Endometrial vezatin and its association with endometriosis risk

Sarah J. Holdsworth-Carson†, Jenny N. Fung‡, Hien T.T. Luong‡, Yadav Sapkota‡, Lisa M. Bowdler‡, Leanne Wallace‡, Wan Tinn Teh§, Joseph E. Powell#, Jane E. Girling†, Martin Healey§, Grant W. Montgomery‡, and Peter A.W. Rogers†

1Department of Obstetrics and Gynaecology, University of Melbourne, Gynaecology Research Centre, Royal Women's Hospital, Level 7, 20 Flemington Road, Parkville, Victoria 3052, Australia 2Molecular Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, Queensland 4029, Australia 3The Institute for Molecular Bioscience, University of Queensland, Brisbane, Queensland 4072, Australia 4Centre for Neurogenetics and Statistical Genomics, Queensland Brain Institute, University of Queensland, St Lucia, Queensland 4072, Australia

*Correspondence address. E-mail: parogers@unimelb.edu.au

Submitted on January 12, 2016; resubmitted on February 11, 2016; accepted on February 23, 2016

STUDY QUESTION: Do endometriosis risk-associated single nucleotide polymorphisms (SNPs) found at the 12q22 locus have effects on vezatin (VEZT) expression?

SUMMARY ANSWER: The original genome-wide association study (GWAS) SNP (rs10859871), and other newly identified association signals, demonstrate strong evidence for cis-expression quantitative trait loci (eQTL) effects on VEZT expression.

WHAT IS KNOWN ALREADY: GWAS have identified several disease-risk loci (SNPs) associated with endometriosis. The SNP rs10859871 is located within the VEZT gene. VEZT expression is altered in the endometrium of endometriosis patients and is an excellent candidate for having a causal role in endometriosis. Most of the SNPs identified from GWAS are not located within the coding region of the genome. However, they are likely to have an effect on the regulation of gene expression. Genetic variants that affect levels of gene expression are called expression quantitative trait loci (eQTL).

STUDY DESIGN, SIZE, DURATION: Samples for genotyping and VEZT variant screening were drawn from women recruited for genetic studies in Australia/New Zealand and women undergoing surgery in a tertiary care centre. Coding variants for VEZT were screened in blood from 100 unrelated individuals (endometriosis-dense families) from the QIMR Berghofer Medical Research Institute dataset. SNPs at the 12q22 locus were imputed and reanalysed for their association with endometriosis. Reanalysis of endometriosis risk-association was performed on a final combined Australian dataset of 2594 cases and 4496 controls. Gene expression was performed on 136 endometrial samples. eQTL analysis in whole blood was performed on 862 individuals from the Brisbane Systems Genetics Study. Endometrial tissue-specific eQTL analysis was performed on 122 samples (eutopic endometrium) collected following laparoscopic surgery. VEZT protein expression studies employed n = 56 (western blotting) and n = 42 (immunohistochemistry) endometrial samples.

PARTICIPANTS/MATERIALS, SETTING, METHODS: The women recruited for this study provided blood and/or endometrial tissue samples in a hospital setting. Genomic DNA was screened for common and coding variants. SNPs of interest in the 12q22 region were genotyped using Agena MassARRAY technology or Taqman SNP genotyping assay. Gene expression profiles from RNA extracted from blood and endometrial tissue samples were generated using Illumina whole-genome expression chips (Human HT-12 v4.0). Whole protein extracted from endometrium was used for VEZT western blots, and paraffin sections of endometrium were employed for VEZT immunohistochemistry semi-quantitative analysis.

MAIN RESULTS AND THE ROLE OF CHANCE: A total of 11 coding variants of VEZT (including one novel variant) were identified from an endometriosis-dense cohort. Polymorphic coding and imputed SNPs were combined with previous GWAS data to reanalyse the endometriosis risk association of the 12q22 region. The disease association signal at 12q22 was due to coding variants in VEZT or FGD6 (FYVE, RhoGEF and PH domain-containing 6) and SNPs with the strongest signals were either intronic or intergenic. We found strong evidence for VEZT cis-eQTLs with the sentinel SNP (rs10859871) in blood and endometrium, where the endometriosis risk allele (C) was associated with an increase in VEZT

† These authors contributed equally to this work.
‡ These authors contributed equally to this work.
§ The Author 2016. Published by Oxford University Press on behalf of the European Society of Human Reproduction and Embryology. All rights reserved. For Permissions, please email: journals.permissions@oup.com
expression. We could not demonstrate this genotype-specific effect on VEZT protein expression in endometrium. However, we did observe a menstrual cycle stage specific increase in VEZT protein expression in endometrial glands, specific to the secretory phase ($P = 2.0 \times 10^{-5}$).

**LIMITATIONS, REASONS FOR CAUTION:** In comparison to the blood sample datasets, the study numbers of endometrial tissues were substantially reduced. Protein studies failed to complement RNA results, also likely a reflection of the low study numbers in these experiments. In silico prediction tools used in this investigation are typically based on cell lines different to our tissues of interest, thus any functional annotations drawn from these approaches should be considered carefully. Therefore, functional studies on VEZT and related pathway components are still warranted to unequivocally implicate a causal role for VEZT in endometriosis pathophysiology.

**WIDER IMPLICATIONS OF THE FINDINGS:** GWAS have proven to be very valuable tools for deciphering complex diseases. Endometriosis is a text-book example of a complex disease, involving genetic, lifestyle and environmental influences. Our focused investigation of the 12q22 region validates an association with increased endometriosis risk. Endometriosis risk SNPs (including rs10859871) located within this locus demonstrated evidence for cis-eQTLs on VEZT expression. By examining women who possess an enhanced genetic risk of developing endometriosis, we have identified an effect on VEZT expression and therefore a potential gene/gene pathway in endometriosis disease establishment and development.

**STUDY FUNDING/COMPETING INTEREST(S):** Funding for this work was provided by NHMRC Project Grants GNT1012245, GNT1026033, GNT1049472 and GNT1046880. G.W.M. is supported by the NHMRC Fellowship scheme (GNT1078399). S.J.H.-C. is supported by the J.N. Peters Bequest Fellowship. The authors declare no competing interests.

**TRIAL REGISTRATION NUMBER:** N/A.

**Key words:** endometriosis / vezatin / genetics / expression quantitative trait loci / GWAS

### Introduction

Endometriosis is a heterogenic gynaecological disorder, with a wide range of symptoms and often differing impact on a women’s quality of life. It is the most common cause of chronic pelvic pain (Hickey et al., 2014); however, except for cases of deep infiltrating endometriosis, there is little correlation between endometriosis disease severity and degree of pain symptoms (Gruppo Italiano per lo Studio dell’Endometriosi, 2001; Fauconnier and Chapron, 2005; Vercellini et al., 2007). Endometriosis affects up to 10% of reproductive-aged women (Burney and Giudice, 2012), but with varying symptoms, delayed diagnosis and misdiagnosis, the true prevalence of endometriosis is unknown. Endometriosis is a complex disease with genetic, environmental and lifestyle factors all contributing to the disorder. The heritability of endometriosis from twin studies has been estimated at 0.51 (Treloar et al., 1999) and 0.47 (Saha et al., 2015).

Genome-wide association studies (GWASs) have revealed the genetic basis of many complex traits (including Type 1 and 2 diabetes, obesity and systemic lupus erythematosus). Specific to endometriosis, GWASs and meta-analyses have led to the identification of disease-risk loci that increase a woman’s risk of developing the disease (Uno et al., 2010; Painter et al., 2011; Nyholt et al., 2012; Albertsen et al., 2013; Rahmioglu et al., 2014; Pagliardini et al., 2015; Sapkota et al., 2015a). Furthermore, the effect size of these loci tends to increase as the proportion of cases analysed are limited to more severe disease stages (Stage III/IV endometriosis), indicating that moderate to severe endometriosis cases have a greater genetic burden relative to minimal or mild disease (Rahmioglu et al., 2014; Sapkota et al., 2015a). The single nucleotide polymorphism (SNP) rs10859871 at locus 12q22 is highly ranked and has one of the strongest associations among the SNPs identified by GWASs (Nyholt et al., 2012; Rahmioglu et al., 2014; Pagliardini et al., 2015). A recent large meta-analysis confirms a genome-wide significant association for SNP rs10859871 (odds ratio (OR) = 1.19, $P = 7.9 \times 10^{-20}$) (Pagliardini et al., 2015). Comprehensive analysis of the 12q22 region showed a number of SNPs in linkage disequilibrium (LD) with rs10859871, and the strongest signals were located within the 3′ untranslated region (UTR) of the vezatin (VEZT) gene and extend across the adjacent gene FYVE, RhoGEF and PH domain-containing 6 (FGD6) (Nyholt et al., 2012). To begin to better understand a potential role for this risk variant association in endometriosis pathophysiology, we have undertaken a focused investigation of the rs10859871 region in human endometrial tissues.

The function of individual trait-associated SNPs or risk variants in disease largely remains unknown. However, it is widely accepted that risk loci contain variants which can modulate the expression levels of other genes (Westra and Franke, 2014). This is an important mechanism given that many risk-associated SNPs do not alter final protein sequence or structure and are often located in non-coding regions of the genome. To see which genes are regulated by genetic variants, a correlation between the genetic variant(s) and a molecular quantitative trait (for example, gene expression) can be undertaken to identify possible expression quantitative trait loci (eQTLs) (Westra and Franke, 2014). eQTLs can have nearby effects (cis), where the genetic variant is found close to the affected gene (within 1 mb) or distant effects (trans), where the variant is found far from the influenced gene (>5 mb away). Trait-associated SNPs are enriched for cis-eQTLs and furthermore, cis-eQTLs are highly tissue-specific (Westra and Franke, 2014; Xu et al., 2014). Therefore, our study utilized eutopic endometrium (disease-specific tissue) to undertake a cis-eQTL analysis of the gene closest to rs10859871.

The SNP rs10859871 is located 17 kb upstream of the VEZT gene (Rahmioglu et al., 2014). VEZT is a transmembrane protein, with a short extracellular domain and long intracellular domain, which anchors to myosin VIIA as part of the adherens junctional complex in epithelial cells (Kussel-Andermann et al., 2000; Blaschuk and Rowlands, 2002). VEZT is also essential for implantation; embryos from VEZT-null mice fail to develop past the blastocyst stage due to a loss of cell–cell adhesion (Hyenne et al., 2007). Given the known physiological roles of
Materials and Methods

Patient recruitment and tissue collection

Samples for genotyping and screening the coding regions of VEZT were drawn from women recruited for genetic studies in Australia and New Zealand (Treloar et al., 2002; Zhao et al., 2006; Painter et al., 2011) and datasets are described below. Samples for gene expression studies in whole blood were drawn from adolescent twins and their families from the Brisbane Systems Genetics Study (BSGS) (Powell et al., 2012) and the data and analyses are described below. Blood, endometrial tissue (curette) and hysterectomy samples were obtained from women recruited through the Royal Women’s Hospital (RWH), Melbourne, Australia following written informed consent. All studies were approved by the Human Research Ethics Committees of the QIMR Berghofer Medical Research Institute, Brisbane, Australia and by the Royal Women’s Hospital, Melbourne, Australia for the endometrial tissue studies.

To examine protein expression and localization of VEZT across the different stages of the menstrual cycle, full thickness uterine tissue (endometrium to myometrium) were obtained from premenopausal women, with no history of endometriosis, who were free from hormone treatment for 3 months prior to surgery, undergoing hysterectomy for fibroids (n = 6 proliferative phase and n = 6 secretory phase). Endometrial cycle stage was determined following histological assessment at pathology. Full thickness tissues were immediately fixed in formalin for immunohistochemistry (IHC).

Women with and without endometriosis undergoing laparoscopy (n = 136) were recruited for inclusion in gene expression analysis. Some of the clinical details of these women are listed in Table I. Out of the 136 samples with gene expression data, 122 samples with sufficient quantity of DNA were genotyped and used for eQTL analyses. Whole blood, eutopic endometrium, detailed patient questionnaires, past and present clinical histories, pathology findings and surgical notes were collected and recorded for each participant. In theatre, endometrium (curette) was collected immediately in RNAlater (Life Technologies, NY, USA) and stored for a minimum of overnight at 4°C before being stored at −80°C for mRNA/protein extraction. Paraffin blocks of curettes were recalled from pathology following completion of diagnostic studies for IHC. Subsets of women who were premenopausal and free from hormone treatment were selected for eQTL and IHC studies.

Table 1 Clinical details of women in study of endometrial vezatin (VEZT) and its association with endometriosis risk.

<table>
<thead>
<tr>
<th>Cycle stage</th>
<th>Age (years) (mean ± SEM)</th>
<th>Menstrual</th>
<th>Proliferative</th>
<th>Secretory</th>
<th>Inactive/unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>31.15 ± 0.68</td>
<td>2.94% (4/136)</td>
<td>54.41% (74/136)</td>
<td>40.44% (55/136)</td>
<td>2.21% (3/136)</td>
</tr>
<tr>
<td>Disease status</td>
<td>No endometriosis (control)</td>
<td>31.62% (43/136)</td>
<td>68.38% (93/136)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disease stage</td>
<td>Endometriosis Stage I</td>
<td>46.24% (45/93)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Endometriosis Stage II</td>
<td>12.90% (12/93)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Endometriosis Stage III</td>
<td>15.05% (14/93)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Endometriosis Stage IV</td>
<td>11.83% (11/93)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deep infiltrating</td>
<td>Unknown stage</td>
<td>13.89% (13/93)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ovarian</td>
<td>28.75% (23/80)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Peritoneum</td>
<td>33.75% (27/80)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

VEZT, the potential for a functional role in endometriosis is plausible. In women with endometriosis, VEZT has been shown to be up-regulated in ectopic endometrium relative to eutopic endometrium (Meola et al., 2010). The aim of this study was to investigate eQTLs for VEZT in endometrium at locus 12q22. We hypothesized that the endometriosis risk-associated SNPs will also affect VEZT gene expression. By examining women who possess an enhanced genetic risk of developing endometriosis, we hope to significantly improve the current understanding of the mechanisms that cause endometriosis.

High-resolution melt (HRM) assay and sequencing

The coding regions of VEZT were screened for variants in 100 unrelated individuals chosen from 100 case–disease families drawn from the QIMR dataset (Treloar et al., 2002). They include 15 families with at least four affected sisters, 71 families with three affected sisters and 14 families with at least two sisters and two other relatives diagnosed with endometriosis. Cases for genotyping were one sample per family, chosen as the case with the most severe stage of disease in each family.

Genomic DNA was extracted from peripheral venous blood samples (Miller et al., 1988). Variants were screened using HRM on a Rotor-Gene 6000 Real-timeRotor Analyser (Corbett Research, QIAGEN, Hilden, Germany). Melting curves were visualized and analysed using the Rotor-Gene 6000 analysis software v 1.7 (Corbett Research, QIAGEN). Samples showing patterns different from wild-type were sequenced to confirm the variant using BigDye 3.0 terminator chemistry (Thermo Fisher Scientific (Applied Biosystems), Scoresby, VIC, Australia) (Luong et al., 2013).

Genotyping and association analysis with coding variants

SNP genotypes in the 12q22 region were extracted from genome-wide data from our previously published GWAS analysis (Painter et al., 2011; Nyholt et al., 2012). We also re-analysed and genotyped part of this dataset on Human CoreExome genotyping chips (Immuna, Inc., San Diego, CA, USA). There were 2213 surgically confirmed endometriosis cases and 1375 controls (out of the 2044 controls in GWAS) with DNA samples available.

Additionally, key imputed SNPs and known coding variants in VEZT and FGDO6 identified from public databases (including dbSNP, 1000 Genomes, the NHLB1 GO Exome Sequencing Project) (Luong et al., 2013) were genotyped in a subset of samples, including 929 of the surgically confirmed cases with a family history of endometriosis and a control group of 959 unrelated women with self-reports of no endometriosis (Zhao et al., 2006). SNPs were genotyped using Agena MassARRAY technology. The genotyping was conducted with iPLEX chemistry on a matrix assisted laser desorption/ionization time-of-flight (MALDI-TOF) Mass Spectrometer following the standard Sequenom procedures (Sequenom, Inc., San Diego, CA, USA).

Standard quality control procedures were applied to individual datasets as outlined previously (Luong et al., 2013). Briefly, SNPs with >5% missing rate, out of Hardy–Weinberg equilibrium (P < 10−4) in controls and minor allele frequency (MAF) < 1% were excluded. The samples overlapping between the three datasets as described above (including samples from GWAS, samples genotyped on Human CoreExome chips and the sample subset with surgically confirmed cases with a family history of endometriosis and a control group of unrelated women with self-reports of no endometriosis), from non-European ancestry and with low call rates (< 95%) were excluded from the downstream analyses. The final combined Australian dataset for the fine-mapping analyses consisted of 2594 cases and 4496 controls. Of the total 7090 individuals in the combined dataset, 6503 are unrelated while 587 are related to some degree.
**In silico analysis**

Potential functional roles and regulatory annotations for identified SNPs were investigated using HaploReg v2 (Broad Institute), UCSC Genome Browser and the ENCODE database. The ENCODE data include areas of open chromatin identified using DNase hypersensitivity (HS), formaldehyde-assisted isolation of regulatory elements (FAIRE) and chromatin immunoprecipitation (ChIP) experiments and the locations of functional regulatory elements (including promoters, enhancers, silencers and insulators). The RegulomeDB programme was employed to rank the potential functional roles for SNPs based on their location across the entire 107 kb region. The scoring system for RegulomeDB ranges from 1 to 6 with 1 being the strongest evidence for functional roles (Boyle et al., 2012). LD for the genotyped SNPs was examined using SNAP Pairwise LD (Broad Institute).

**Gene expression in blood**

We used gene expression data from the BSGS to investigate the effect of SNPs located within chromosome 12q22 on cis-located probes. BSGS comprises 862 individuals of European descent from 274 independent families (Powell et al., 2012). DNA samples from each individual were genotyped on the Illumina 610-Quad Beadchip by the Scientific Services Division at deCODE Genetics Iceland (Medland et al., 2009; Powell et al., 2012). Filtered genotypes were then imputed to the 1000 Genomes Reference Panel (release 3.0) using hapi-ur (Williams et al., 2012) and impute2 (Howie et al., 2009). SNPs with a poor imputation quality score (R² < 0.95) and with a MAF < 0.05 were removed.

Whole blood for expression profiling was collected directly into PAXgene tubes (QIAGEN, Valencia, CA, USA). Total RNA was extracted from PAXgene tubes using the whole blood gene RNA purification kit (QIAGEN). RNA from all samples was run on an Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA) to assess RNA integrities and to estimate RNA concentrations. RNA was amplified and converted to biotinylated cRNA using the Ambion Illumina TotalPrep RNA Amplification Kit (Thermo Fisher Scientific, Scoresby, VIC, Australia).

Expression profiles were generated by hybridizing 750 ng of cRNA to Illumina Human HT-12 v4.0 Beadchips according to Illumina whole-genome gene expression direct hybridization assay guide (Illumina, Inc.). Briefly, 500 ng of total RNA were used to generate biotinylated cRNA, which was fragmented and to estimate RNA quantities. RNA was amplified and converted to biotinylated cRNA, with check for balance across families, sex and generation.

**Gene expression in endometrial tissue**

Total RNA from homogenized endometrial tissues from 136 women was extracted using RNA lysis solution (RLT buffer) and the RNeasy Plus Mini Kit according to the manufacturer’s instructions (QIAGEN). RNA quality was assessed with the Agilent Bioanalyzer 2100 (Agilent Technologies) and concentrations were determined using the NanoDropND-6000 (Thermo Fisher Scientific). RNA (250 ng) was amplified and converted to biotinylated cRNA using the Ambion Illumina TotalPrep RNA Amplification Kit (Thermo Fisher Scientific). Expression profiles in endometrial tissue were generated by hybridizing 750 ng of cRNA to Illumina Human HT-12 v4.0 Beadchips (as described above).

DNA was extracted from samples of whole blood from the same individuals and samples were genotyped for a total of 12 variants located within the 12q22 region (4 top GWAS/imputed SNPs from meta-analysis, rs10859871, rs11107968, rs12298029, rs1387047, top imputed SNP in stage B in meta-analysis, rs4762347, top SNPs from fine-mapping, rs10777670 and rs4128507, top SNPs after conditional analysis, rs7302937 and rs8553 and functionally annotated SNPs, rs6538617 and rs2291332) and strong genotyped expression SNP (eSNP) from BSGS (rs14121) using the Agena MassARRAY platform (Sequenom, Inc.) or Taqman SNP genotyping assay. All SNPs had call rates > 95%.

**Western blotting**

A minimum of 10 mg of frozen endometrium was used per patient for whole cell protein extraction. Protein lysates were prepared as described previously (Zaitseva et al., 2013; Fung et al., 2015). Protein concentration was determined using the Pierce BCA Protein Assay following the manufacturer’s instructions (Thermo Fisher Scientific). Probing of western blots was carried out as described previously (Zaitseva et al., 2013; Fung et al., 2015) with the following primary and secondary antibodies; mouse monoclonal IgG2b VEZT (A-3) (Santa Cruz Biotechnology, Inc.) (1:10 000 dilution), rabbit polyclonal IgG VEZT (N3C3) (GeneTex, Irvine, CA, USA) (1:1000), mouse monoclonal β-actin (Sigma-Aldrich) (1:10 000 dilution), and HRP-conjugated goat anti-rabbit or mouse secondary antibodies (1:2000) (Bio-Rad Laboratories, Hercules, CA, USA). β-Actin was detected at 42 kDa and VEZT between 90 and 100 kDa. VEZT immunoblots were normalized using an internal protein lysate control and β-actin, with values displayed as mean optical densities (OD/mm²).

**Statistical analysis**

Fine-mapping and association analyses

The final merged data were imputed using the MACH program to impute missing genotypes (Li et al., 2009, 2010). Quality of the imputed genotypes was assessed by R² metric. All SNPs passed the standard imputation quality control threshold (R² > 0.3). Association analysis of imputed genotype dosage scores with endometriosis was performed using PLINK software (http://pngu.mgh.harvard.edu/purcell/plink/) (Purcell et al., 2007). To account for relatedness in the dataset, the analysis was conducted using a robust variance estimation approach available in PLINK (Barlow, 1994; Williams, 2000).

eQTL mapping in whole blood

Normalization procedures included pre-processing of data generated by the Illumina Bead Array Reader using Illumina software, GenomeStudio (Illumina, Inc.). Pre-processing included; correction for chip background effects, removal of
outlier beads, computation of average bead signal and calculation of detection P-values using negative controls present on the array. Removal of chip background effects can lead to negative expression levels for transcripts with low levels of measured expression. To avoid problems with further normalization procedures, negative values were denoted as missing data identifiers. Thus, in subsequent normalization procedures and analyses samples with probes coded as missing were ignored.

The Illumina HT-12 v4.0 chip contains 19 probes that tag transcripts located within the 12q22 region. To avoid spurious associations, we removed any probes where <10% of samples had a detection P-value <0.05 (11 probes), leaving 8 probes for analysis.

The gene expression normalization and eQTL mapping have been described in detail elsewhere (Powell et al., 2012, 2013). However, we will briefly describe the methods here. In order to minimize the influence of overall signal levels, which may reflect RNA quantity and quality rather than a true biological difference between individuals, the following standardization procedures were applied. Adjusted expression levels for each probe were transformed using a Quantile transformation (Bolstad et al., 2003; Smyth and Speed, 2003) to achieve a stabilized distribution across average expression levels. Further normalization was performed to allow expression levels to be compared across chips and genes. This was achieved by fitting a linear mixed model:

\[ y = X\beta + Z\gamma + \epsilon \]  

where \( y \) is a vector of log-transformed probe expression levels, \( \beta \) is an unknown vector of fixed effects of batch and blood cell counts (extraction date, gender, red blood cells, platelets, neutrophils, monocytes, eosinophils, basophils, CD4, CD8, CD19, CD56). \( \gamma \) is an unknown vector of random effects of batch (chip and chip position, age) with known design matrix \( Z \), and \( \epsilon \) is a vector of residual errors. The residuals from this model were standardized to z-scores and used in all further analyses.

The relationship between SNP genotypes and normalized probe expression levels had been tested for using a linear mixed model (—assoc command) implemented in MERLIN (Abecasis et al., 2002). SNP genotype effects were estimated assuming an additive genetic model.

eQTL mapping in endometrial tissue

Out of the 136 samples with gene expression array data, the eQTL analysis was performed on the total of 122 samples that had both gene expression and genotyping data. Adjusted expression levels for each probe were transformed as described above. Further normalization was performed to allow expression levels to be compared across chips and genes. This was achieved as described above.

Logistic regression was used to test for differential gene expression between cases and controls and between phases of menstrual cycle of the tissue samples (proliferative and secretory phases determined from histological evaluation), with and without adjusting for phases of menstrual cycle and case/control status, respectively. An interaction term in the logistic model was also included to assess for possible interaction between phases of the menstrual cycle and case/control status.

For each of the 12 variants examined, a cis-eQTL analysis was performed to investigate putative association between the variant and expression levels of nearby transcripts. The eQTL analysis was performed on the total of 122 tissue samples with recoded SNP genotypes based on minor allele dosage and fitted linear regression models, with phases of the menstrual cycle included as covariates in the model. Study-wide significance (6.11 × 10^{-40}) was determined using a Bonferroni adjustment (0.05/number of tests performed).

**Protein studies**

Results for protein studies were analysed using Graph Pad Prism™ software (version 5, Graph Pad software, CA, USA). Data were normally distributed (by Bartlett’s test). IHC scoring data were analysed by two-way analysis of variance (ANOVA) and western blot densitometry was analysed by one-way ANOVA (Tukey) or unpaired t tests (with Welch-correction when necessary) with significance at \( p < 0.05 \).

**Results**

Different VEZT coding variants identified in an endometriosis-dense cohort

The coding region of VEZT was screened to identify coding variants in an endometriosis-dense cohort. Twelve coding exons (exons 2–13), the 5′UTR, the 3′UTR and the intron/exon boundaries for VEZT were screened by HRM analysis in 100 endometriosis cases with results confirmed by sequencing. Eleven coding variants were detected, including ten known variants and one novel variant (Table II). Nine of the ten known variants were common, with one rare synonymous change observed in exon 9 (rs61735395). Of five missense variants identified (rs17855933, rs10507051, rs17344738, rs17855934 and rs14121), three variants were predicted to be deleterious (D) (rs17340751, rs17344738 and rs14121), and two variants, rs17855933 and rs10507051, were predicted to have an effect on splicing (Table II).

Two common coding variants in the 3′UTR of VEZT, rs1046007 and rs4468424, were located at binding sites of human micro-RNAs (miRNAs), including hsa-miR-448 for rs1046007, and hsa-miR-552 and hsa-miR-606 for rs4468424 (Table II).

We identified a novel variant in the 5′UTR of VEZT with a G to A change 26 bp downstream from the transcription start site (Table II). This novel variant was predicted to be at the binding site for the transcription factor MyoD. The novel variant was detected in one endometriosis patient in heterozygous form. Other members of the family for whom DNA samples were available, including the proband individual, were screened via HRM and/or sequencing for the presence of the relevant variant. The 5′UTR novel variant was present in the proband, mother (who does not have endometriosis) and an affected sister, but was not detected in another affected sister (Supplementary data, Fig. S1).

Association analysis with endometriosis risk for common and coding variants

Forty-five VEZT coding variants including those identified by sequencing (Table II) and 62 coding variants with functional predictions in silico in the adjacent gene FGD6 were selected from public databases and genotyped in 929 endometriosis cases and 958 controls (Zhao et al., 2006; Luong et al., 2013). Most variants were non-polymorphic in our sample (38/45 variants for VEZT and 53/62 variants for FGD6).

Association signals for endometriosis risk in the 12q22 region were compared in combined data for Australian cases and controls (total 7090 individuals), including polymorphic coding SNPs, additional genotyping for imputed SNPs with potential functional effects (Luong et al., 2013) and previous GWAS data (Table III). In the combined data, rs10777670 [A/G] located at bp 95574831 in intron two of FGD6 gave a slightly stronger signal than rs10859871 (\( P = 7.44 \times 10^{-4} \), OR = 1.17 (95% CI: 1.07–1.28)) (Table III). This SNP is in moderate LD with the published GWAS signal rs10859871 (\( r^2 = 0.49, D' = 1 \)). In further analysis and conditioning on the sentinel SNPs rs10859871 or rs10777670, there were no other SNPs showing stronger association and no evidence for secondary association signals in the VEZT region. This suggests that the endometriosis association signal is not explained...
by any coding variants in VEZT or FGD6. The functional annotation(s) of the SNPs with nominally significant \( (P < 2 \times 10^{-3}) \) association with endometriosis risk are listed in Supplementary data, Table SI.

### Localization of VEZT protein in human endometrium

Before investigating a possible eQTL in VEZT expression between women with different genotypes at the sentinel SNP, we first verified the presence, localization and effect of menstrual cycle stage on basal VEZT protein expression in human endometrium and myometrium (from non-genotyped patients).

VEZT protein was identified in the endometrium (Fig. 1a–f) and myometrium (Fig. 1g) from premenopausal cycling women by IHC. VEZT cellular localization was both nuclear and cytoplasmic. Not all

<table>
<thead>
<tr>
<th>Table 2 Variants identified by screening VEZT in 100 endometriosis cases.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Variants</strong></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Novel</td>
</tr>
<tr>
<td>rs3751272</td>
</tr>
<tr>
<td>rs1785933</td>
</tr>
<tr>
<td>rs61735395</td>
</tr>
<tr>
<td>rs12424550</td>
</tr>
<tr>
<td>rs10507501</td>
</tr>
<tr>
<td>rs17344738</td>
</tr>
<tr>
<td>rs1785934</td>
</tr>
<tr>
<td>rs14121</td>
</tr>
<tr>
<td>rs1046007</td>
</tr>
<tr>
<td>rs4648424</td>
</tr>
</tbody>
</table>

**Table 3 Endometriosis association information for common single nucleotide polymorphisms (SNPs) with the key SNP (rs10859871) and significant association with endometriosis risk \( (P < 2 \times 10^{-3}) \).**

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Position (Hg19)</th>
<th>LD with rs10859871 (( r^2 ))</th>
<th>RA</th>
<th>OA</th>
<th>RA case</th>
<th>RA control</th>
<th>OR (95% CI)</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs10777670</td>
<td>95578431</td>
<td>0.49</td>
<td>A</td>
<td>G</td>
<td>0.197</td>
<td>0.174</td>
<td>1.17 (1.07–1.28)</td>
<td>7.44 ( \times 10^{-4} )</td>
</tr>
<tr>
<td>rs10859871</td>
<td>95711876</td>
<td>I</td>
<td>C</td>
<td>A</td>
<td>0.330</td>
<td>0.303</td>
<td>1.14 (1.05–1.23)</td>
<td>1.01 ( \times 10^{-3} )</td>
</tr>
<tr>
<td>rs2290829</td>
<td>95695804</td>
<td>I</td>
<td>A</td>
<td>C</td>
<td>0.330</td>
<td>0.303</td>
<td>1.13 (1.05–1.22)</td>
<td>1.21 ( \times 10^{-3} )</td>
</tr>
<tr>
<td>rs4128507</td>
<td>95577884</td>
<td>0.49</td>
<td>G</td>
<td>A</td>
<td>0.200</td>
<td>0.177</td>
<td>1.16 (1.06–1.27)</td>
<td>1.42 ( \times 10^{-3} )</td>
</tr>
<tr>
<td>rs11107968</td>
<td>95690444</td>
<td>I</td>
<td>G</td>
<td>A</td>
<td>0.331</td>
<td>0.306</td>
<td>1.13 (1.05–1.23)</td>
<td>1.67 ( \times 10^{-3} )</td>
</tr>
<tr>
<td>rs4762342</td>
<td>95690996</td>
<td>I</td>
<td>A</td>
<td>G</td>
<td>0.336</td>
<td>0.310</td>
<td>1.14 (1.05–1.23)</td>
<td>1.82 ( \times 10^{-3} )</td>
</tr>
<tr>
<td>rs6538617</td>
<td>95611141</td>
<td>0.51</td>
<td>T</td>
<td>C</td>
<td>0.226</td>
<td>0.204</td>
<td>1.15 (1.05–1.26)</td>
<td>1.93 ( \times 10^{-3} )</td>
</tr>
</tbody>
</table>

Fine-mapping results were from 7090 individuals (2594 cases and 4496 controls) in the combined Australian dataset. CI, confidence interval; RA, risk allele; OA, other non-risk allele.

\( ^1 \)LD (linkage disequilibrium) from Broad SNAP results 1000G Pilot.

\( ^2 \)Risk allele frequency.

\( ^3 \)Odd ratios (OR) were calculated for the risk allele.
There was strong evidence for significant association between SNPs in the region affecting expression of VEZT, and the pattern of association completely overlapped our association signal for endometriosis risk in this region (Fig. 2a and b). The sentinel SNP for the endometriosis association rs10859871 showed evidence for association with VEZT gene expression (Table IV, \( P = 2.16 \times 10^{-4} \)), where the risk allele (C) increased the expression levels of VEZT (Fig. 3a). There was little evidence for effects of rs10859871 on the expression levels of other transcripts located within the 12q22 locus (including ubiquitin-conjugating enzyme E2N (UBE2N), coiled-coil domain-containing protein 41 (CCDC41), nuclear receptor subfamily 2, group C, member 1 (NR2C1) and methionyl aminopeptidase 2 (METAP2)) (Table IV).
SNPs within the 12q22 locus are associated with both endometriosis risk and the expression levels of VEZT. Association results for individual SNPs are plotted by position on chromosome 12 (hg19; X-axis) and as $-\log_{10} P$-values (Y-axis) for endometriosis risk (a) and VEZT expression in the BSGS (b). The relative locations of genes (RefSeq genes) within the 12q22 locus are shown in (c).
The genotyped SNP with the strongest association on VEZT expression (Fig. 2b) was rs2367305 (VEZT probe ILMN_1774828; \( P = 3.3 \times 10^{-11} \)). With respect to rs2367305, there was little association with endometriosis risk (\( P = 0.102 \)). We also examined the SNP rs14121, in high LD with rs2367305 (LD: \( r^2 = 0.947 \), SNP Annotation and Proxy Search, Broad Institute), which was a strong eSNP for the same probe (\( P = 2.85 \times 10^{-10} \)). This SNP did not have a strong association with endometriosis risk (\( P = 0.076 \)) from fine-mapping results described above; however, it was a coding variant of VEZT (see Table II). The minor allele (G) of rs14121 was associated with a decrease in VEZT expression (Fig. 3b).

### Effect of endometriosis SNPs at 12q22 on VEZT transcription in endometrium

For SNPs in the region of chromosome 12q22 typed on the Agena MassARRAY or TaqMan SNP genotyping assays, effects of genotype on endometrial tissue-specific gene expression were tested after fitting stage of the cycle and case/control status as covariates. Expression levels for the VEZT probe (ILMN_1774828) showed strongest evidence for eQTLs with sentinel SNP rs10859871 and rs12298029 (\( P = 7.19 \times 10^{-4} \)) (Table V). Figure 3c shows a box plot for the effect of rs10859871 on endometrial VEZT expression, the risk allele (C) was associated with an increase in VEZT. We were unable to genotype SNP rs2367305 in our tissue samples as the Taqman assay failed the manufacturer’s quality control. Instead we genotyped the SNP rs14121 in the blood samples from the same individuals who we collected endometrial tissue samples from. The minor allele (G) of rs14121 was associated with a decrease in VEZT expression (\( P = 4.62 \times 10^{-7} \)) (Fig. 3d). Both SNPs rs10859871 and rs14121 showed the same directional effects in endometrium as observed in whole blood (Fig. 3). No other genes in the region showed any evidence for effects of genotype on endometrial expression with any of the SNPs tested.

### Effect of endometriosis SNP rs10859871 at 12q22 on VEZT protein expression in endometrial tissues

Two commercial antibodies raised against VEZT were employed to quantify VEZT protein expression in endometrium. The mouse antibody was raised against the first 160 amino acids of VEZT isoform 1, which shares sequence in common with all VEZT isoforms. Using the mouse monoclonal VEZT antibody, a single VEZT band (100 kDa) was detected by western blotting (Fig. 4a and b). The rabbit antibody was raised against the central region of VEZT isoform 1: this region of the protein is the same for VEZT isoform 2. The rabbit polyclonal VEZT antibody resulted in the detection of two bands in this same region, one faint band also at 100 kDa (Fig. 4a and d), and a second band at 90 kDa (Fig. 4a and c). All bands were analysed for an effect of rs10859871 on endometrial VEZT expression. VEZT protein expression in the endometrium was not different between women homozygous for the risk allele (C), the other non-risk allele (A) and heterozygous (CA) genotypes for rs10859871 (Fig. 4b–d). When rs10859871 genotyped women were grouped according to menstrual cycle stage (proliferative or secretory), no significant change in VEZT expression was observed by western blot (Fig. 4e).

Furthermore, we were unable to detect any difference in VEZT protein expression caused by rs14121 genotype (\( P = 0.6009 \)) (data not shown).

Women were also grouped into disease state (regardless of rs10859871 genotype); confirmed endometriosis (cases) and controls. Western blots showed that endometrial VEZT protein expression was not affected by endometriosis disease status (\( P = 0.1400 \)) (Fig. 4f).

Western blotting with whole cell protein extracts did not reveal any difference in VEZT protein expression in response to endometriosis SNP rs10859871. Therefore, we performed IHC on endometrial curette sections collected at the time of laparoscopic surgery to assess if changes in VEZT expression are occurring in a cell-specific manner, and were subsequently missed by western blot analysis. IHC data for VEZT in endometrial curettes were typical of those seen in full thickness hysterectomy sections (Figs 1 and 4a–d), for example, in the secretory phase VEZT staining in the glands was stronger relative to stromal cells (Fig. 4a–d). Semi-quantitative analysis of VEZT staining intensity validated a significant increase in VEZT staining in the glands of both risk allele (C) and other allele (A) women compared with risk allele stromal cells, in the secretory phase (\( P = 1 \times 10^{-4} \)) (Fig. 4e). In risk allele (C) endometrium, proliferative phase glands had significantly increased VEZT present relative to secretory phase stromal cells (Fig. 4a, b and e). This cycle–stage specific increase was not observed in not-at-risk (A) endometrial glands (Fig. 4e). There was however, no significant difference between

### Table 4 Effect of the original genome-wide association study sentinel (rs10859871) endometriosis SNPs on the expression levels of transcripts located within 12q22 locus.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Probe (bp, hg19)</th>
<th>P-Value rs10859871</th>
<th>Effect (SE) rs10859871</th>
</tr>
</thead>
<tbody>
<tr>
<td>UBE2N</td>
<td>ILMN_1793651</td>
<td>0.879</td>
<td>-0.01 (0.06)</td>
</tr>
<tr>
<td>CCDC41</td>
<td>ILMN_1799113</td>
<td>0.323</td>
<td>-0.08 (0.08)</td>
</tr>
<tr>
<td>NR2C1</td>
<td>ILMN_1728983</td>
<td>0.636</td>
<td>0.03 (0.06)</td>
</tr>
<tr>
<td>NR2C1</td>
<td>ILMN_1734867</td>
<td>9.14 x 10^-2</td>
<td>-0.12 (0.07)</td>
</tr>
<tr>
<td>NR2C1</td>
<td>ILMN_2326675</td>
<td>0.192</td>
<td>0.07 (0.06)</td>
</tr>
<tr>
<td>NR2C1</td>
<td>ILMN_1774382</td>
<td>0.972</td>
<td>0.06</td>
</tr>
<tr>
<td>VEZT</td>
<td>ILMN_1774828</td>
<td>2.16 x 10^-4</td>
<td>0.21 (0.06)</td>
</tr>
<tr>
<td>METAP2</td>
<td>ILMN_1670420</td>
<td>0.243</td>
<td>-0.07 (0.06)</td>
</tr>
</tbody>
</table>

Expression levels of probes were measured in whole blood for 862 individuals from the Brisbane Systems Genetics Study. UBE2N, ubiquitin-conjugating enzyme E2N; CCDC41, coiled-coil domain-containing protein 41; NR2C1, nuclear receptor subfamily 2, group C, member 1; VEZT, vezatin, and METAP2: methionyl aminopeptidase 2.
women homozygous for the risk allele (C) and other allele (A) genotypes overall (for rs10859871) \( P = 0.5638 \).

**Discussion**

In this study, we undertook a detailed analysis of chromosome region 12q22, which is associated with an increased risk of developing endometriosis (Nyholt et al., 2012; Rahmioglu et al., 2014). The sentinel SNP rs10859871 lies in close proximity to genes VEZT and FGD6 (Nyholt et al., 2012). VEZT and FGD6-associated pathway components have demonstrated changes in expression in association with endometriosis (Meola et al., 2010; Luong et al., 2013). Both of these genes physiologically play a role in plasma membrane and cytoskeletal remodelling, and are thus strong candidates for having a causal role in endometriosis and lesion establishment. We identified 11 coding variants of VEZT in an endometriosis-dense cohort, of which one was novel and three were predicted to be deleterious. Results from additional genotyping in this study demonstrate that the association signal at 12q22 is not accounted for by coding variants in VEZT or FGD6. We found strong evidence for cis-eQTLs, with SNPs in the 12q22 region affecting VEZT expression in blood and endometrium that overlapped the association signal for endometriosis risk. Both blood and endometrial VEZT showed strong evidence for eQTLs with SNP rs10859871. However, the new sentinel SNP rs10777670 did not show strong eQTLs for VEZT in both blood and endometrial tissue. These eQTLs were not replicated at the protein level, nevertheless, we determined that VEZT expression is cycle stage and cell-type specific in the endometrium. Therefore following our cis-eQTL of the 12q22 region, VEZT remains a significant

![Figure 3 VEZT expression quantitative trail loci (eQTL) effects in blood and eutopic endometrium.](https://academic.oup.com/humrep/article-abstract/31/5/999/1750240/1008)

(a) Blood VEZT (ILMN_1774828) expression with eQTL effects for the relationship between rs10859871 (a and c) and rs14121 (b and d) and gene expression in blood (top panels) and in endometrium (bottom panels). Risk allele (CC), heterozygotes (CA) and other not-at-risk allele (AA).
candidate for enhancing genetic risk of endometriosis as well as having a plausible functional role in the pathogenesis of the disease. Detailed screening of SNP variants within the coding region of VEZT allowed us to re-evaluate endometriosis risk association for this region. Nine common VEZT variants, one rare variant and one novel VEZT variant were identified. In silico functional tools predict that common variants rs10507051, rs17344738 and rs14121 lead to amino acid changes that are potentially deleterious to VEZT function. Two other common variants were identified in the 3′UTR of VEZT (rs1046007 and rs4468424) and were predicted to be related to miRNA binding sites, thus potentially effecting regulation of VEZT expression. Association signals for endometriosis risk took into consideration the eleven coding variants identified in this investigation, plus 34 other known VEZT coding variants. In this study, the strongest signals for association with endometriosis were the sentinel GWAS SNP rs10859871 and rs10777670, SNPs located in the intergenic region downstream of VEZT and within an intron of FGD6, respectively (Fig. 5).

Gene expression data from whole blood showed strong evidence for an eQTL for VEZT expression with the signal completely overlapping the association signal for endometriosis risk on chromosome 12q22. The genotyped SNP rs2367305 with the greatest effect on VEZT expression did not have strong association with endometriosis risk. The sentinel GWAS SNP rs10859871 demonstrated an effect on VEZT expression in blood, but the eQTL was not significant after multiple testing. While the sample size for endometrial tissue was smaller than for blood samples, the risk allele frequency for rs10859871 is common (allele frequency 0.303 in controls), the risk allele increased expression in both endometrial tissue and blood RNA and the effect sizes were similar. However, the results for gene expression and disease-risk were generated in different samples and future studies will need to determine if the same causal variants affect both VEZT expression and endometriosis risk. Our eQTL results are supported by a meta-analysis in the world’s largest cis-eQTL study on peripheral blood samples from ~5000 individuals, which identified a strong cis-association mapped to endometriosis risk locus at 12q22 with target gene VEZT (for rs10859871 P = 2.78 × 10^-22) (Westra et al., 2013).

Our results and publicly available data provide evidence that many disease eQTLs are tissue or cell-type specific (McKenzie et al., 2014; GTEx Consortium, 2015). Therefore, it is important to map eQTLs in both whole blood and disease-specific tissues (for example, eutopic endometrial tissue), as we have undertaken in this study. In endometrial tissues, VEZT expression was effected by rs10859871. Mirroring the blood eQTL in BSGS, individuals homozygous for the risk allele (C) showed evidence for increased endometrial VEZT expression. The pilot analysis genotype-tissue expression (GTEx) study did not provide evidence for an effect of SNP rs10859871 on VEZT expression, but there was a significant eQTL for nearby gene FGD6 in tibial artery tissue (P = 2.4 × 10^-7) (GTEx Consortium, 2015), where the minor allele (C, the risk allele for endometriosis) showed a decrease in FGD6 expression. However, FGD6 was expressed in <90% of the endometrial tissue samples tested in this study and expressed at low levels in the whole blood in BSGS and GTEx, therefore this gene was not included in our eQTL analyses. Given the sample sizes of the BSGS and study of Westra et al. (2013) were much greater than the endometrial tissue samples available for this study, future studies with increased sample sizes will be required to increase the statistical power of disease-specific eQTLs. We estimate that with a sample size of 100, we have 80% power to detect an eQTL explaining 10% of the variation in VEZT expression. A substantial increase in sample size will be needed to detect more subtle effects. The gene expression in the endometrium varies throughout the menstrual cycle, therefore even larger sample sizes with well-defined stages of the menstrual cycle will be required to analyse gene regulation in the endometrium and fully account for effects of the stage of the menstrual cycle.

Our protein studies did not demonstrate any significant changes in endometrial VEZT expression with respect to rs10859871, and the results may also be a consequence of small sample size. However, we observed an increase in VEZT expression specific to proliferative phase glandular epithelium in risk allele (C) individuals. VEZT has fundamental roles in epithelial cell adherens junctions and since this difference was not observed in non-risk endometrium (A), our finding supports a possible functional difference in VEZT between at-risk and not-at-risk women (specific to proliferative phase epithelial glandular cells). Elevated VEZT had been reported previously in association with endometriosis (in ectopic versus eutopic endometrium of women with endometriosis) (Meola et al., 2010). We suggest that an increase in VEZT might contribute to greater tension between the plasma membrane and the actin cytoskeleton (cell–cell adhesion), enhancing the capacity of lesion epithelia to establish and maintain connection to the basement membrane. Loss of VEZT is deleterious in conditional knock-out mice (Hyenne et al.,

### Table 5  SNP effects for transcription level measured in endometrial tissue.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Position (hg19)</th>
<th>LD * (r²)</th>
<th>VEZT (ILMN_1774828)</th>
<th>eQTL P-value</th>
<th>Dist from SNP (bp)</th>
<th>Endo P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs10859871</td>
<td>95711876</td>
<td>I</td>
<td>C A</td>
<td>0.20</td>
<td>7.19 × 10^-6</td>
<td>17539</td>
</tr>
<tr>
<td>rs12298029</td>
<td>95695804</td>
<td>I</td>
<td>A C</td>
<td>0.20</td>
<td>7.19 × 10^-6</td>
<td>1468</td>
</tr>
<tr>
<td>rs11107968</td>
<td>95690444</td>
<td>I</td>
<td>G A</td>
<td>0.20</td>
<td>9.10 × 10^-6</td>
<td>1034</td>
</tr>
<tr>
<td>rs6538617</td>
<td>9561141I</td>
<td>0.51</td>
<td>T C</td>
<td>0.16</td>
<td>6.24 × 10^-4</td>
<td>-83 195</td>
</tr>
<tr>
<td>rs4762342</td>
<td>95690996</td>
<td>I</td>
<td>A G</td>
<td>0.16</td>
<td>8.44 × 10^-4</td>
<td>-38 93</td>
</tr>
<tr>
<td>rs10777670</td>
<td>95574831</td>
<td>0.49</td>
<td>A G</td>
<td>0.14</td>
<td>7.65 × 10^-3</td>
<td>-119 505</td>
</tr>
<tr>
<td>rs4128507</td>
<td>95577884</td>
<td>0.49</td>
<td>G A</td>
<td>0.12</td>
<td>1.59 × 10^-2</td>
<td>-116 452</td>
</tr>
</tbody>
</table>

*LD (r²) with the sentinel SNP rs10859871 showing the strongest association with endometriosis risk eQTL, expression quantitative trait loci.

Downloaded from https://academic.oup.com/humrep/article-abstract/31/5/999/1750240 by guest on 03 January 2019
and reduced expression is linked to gastric cancer (Guo et al., 2011; Miao et al., 2013), however, the functional significance of enhanced VEZT expression is not known.

Our endometrial cis-eQTL data identified significant associations for VEZT expression with women homozygous for risk alleles (rs10859871 and rs14121). To improve our understanding of the role of VEZT in endometrium, we demonstrated that VEZT protein is widely expressed in human endometrium and myometrium. The literature lacks information describing adherens junctional complex member VEZT in gynaecological tissues and, to date, it is not known if this protein regulates or is regulated by steroid hormones. We show that VEZT expression significantly increases in glandular epithelium during the secretory phase of the menstrual cycle. In support of our finding, adherens junction members (E-cadherin, α- and β-catenin) mRNA expression is also increased in the secretory phase relative to the proliferative phase indicating that progesterone may activate cell to

Figure 4 VEZT protein expression in eutopic endometrium from women homozygous for the risk allele (RA), homozygous for the other non-risk allele (OA) or heterozygous (Het.). Western blots for VEZT were performed on endometrial whole cell protein lysates from RA (CC) (n = 8), Het. (CA) (n = 22) and OA (AA) women (n = 26). (a) Top panel, a representative western blot of VEZT using a mouse monoclonal VEZT antibody (mVEZT, 100 kDa). Middle panel, VEZT bands detected using a rabbit polyclonal antibody (rVEZT), weak band detected at 100 kDa and a stronger band detected at 90 kDa. Lower panel, western blots were normalized for β-actin (42 kDa). (b) Relative protein expression of VEZT (100 kDa; mouse antibody) is displayed as the mean optical density (OD/mm²) ± SEM. (c and d) Relative protein expression of VEZT (100 and 90 kDa; rabbit antibody) are displayed as the mean optical density (OD/mm²) ± SEM. (e) VEZT protein expression (mouse antibody) from RA, Het. and OA individuals relative to menstrual cycle stage (proliferative (Pro) and secretory (Sec) phases). RA white bars, Het. diagonal striped bars and OA grey bars. (f) Relative VEZT protein expression (mouse antibody) in endometrium from endometriosis cases (white, n = 42) and controls (grey, n = 15).
The VEZT promoter does not contain a response element for the progesterone receptor (PR), but it does contain an NF-kappaB (NF-κB) binding site. NF-κB, a pro-inflammatory transcription factor, is implicated in the pathogenesis of endometriosis, demonstrates cycle stage-dependent regulation in the endometrium and mutual regulation with PR (Guo, 2007). Therefore, the observed changes in VEZT levels in endometrial glandular cells may be occurring in response to the dynamic fluctuation in progesterone and associated-NF-κB modifications.

The strengths of GWAS are numerous; they have discovered hundreds of genetic variants associated with many complex diseases (for example, Crohn’s disease, Type 2 diabetes and coronary disease) and have allowed a better understanding of the complexities of genetic architecture (Manolio, 2009). However, most identified variants explain only a minor proportion of the estimated heritability for most complex diseases, and the effect sizes for most individual SNP variants are relatively small (ORs of risk alleles from 1.1 to 1.5) (Visscher and Montgomery, 2009; Lee et al., 2013). Given that most causal variants have a small effect size, multiple loci must be acting together to cause the genetic variation that leads to disease-risk and therefore, individuals must carry their own novel assortment of risk variants (Visscher and Montgomery, 2009). Hence examining individual SNP variants, independently and one at a time, may not be enough to fully understand the genetic basis of diseases.

**Figure 5** Representative micrographs of the immunolocalization of VEZT in human endometrium from women homozygous for the RA or OA. Representative micrographs of VEZT immunolocalization in women homozygous for RA (CC) (a) \( n = 5 \) and (b) \( n = 5 \) and homozygous for OA (AA) (c) \( n = 10 \) and (d) \( n = 10 \). Proliferative phase endometrium (a) and (c). Secretory phase endometrium (b) and (d). Scale bar is equal to 100 μm. (e) VEZT staining intensity was scored between 0 and 3 (where 0 = none and 3 = strong) by three independent scorers separately for endometrial stroma and glands. Semi-quantitative data were analysed by two-way ANOVA (significance denoted by \(* P < 0.05\).
time, may not be the most effective method for validating a role for highly ranked SNPs associated with endometriosis risk. Future approaches should examine various significant candidates in combination with other candidate SNPs.

We hypothesized that the endometriosis risk-associated SNP, rs10859871, would have an eQTL for the VEZT gene. In addition to the original sentinel SNP rs10859871, we also identified rs14121, which both show evidence of eQTLs with VEZT in blood and endometrial tissue. While further work is necessary to provide functional insight into the role of VEZT in endometriosis, we have undertaken the essential next step in validating the important data assembled from large-scale innovative GWAS aimed at better understanding endometriosis (Rogers et al., 2013).

Supplementary data
Supplementary data are available at http://humrep.oxfordjournals.org/.

Acknowledgements
The authors acknowledge the hard work of Research Nurses Ranita Charitra, Tracy Middleton and Irene Bell from the Royal Women’s Hospital. We would like to acknowledge Ashleigh Henders for sample processing, RNA/DNA extraction and genotyping. Thank to Matthew Chapman and Leanne Taylor for access to the Aperio ScanScopeXT system at the Melbourne Health Tissue Bank (Royal Melbourne Hospital). We would also like to thank the staff from Anatomical Pathology at the Royal Women’s Hospital and Royal Children’s Hospital sites. Thank, most of all, to the staff and patients of Gynaecology 1 Clinic (Royal Women’s Hospital) for donating their time and tissues to this project.

Authors’ roles
All authors played a role in the concept/design of the study or acquisition and or analysis/interpretation of data, drafted or critically revised the article and gave their final approval of the article.

Funding
Funding for this work was provided by National Health and Medical Research Council (NHMRC, Australia) Project Grants GNT1012245, GNT1026033, GNT1049472 and GNT1046880. G.W.M. is supported by the NHMRC Fellowship scheme (GNT1078399). S.J.H.-C. is supported by the J.N. Peters Bequest Fellowship (University of Melbourne).

Conflict of interest
None declared.

References
Overlap of endometrial vezatin and endometriosis risk

Endometriosis in Montpellier, France.

Research: workshop report from the 2011 World Congress of Endometriosis and Endometriosis Risk Replication Datasets.

Insights from meta-analysis of eight genome-wide association and population-based linkage analyses.

Hum Genet 2007;112:1750–1773.


Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, De Bakker PI, Daly MJ et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–575.

