medRxiv preprint doi: https://doi.org/10.1101/2020.12.03.20243758; this version posted December 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Multi-tissue transcriptome-wide association study identifies genetic mechanisms underlying endometrial cancer susceptibility

Pik-Fang Kho[1,2], Gabriel Cuellar Partida[1,#], Thilo Dörk[3], Ellen L. Goode[4], Diether Lambrechts[5,6], Rodney J. Scott[7,8], Endometrial Cancer Association Consortium, Amanda B. Spurdle[1], Tracy A. O'Mara[1*], Dylan M. Glubb[1*]

[1]Department of Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia.

[2]School of Biomedical Science, Queensland University of Technology, Brisbane, Queensland, Australia.

[3]Gynaecology Research Unit, Hannover Medical School, Hannover, Germany.

[4]Department of Health Science Research, Division of Epidemiology, Mayo Clinic, Rochester, MN, USA.

[5]Laboratory for Translational Genetics, Department of Human Genetics, KU Leuven, Leuven, Belgium.

[6]VIB Center for Cancer Biology, Leuven, Belgium.

[7]Division of Molecular Medicine, Pathology North, John Hunter Hospital, Newcastle, New South Wales, Australia.

[8]Discipline of Medical Genetics, School of Biomedical Sciences and Pharmacy, Faculty of Health, University of Newcastle, Callaghan, New South Wales, Australia.

[#]Current address: 23andMe Inc, Sunnyvale, CA, USA.

[*]These authors contributed equally to the work.

Corresponding Author

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.03.20243758; this version posted December 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Dr Tracy O'Mara, PhD, Molecular Cancer Epidemiology Group, QIMR Berghofer Medical Research Institute, 300 Herston Road, Brisbane QLD Australia 4006. Phone: +61 7 3362 0389. Email: Tracy.OMara@qimrberghofer.edu.au

Abstract

Genome-wide association studies (GWAS) of endometrial cancer have identified 16 genetic susceptibility loci. To identify candidate endometrial cancer susceptibility genes, we have performed the first transcriptome-wide association study (TWAS) of endometrial cancer. For this analysis, we have used the largest endometrial cancer GWAS and gene expression from 48 tissues to maximise statistical power. We used colocalisation analysis to prioritise seven candidate susceptibility genes, including *CYP19A1*, which we previously established as an endometrial cancer susceptibility gene. Notably, one of the candidate susceptibility genes was located at 17q24.2, a potentially novel risk locus for endometrial cancer. Using phenome-wide association analysis, candidate susceptibility genes were found to associate with traits related to endometrial cancer risk factors, in addition to other clinical phenotypes that may provide novel risk factors. TWAS data were also used to perform drug repurposing analysis and identified 14 compounds; two of these are tubulin inhibitors, a drug class used to treat advanced endometrial cancer. In summary, this study has revealed biologically relevant endometrial cancer susceptibility genes, providing insights into endometrial cancer aetiology and avenues for the development of new treatments.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.03.20243758; this version posted December 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Introduction

Endometrial cancer is the most commonly diagnosed gynaecological cancer in developed countries, with incidence expected to continue to increase over the next decade (reviewed in (Morice, Leary, Creutzberg, Abu-Rustum, & Darai, 2016)). To explore the role of common genetic variation in endometrial cancer susceptibility, we recently conducted a genome-wide association study (GWAS) meta-analysis and identified 16 susceptibility loci (O'Mara et al., 2018). Despite the success of GWAS in identifying genetic variants underlying endometrial cancer risk, biological interpretation remains challenging given that the majority of the GWAS risk variation is located in non-protein-coding regions (reviewed in (O'Mara, Glubb, Kho, Thompson, & Spurdle, 2019)).

Trait-associated loci identified by GWAS have demonstrated enrichment for expression quantitative trait loci (eQTLs), suggesting that changes in gene expression link genetic variation and phenotypes (Consortium et al., 2017). As eQTL data are not typically available for GWAS samples, the transcriptome-wide association study (TWAS) approach has been developed to integrate eQTL and GWAS data from independent sample sets (Gamazon et al., 2015). TWAS uses effect estimates from eQTLs to impute gene expression in GWAS samples, and subsequently associate the imputed gene expression with phenotypes. TWAS has an additional advantage of alleviating the multiple testing penalty of GWAS in statistical inference by testing the imputed expression of thousands of genes rather than millions of genetic variants.

TWAS has identified candidate causal genes for many complex traits, using eQTL data from disease-relevant tissues or cell types (Bien et al., 2019; Gusev et al., 2019; Lamontagne et al., 2018; Wu et al., 2018). However, endometrial or uterine eQTL data are less commonly available than for other tissues or cell types (e.g. whole blood). Nevertheless, Mortlock et al. (2020) have shown that 85% of eQTLs found in the endometrium are shared across multiple tissues, suggesting that data from other tissues could be incorporated to perform endometrial cancer TWAS. Recently, innovative multi-tissue TWAS approaches that leverage shared eQTLs across multiple tissues have become available, and provide increased statistical power to identify

medRxiv preprint doi: https://doi.org/10.1101/2020.12.03.20243758; this version posted December 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

candidate causal genes (Barbeira et al., 2019; Hu et al., 2019). These multi-tissue TWAS approaches account for the correlation of gene expression between tissues and perform multivariate regression to predict gene expression from multiple tissues simultaneously.

In the current study, we identified candidate endometrial cancer susceptibility genes by conducting multi-tissue TWAS analysis, integrating GWAS data from the Endometrial Cancer Association Consortium (O'Mara et al., 2018) and a large-scale cis-eQTL mapping study across 48 tissues from the Genotype-Tissue Expression (GTEx) project (Consortium et al., 2017). We also performed a phenome-wide association analysis to uncover traits that relate to candidate endometrial cancer susceptibility genes and provide insights into endometrial cancer aetiology. Finally, we explored drug repurposing opportunities for endometrial cancer by analysing TWAS findings using the Connectivity Map database (Subramanian et al., 2017) and Open Target platform (Carvalho-Silva et al., 2019).

Results

Multi-tissue TWAS analysis using S-MultiXcan revealed 20 genes whose imputed expression associated with endometrial cancer susceptibility (FDR<0.05), five of which (*BHLHE41*, *CYP19A1*, *EIF2AK4*, *IQSEC1* and *SRCIN1*) also passed a Bonferroni correction threshold ($P<2\times10^{-6}$) (**Table 1, Figures 1 and 2**). Given that the TWAS associations could result from linkage disequilibrium (LD) between genetic variants independently affecting gene expression and endometrial cancer risk, we performed colocalization analysis to reduce LD contamination (Giambartolomei et al., 2014). The associations of *AC145343.2*, *BHLHE41*, *CYP19A1*, *RAB11FIP4*, *RP11-521C20.2*, *SRCIN1* and *SKAP1* were supported by colocalization of eQTL and GWAS signals (**Table 1 and Figure 3**); thus, these seven genes were prioritised as candidate endometrial cancer susceptibility genes. Notably, *AC145343.2* is located further than 1 Mb from any previously identified endometrial cancer GWAS risk locus and may provide a novel risk region.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.03.20243758; this version posted December 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Secondary TWAS analysis of endometrioid endometrial cancer risk identified eight genes (FDR<0.05), seven of which were identified in the main analysis (**Supplementary Table 1 and Supplementary Figure 1**). Additionally, this analysis identified a long non-coding RNA (lncRNA), *RP11-358N4.6*, which was supported by eQTL colocalization but was not found in the main analysis. *RP11-358N4.6* is also located further than 1Mb from any previously identified endometrial cancer GWAS risk locus but was not prioritised for further follow-up as it was only identified through a male-specific tissue (i.e. testis).

Phenome-wide analysis of the seven candidate susceptibility genes highlighted three genes (*CYP19A1*, *AC145343.2* and *RAB11FIP4*) with effects on traits which were grouped into five phenotypic categories: bone, cardiovascular, haematopoiesis, diabetes, sex hormones and liver function (**Supplementary Table 2**). Three of these categories (bone, hormone and diabetes) contain traits that relate to or include genetically established endometrial cancer risk factors (e.g. hyperinsulinemia, circulating estrogen and age at menopause) (O'Mara, Glubb, et al., 2019).

We integrated data from the main TWAS analysis with drug-induced gene expression profiles from the Connectivity Map database and identified 14 compounds with profiles opposing the expression of genes associated with endometrial cancer risk (**Table 2**). These drugs included two tubulin inhibitors, ten ATPase inhibitors, a calcium channel activator and an opioid receptor antagonist. None of the 14 drugs have been approved for endometrial cancer treatment but the tubulin inhibitor vincristine has shown modest activity in a Phase II endometrial cancer study (Broun, Blessing, Eddy, & Adelson, 1993). We also explored drug repurposing opportunities for endometrial cancer using the Open Targets platform to assess the seven candidate susceptibility genes. Only the protein encoded by *CYP19A1* (aromatase), is targeted by drugs that have received clinical approval, with indications for breast cancer and Cushing syndrome (**Table 3**).

Discussion

We prioritised seven candidate susceptibility genes for endometrial cancer using a multi-tissue TWAS approach followed by colocalisation analysis. These genes included *AC145343.2*, located

medRxiv preprint doi: https://doi.org/10.1101/2020.12.03.20243758; this version posted December 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

at a potentially novel endometrial cancer risk region which may be revealed by future better-powered GWAS. A secondary analysis revealed a further potential candidate susceptibility gene through association of imputed expression of *RP11-358N4.6* from testis with endometrioid endometrial cancer risk; however, this finding may not be robust because only a single, male-specific, tissue could genetically predict expression of this gene.

Two of the candidate susceptibility genes encode lncRNAs. Of these only *AC145343.2* appears to have been previously studied and is predicted to regulate apoptotic genes such as *CASP8* and *TNFRSF1A* (Liang et al., 2020). Consistent with an association between increased imputed expression of *AC145343.2* and increased endometrial cancer risk, we observed that *AC145343.2* is upregulated in The Cancer Genome Atlas endometrial tumours compared to normal endometrial or uterine tissue (http://gepia2.cancer-pku.cn/#analysis). Recently, we have identified the candidate susceptibility genes *BHLHE41*, *RAB11FIP4*, *SKAP1* and *SRCIN1* as potential regulatory targets of endometrial cancer GWAS risk variation through enhancer-promoter chromatin looping studies (O'Mara, Spurdle, Glubb, & Endometrial Cancer Association, 2019). Additionally, the TWAS-identified genes *EIF2AK4* and *SNX11*, which were not prioritised by colocalisation, had also been established as potential regulatory targets of endometrial cancer GWAS risk variation (O'Mara, Spurdle, Glubb, & Endometrial Cancer Association, 2019). The current study thus provides a crucial link between candidate GWAS target genes and endometrial cancer risk through their imputed expression.

Among these candidate susceptibility genes, *SRCIN1* (Src kinase signalling inhibitor 1) is a compelling biological candidate. Decreased *SRCIN1* expression was associated with increased endometrial cancer risk in our TWAS analysis. This finding is consistent with the tumour suppressor function of the protein encoded by *SRCIN1* which suppresses Src kinase activity (Cao et al., 2014; Wang et al., 2016; Xu et al., 2015; F. Yang et al., 2017) and the observation that Src kinase expression is increased in endometrial cancer (reviewed in (Grundker, Gunthert, & Emons, 2008)). Consistent with these observations, dasatinib, a Src kinase inhibitor, is being studied in clinical trials of recurrent endometrial cancer (ClinicalTrials.gov identifiers: NCT02059265 and NCT01440998).

medRxiv preprint doi: https://doi.org/10.1101/2020.12.03.20243758; this version posted December 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

In terms of the relationships between the candidate susceptibility genes and endometrial cancer, the best characterised gene is *CYP19A1*. We have previously linked *CYP19A1* to endometrial cancer susceptibility through genetic association with circulating estrogen (Thompson et al., 2016). *CYP19A1* encodes the aromatase enzyme, responsible for a rate-limiting step in the synthesis of estrogen and unopposed estrogen exposure is one of the most well established risk factors for endometrial cancer (Ignatov & Ortmann, 2020). Consistent with its function, increased *CYP19A1* expression was associated with increased endometrial cancer risk in our study. Furthermore, drug repurposing analysis linked *CYP19A1* to aromatase inhibitors (letrozole, anastrozole, exemestane and aminoglutethimide) which are used to treat breast cancer. Notably, letrozole has shown efficacy in treating endometrial hyperplasia (Moradan, Nikkhah, & Mirmohammadkhanai, 2017; Tabatabaie et al., 2013), a precursor to endometrial cancer; however, there is less evidence for the use of aromatase inhibitors to treat advanced endometrial cancer (Lindemann et al., 2014; Ma et al., 2004; Rose et al., 2000).

Phenome-wide lookup of candidate endometrial cancer susceptibility genes revealed associations with traits belonging to bone, hormone and diabetes phenotypic categories, which contain traits already related to endometrial cancer risk (O'Mara, Glubb, et al., 2019). Another phenotypic category contained cardiovascular traits (e.g. ischaemic heart disease and systolic blood pressure) which are associated with genetically established endometrial cancer risk factors such as body mass index, LDL and HDL cholesterol (Kho et al., 2020; O'Mara, Glubb, et al., 2019). These findings suggest that the identified traits belonging to the remaining phenotypic categories (liver function and haematopoiesis) may provide novel insights into endometrial cancer aetiology. GWAS data for these traits could be used to assess their effects on endometrial cancer risk using similar Mendelian randomisation approaches to our studies of body mass index and blood lipids (Kho et al., 2020; O'Mara, Glubb, et al., 2019).

We found 14 compounds that may counteract gene expression changes associated with endometrial cancer susceptibility. As drug targets supported by genetic evidence of disease

medRxiv preprint doi: https://doi.org/10.1101/2020.12.03.20243758; this version posted December 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

association are more likely to receive clinical approval (King, Davis, & Degner, 2019; Nelson et al., 2015), there is a strong rationale for future pre-clinical endometrial cancer sensitivity studies of candidate compounds identified in this analysis. Six of the 14 compounds have either been clinically studied or approved for disease treatment and thus may have potential for drug repurposing. It is noteworthy that the two drugs with the strongest opposing gene expression profiles were the tubulin inhibitors nocodazole and vincristine, as the tubulin inhibitor paclitaxel is currently used to treat endometrial cancer patients with advanced disease (Burke et al., 2014). Although nocodazole does not appear to have been studied in humans, vincristine has been shown to have a limited clinical benefit on endometrial cancer but is associated with significant toxicity (Broun et al., 1993). Nine of the compounds were $Na^{+/}$ K^{+}-ATPase inhibitors, of which one, digoxin, has been approved for treatment of cardiac disease. Na^{+}/K^{+}-ATPase inhibitors have demonstrated anti-cancer activity in experimental models, including bufalin which was identified in our analysis and has shown effects specific to ovarian and endometrial cancer cell lines (Takai, Ueda, Nishida, Nasu, & Narahara, 2008). Although clinical studies have shown limited efficacy of Na^{+}/K^{+}-ATPase in the treatment of cancer, effects on endometrial cancer do not appear to have been studied yet (Durlacher et al., 2015). Cinacalcet, a calcium channel activator, approved for treatment of hypercalcemia and secondary hyperparathyroidism, has demonstrated anti-cancer activity in animal models (Rodríguez-Hernández et al., 2016) but has not yet been clinically studied as a cancer treatment. The Ca^{2+}-ATPase inhibitor thapsigargin has demonstrated strong anti-proliferative effects in endometrial cancer cell lines in the Genomics of Drug Sensitivity in Cancer database (W. Yang et al., 2013). However, thapsigargin is not selective for cancer cells and thus a tumour-targeted derivative (mipsagargin) has been developed (Cui, Merritt, Fu, & Pan, 2017). Mipsagargin has been clinically studied for treatment of several tumour types (not including endometrial cancer) and shown modest activity (Mahalingam et al., 2019; Mahalingam et al., 2016).

We acknowledge some limitations to this study. As our TWAS analysis was performed using eQTL data from multiple tissues to maximise our power to detect candidate susceptibility genes, genes with tissue-specific expression may have been discounted. However, we have analysed the relatively small GTEx uterine eQTL sample set, arguably the most relevant tissue, and failed to

medRxiv preprint doi: https://doi.org/10.1101/2020.12.03.20243758; this version posted December 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

identify genes that passed the TWAS multiple testing correction threshold. Furthermore, although we have applied colocalization analysis to identify concordant GWAS and eQTL signals and reduce false positive findings, the extent to which colocalization results represent genuine causal relationships remains unclear. Indeed, while we had previously established five of the seven genes with concordant signals as candidate endometrial cancer susceptibility genes or GWAS targets, two genes with discordant GWAS and eQTL signals also had evidence of targeting by endometrial cancer risk GWAS variation. Further functional studies will be necessary to determine the biological effects of the candidate susceptibility genes identified in this study on experimental phenotypes related to endometrial cancer risk.

In summary, our study has highlighted the value of using a multi-tissue transcriptome-wide association study and colocalisation analysis to prioritise biologically relevant candidate endometrial cancer susceptibility genes. As well as generating insights into biological mechanisms underlying endometrial cancer risk, we have identified compounds with potential for endometrial cancer repurposing and provided avenues for future studies of endometrial cancer aetiology.

Methods

Endometrial cancer risk GWAS data

We used summary data from the largest endometrial cancer GWAS meta-analysis to date, comprising 12,906 cases and 108,979 controls of European ancestry (O'Mara et al., 2018). Endometrial cancer subtype analyses were additionally performed using GWAS summary data restricting cases to endometrioid endometrial cancer patients (8,758 cases and 46,126 controls) (https://www.ebi.ac.uk/gwas/publications/30093612). For analysis, only variants located outside major histocompatibility complex region (chr6:26Mb–34Mb), were used.

eQTL data

medRxiv preprint doi: https://doi.org/10.1101/2020.12.03.20243758; this version posted December 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

We used cis-eQTL summary statistics from 48 tissues that are available in GTEx Project (version 7). Cis-eQTL summary statistics for 48 tissues and information about respective sample sizes are available from the GTEx website (https://gtexportal.org/home/).

Multi-tissue TWAS analysis

Multi-tissue TWAS analysis was performed in two steps. Firstly, we used S-PrediXcan (Barbeira et al., 2019) to incorporate endometrial cancer GWAS and eQTL data from each of the 48 GTEx tissues (version 7) and identify genes associated with endometrial cancer risk. Pre-specified weights based on covariance of genetic variants were used in S-PrediXcan analysis and obtained from PredictDB data repository (http://predictdb.org/). Secondly, S-PrediXcan data from all tissues were jointly analysed using multivariate regression in S-MultiXcan (Barbeira et al., 2019). To identify statistically significant TWAS genes from S-MultiXcan, a Bonferroni correction was considered. However, this approach would assume independent gene expression, whereas genes tend to be co-expressed (Wainberg et al., 2019). Therefore, we used a Benjamini-Hochberg false discovery rate (FDR) of less than 0.05 to correct for multiple comparisons.

Colocalization analysis

Colocalization analysis was performed for each S-MultiXcan identified gene to assess the co-occurrence of endometrial cancer risk and eQTL signals from the best performing tissue, i.e. the most-significant tissue identified by S-PrediXcan. This approach enabled us to tease apart pleiotropic associations (where the same genetic variant is associated with gene expression and disease risk) from linkage (where correlated genetic variants are associated with gene expression and disease risk independently). Colocalization analysis was performed using the COLOC package with default parameters in R (Giambartolomei et al., 2014) and the input was identical to the TWAS analysis but restricted to variants within 1 Mb of the S-MultiXcan identified genes. COLOC provided posterior probabilities for the GWAS and eQTL signals to be explained by the same genetic variant. We considered a posterior probability greater than 0.75 for pleiotropic association as evidence of colocalization within a locus.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.03.20243758; this version posted December 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Phenome-wide association analysis

We assessed pleiotropic associations of the TWAS-identified genes that showed evidence of colocalization by performing a phenome-wide lookup in the CTG-VIEW database (Cuéllar-Partida et al., 2019). The phenome-wide lookup was performed using summary-data-based Mendelian randomization (SMR) (Zhu et al., 2016) results from ~1,600 phenotypes that are available from this platform. SMR analyses were performed using eQTL data from relevant or well-powered tissues (i.e. adipose subcutaneous, adipose visceral momentum, uterus and whole blood) and included a heterogeneity test (HEIDI) to identify candidate genes that are affected by the same GWAS risk variant, comparable to our colocalization analysis. Genes passing a Bonferroni correction ($P_{SMR}<7.2\times10^{-3}$) and no evidence of heterogeneity ($P_{HEIDI} > 0.05$) were defined as having pleiotropic associations with candidate susceptibility genes.

Computational drug repurposing analysis

We used the Connectivity Map platform (Subramanian et al., 2017) to identify compounds with gene expression profiles opposing the endometrial cancer risk TWAS data. As Connectivity Map analysis requires at least ten upregulated and ten downregulated coding genes as input, we used the identified TWAS genes from S-MultiXcan at FDR<0.15 (**Supplementary Table 3**). Connectivity scores were generated based on a modified Kolmogorov-Smirnov score, which summarises the relationship of the endometrial cancer TWAS gene expression profile with the drug-induced gene expression profile across cancer cell types in the Connectivity Map database. A connectivity score ≤ -90 suggest that the expression of the query genes opposes the drug-induced gene expression profile. We also used the Open Targets platform (Carvalho-Silva et al., 2019) to determine if candidate endometrial cancer susceptibility genes encode known targets for drugs that have been clinically studied.

Acknowledgements

We thank the many women who participated in the Endometrial Cancer Association Consortium, and the numerous institutions and their staff who supported recruitment. A full list of consortium members and acknowledgements can be found in the Supplementary Note.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.03.20243758; this version posted December 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

References

Barbeira, A. N., Pividori, M., Zheng, J., Wheeler, H. E., Nicolae, D. L., & Im, H. K. (2019). Integrating predicted transcriptome from multiple tissues improves association detection. *PLoS Genet, 15*(1), e1007889. doi:10.1371/journal.pgen.1007889

Bien, S. A., Su, Y. R., Conti, D. V., Harrison, T. A., Qu, C., Guo, X., . . . Peters, U. (2019). Genetic variant predictors of gene expression provide new insight into risk of colorectal cancer. *Hum Genet, 138*(4), 307-326. doi:10.1007/s00439-019-01989-8

Block, G. A., Martin, K. J., de Francisco, A. L., Turner, S. A., Avram, M. M., Suranyi, M. G., . . . Goodman, W. G. (2004). Cinacalcet for secondary hyperparathyroidism in patients receiving hemodialysis. *N Engl J Med, 350*(15), 1516-1525. doi:10.1056/NEJMoa031633

Broun, G. O., Blessing, J. A., Eddy, G. L., & Adelson, M. D. (1993). A phase II trial of vincristine in advanced or recurrent endometrial carcinoma. A Gynecologic Oncology Group Study. *Am J Clin Oncol, 16*(1), 18-21. doi:10.1097/00000421-199302000-00005

Burke, W. M., Orr, J., Leitao, M., Salom, E., Gehrig, P., Olawaiye, A. B., . . . Shahin, F. A. (2014). Endometrial cancer: A review and current management strategies: Part II. *Gynecologic Oncology, 134*(2), 393-402. doi:https://doi.org/10.1016/j.ygyno.2014.06.003

Cao, M., Hou, D., Liang, H., Gong, F., Wang, Y., Yan, X., . . . Chen, X. (2014). miR-150 promotes the proliferation and migration of lung cancer cells by targeting SRC kinase signalling inhibitor 1. *Eur J Cancer, 50*(5), 1013-1024. doi:10.1016/j.ejca.2013.12.024

Carvalho-Silva, D., Pierleoni, A., Pignatelli, M., Ong, C., Fumis, L., Karamanis, N., . . . Dunham, I. (2019). Open Targets Platform: new developments and updates two years on. *Nucleic Acids Res, 47*(D1), D1056-D1065. doi:10.1093/nar/gky1133

Consortium, G. T., Laboratory, D. A., Coordinating Center -Analysis Working, G., Statistical Methods groups-Analysis Working, G., Enhancing, G. g., Fund, N. I. H. C., . . . Montgomery, S. B. (2017). Genetic effects on gene expression across human tissues. *Nature, 550*(7675), 204-213. doi:10.1038/nature24277

Cuéllar-Partida, G., Lundberg, M., Kho, P. F., D'Urso, S., Gutiérrez-Mondragón, L. F., Ngo, T. T., & Hwang, L.-D. (2019). Complex-Traits Genetics Virtual Lab: A community-driven web platform for post-GWAS analyses. *bioRxiv*, 518027. doi:10.1101/518027

Cui, C., Merritt, R., Fu, L., & Pan, Z. (2017). Targeting calcium signaling in cancer therapy. *Acta Pharm Sin B, 7*(1), 3-17. doi:10.1016/j.apsb.2016.11.001

Durlacher, C. T., Chow, K., Chen, X. W., He, Z. X., Zhang, X., Yang, T., & Zhou, S. F. (2015). Targeting Na(+)/K(+) -translocating adenosine triphosphatase in cancer treatment. *Clin Exp Pharmacol Physiol, 42*(5), 427-443. doi:10.1111/1440-1681.12385

Gamazon, E. R., Wheeler, H. E., Shah, K. P., Mozaffari, S. V., Aquino-Michaels, K., Carroll, R. J., . . . Im, H. K. (2015). A gene-based association method for mapping traits using reference transcriptome data. *Nat Genet, 47*(9), 1091-1098. doi:10.1038/ng.3367

Giambartolomei, C., Vukcevic, D., Schadt, E. E., Franke, L., Hingorani, A. D., Wallace, C., & Plagnol, V. (2014). Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet, 10*(5), e1004383. doi:10.1371/journal.pgen.1004383

Grundker, C., Gunthert, A. R., & Emons, G. (2008). Hormonal heterogeneity of endometrial cancer. *Adv Exp Med Biol, 630*, 166-188. doi:10.1007/978-0-387-78818-0_11

medRxiv preprint doi: https://doi.org/10.1101/2020.12.03.20243758; this version posted December 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Gusev, A., Lawrenson, K., Lin, X., Lyra, P. C., Jr., Kar, S., Vavra, K. C., . . . Gayther, S. A. (2019). A transcriptome-wide association study of high-grade serous epithelial ovarian cancer identifies new susceptibility genes and splice variants. *Nat Genet, 51*(5), 815-823. doi:10.1038/s41588-019-0395-x

Hu, Y., Li, M., Lu, Q., Weng, H., Wang, J., Zekavat, S. M., . . . Zhao, H. (2019). A statistical framework for cross-tissue transcriptome-wide association analysis. *Nat Genet, 51*(3), 568-576. doi:10.1038/s41588-019-0345-7

Ignatov, A., & Ortmann, O. (2020). Endocrine Risk Factors of Endometrial Cancer: Polycystic Ovary Syndrome, Oral Contraceptives, Infertility, Tamoxifen. *Cancers (Basel), 12*(7). doi:10.3390/cancers12071766

Kho, P. F., Amant, F., Annibali, D., Ashton, K., Attia, J., Auer, P. L., . . . O'Mara, T. A. (2020). Mendelian randomization analyses suggest a role for cholesterol in the development of endometrial cancer. *Int J Cancer*. doi:10.1002/ijc.33206

King, E. A., Davis, J. W., & Degner, J. F. (2019). Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval. *PLoS Genet, 15*(12), e1008489. doi:10.1371/journal.pgen.1008489

Lamontagne, M., Berube, J. C., Obeidat, M., Cho, M. H., Hobbs, B. D., Sakornsakolpat, P., . . . Bosse, Y. (2018). Leveraging lung tissue transcriptome to uncover candidate causal genes in COPD genetic associations. *Hum Mol Genet, 27*(10), 1819-1829. doi:10.1093/hmg/ddy091

Liang, Q., Guan, G., Li, X., Wei, C., Wu, J., Cheng, P., . . . Cheng, W. (2020). Profiling pro-neural to mesenchymal transition identifies a lncRNA signature in glioma. *J Transl Med, 18*(1), 378. doi:10.1186/s12967-020-02552-0

Lindemann, K., Malander, S., Christensen, R. D., Mirza, M. R., Kristensen, G. B., Aavall-Lundqvist, E., . . . Nordstrom, B. (2014). Examestane in advanced or recurrent endometrial carcinoma: a prospective phase II study by the Nordic Society of Gynecologic Oncology (NSGO). *BMC Cancer, 14*, 68. doi:10.1186/1471-2407-14-68

Ma, B. B., Oza, A., Eisenhauer, E., Stanimir, G., Carey, M., Chapman, W., . . . Fyles, A. (2004). The activity of letrozole in patients with advanced or recurrent endometrial cancer and correlation with biological markers--a study of the National Cancer Institute of Canada Clinical Trials Group. *Int J Gynecol Cancer, 14*(4), 650-658. doi:10.1111/j.1048-891X.2004.14419.x

Mahalingam, D., Peguero, J., Cen, P., Arora, S. P., Sarantopoulos, J., Rowe, J., . . . Campos, L. (2019). A Phase II, Multicenter, Single-Arm Study of Mipsagargin (G-202) as a Second-Line Therapy Following Sorafenib for Adult Patients with Progressive Advanced Hepatocellular Carcinoma. *Cancers (Basel), 11*(6). doi:10.3390/cancers11060833

Mahalingam, D., Wilding, G., Denmeade, S., Sarantopoulas, J., Cosgrove, D., Cetnar, J., . . . Carducci, M. (2016). Mipsagargin, a novel thapsigargin-based PSMA-activated prodrug: results of a first-in-man phase I clinical trial in patients with refractory, advanced or metastatic solid tumours. *Br J Cancer, 114*(9), 986-994. doi:10.1038/bjc.2016.72

Moradan, S., Nikkhah, N., & Mirmohammadkhanai, M. (2017). Comparing the Administration of Letrozole and Megestrol Acetate in the Treatment of Women with Simple Endometrial Hyperplasia without Atypia: A Randomized Clinical Trial. *Adv Ther, 34*(5), 1211-1220. doi:10.1007/s12325-017-0509-8

medRxiv preprint doi: https://doi.org/10.1101/2020.12.03.20243758; this version posted December 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Morice, P., Leary, A., Creutzberg, C., Abu-Rustum, N., & Darai, E. (2016). Endometrial cancer. *Lancet, 387*(10023), 1094-1108. doi:10.1016/S0140-6736(15)00130-0

Mortlock, S., Kendarsari, R. I., Fung, J. N., Gibson, G., Yang, F., Restuadi, R., . . . Montgomery, G. W. (2020). Tissue specific regulation of transcription in endometrium and association with disease. *Hum Reprod, 35*(2), 377-393. doi:10.1093/humrep/dez279

Nelson, M. R., Tipney, H., Painter, J. L., Shen, J., Nicoletti, P., Shen, Y., . . . Sanseau, P. (2015). The support of human genetic evidence for approved drug indications. *Nat Genet, 47*(8), 856-860. doi:10.1038/ng.3314

O'Mara, T. A., Glubb, D. M., Amant, F., Annibali, D., Ashton, K., Attia, J., . . . Thompson, D. J. (2018). Identification of nine new susceptibility loci for endometrial cancer. *Nat Commun, 9*(1), 3166. doi:10.1038/s41467-018-05427-7

O'Mara, T. A., Glubb, D. M., Kho, P. F., Thompson, D. J., & Spurdle, A. B. (2019). Genome-Wide Association Studies of Endometrial Cancer: Latest Developments and Future Directions. *Cancer Epidemiol Biomarkers Prev, 28*(7), 1095-1102. doi:10.1158/1055-9965.EPI-18-1031

O'Mara, T. A., Spurdle, A. B., Glubb, D. M., & Endometrial Cancer Association, C. (2019). Analysis of Promoter-Associated Chromatin Interactions Reveals Biologically Relevant Candidate Target Genes at Endometrial Cancer Risk Loci. *Cancers (Basel), 11*(10). doi:10.3390/cancers11101440

Rodríguez-Hernández, C. J., Mateo-Lozano, S., García, M., Casalà, C., Briansó, F., Castrejón, N., . . . de Torres, C. (2016). Cinacalcet inhibits neuroblastoma tumor growth and upregulates cancer-testis antigens. *Oncotarget, 7*(13).

Rose, P. G., Brunetto, V. L., VanLe, L., Bell, J., Walker, J. L., & Lee, R. B. (2000). A phase II trial of anastrozole in advanced recurrent or persistent endometrial carcinoma: a Gynecologic Oncology Group study. *Gynecol Oncol, 78*(2), 212-216. doi:10.1006/gyno.2000.5865

Shoback, D. M., Bilezikian, J. P., Turner, S. A., McCary, L. C., Guo, M. D., & Peacock, M. (2003). The calcimimetic cinacalcet normalizes serum calcium in subjects with primary hyperparathyroidism. *J Clin Endocrinol Metab, 88*(12), 5644-5649. doi:10.1210/jc.2002-021597

Subramanian, A., Narayan, R., Corsello, S. M., Peck, D. D., Natoli, T. E., Lu, X., . . . Golub, T. R. (2017). A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell, 171*(6), 1437-1452.e1417. doi:10.1016/j.cell.2017.10.049

Tabatabaie, A., Karimi Zarchi, M., Dehghani-Tafti, M., Miratashi-Yazdi, A., Teimoori, S., & Dehghani, A. (2013). Comparing letrozole with medroxyprogesterone acetate (MPA) as hormonal therapy for simple endometrial hyperplasia without atypia in adult and middle-aged women. *Eur J Gynaecol Oncol, 34*(6), 552-555.

Takai, N., Ueda, T., Nishida, M., Nasu, K., & Narahara, H. (2008). Bufalin induces growth inhibition, cell cycle arrest and apoptosis in human endometrial and ovarian cancer cells. *Int J Mol Med, 21*(5), 637-643. doi:10.3892/ijmm.21.5.637

Thompson, D. J., O'Mara, T. A., Glubb, D. M., Painter, J. N., Cheng, T., Folkerd, E., . . . Spurdle, A. B. (2016). CYP19A1 fine-mapping and Mendelian randomization: estradiol is causal for endometrial cancer. *Endocr Relat Cancer, 23*(2), 77-91. doi:10.1530/ERC-15-0386

Wainberg, M., Sinnott-Armstrong, N., Mancuso, N., Barbeira, A. N., Knowles, D. A., Golan, D., . . . Kundaje, A. (2019). Opportunities and challenges for transcriptome-wide association studies. *Nat Genet, 51*(4), 592-599. doi:10.1038/s41588-019-0385-z

medRxiv preprint doi: https://doi.org/10.1101/2020.12.03.20243758; this version posted December 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Wang, P., Wang, H., Li, X., Liu, Y., Zhao, C., & Zhu, D. (2016). SRCIN1 Suppressed Osteosarcoma Cell Proliferation and Invasion. *PLoS One, 11*(8), e0155518. doi:10.1371/journal.pone.0155518

Wu, L., Shi, W., Long, J., Guo, X., Michailidou, K., Beesley, J., . . . Zheng, W. (2018). A transcriptome-wide association study of 229,000 women identifies new candidate susceptibility genes for breast cancer. *Nat Genet, 50*(7), 968-978. doi:10.1038/s41588-018-0132-x

Xu, X., Wang, W., Su, N., Zhu, X., Yao, J., Gao, W., . . . Sun, Y. (2015). miR-374a promotes cell proliferation, migration and invasion by targeting SRCIN1 in gastric cancer. *FEBS Lett, 589*(3), 407-413. doi:10.1016/j.febslet.2014.12.027

Yang, F., Luo, L. J., Zhang, L., Wang, D. D., Yang, S. J., Ding, L., . . . Tang, J. H. (2017). MiR-346 promotes the biological function of breast cancer cells by targeting SRCIN1 and reduces chemosensitivity to docetaxel. *Gene, 600*, 21-28. doi:10.1016/j.gene.2016.11.037

Yang, W., Soares, J., Greninger, P., Edelman, E. J., Lightfoot, H., Forbes, S., . . . Garnett, M. J. (2013). Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Research, 41*(D1), D955-D961. doi:10.1093/nar/gks1111

Zhu, Z., Zhang, F., Hu, H., Bakshi, A., Robinson, M. R., Powell, J. E., . . . Yang, J. (2016). Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet, 48*(5), 481-487. doi:10.1038/ng.3538

medRxiv preprint doi: https://doi.org/10.1101/2020.12.03.20243758; this version posted December 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Figures

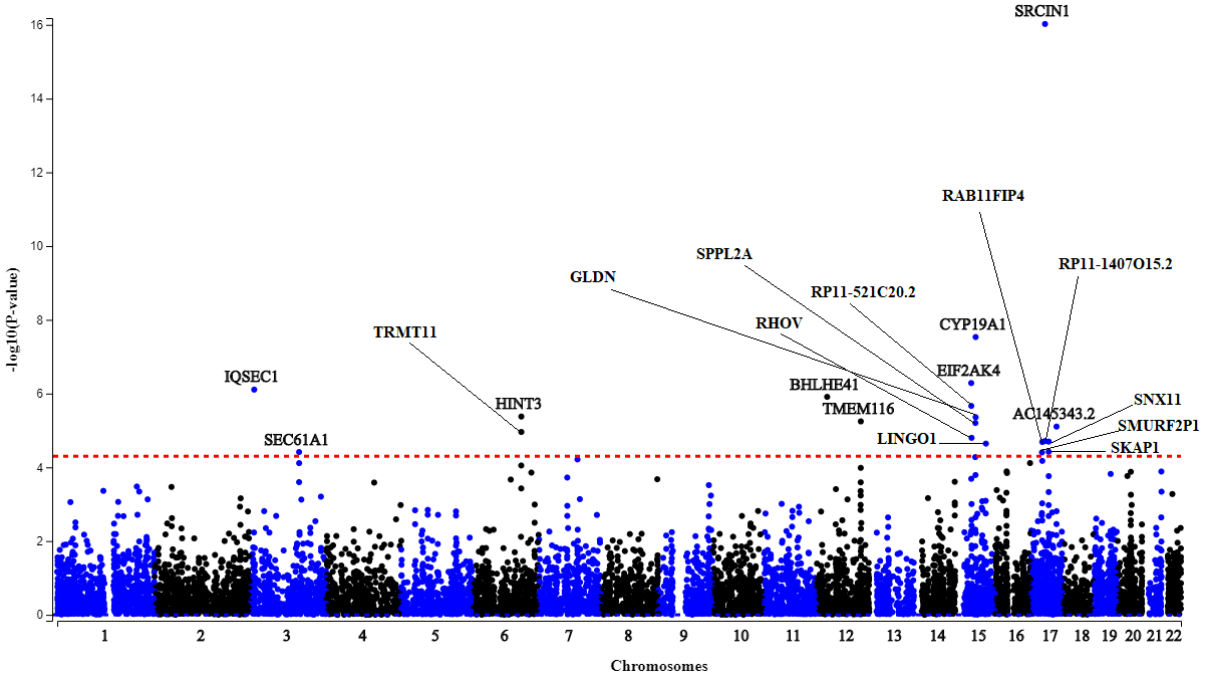

Figure 1. Manhattan plot of TWAS analysis of endometrial cancer risk. Genes are plotted according to chromosome and position against $-\log_{10}$(P value) of the TWAS association. Red dotted lines represent the Benjamini-Hochberg significance threshold (FDR<0.05), which is equivalent to $P<5 \times 10^{-5}$.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.03.20243758; this version posted December 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

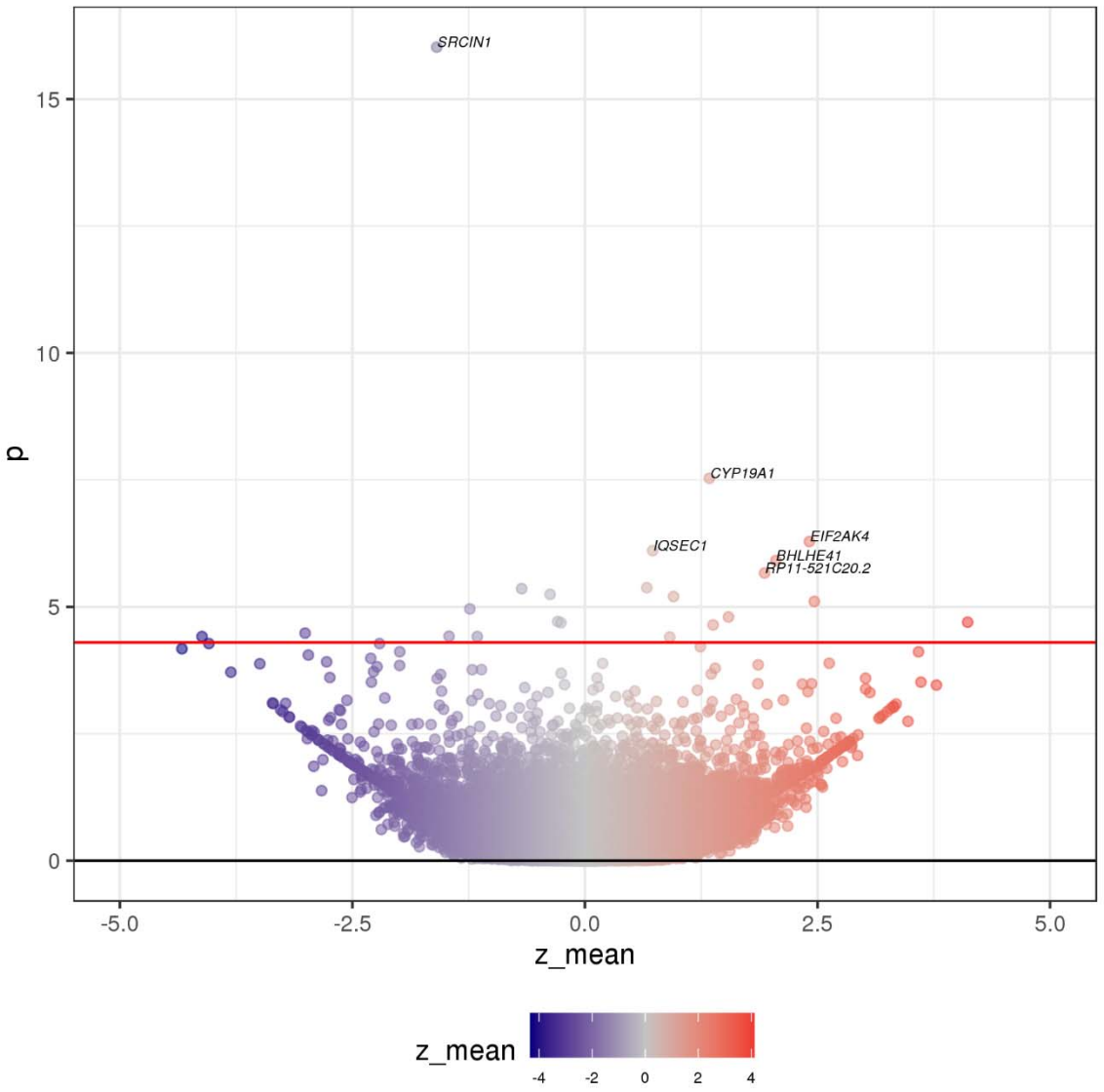

Figure 2. Volcano plot from TWAS analysis of endometrial cancer risk. Each dot represents the effect of the average predicted expression of each gene on endometrial cancer risk. Red line represents the Benjamini-Hochberg significance threshold (FDR< 0.05). Genes with FDR<0.01 are labelled.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.03.20243758; this version posted December 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

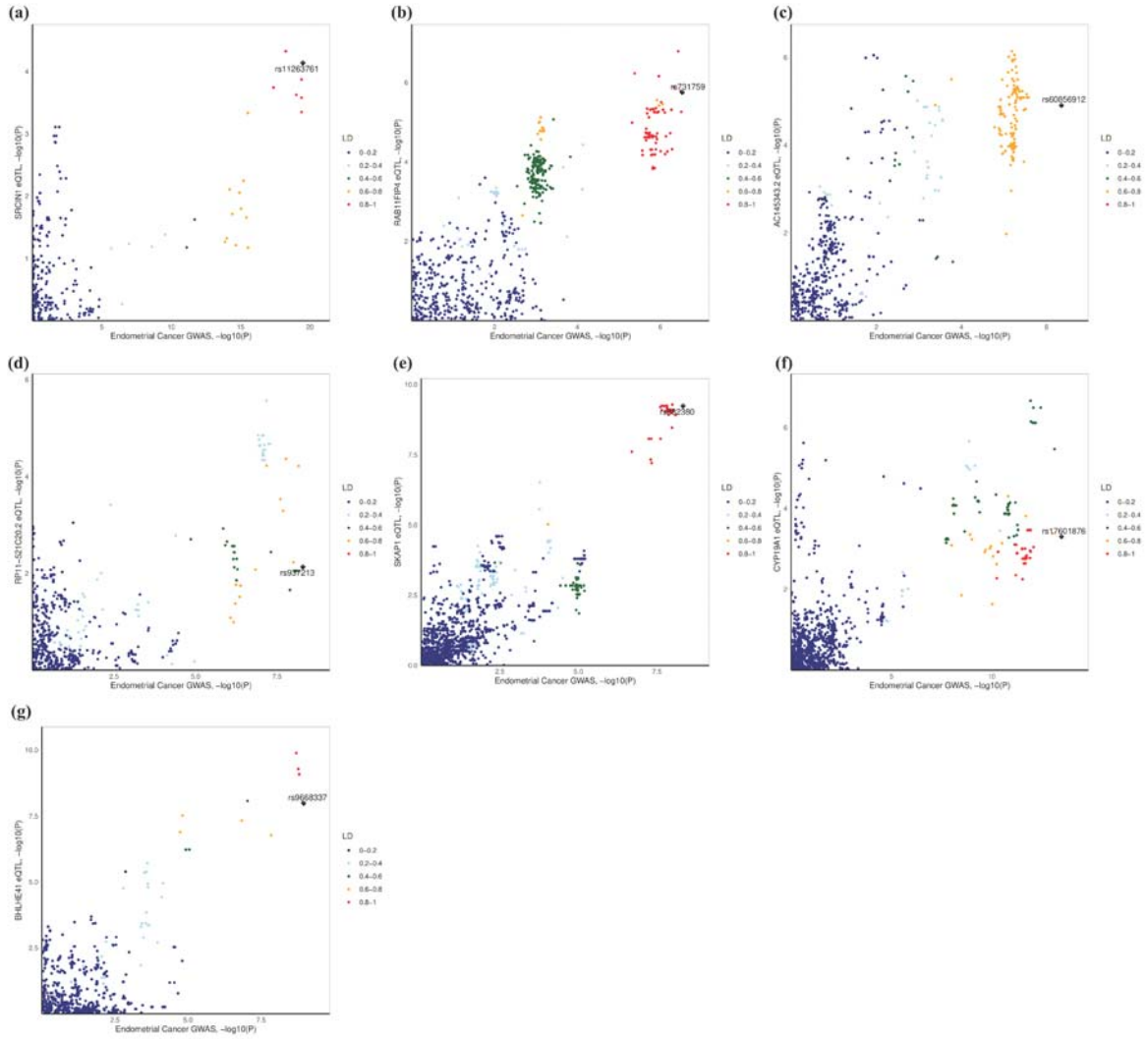

Figure 3. Colocalization of association signals between endometrial cancer GWAS and eQTL data from the corresponding best performing tissue in the TWAS analysis. -log10(P) of endometrial cancer GWAS was plotted against -log10(P) of eQTLs for (a) *SRCIN1*, (b) *RAB11FIP4*, (c) *AC145343.2*, (d) *RP11-521C20.2*, (e) *SKAP1*, (f) *CYP19A1* and (g) *BHLHE41*. The lead variant of each locus identified in endometrial cancer GWAS is shown as a purple diamond. Other variants are coloured according to their degree of LD with lead variant in the locus.

Tables

Table 1. Genes associated with endometrial cancer risk as identified by TWAS (FDR<0.05)

Region	Gene	Best performing tissue	N_{tissue}	P	FDR	Z_{min}	Z_{max}	Z_{mean}	Z_{SD}	PP_{COLOC}
3p25.1	*IQSEC1*[#]	Esophagus Mucosa	11	7.81E-07	4.97E-03	-1.92	4.77	0.73	1.92	0.701
3q21.3	*SEC61A1*	Cells Transformed fibroblasts	5	3.85E-05	0.046	-4.09	1.40	-1.16	2.17	0.398
6q22.32	*HINT3*	Esophagus Muscularis	3	4.20E-06	0.014	-4.21	3.36	0.66	4.23	0.026
6q22.32	*TRMT11*	Stomach	5	1.10E-05	0.023	-3.24	5.04	-1.24	3.53	0.168
12p12.1	***BHLHE41***[#]	Thyroid	8	1.23E-06	6.25E-03	-1.09	5.82	2.05	2.28	0.993
12q24.13	*TMEM116*	Lung	24	5.68E-06	0.016	-4.19	1.18	-0.38	1.27	0.001
15q15.1	*EIF2AK4*[#]	Colon Transverse	31	5.19E-07	4.41E-03	-2.56	5.44	2.41	1.97	0.500
15q15.1	***RP11-521C20.2***	Heart Left Ventricle	6	2.16E-06	9.18E-03	-1.30	5.39	1.93	2.74	0.882
15q15.1	*RHOV*	Brain Cerebellum	6	1.59E-05	0.031	-1.91	3.07	1.54	2.25	0.000
15q21.2	***CYP19A1***[#]	Skin Sun Exposed Lower leg	8	2.93E-08	3.73E-04	-2.31	5.76	1.34	2.89	0.942
15q21.2	*GLDN*	Adipose Subcutaneous	21	4.38E-06	0.014	-2.40	3.75	-0.68	1.43	0.001
15q21.2	*SPPL2A*	Stomach	30	6.27E-06	0.016	-3.05	2.75	0.95	1.11	0.000
15q24.3	*LINGO1*	Esophagus Muscularis	16	2.27E-05	0.034	-1.61	4.78	1.38	1.62	0.552
17q11.2	***RAB11FIP4***	Esophagus Mucosa	12	2.05E-05	0.033	-2.72	3.51	-0.26	2.40	0.957
17q11.2	*SMURF2P1*	Cells Transformed fibroblasts	4	3.93E-05	0.046	-2.58	2.58	0.91	2.35	0.004
17q12	***SRCIN1***[#]	Esophagus Gastroesophageal Junction	6	9.49E-17	2.42E-12	-8.96	2.00	-1.59	4.10	0.875
17q12	*RP11-1407O15.2*	Brain Anterior cingulate cortex BA24	10	1.95E-05	0.033	-5.84	1.95	-0.29	2.06	0.235
17q21.32	*SNX11*	Adipose Subcutaneous	26	2.01E-05	0.033	1.27	5.59	4.11	0.98	0.026
17q21.32	***SKAP1***	Liver	8	3.79E-05	0.046	-5.68	0.98	-1.46	1.91	0.940
17q24.2	***AC145343.2***	Esophagus Muscularis	11	7.85E-06	0.018	-0.61	4.55	2.47	1.91	0.800

medRxiv preprint doi: https://doi.org/10.1101/2020.12.03.20243758; this version posted December 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

N_{tissue}: number of tissues available for this gene; FDR: false discovery rate; Z_{min}: minimum Z score from single tissue S-PrediXcan result; Z_{max}: maximum Z score from single tissue S-PrediXcan result; Z_{mean}: mean Z score from single tissue S-PrediXcan result; Z_{SD}: standard deviation of Z score from single tissue S-PrediXcan result; PP_{COLOC}: posterior probability that endometrial cancer risk and eQTL variants from best performing tissue colocalize. [#]TWAS association passed Bonferroni correction threshold ($P<2 \times 10^{-6}$). Genes in bold showed evidence of colocalization ($PP_{COLOC}>0.75$).

medRxiv preprint doi: https://doi.org/10.1101/2020.12.03.20243758; this version posted December 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

Table 2. Connectivity Map compounds with gene expression profiles opposing endometrial cancer risk TWAS associations.

Connectivity score*	Drug	Drug class	Clinically tested (highest phase) or FDA approved	Indication
-95.60	nocodazole	Tubulin inhibitor	No	-
-95.34	vincristine	Tubulin inhibitor	Approved	Acute lymphocytic leukemia, acute myeloid leukemia, Hodgkin and non-Hodgkin lymphoma
-94.75	strophanthidin	Na+/ K+ -ATPase inhibitor	No	-
-94.45	cinacalcet	Calcium channel activator	Approved	Hypercalcemia and secondary hyperparathyroidism
-94.40	digitoxigenin	Na+/ K+- ATPase inhibitor	No	-
-93.97	periplocymarin	Na+/ K+ -ATPase inhibitor	No	-
-93.34	proscillaridin	Na+/ K+ -ATPase inhibitor	No	-
-93.06	thapsigargin	Ca^{2+}- ATPase inhibitor	No	-
-92.92	cinobufagin	Na+/ K+- ATPase inhibitor	Yes (Phase IV)	Liver cancer (NCT03843229) and gastrointestinal neoplasm (NCT02860429)
-92.46	digitoxin	Na+/ K+- ATPase inhibitor	Yes (Phase II)	Cystic fibrosis (NCT00782288) and sarcoma (NCT00017446)
-92.35	digoxin	Na+/ K+- ATPase inhibitor	Approved	Heart failure and atrial fibrillation
-92.33	ouabain	Na+/ K+- ATPase inhibitor	No	-
-92.21	bufalin	Na+/ K+- ATPase inhibitor	Yes (Phase II)	Pancreatic cancer (NCT00837239)
-90.84	BNTX	Opioid receptor antagonist	No	-

*Connectivity scores ≤ -90 provides compounds with opposing gene expression profiles.
FDA: Federal Drug Administration (USA); NCT numbers were obtained from ClinicalTrials.gov.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.03.20243758; this version posted December 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

Table 3. Drug targets of TWAS-identified genes as identified by Open Targets Database

Drug	Target protein	Indication	Status
Letrozole	CYP19A1	Breast cancer	Approved
Anastrozole	CYP19A1	Breast cancer	Approved
Exemestane	CYP19A1	Breast cancer	Approved
Aminoglutethimide	CYP19A1	Breast cancer	Approved
Aminoglutethimide	CYP19A1	Cushing syndrome	Approved

medRxiv preprint doi: https://doi.org/10.1101/2020.12.03.20243758; this version posted December 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.