance of PRS to clinical risk factors for many diseases.[58]

Theoretical PRS including variously 19 and 24 curated endometrial cancer risk variants predicted a 1.94 to 3.16-fold increase in endometrial cancer risk between women in the top 1% of the PRS distribution compared with the mean,[59] highlighting the potential of this approach. Further, a PRS generated using nine curated endometrial cancer risk variants and tested in the UK Biobank and the Genetic Epidemiology Research on Adult Health and Aging (GERA) cohorts reported that one standard deviation increase in PRS is associated with a 19% increased risk of endometrial cancer.[60] These results are similar to a PRS tested in the UK Biobank dataset only, using the same endometrial cancer risk variants and weights.[61]

PRS generated from currently known endometrial cancer risk variants are limited by their low discriminatory power, with area under the curve (AUC) results for PRS predictability reported around 0.56-0.57.[62,63] Including risk factors, such as BMI and parity, into the PRS, has led to a slight improvement in distinguishing cancer cases from cancer-free individuals.[61] This suggests that risk prediction for endometrial cancer could be further improved by the integration of disease PRS with risk factors to generate a multi-trait PRS (multi-PRS), as has been shown for other complex traits, including ischemic stroke[58] and type 2 diabetes.[64]

There is scope to integrate an endometrial cancer PRS in the familial cancer setting to identify women that are more likely to develop endometrial cancer and would benefit from prevention interventions. For example, while women from Lynch Syndrome families are at a very high risk of endometrial cancer compared to the general population, only 50% of them will develop the cancer. This approach would be similar to that used by a study in prostate cancer, demonstrating the benefit of combining common and rare genetic variants in stratifying men into low and high risk groups to inform clinical management.[65]

## Outstanding questions

Several areas need specific attention to advance progress in this field. Firstly, similar to other GWAS studies, the majority of endometrial cancer GWAS have been performed using European ancestry populations, with only a few non-European GWAS studies (a Chinese[66] and two Japanese[19,67]) conducted to date. This limitation impacts our ability to identify non-European-specific endometrial cancer risk loci, but perhaps more importantly also affects the potential to translate results from current endometrial cancer GWAS into PRS for diverse populations. Secondly, no progress has been made in the identification of risk loci for non-endometrioid endometrial cancers. Acquisition of such sample sets should be a focus of future endometrial cancer GWAS to understand these less common, but more aggressive cancers. Lastly, functional genomic studies have been performed in endometrial cell lines to infer targeting by endometrial cancer GWAS risk variation. However, cell lines by their nature have limitations as experimental models and it is not known how well cell lines model gene regulation in endometrial tissue or tumours. Recently, organoids derived from a range of endometrial tumour types, as well as normal and hyperplastic endometrium, have been established. These organoids accurately recapitulate the morphological and molecular features of the corresponding tissues from which they were derived,[68] providing highly relevant experimental systems for future functional genomic studies.

## Conclusions

It is evident that great strides have been made in understanding the role of common genetic variation in endometrial cancer risk since the first GWAS performed for this disease in 2011. Endometrial cancer GWAS data have provided insights into its aetiology by demonstrating that a number of related traits or established risk factors share a common genetic background, teasing apart relationships that may be mediated by BMI and providing evidence of causality for novel potential risk factors such as LDL cholesterol. Moreover, traits that are genetically related to endometrial cancer can be used to identify novel risk loci in cross-trait GWAS. Functional analyses of endometrial cancer GWAS loci have prioritized candidate susceptibility genes and enable insight into endometrial cancer biology through their function. In the next 10 years, significant progress should be made in all these areas in addition to the growth in the size and diversity of endometrial cancer GWAS datasets. This outcome may contribute to the most immediate avenue for clinical translation, which is the development of clinical endometrial cancer PRS studies that would establish screening for women at high disease risk.

## Search strategy and selection criteria

Data for this Review were identified by searches of PubMed, Google Scholar and references from relevant

articles using the search terms "endometrial cancer GWAS" and "endometrial cancer genome-wide association study". Only articles published in English between 1 April 2011 and 31 October 2021 were included.

## Contributors

Conceptualisation: TOM, DG, XW; literature search: TOM, DG, XW; data analysis: DG, XW; data interpretation: TOM, DG, XW; original draft: TOM, DG, XW; figures and visualization: DG, XW; review and editing: TOM, DG, XW. All authors read and approved the final version of the manuscript.

## Declaration of interests

Dr. O'Mara reports grants from National Health and Medical Research Council of Australia during the conduct of the study.

## Acknowledgements

Tracy O'Mara is supported by a National Health and Medical Research Council (NHMRC) of Australia Investigator Fellowship (APP1173170). Funding sources had no role in study design, data curation and analysis, data interpretation, report writing and submission for publication. Funders had no role in paper design, data collection, data analysis, interpretation, or writing of the paper.

### References

1 Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394–424.
2 Lortet-Tieulent J, Ferlay J, Bray F, Jemal A. International patterns and trends in endometrial cancer incidence, 1978–2013. *JNCI J Natl Cancer Inst*. 2018;110(4):354–361.
3 Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2021. *CA Cancer J Clin*. 2021;71(1):7–33.
4 Amant F, Moerman P, Neven P, Timmerman D, Van Limbergen E, Vergote I. Endometrial cancer. *Lancet N Am Ed*. 2005;366 (9484):491–505.
5 Clarke MA, Devesa SS, Harvey SV, Wentzensen N. Hysterectomy-corrected uterine corpus cancer incidence trends and differences in relative survival reveal racial disparities and rising rates of nonendometrioid cancers. *J Clin Oncol*. 2019;37(22):1895–1908.
6 Johnatty SE, Tan YY, Buchanan DD, et al. Family history of cancer predicts endometrial cancer risk independently of lynch syndrome: implications for genetic counselling. *Gynecol Oncol*. 2017;147 (2):381–387.
7 Ryan NAJ, Glaire MA, Blake D, Cabrera-Dandy M, Evans DG, Crosbie EJ. The proportion of endometrial cancers associated with Lynch syndrome: a systematic review of the literature and meta-analysis. *Genet Med*. 2019;21(10):2167–2180.
8 Dörk T, Hillemanns P, Tempfer C, Breu J, Fleisch MC. Genetic susceptibility to endometrial cancer: risk factors and clinical management. *Cancers (Basel)*. 2020;12(9):2407.
9 Finan C, Gaulton A, Kruger FA, et al. The druggable genome and support for target identification and validation in drug development. *Sci Transl Med*. 2017;9(383):eaag1166.
10 Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. Benefits and limitations of genome-wide association studies. *Nat Rev Genet*. 2019;20(8):467–484.
11 Spurdle AB, Thompson DJ, Ahmed S, et al. Genome-wide association study identifies a common variant associated with risk of endometrial cancer. *Nat Genet*. 2011;43(5):451–454.
12 O'Mara TA, Glubb DM, Kho PF, Thompson DJ, Spurdle AB. Genome-wide association studies of endometrial cancer: latest developments and future directions. *Cancer Epidemiol Biomark Prev*. 2019;28(7):1095–1102.
13 O'Mara TA, Glubb DM, Amant F, et al. Identification of nine new susceptibility loci for endometrial cancer. *Nat Commun*. 2018;9 (1):3166.
14 Raglan O, Kalliala I, Markozannes G, et al. Risk factors for endometrial cancer: an umbrella review of the literature. *Int J Cancer*. 2019;145(7):1719–1730.
15 Mogensen JB, Kjaer SK, Mellemkjaer L, Jensen A. Endometriosis and risks for ovarian, endometrial and breast cancers: a nationwide cohort study. *Gynecol Oncol*. 2016;143(1):87–92.
16 Yu HC, Lin CY, Chang WC, et al. Increased association between endometriosis and endometrial cancer a nationwide population-based retrospective cohort study. *Int J Gynecol Cancer*. 2015;25 (3):447–452.
17 Yin WY, Falconer H, Yin L, Xu LZ, Ye W. Association between polycystic ovary syndrome and cancer risk. *JAMA Oncol*. 2019;5 (1):106–107.
18 Barry JA, Azizia MM, Hardiman PJ. Risk of endometrial, ovarian and breast cancer in women with polycystic ovary syndrome: a systematic review and meta-analysis. *Hum Reprod Update*. 2014;20 (5):748–758.
19 Masuda T, Low SK, Akiyama M, et al. GWAS of five gynecologic diseases and cross-trait analysis in Japanese. *Eur J Hum Genet*. 2020;28(1):95–107.
20 Kho PF, Mortlock S, Rogers PA, et al. Genetic analyses of gynecological disease identify genetic relationships between uterine fibroids and endometrial cancer, and a novel endometrial cancer genetic risk region at the WNT4 1p36. 12 locus. *Hum Genet*. 2021;140(9):1353–1365.
21 Painter JN, O'Mara TA, Morris AP, et al. Genetic overlap between endometriosis and endometrial cancer: evidence from cross-disease genetic correlation and GWAS meta-analyses. *Cancer Med*. 2018;7 (5):1978–1987.
22 Poole EM, Lin WT, Kvaskoff M, De Vivo I, Terry KL, Missmer SA. Endometriosis and risk of ovarian and endometrial cancers in a large prospective cohort of U.S. nurses. *Cancer Causes Control*. 2017;28(5):437–445.
23 Cramer DW. The epidemiology of endometrial and ovarian cancer. *Hematol Oncol Clin N Am*. 2012;26(1):1–12.
24 Nead KT, Sharp SJ, Thompson DJ, et al. Evidence of a causal association between insulinemia and endometrial cancer: a mendelian randomization analysis. *JNCI J Natl Cancer Inst*. 2015;107(9).
25 Painter JN, O'Mara TA, Marquart L, et al. Genetic risk score mendelian randomization shows that obesity measured as body mass index, but not waist: hip ratio, is causal for endometrial cancer. *Cancer Epidemiol Biomark*. 2016;25(11):1503–1510.
26 Masuda T, Ogawa K, Kamatani Y, Murakami Y, Kimura T, Okada Y. A mendelian randomization study identified obesity as a causal risk factor of uterine endometrial cancer in Japanese. *Cancer Sci*. 2020;111(12):4646–4651.
27 Dimou NL, Papadimitriou N, Mariosa D, et al. Circulating adipokine concentrations and risk of five obesity-related cancers: a mendelian randomization study. *Int J Cancer*. 2021;148(7):1625–1636.
28 Yuan S, Carter P, Bruzelius M, et al. Effects of tumour necrosis factor on cardiovascular disease and cancer: a two-sample Mendelian randomization study. *eBioMedicine*. 2020;59: 102956.
29 Yuan S, Kar S, Carter P, et al. Is type 2 diabetes causally associated with cancer risk? Evidence from a two-sample mendelian randomization study. *Diabetes*. 2020;69(7):1588–1596.
30 Kho PF, Amant F, Annibali D, et al. Mendelian randomization analyses suggest a role for cholesterol in the development of endometrial cancer. *Int J Cancer*. 2020.
31 Day FR, Thompson DJ, Helgason H, et al. Genomic analyses identify hundreds of variants associated with age at menarche and support a role for puberty timing in cancer risk. *Nat Genet*. 2017;49 (6):834–841.
32 Ruth KS, Day FR, Hussain J, et al. Genetic insights into biological mechanisms governing human ovarian ageing. *Nature*. 2021;596 (7872):393–397.
33 Rodriguez AC, Blanchard Z, Maurer KA, Gertz J. Estrogen signaling in endometrial cancer: a key oncogenic pathway with several open questions. *Horm Cancer*. 2019;10(2-3):51–63.

34 Thompson DJ, O'Mara TA, Glubb DM, et al. CYP19A1 fine-mapping and mendelian randomization: estradiol is causal for endometrial cancer. *Endocr Relat Cancer*. 2016;23(2):77–91.
35 Ruth KS, Day FR, Tyrrell J, et al. Using human genetics to understand the disease impacts of testosterone in men and women. *Nat Med*. 2020;26(2):252–258.
36 Mullee A, Dimou N, Allen N, O'Mara T, Gunter MJ, Murphy N. Testosterone, sex hormone-binding globulin, insulin-like growth factor-1 and endometrial cancer risk: observational and Mendelian randomization analyses. *Br J Cancer*. 2021;125(9):1308–1317.
37 Stewart EA, Cookson CL, Gandolfo RA, Schulze-Rath R. Epidemiology of uterine fibroids: a systematic review. *BJOG*. 2017;124(10):1501–1512.
38 Glubb DM, Thompson DJ, Aben KKH, et al. Cross-cancer genome-wide association study of endometrial cancer and epithelial ovarian cancer identifies genetic risk regions associated with risk of both cancers. *Cancer Epidemiol Biomark Prev*. 2021;30(1):217–228.
39 Kar SP, Lindstroem S, Hung R, et al. Combining genome-wide studies of breast, prostate, ovarian and endometrial cancers maps cross-cancer susceptibility loci and identifies new genetic associations. bioRxiv 2020. https://doi.org/10.1101/2020.06.16.146803.
40 Zhang YD, Hurson AN, Zhang H, et al. Assessment of polygenic architecture and risk prediction based on common variants across fourteen cancers. *Nat Commun*. 2020;11(1):3353.
41 Gallagher MD, Chen-Plotkin AS. The post-GWAS era: from association to function. *Am J Hum Genet*. 2018;102(5):717–730.
42 Cheng TH, Thompson DJ, O'Mara TA, et al. Five endometrial cancer risk loci identified through genome-wide association analysis. *Nat Genet*. 2016;48(6):667–674.
43 Painter JN, O'Mara TA, Batra J, et al. Fine-mapping of the HNF1B multicancer locus identifies candidate variants that mediate endometrial cancer risk. *Hum Mol Genet*. 2015;24(5):1478–1492.
44 Oh GS, Yoon J, Kim G, et al. Regulation of adipocyte differentiation by clusterin-mediated Krüppel-like factor 5 stabilization. *FASEB J*. 2020;34(12):16276–16290.
45 Quilichini E, Fabre M, Nord C, et al. Insights into the etiology and physiopathology of MODY5/HNF1B pancreatic phenotype with a mouse model of the human disease. *J Pathol*. 2021;254(1):31–45.
46 O'Mara TA, Spurdle AB, Glubb DM, Endometrial Cancer Association Consortium. Analysis of promoter-associated chromatin interactions reveals biologically relevant candidate target genes at endometrial cancer risk loci. *Cancers*. 2019;11(10):1440.
47 Wang X, Goldstein DB. Enhancer domains predict gene pathogenicity and inform gene discovery in complex disease. *Am J Hum Genet*. 2020;106(2):215–233.
48 Wainberg M, Sinnott-Armstrong N, Mancuso N, et al. Opportunities and challenges for transcriptome-wide association studies. *Nat Genet*. 2019;51(4):592–599.
49 Kho PF, Wang X, Cuéllar-Partida G, et al. Multi-tissue transcriptome-wide association study identifies eight candidate genes and tissue-specific gene expression underlying endometrial cancer susceptibility. *Commun Biol*. 2021;4(1):1211.
50 Ma CX, Reinert T, Chmielewska I, Ellis MJ. Mechanisms of aromatase inhibitor resistance. *Nat Rev Cancer*. 2015;15(5):261–275.
51 Zhu J, O'Mara TA, Liu D, et al. Associations between genetically predicted circulating protein concentrations and endometrial cancer risk. *Cancers (Basel)*. 2021;13(9).
52 Stacey D, Fauman EB, Ziemek D, et al. ProGeM: a framework for the prioritization of candidate causal genes at molecular quantitative trait loci. *Nucleic Acids Res*. 2019;47(1):e3.
53 Barbeira AN, Bonazzola R, Gamazon ER, et al. Exploiting the GTEx resources to decipher the mechanisms at GWAS loci. *Genome Biol*. 2021;22(1):49.
54 Chen Z, Wen W, Beeghly-Fadiel A, et al. Identifying putative susceptibility genes and evaluating their associations with somatic mutations in human cancers. *Am J Hum Genet*. 2019;105(3):477–492.
55 Michailidou K, Lindstrom S, Dennis J, et al. Association analysis identifies 65 new breast cancer risk loci. *Nature*. 2017;551(7678):92–94.
56 Bailey MH, Tokheim C, Porta-Pardo E, et al. Comprehensive characterization of cancer driver genes and mutations. *Cell*. 2018;173(2):371. -85 e18.
57 Clarke MA, Long BJ, Del Mar Morillo A, Arbyn M, Bakkum-Gamez JN, Wentzensen N. Association of endometrial cancer risk with postmenopausal bleeding in women: a systematic review and meta-analysis. *JAMA Intern Med*. 2018;178(9):1210–1222.
58 Abraham G, Malik R, Yonova-Doing E, et al. Genomic risk score offers predictive performance comparable to clinical risk factors for ischaemic stroke. *Nat Commun*. 2019;10(1):5819.
59 Bafligil C, Thompson DJ, Smith MJ, Aj Ryan N, Gareth D, Crosbie EJ. Association between genetic polymorphisms and endometrial cancer risk: a systematic review. *J Med Genet*. 2020;0:1–10.
60 Graff RE, Cavazos TB, Thai KK, et al. Cross-cancer evaluation of polygenic risk scores for 16 cancer types in two large cohorts. *Nat Commun*. 2021;12(1):970.
61 Kachuri L, Graff RE, Smith-Byrne K, et al. Pan-cancer analysis demonstrates that integrating polygenic risk scores with modifiable risk factors improves risk prediction. *Nat Commun*. 2020;11(1):6084.
62 Choi J, Jia G, Wen W, Long J, Zheng W. Evaluating polygenic risk scores in assessing risk of nine solid and hematologic cancers in European descendants. *Int J Cancer*. 2020;147(12):3416–3423.
63 Fritsche LG, Patil S, Beesley LJ, et al. Cancer PRSweb: an online repository with polygenic risk scores for major cancer traits and their evaluation in two independent biobanks. *Am J Hum Genet*. 2020;107(5):815–836.
64 Sinnott-Armstrong N, Tanigawa Y, Amar D, et al. Genetics of 35 blood and urine biomarkers in the UK Biobank. *Nat Genet*. 2021;53(2):185–194.
65 Darst BF, Sheng X, Eeles RA, Kote-Jarai Z, Conti DV, Haiman CA. Combined effect of a polygenic risk score and rare genetic variants on prostate cancer risk. *Eur Urol*. 2021;80(2):134–138.
66 Long J, Zheng W, Xiang YB, et al. Genome-wide association study identifies a possible susceptibility locus for endometrial cancer. *Cancer Epidemiol Biomark Prev*. 2012;21(6):980–987.
67 Ishigaki K, Akiyama M, Kanai M, et al. Large-scale genome-wide association study in a Japanese population identifies novel susceptibility loci across different diseases. *Nat Genet*. 2020;52(7):669–679.
68 Boretto M, Maenhoudt N, Luo X, et al. Patient-derived organoids from endometrial disease capture clinical heterogeneity and are amenable to drug screening. *Nat Cell Biol*. 2019;21(8):1041–1051.
69 Prescott J, Setiawan VW, Wentzensen N, et al. Body mass index genetic risk score and endometrial cancer risk. *PLoS One*. 2015;10(11):e0143256.
70 Seyed Khoei N, Carreras-Torres R, Murphy N, et al. Genetically raised circulating bilirubin levels and risk of ten cancers: a mendelian randomization study. *Cells*. 2021;10(2):394.
71 Kho PF, Glubb DM, Thompson DJ, Spurdle AB, O'Mara TA. Assessing the role of selenium in endometrial cancer risk: a mendelian randomization study. *Front Oncol*. 2019;9:182.
72 Ong JS, Dixon-Suen SC, Han X, et al. A comprehensive re-assessment of the association between vitamin D and cancer susceptibility using Mendelian randomization. *Nat Commun*. 2021;12(1):246.
73 Senkevich K, Bandres-Ciga S, Yu E, et al. No evidence for a causal relationship between cancers and parkinson's disease. *J Parkinsons Dis*. 2021;11(2):801–809.
74 Larsson SC, Carter P, Kar S, et al. Smoking, alcohol consumption, and cancer: a mendelian randomisation study in UK Biobank and international genetic consortia participants. *PLoS Med*. 2020;17:(7)e1003178.
75 Telomeres Mendelian Randomization CollaborationHaycock PC, Burgess S, et al. Association between telomere length and risk of cancer and non-neoplastic diseases: a mendelian randomization study. *JAMA Oncol*. 2017;3(5):636–651.
76 Chen MM, O'Mara TA, Thompson DJ, et al. GWAS meta-analysis of 16 852 women identifies new susceptibility locus for endometrial cancer. *Hum Mol Genet*. 2016;25(12):2612–2620.
77 Cheng TH, Thompson D, Painter J, et al. Meta-analysis of genome-wide association studies identifies common susceptibility polymorphisms for colorectal and endometrial cancer near SH2B3 and TSHZ1. *Sci Rep*. 2015;5:17369.