Genome-wide enrichment analysis between endometriosis and obesity-related traits reveals novel susceptibility loci

Nilufer Rahmioglu1, Stuart Macgregor2, Alexander W. Drong1, Åsa K. Hedman1,5, Holly R. Harris6,7, Joshua C. Randell8, Inga Prokopenko1,9,10, The International Endogene Consortium (IEC), The GIANT Consortium, Dale R. Nyholt3, Andrew P. Morris1,11, Grant W. Montgomery4,†, Stacey A. Missmer6,†, Cecilia M. Lindgren1,12,† and Krina T. Zondervan1,13,∗,†

1Wellcome Trust Center for Human Genetics, University of Oxford, Oxford OX3 7BN, UK, 2Statistical Genetics, 3Neurogenetics, 4Molecular Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4029, Australia, 5Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden, 6Department of Obstetrics, Gynecology and Reproductive Biology, Brigham and Women’s Hospital and Harvard Medical School, 75 Francis Street, Boston, MA 02115, USA, 7Unit of Nutritional Epidemiology, Institute for Environmental Medicine, Karolinska Institutet, PO Box 210, SE-171 77 Stockholm, Sweden, 8Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK, 9Department of Genomics of Common Disease, Imperial College London, London W12 0NN, UK, 10Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford OX3 7LJ, UK, 11Department of Biostatistics, University of Liverpool, Duncan Building, Daulby Street, Liverpool L69 3GA, UK, 12Broad Institute of the Massachusetts Institute of Technology and Harvard University, Cambridge 02142 MA, USA and 13Nuffield Department of Obstetrics and Gynaecology & Endometriosis CaRe Centre, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, UK

Received April 4, 2014; Revised and Accepted October 6, 2014

Endometriosis is a chronic inflammatory condition in women that results in pelvic pain and subfertility, and has been associated with decreased body mass index (BMI). Genetic variants contributing to the heritable component have started to emerge from genome-wide association studies (GWAS), although the majority remain unknown. Unexpectedly, we observed an intergenic locus on 7p15.2 that was genome-wide significantly associated with both endometriosis and fat distribution (waist-to-hip ratio adjusted for BMI; WHRadjBMI) in an independent meta-GWAS of European ancestry individuals. This led us to investigate the potential overlap in genetic variants underlying the aetiology of endometriosis, WHRadjBMI and BMI using GWAS data. Our analyses demonstrated significant enrichment of common variants between fat distribution and endometriosis \((P = 3.7 \times 10^{-3})\), which was stronger when we restricted the investigation to more severe (Stage B) cases \((P = 4.5 \times 10^{-4})\). However, no genetic enrichment was observed between endometriosis and BMI \((P = 0.79)\). In addition to 7p15.2, we identify four more variants with statistically significant evidence of involvement in both endometriosis and WHRadjBMI \((\text{in/near KIFAP3, CAB39L, WNT4, GRB14})\); two of these, KIFAP3 and CAB39L, are novel associations for both traits. KIFAP3, WNT4 and 7p15.2 are associated with the WNT signalling pathway; formal pathway analysis confirmed a statistically significant \((P = 6.41 \times 10^{-5})\) overrepresentation of shared associations in developmental processes/WNT signalling between the two traits. Our results

∗To whom correspondence should be addressed at: Wellcome Trust Centre for Human Genetics/Nuffield Department of Obstetrics & Gynaecology, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK. Tel: +44 1865 287627; Email: krinaz@well.ox.ac.uk
†These authors jointly directed this work.

© The Author 2014. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
INTRODUCTION

Endometriosis is a common condition in premenopausal women characterized by chronic pelvic inflammation causing pain and subfertility (1), and has an estimated heritability of 51% (2). The International Endogene Consortium (IEC) performed the largest endometriosis GWAS to date in 3194 surgically confirmed cases (including 1364 moderate–severe—Stage B—cases) and 7060 controls of European ancestry, with replication in a further 2392 cases and 2271 controls (3). One genome-wide significant locus was observed in an intergenic region on chromosome 7p15.2 (rs12700667), primarily associated with Stage B disease \( (P = 1.5 \times 10^{-5}, OR = 1.38, 95\% CI 1.24–1.53) \). A second locus near WNT4 (rs7521902) was found after meta-analysis with published results from a Japanese GWAS of 1423 cases and 1318 controls (4); a genome-wide meta-analysis confirmed the two loci and found a further five (5).

Rs12700667 on 7p15.2 also marked 1 of 16 reported genome-wide significant loci associated with waist-to-hip ratio adjusted for BMI (WHRadjBMI) in an independent GWAS meta-analysis by the GIANT Consortium involving 77 167 individuals of European ancestry with replication in a further 113 636 individuals (rs1055144: discovery \( P = 1.5 \times 10^{-8} \); meta-analysis \( P = 1.0 \times 10^{-24}; r^2 = 0.5 \) with rs12700667 in 1000G pilot CEU data) (6,7). This was surprising, as prospective epidemiological studies have suggested consistently that reduced BMI—a measure of overall adiposity—is associated with increased risk of endometriosis, but there is relatively limited evidence for an association with WHRadjBMI—a measure of fat distribution (8,9). We conducted a logistic regression analysis in the IEC dataset of rs1055144 on endometriosis disease status, conditioning on rs12700667, which demonstrated that the SNPs reflected the same association signal (unpublished data; conditional \( P = 0.65 \)).

The epidemiological evidence of an association between endometriosis and BMI, together with the observed GWAS locus in common between endometriosis and WHRadjBMI, led us to conduct a systematic investigation of overlap in association signals between the IEC endometriosis GWAS and GIANT Consortium WHRadjBMI (\( N = 77 167 \)) (6,7) and BMI (\( N = 123 865 \)) (7,10) meta-GWAS datasets through genetic enrichment analyses.

RESULTS

Genetic enrichment analysis of endometriosis with overall adiposity and fat distribution

Using independent, imputed (1000 Genomes pilot reference panel) GWAS datasets of endometriosis (IEC; 3194 cases including 1364 Stage B cases, 7060 controls), BMI (GIANT; 123 865 individuals) and WHRadjBMI (GIANT: 77 167 individuals), we first considered loci genome-wide significantly associated with endometriosis, BMI or WHRadjBMI in each of the individual GWAS. The two genome-wide significant endometriosis loci (intergenic 7p15.2 and WNT4) had significantly lower \( P \)-values of association than expected by chance in the WHRadjBMI GWAS (Table 1: \( rs12700667, P = 4.4 \times 10^{-5} \) and \( rs7521902, P = 1.3 \times 10^{-3} \); binomial \( P = 1.0 \times 10^{-4} \)), while 2 of the 16 genome-wide significant WHRadjBMI loci (intergenic 7p15.2 and GRB14) had \( P < 0.01 \) in the endometriosis GWAS (binomial \( P = 0.011 \)). No enrichment between genome-wide significantly associated loci was observed for endometriosis versus BMI (Supplementary Material, Table S1: \( rs12700667, P = 0.27 \) and \( rs7521902, P = 0.92 \)).

To investigate whether statistical enrichment extended beyond genome-wide significant loci, we investigated the most significant \( (P < 1 \times 10^{-3}) \) independent \( (r^2 < 0.2) \) endometriosis GWAS signals for enrichment of WHRadjBMI or BMI signals with \( P < 0.05 \) and vice versa (number of lookup SNPs per dataset: \( n = 717 \) to 748; see Supplementary Material, Methods). We observed statistically significant enrichment between variants associated with endometriosis (particularly Stage B) and WHRadjBMI (all endometriosis versus WHRadjBMI: \( P = 3.7 \times 10^{-3} \); Stage B endometriosis versus WHRadjBMI: \( P = 4.5 \times 10^{-4} \)), but not between endometriosis and BMI (all endometriosis versus BMI: \( P = 0.79 \); Stage B endometriosis versus BMI: \( P = 0.85 \)) (Fig. 1; Supplementary Material, Table S2). Results were similar when using female-limited WHRadjBMI (\( N = 42 969 \) women) and BMI (\( N = 73 137 \) women) GWAS summary statistics (7); to optimize power, in the remainder of the paper we therefore focus on sex-combined WHRadjBMI and BMI datasets (Supplementary Material, Fig. S1). Empirical testing of statistical enrichment through permutation (see Supplementary Material, Methods) provided near-identical results (Fig. 1; Supplementary Material, Fig. S1).

The choice of significance thresholds in the discovery and lookup datasets was based on a balance between applying a sufficiently stringent significance threshold in the discovery dataset that would minimize the proportion of false-positive association signals, while still having sufficient numbers of loci in each of the phenotypic association strata to investigate statistical enrichment, and allow the pursuit of meaningful biological pathway analyses subsequently. We considered the effect of different significance thresholds, for both discovery and lookup, which confirmed results showing enrichment of association signals between endometriosis and WHRadjBMI (Supplementary Material, Table S3), but no enrichment between endometriosis and BMI (Supplementary Material, Table S4).

To investigate potential genome-wide sharing of loci between endometriosis and WHRadjBMI or BMI, we performed polygenic prediction analyses (11) evaluating whether the aggregate effect of many variants of small effect in the WHRadjBMI and BMI GWAS could predict endometriosis status in the IEC GWAS (see Supplementary Material, Methods). There was no significant association between the WHRadjBMI- or BMI-
Table 1. Association results of published IEC genome-wide significant endometriosis loci (3) in the GIANT WHRadjBMI GWAS, and of WHRadjBMI loci (6,7) in endometriosis GWAS (lookup results are shown in bold).

<table>
<thead>
<tr>
<th>GWAS</th>
<th>SNP (proxy; r²)</th>
<th>Ch</th>
<th>Location (B36)</th>
<th>RAF (allele)</th>
<th>Status</th>
<th>Endometriosis all cases P-value</th>
<th>OR (95% CI)</th>
<th>Endometriosis Stage B only P-value</th>
<th>OR (95% CI)</th>
<th>Overall WHRadjBMI P-value</th>
<th>WHRadjBMI Effect (SE)</th>
<th>Female-limited WHRadjBMI Effect (SE)</th>
<th>Nearest gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endometriosis</td>
<td>rs12700667</td>
<td>7</td>
<td>25 868 164</td>
<td>0.74 (A)</td>
<td>G</td>
<td>5.1 × 10⁻⁷</td>
<td>1.21 (1.12–1.31)</td>
<td>3.3 × 10⁻⁴</td>
<td>1.36 (1.23–1.50)</td>
<td>4.4 × 10⁻⁵</td>
<td>-0.023 (0.005)</td>
<td>3.3 × 10⁻⁸</td>
<td>Intergenic</td>
</tr>
<tr>
<td>Endometriosis</td>
<td>rs7521902</td>
<td>1</td>
<td>22 363 311</td>
<td>0.25 (A)</td>
<td>G</td>
<td>8.9 × 10⁻⁵</td>
<td>1.16 (1.08–1.25)</td>
<td>7.5 × 10⁻⁴</td>
<td>1.26 (1.14–1.39)</td>
<td>1.3 × 10⁻³</td>
<td>-0.020 (0.006)</td>
<td>6.1 × 10⁻⁶</td>
<td>Intergenic</td>
</tr>
<tr>
<td>WHRadjBMI</td>
<td>rs1055144</td>
<td>7</td>
<td>25 837 634</td>
<td>0.19 (T)</td>
<td>G</td>
<td>3.7 × 10⁻⁵</td>
<td>0.84 (0.77–0.91)</td>
<td>3.1 × 10⁻⁴</td>
<td>0.78 (0.70–0.88)</td>
<td>1.5 × 10⁻⁸</td>
<td>0.034 (0.006)</td>
<td>2.3 × 10⁻¹⁵</td>
<td>Intergenic</td>
</tr>
<tr>
<td>WHRadjBMI</td>
<td>rs10195252</td>
<td>2</td>
<td>165 221 337</td>
<td>0.41 (C)</td>
<td>G</td>
<td>9.8 × 10⁻⁷</td>
<td>0.92 (0.85–0.98)</td>
<td>0.56</td>
<td>0.92 (0.84–1.00)</td>
<td>3.2 × 10⁻¹⁰</td>
<td>-0.031 (0.005)</td>
<td>6.3 × 10⁻¹⁵</td>
<td>Intergenic</td>
</tr>
<tr>
<td>Female WHRadjBMI</td>
<td>rs4648845</td>
<td>3</td>
<td>12 463 882</td>
<td>0.43 (C)</td>
<td>1 (0.98)</td>
<td>0.07</td>
<td>1.06 (0.99–1.14)</td>
<td>0.14</td>
<td>1.07 (0.98–1.17)</td>
<td>1.0 × 10⁻⁴</td>
<td>0.019 (0.005)</td>
<td>2.3 × 10⁻⁸</td>
<td>Intergenic</td>
</tr>
<tr>
<td>WHRadjBMI</td>
<td>rs718314</td>
<td>12</td>
<td>26 344 530</td>
<td>0.24 (G)</td>
<td>G</td>
<td>0.11</td>
<td>1.06 (0.99–1.15)</td>
<td>0.054</td>
<td>1.10 (0.99–1.22)</td>
<td>2.4 × 10⁻⁸</td>
<td>0.031 (0.005)</td>
<td>8.2 × 10⁻¹⁰</td>
<td>Intergenic</td>
</tr>
<tr>
<td>WHRadjBMI</td>
<td>rs6861681</td>
<td>5</td>
<td>173 362 458</td>
<td>0.32 (A)</td>
<td>1 (0.96)</td>
<td>0.15</td>
<td>0.95 (0.86–1.04)</td>
<td>0.11</td>
<td>0.93 (0.85–1.00)</td>
<td>1.4 × 10⁻⁶</td>
<td>0.026 (0.005)</td>
<td>2.1 × 10⁻⁹</td>
<td>Intergenic</td>
</tr>
<tr>
<td>WHRadjBMI</td>
<td>rs6795735</td>
<td>3</td>
<td>64 680 405</td>
<td>0.41 (T)</td>
<td>G</td>
<td>0.21</td>
<td>1.04 (0.98–1.12)</td>
<td>0.32</td>
<td>1.04 (0.96–1.14)</td>
<td>2.5 × 10⁻⁵</td>
<td>-0.025 (0.005)</td>
<td>7.8 × 10⁻¹⁷</td>
<td>Intergenic</td>
</tr>
<tr>
<td>WHRadjBMI</td>
<td>rs2820446</td>
<td>1</td>
<td>21 974 881</td>
<td>0.71 (C)</td>
<td>1 (0.99)</td>
<td>0.31</td>
<td>1.04 (0.97–1.12)</td>
<td>0.22</td>
<td>1.06 (0.97–1.17)</td>
<td>5.1 × 10⁻¹²</td>
<td>0.037 (0.005)</td>
<td>8.5 × 10⁻¹⁸</td>
<td>Intergenic</td>
</tr>
</tbody>
</table>

*Table legend:*
- SNP: Single Nucleotide Polymorphism
- Ch: Chromosome
- Location: Base Pair Reference
- RAF: Reference Allele
- Status: Genetic Variation
- Endometriosis all cases: P-value
- OR: Odds Ratio
- Endometriosis Stage B only: P-value
- OR: Odds Ratio
- Overall WHRadjBMI: P-value
- WHRadjBMI Effect: Effect Size
- Female-limited WHRadjBMI: Effect Size
- Nearest gene:

1) Logistic regression analysis in the IEC GWAS shows that rs1055144 marks the same locus as rs12700667 (conditional P = 0.65; r² = 0.8).
2) SNP was not genotyped in the endometriosis GWAS dataset; result shown is of proxy SNP.
3) Results are based on an updated GWAS performed using genotype data imputed up to 1000 Genomes pilot reference panel (B36, June 2010).
4) Results are from the GIANT WHRadjBMI discovery GWAS dataset (N = 77 167); 3 of the 14 WHRadjBMI loci have P > 5.0 × 10⁻⁸, however, they reached genome-wide significance combined with replication analyses in up to a further 113 636 individuals (6).
5) Results from the GIANT WHRadjBMI discovery female-limited GWAS dataset (N = 42 969); one of the two female-limited WHRadjBMI loci have P > 5.0 × 10⁻⁸, however, they reached genome-wide significance combined with replication analyses in up to a further 71 295 individuals (7).
derived profile scores (overall or female limited) and all/Stage B endometriosis (Supplementary Material, Tables S5–S8), suggesting no evidence for a directionally consistent en masse, genome-wide, shared common genetic component.

We next investigated the variants with most significant evidence for association with both endometriosis ($P < 0.05$) and WHRadjBMI ($P < 0.05$) for predominance in direction of phenotypic effects (Supplementary Material, Tables S9 and S10 and Fig. S2). No statistically significant directional consistency was observed for these variants ($P > 0.47$), nor for the 17 loci (Table 1) that were genome-wide significantly associated with either trait (Fig. 2, $P > 0.44$). Intergenic 7p15.2 and WNT4 showed discordant directions of effect, while the effect of GRB14 was concordant (Fig. 2). This could suggest the presence of multiple biological pathways through which the variants influence the two phenotypes. We next set out to investigate the common biology suggested by genetic variants associated with both endometriosis and WHRadjBMI.

Figure 1. Genetic enrichment analyses between endometriosis, BMI and WHRadjBMI GWAS datasets, using independent ($r^2 < 0.2$) SNPs. The panels show (i) The proportion of SNPs nominally associated ($P < 0.05$) with WHRadjBMI (A) or BMI (B) by significance of overall and Stage B endometriosis association ($P < 1 \times 10^{-3}$ versus $P \geq 1 \times 10^{-3}$); (ii) The proportion of SNPs nominally associated ($P < 0.05$) with overall and Stage B endometriosis by significance of WHRadjBMI (C) and BMI (D) association ($P < 1 \times 10^{-3}$ versus $P \geq 1 \times 10^{-3}$). P-values of $\chi^2$ tests assessing statistical difference between proportions are shown above each set of bars, and 95% confidence intervals of the proportions are given on each bar. For differences with $P_{\text{chiq}} < 0.2$, empirical P-values are given in brackets (see Supplementary Material, Methods).
Biology of the loci shared between endometriosis and fat distribution

Our analysis showing significant enrichment between SNPs associated with all or Stage B endometriosis \((P < 1 \times 10^{-3})\) and WHRadjBMI \((P < 0.05)\) shown in Figure 1 involved 1284 independent \((r^2 > 0.2)\) loci. We explored the biological function of the loci most strongly associated with WHRadjBMI, at nominal \(P < 0.005\) \((n = 16, \text{Table 2}; \text{see Supplementary Material})\), Tables S11 and S12 for all variants associated at \(P < 0.05\). Two novel loci, rs560584 near KIFAP3 (all endometriosis) and rs11619804 in CAB39L (Stage B endometriosis), were significantly associated with WHRadjBMI after Bonferroni correction allowing for 1284 independent tests \((P < 3.89 \times 10^{-5})\).

The endometriosis risk allele T of rs560584 \((OR = 1.14 (1.07–1.22), P = 1.42 \times 10^{-5})\) was associated with lower WHRadjBMI \((\beta = -0.021, P = 1.47 \times 10^{-5})\), and located in an intergenic region 46 kb downstream of KIFAP3 \((\text{Kinesin-associated protein 3})\). Together with KIF3A and KIF3B, KIFAP3 forms a kinesin motor complex, KIF3, that mediates cellular transport of N-cadherin and \(\beta\)-catenins \((12)\), which are involved in cell adhesion, the Wnt canonical pathway and cell cycle progression \((13)\). The Wnt/\(\beta\)-catenin signalling pathway acts as a molecular switch for adipogenesis \((14)\) and has multiple suggested roles in endometriosis through sex hormone homeostasis regulation \((15)\), its role in development of female reproductive organs \((16)\), molecular mechanisms of infertility \((17)\) and mediation of fibrogenesis \((18)\).

The Stage B endometriosis risk allele C of rs11619804 \((OR = 1.17 (1.07–1.28), P = 4.88 \times 10^{-4})\), located in CAB39L \((\text{Calcium-Binding Protein 39-Like})\), was associated with increased WHRadjBMI \((\beta = 0.022, P = 1.06 \times 10^{-5})\; \text{Table 2}\). The function of this gene is not well characterized but the encoded protein interacts with a serine threonine kinase \((\text{STK11})\) that functions as a tumour suppressor \((19)\).

Rs12700667 in the intergenic region 7p15.2 remained among the strongest associated shared signals, with the endometriosis risk allele A associated with reduced WHRadjBMI \((\beta = -0.023, P = 4.4 \times 10^{-5})\). The association maps to an intergenic high LD region of 48 kb \((r^2 > 0.8)\) of unknown functionality.

Further interesting nearby loci include the miRNA hsa-mir-148a, with a purported role in Wnt/\(\beta\)-catenin signalling \((14)\); NFE2L3 \((\text{nuclear factor (erythroid-derived 2)-like 3})\), a transcription factor suggested to be involved in cell differentiation, inflammation and carcinogenesis \((20)\). The WNT signalling pathway was further highlighted by the nominal association of two independent \((r^2 = 0.06)\) endometriosis risk variants near WNT4 \((\text{wingless-type MMTV integration site family})\), rs3820282 \((\text{genic})\) and rs2807357 \((22.4 \text{~kb downstream})\), with reduced WHRadjBMI \((\beta = -0.019, P = 5.0 \times 10^{-3}; \beta = -0.015, P = 3.7 \times 10^{-3})\; \text{Table 2}\). Of note is that all shared variants implicated in WNT signalling \((\text{in/near intergenic} \ 7\text{p15.2, WNT4, KIFAP3})\) showed consistent—discordant—phenotypic directions of effect.

Risk variant rs10195252, 34.6 kb downstream of GRB14 \((\text{growth factor receptor-bound protein 14})\) was the third locus with significant evidence for association with both overall (not Stage B) endometriosis and WHRadjBMI \((\text{Table 1})\). GRB14 has an inhibitory effect on insulin receptor signalling \((21)\), may have a role in signalling pathways that regulate growth and metabolism and has been shown to interact with fibroblast growth factor receptors \((22)\). This shared variant is also in high LD \((r^2 = 0.93 \text{~and} = 0.87, \text{respectively})\) with a type 2 diabetes risk variant rs13389219 \((23)\) and fasting insulin risk variant rs6717858 \((24)\).

Other loci of interest include rs2921188 in PPARG \((\text{peroxisome proliferator-activated receptor gamma})\) which is a nuclear hormone receptor that regulates fatty acid storage and glucose metabolism \((18)\). Synthetic ligands, such as insulin sensitizing drugs, target PPARG in treatment of diabetes to lower serum glucose levels \((25)\) and are also documented to have anti-inflammatory, anti-angiogenic and anti-proliferative effects on endometrium, with baboon models suggesting a role in targeting endometriotic disease \((26)\). Stage B endometriosis risk allele G of rs6556301 near FGFR4 \((\text{fibroblast growth factor receptor})\), OR \(= 1.17 (1.07–1.28), P = 7.4 \times 10^{-4}\) is associated with reduced WHRadjBMI \((\beta = -0.021, P = 1.9 \times 10^{-4})\). FGFR4 interacts
Table 2. Results of the top all/Stage B endometriosis loci ($P < 1 \times 10^{-5}$) associated with WHRadjBMI at $P < 0.005$

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chr</th>
<th>Position (B36)</th>
<th>RAF (allele)</th>
<th>Endometriosis $P$-value OR (95% CI)</th>
<th>Overall WHRadjBMI $P$-value Effect</th>
<th>SE</th>
<th>Female-limited WHRadjBMI $P$-value Effect</th>
<th>SE</th>
<th>Nearest loci (distance)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All cases</td>
<td>1</td>
<td>168 357 136</td>
<td>0.41 (T)</td>
<td>$1.4 \times 10^{-4}$ 1.14 (1.07–1.22)</td>
<td>$1.4 \times 10^{-3}$ -0.021</td>
<td>0.005</td>
<td>$1.1 \times 10^{-3}$ -0.022</td>
<td>0.677</td>
<td>KIFAP3 (46 632)</td>
</tr>
<tr>
<td>rs560584</td>
<td>7</td>
<td>25 868 164</td>
<td>0.74 (A)</td>
<td>$5.1 \times 10^{-7}$ 1.22 (1.13–1.32)</td>
<td>$4.4 \times 10^{-5}$ -0.023</td>
<td>0.005</td>
<td>$3.4 \times 10^{-4}$ -0.028</td>
<td>0.284</td>
<td>NF2L3 (290 221)</td>
</tr>
<tr>
<td>rs2921188</td>
<td>3</td>
<td>12 387 115</td>
<td>0.64 (A)</td>
<td>$5.9 \times 10^{-4}$ 1.13 (1.05–1.21)</td>
<td>$1.1 \times 10^{-3}$ -0.017</td>
<td>0.005</td>
<td>$1.8 \times 10^{-4}$ -0.026</td>
<td>0.054</td>
<td>PPARG (0)</td>
</tr>
<tr>
<td>rs1250248</td>
<td>2</td>
<td>215 995 338</td>
<td>0.27 (A)</td>
<td>$1.6 \times 10^{-5}$ 1.17 (1.09–1.26)</td>
<td>$1.0 \times 10^{-3}$ 0.018</td>
<td>0.005</td>
<td>9.9 $\times 10^{-4}$ 0.025</td>
<td>0.242</td>
<td>FN1 (0)</td>
</tr>
<tr>
<td>rs2630787</td>
<td>3</td>
<td>21 847 339</td>
<td>0.52 (C)</td>
<td>$9.2 \times 10^{-4}$ 1.12 (1.05–1.19)</td>
<td>$1.9 \times 10^{-3}$ -0.015</td>
<td>0.004</td>
<td>0.38</td>
<td>-0.006</td>
<td>ZNF569 (79 518)</td>
</tr>
<tr>
<td>rs1430788</td>
<td>2</td>
<td>67 721 916</td>
<td>0.31 (C)</td>
<td>$9.3 \times 10^{-5}$ 1.15 (1.07–1.23)</td>
<td>$2.7 \times 10^{-3}$ 0.016</td>
<td>0.005</td>
<td>3.1 $\times 10^{-3}$ 0.022</td>
<td>0.330</td>
<td>ETA1A (230 878)</td>
</tr>
<tr>
<td>rs906721</td>
<td>3</td>
<td>184 687 691</td>
<td>0.41 (A)</td>
<td>$6.1 \times 10^{-5}$ 1.16 (1.08–1.24)</td>
<td>$4.2 \times 10^{-2}$ 0.015</td>
<td>0.005</td>
<td>1.7 $\times 10^{-3}$ 0.023</td>
<td>0.140</td>
<td>KLHL6 (322)</td>
</tr>
<tr>
<td>rs1868894</td>
<td>4</td>
<td>187 606 728</td>
<td>0.80 (C)</td>
<td>$2.3 \times 10^{-4}$ 1.16 (1.07–1.26)</td>
<td>$4.9 \times 10^{-2}$ -0.018</td>
<td>0.006</td>
<td>0.13</td>
<td>-0.013</td>
<td>0.524</td>
</tr>
<tr>
<td>rs3820282</td>
<td>1</td>
<td>22 340 802</td>
<td>0.16 (T)</td>
<td>$3.3 \times 10^{-7}$ 1.26 (1.15–1.37)</td>
<td>$5.0 \times 10^{-3}$ -0.019</td>
<td>0.007</td>
<td>0.09</td>
<td>-0.016</td>
<td>0.749</td>
</tr>
<tr>
<td>Stage B cases</td>
<td>13</td>
<td>49 888 131</td>
<td>0.53 (C)</td>
<td>$4.8 \times 10^{-4}$ 1.17 (1.07–1.28)</td>
<td>$1.1 \times 10^{-5}$ 0.022</td>
<td>0.005</td>
<td>2.2 $\times 10^{-2}$ 0.016</td>
<td>0.022</td>
<td>CAB39L (0)</td>
</tr>
<tr>
<td>rs1619804</td>
<td>7</td>
<td>25 868 164</td>
<td>0.74 (A)</td>
<td>$3.3 \times 10^{-9}$ 1.36 (1.23–1.50)</td>
<td>$4.4 \times 10^{-2}$ -0.023</td>
<td>0.005</td>
<td>3.4 $\times 10^{-4}$ -0.028</td>
<td>0.284</td>
<td>NF2L3 (290 221)</td>
</tr>
<tr>
<td>rs2782659</td>
<td>6</td>
<td>45 794 768</td>
<td>0.33 (G)</td>
<td>$4.2 \times 10^{-4}$ 1.18 (1.08–1.30)</td>
<td>$9.2 \times 10^{-5}$ 0.020</td>
<td>0.005</td>
<td>1.7 $\times 10^{-4}$ 0.027</td>
<td>0.108</td>
<td>RUNX2 (167 970)</td>
</tr>
<tr>
<td>rs6556301</td>
<td>5</td>
<td>176 460 183</td>
<td>0.63 (G)</td>
<td>$7.4 \times 10^{-4}$ 1.17 (1.07–1.28)</td>
<td>$1.9 \times 10^{-4}$ -0.021</td>
<td>0.005</td>
<td>7.8 $\times 10^{-3}$ -0.021</td>
<td>0.845</td>
<td>FGFR4 (2450)</td>
</tr>
<tr>
<td>rs1250248</td>
<td>2</td>
<td>215 995 338</td>
<td>0.27 (A)</td>
<td>$2.9 \times 10^{-4}$ 1.32 (1.19–1.45)</td>
<td>$1.2 \times 10^{-3}$ 0.018</td>
<td>0.005</td>
<td>9.9 $\times 10^{-4}$ 0.025</td>
<td>0.242</td>
<td>FN1 (0)</td>
</tr>
<tr>
<td>rs4318206</td>
<td>1</td>
<td>161 662 648</td>
<td>0.85 (T)</td>
<td>$5.4 \times 10^{-5}$ 1.24 (1.10–1.41)</td>
<td>$1.5 \times 10^{-3}$ 0.022</td>
<td>0.007</td>
<td>0.25</td>
<td>0.011</td>
<td>0.072</td>
</tr>
<tr>
<td>rs9912335</td>
<td>17</td>
<td>77 552 948</td>
<td>0.69 (T)</td>
<td>$3.1 \times 10^{-4}$ 1.19 (1.08–1.31)</td>
<td>$3.5 \times 10^{-3}$ -0.021</td>
<td>0.007</td>
<td>0.10</td>
<td>-0.016</td>
<td>0.454</td>
</tr>
<tr>
<td>rs10878362</td>
<td>12</td>
<td>64 703 760</td>
<td>0.69 (C)</td>
<td>$4.9 \times 10^{-4}$ 1.19 (1.08–1.31)</td>
<td>$3.6 \times 10^{-3}$ 0.015</td>
<td>0.005</td>
<td>3.1 $\times 10^{-3}$ 0.022</td>
<td>0.204</td>
<td>HMCGA2 (57 421)</td>
</tr>
<tr>
<td>rs2807357</td>
<td>1</td>
<td>22 364 571</td>
<td>0.64 (A)</td>
<td>$9.7 \times 10^{-4}$ 1.16 (1.06–1.27)</td>
<td>$3.7 \times 10^{-3}$ -0.015</td>
<td>0.005</td>
<td>1.0 $\times 10^{-3}$ -0.024</td>
<td>0.081</td>
<td>WNT4 (22 373)</td>
</tr>
<tr>
<td>rs906721</td>
<td>3</td>
<td>184 687 691</td>
<td>0.41 (A)</td>
<td>$1.4 \times 10^{-4}$ 1.20 (1.09–1.32)</td>
<td>$4.2 \times 10^{-3}$ 0.015</td>
<td>0.005</td>
<td>1.7 $\times 10^{-3}$ 0.023</td>
<td>0.140</td>
<td>KLHL6 (322)</td>
</tr>
<tr>
<td>rs12267660</td>
<td>10</td>
<td>4 419 530</td>
<td>0.85 (G)</td>
<td>$7.9 \times 10^{-4}$ 1.24 (1.09–1.40)</td>
<td>$4.6 \times 10^{-3}$ 0.021</td>
<td>0.007</td>
<td>8.0 $\times 10^{-3}$ 0.030</td>
<td>0.133</td>
<td>CR749391 (191 913)</td>
</tr>
<tr>
<td>rs1105481</td>
<td>2</td>
<td>67 590 253</td>
<td>0.15 (C)</td>
<td>$8.4 \times 10^{-4}$ 1.23 (1.09–1.38)</td>
<td>$4.8 \times 10^{-3}$ 0.018</td>
<td>0.006</td>
<td>1.1 $\times 10^{-2}$ 0.022</td>
<td>0.451</td>
<td>ETA1A (99 215)</td>
</tr>
</tbody>
</table>
with fibroblast growth factors, which have roles in angiogenesis, wound healing and cell migration (27).

**Expression quantitative trait loci analysis of the shared endometriosis and fat distribution loci**

We investigated the potential impact of the described 16 genes (Table 2) shared between endometriosis and WHRadjBMI on transcriptional function using three public expression data resources; (i) the Mammalian Gene Expression Uterus database (MGEx-Udb) (28) containing published information on transcriptional activity of specific genes in human endometrial tissue from individuals with and without endometriosis; (ii) the MuTHER study which collected expression and eQTL data from 776 abdominal fat tissues (29); and (iii) the MOLOBB dataset of differential expression levels between abdominal and gluteal fat from 49 individuals (30). Based on the limited available evidence in the MGEx-Udb database, two genes are transcribed in endometrial tissue of women with endometriosis but dormant in those without endometriosis: PPARG and FGFR4 (Supplementary Material, Table S13). Of the 16 genes, 15 had probes present within 1 Mb either side of the SNP in the MuTHER database; however, none showed significant association with nearby transcripts in abdominal fat tissue (Supplementary Material, Table S14). The MOLOBB study data showed cis-eQTL evidence for differential expression of two genes; KIFAP3 (rs560584; fold change = 0.14, adjusted \( P = 0.04 \)) (Supplementary Material, Table S15). Additional transcriptional evidence relevant to the intergenic 7p15.2 locus includes the presence of an expression QTL associated with a transcript of unknown function, AAS53656, in subcutaneous abdominal fat tissue (6), and the differential expression of nearby hsa-miR-148a between gluteal and abdominal fat tissue samples (31).

**Pathway analysis**

To identify potential common biological pathways involved in the aetiology of endometriosis and the variability of fat distribution, we conducted pathway analyses using genes with evidence for enrichment between the traits using (i) the PANTHER database (32) and (ii) GRAIL (33). For the PANTHER analysis, we selected the 91 and 108 genes located in a 1 Mb interval surrounding each independent SNP associated with all endometriosis \((P < 1.0 \times 10^{-5})\) and WHRadjBMI \((P < 0.05)\), and Stage B endometriosis \((P < 1.0 \times 10^{-5})\) and WHRadjBMI \((P < 0.05)\), respectively (see Supplementary Material, Methods). This excluded intergenic loci without a gene within 1 Mb, such as our top shared locus at 7p15.2. We tested whether the two sets of genes showed significant overrepresentation of a particular pathway, for each of 176 curated pathways and 241 biological processes. The top enriched pathways were ‘developmental processes’ (all endometriosis: \(P = 1.2 \times 10^{-5}\); Stage B: \(P = 1.25 \times 10^{-4}\)), ‘WNT signalling’ (all endometriosis: \(P = 1.07 \times 10^{-4}\)), ‘gonadotropin-releasing hormone receptor’ (all endometriosis: \(P = 1.48 \times 10^{-5}\)), ‘cadherin signalling’ (Stage B: \(P = 6.42 \times 10^{-4}\), ‘FGF signalling’ (Stage B: \(P = 2.96 \times 10^{-3}\) and ‘TGF-beta signalling’ (Stage B: \(P = 1.48 \times 10^{-4}\) pathways (Supplementary Material, Tables S16 and S17). Bonferroni correction for the number of pathways tested (see Supplementary Material, Methods) rendered ‘WNT signalling’, ‘developmental processes’, ‘cellular processes’ and ‘cell communication’ significantly enriched; however, this adjustment is conservative, as exemplified by ‘cadherin signalling’ genes being a subset of those in the ‘WNT signalling’ pathway. Sensitivity analyses exploring the effect of different endometriosis association thresholds on pathway analyses showed very consistent results for threshold \(P < 1.0 \times 10^{-2}\), with the same top three enriched pathways—‘WNT signalling’, Cadherin signalling and Gonadotropin-releasing hormone receptor pathway. No meaningful pathway analyses could be conducted on the limited number of genes passing association threshold \(P < 1 \times 10^{-4}\) (Supplementary Material, Table S18).

We used GRAIL (33) to search for connectivity between the 91 and 108 genes all/Stage B endometriosis and WHRadjBMI-associated genes and specific keywords from the published literature that describe potential functional connections. We identified 17 genes with nominal significance \((P < 0.05)\) for potential functional connectivity for ‘all’ endometriosis and WHRadjBMI and six genes for Stage B endometriosis and WHRadjBMI (Supplementary Material, Fig. S3 and Tables S19 and S20). The keywords associated with these connections included ‘cadherin’, ‘differentiation’, ‘development’ and ‘insulin’ for ‘all’ endo, and ‘development’ and ‘embryos’ for Stage B endometriosis, marking again developmental processes and cadherin signalling as biological pathways shared in the origins of endometriosis and fat distribution.

**DISCUSSION**

In this study, we have investigated the overlap in genetic association signals from the largest GWA studies to date of endometriosis, overall adiposity (BMI) and fat distribution (WHRadjBMI). Our results demonstrated that there is a shared genetic basis between endometriosis and fat distribution that extends over and above the single genome-wide significant locus that has been reported in GWAS of the separate traits. Our analyses highlight novel loci in/near KIFAP3 and CAB39L, which together with intergenic 7p15.2, WNT4 and GRB14, showed significant evidence of trait association sharing. The strength of evidence of enrichment was similar for overall versus female-limited WHRadjBMI loci, which may be unexpected, given that endometriosis is a female condition. However, the lack of a stronger enrichment between female-specific WHRadjBMI GWAS results and endometriosis, compared with all WHRadjBMI results should be considered against the effects of a reduced sample size used for female-specific WHRadjBMI analyses on power of association detection.

The enrichment of associated variants was generally stronger when the endometriosis cases were restricted to moderate—severe (Stage B) disease, despite the smaller sample size. Indeed, the association of the top intergenic GWAS locus on 7p15.2, also genome-wide significantly associated with WHRadjBMI, is limited to Stage B endometriosis. Stage B—or ASRM Stages III/IV disease (34)—is typically characterized by ovarian (endometrioma) or deep infiltrating (rectovaginal) lesions, which were shown to have a substantially greater underlying genetic contribution than milder, peritoneal disease (ASRM Stage I/II) (3). The particular enrichment between
WHRRadjBMI and Stages III/IV endometriosis is intriguing, and another reason for further functional work to concentrate on this endometriosis sub-type. There are, however, specific loci that show enrichment of association with WHRRadjBMI and overall endometriosis, the analysis of which therefore remains of interest. An example is GRB14, which did not show significant association with Stage B disease, displayed a concordant direction of effect between endometriosis and WHRRadjBMI, and the biological function of which also seems to suggest an entirely different contribution to the origins of both phenotypes than the 7p15.2 and WNT4 loci.

The limited available eQTL data showed significant evidence for differential expression of KIFAP3 between different fat depots. The variants with most evidence for enrichment between the traits, in/near intergenic 7p15.2, WNT7A, and WNT4, were all implicated in WNT signalling and had consistent—discordant—directions of effect, with endometriosis risk alleles associated with a decreased WHRRadjBMI. Indeed, biological pathway analyses showed significant evidence for the involvement of developmental processes and WNT signalling in endometriosis aetiology and regulation of fat distribution, a potential pleiotropic connection that has not been reported to date.

The relatively limited epidemiological evidence of phenotypic correlation between endometriosis and WHRRadjBMI (8,9) is consistent with the absence of strong directional consistency of phenotypic effects of genetic variants underlying both traits at a genome-wide level. Most studies of genetic pleiotropy between traits to date have focused on genome-wide directional consistency between epidemiologically or clinically (postulated) correlated traits, such as different metabolic traits (6,35) or psychiatric conditions (36). However, genome-wide consistency in directionality of phenotypic effects would most likely apply to traits that share a large proportion of causality, and that epidemiologically lie on the same causal pathway(s) and are thus more likely to be examples of mediated (genetic variants influencing one phenotype indirectly through association with a second phenotype) rather than biological (genetic variants exerting a direct biological influence on more than one phenotype) pleiotropy (37). Thus, our results of genetic enrichment between endometriosis and WHRRadjBMI demonstrate an example of the biological complexity of aetiological associations between complex traits, and suggest that the underlying shared loci are potentially biologically pleiotropic, given the absence of phenotypic correlation between endometriosis and WHRRadjBMI and absence of en masse directional consistency of shared genetic variants on the phenotypes (37,38). It also demonstrates more generally how potential perturbation of a causal pathway through, for example, drug treatment targeting one trait could have unexpected effects on another, even when there is no clear evidence that these traits are associated clinically or epidemiologically—a problem often encountered in drug development. Systematic exploration of biological pleiotropy of genetic variants marking potential drug targets may help in highlighting the potential of such unwanted or unexpected effects.

While the observed genetic enrichment between endometriosis and WHRRadjBMI presents new avenues for exploring common biology, the total absence of any genetic enrichment between endometriosis and BMI (within the limits of power presented by these large datasets) is intriguing given the consistent, prospective, observational epidemiological evidence of phenotypic association between reduced BMI and endometriosis risk (8). Our analyses represent an adaptation of Mendelian randomization analyses (39,40), in which genetic variants robustly associated with BMI in the largest GWAS analyses to date (10) are investigated for association with endometriosis. The total lack of genetic enrichment suggests that reduced BMI is not causally related to endometriosis risk. Rather, it suggests that the observed phenotypic association (8) is either driven by shared environmental factors, or is due to confounding factors related to BMI affecting, for example, diagnostic opportunity for endometriosis.

These novel findings present an entirely new opportunity for functional targeted follow-up of pleiotropic loci between endometriosis and WHRRadjBMI in relevant disease tissues such as endometrium and fat tissue, cellular systems, animal models and further cross-trait comparisons, to uncover their biological functions and to assess how studies in the fat distribution research field can inform research into endometriosis pathogenesis, biomarker identification and drug target discovery and validation.

### MATERIALS AND METHODS

#### Genome-wide association studies

**IEC endometriosis GWAS**

This GWAS included 3194 surgically confirmed endometriosis cases and 7060 controls from Australia and the UK. Disease severity of the endometriosis cases was assessed retrospectively from surgical records using the rAFS classification system and grouped into two phenotypes: Stage A (Stage I or II disease or some ovarian disease with a few adhesions; n = 1686) or Stage B (Stage III or IV disease; n = 1364). We previously showed an increased genetic loading among 1364 cases with Stage B endometriosis compared with 1666 with Stage A disease (3), which led to two GWAS analyses, including (i) 3194 ‘all’ endometriosis case and (ii) 1364 Stage B cases (Table 3). The genotyped data were imputed up to 1000 Genomes pilot reference panel (B36, June 2010) and the GWAS was performed again, using a missing data likelihood in a logistic regression model including

<table>
<thead>
<tr>
<th>GWAS</th>
<th>Consortium</th>
<th>Sample size</th>
<th>No. of SNPs (million)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endometriosis—all cases</td>
<td>IEC</td>
<td>3194 cases, 7060 controls</td>
<td>~12.5</td>
<td>Painter et al. (3)</td>
</tr>
<tr>
<td>Endometriosis—Stage B cases</td>
<td>IEC</td>
<td>1363 cases, 7060 controls</td>
<td>~12.5</td>
<td>Painter et al. (3)</td>
</tr>
<tr>
<td>WHRRadjBMI</td>
<td>GIANT</td>
<td>77 167</td>
<td>~2.85</td>
<td>Heid et al. (6)</td>
</tr>
<tr>
<td>Female-limited WHRRadjBMI</td>
<td>GIANT</td>
<td>42 969</td>
<td>~2.85</td>
<td>Randall et al. (7)</td>
</tr>
<tr>
<td>BMI</td>
<td>GIANT</td>
<td>123 865</td>
<td>~2.85</td>
<td>Speliotes et al. (10)</td>
</tr>
<tr>
<td>Female-limited BMI</td>
<td>GIANT</td>
<td>73 137</td>
<td>~2.85</td>
<td>Randall et al. (7)</td>
</tr>
</tbody>
</table>

IEC, International Endogene Consortium; GIANT, Genetic Investigation of Anthropometric Traits Consortium; BMI, body mass index adjusted for age; WHRRadjBMI, waist to hip ratio adjusted for BMI and age.
a covariate representing the Australian and the UK strata, with the imputed data \((N = 12.5\) million SNPs). The enrichment analysis we present is from this set of results.

**GIANT Consortium**

**WHR GWAS.** A total of 77 167 subjects of European ancestry informative of body fat distribution measurement WHR from 32 GWAS were included \((6).\) The genotype data were imputed up to HapMap 2 CEU reference panel. The associations of 2.85 million SNPs with WHR were examined in a fixed-effects meta-analysis, after inverse normal transformation of WHR and adjusting for BMI and age within each study in an additive genetic model; analyses were conducted for males and females combined \((6)\) and limited to females only \((7)\) (Table 3).

**BMI GWAS.** A total of 123 865 subjects with overall adiposity measurement BMI from 46 GWAS were included \((10).\) The genotype data were imputed up to HapMap two CEU reference panels. The associations of 2.85 million SNPs with BMI were tested in an inverse-variance meta-analysis, after inverse normal transformation of BMI and adjusting for age and other appropriate covariates in an additive genetic model within each study; analyses were conducted for males and females combined \((10)\) and limited to females only \((7)\) (Table 3).

**Genetic enrichment analysis**

With one test of association conducted for each SNP, the GWAS analyses produced a genome-wide distribution of \(P\) -values of individual SNP associations. Prior to testing enrichment: (i) the overlap of SNPs present in endometriosis GWAS versus WHRadjBMI and BMI GWAS was taken, (ii) all SNPs with \(\text{MAF} \leq 0.01\) were removed, (iii) all SNPs with A/T and C/G base pairs were removed, (iv) correlated SNPs \((r^2 > 0.2)\) were removed as previously reported \((41)\) by taking the most significantly associated SNP and eliminating all SNPs that have a HapMap CEU pairwise correlation coefficient \((r^2) > 0.2\) with that SNP, then processing to the next strongly associated SNP remaining. This resulted in 173 157 independent SNPs in endometriosis versus WHRadjBMI and 173 223 in endometriosis versus BMI enrichment analyses.

The independent SNPs in the tails \((P < 1 \times 10^{-3})\) of the association results distribution of the two endometriosis GWAS (all endometriosis and ‘Stage B’ cases) were investigated for enrichment of WHRadjBMI or BMI low \(P\)-value \((P < 0.05)\) association signals; in reversal, SNPs in the tails of WHRadjBMI and BMI GWAS \((P < 1 \times 10^{-3})\) were investigated for evidence of nominal association \((P < 0.05)\) in the two endometriosis GWAS. The threshold of \(P < 1 \times 10^{-3}\) corresponded to the point at which endometriosis GWAS results started to deviate from the null distribution (evidence for association) in the overall and Stage B endometriosis Q–Q plots (Supplementary Material, Fig. S4). Enrichment was assessed in R by means of Pearson’s \(\chi^2\) tests with Yates’ continuity correction, testing for the difference in proportion of SNPs with association \(P < 0.05\) in the lookup dataset according to association in the discovery dataset \((P < 1 \times 10^{-3} \text{ versus } P \geq 1 \times 10^{-3})\). To test for consistency in directionality of phenotypic effects of the SNPs with evidence of enrichment, linear regression analysis was performed on the effect \((\beta)\) of each SNP for WHRadjBMI as predictor variable and the effect \((\beta)\) of endometriosis risk as the outcome variable \((35)\). In addition, a two-sided binomial test was performed with null hypothesis \(P = 0.50\).

**Permutation-based enrichment analysis**

For those results that showed nominally significant \((P < 0.10)\) evidence for enrichment in \(\chi^2\) tests of contingency tables, we performed permutation-based analyses to obtain empirical estimates of significance of enrichment. We (i) randomly picked the same number of independent SNPs ‘associated’ with the discovery trait at \(P < 1 \times 10^{-3}\) (e.g. the number of SNPs associated with all endometriosis at \(P < 1 \times 10^{-3}\) was \(n = 717\)) from the WHRadjBMI dataset; (ii) counted how many of the randomly selected SNPs had \(P\)-values of association with WHRadjBMI < 0.05; (iii) repeated Steps (i) and (ii) 10 000 times; (iv) determined the number of instances among the 10 000 draws in which the number of SNPs associated at \(P < 0.05\) with WHRadjBMI was greater or equal to the number we observed in our original analysis (e.g. \(\geq 52/717\)). For example, for overall endometriosis and overall WHRadjBMI, we observed this in 26/10 000 instances, corresponding to a \(P\)-value of \(2.6 \times 10^{-3}\), which was very similar to the \(P\)-value obtained from the \(\chi^2\) test \((P = 3.7 \times 10^{-3})\).

**Polygenic prediction analysis**

The independent SNPs in both WHRadjBMI and endometriosis GWAS datasets were used to conduct a polygenic prediction analysis \((11)\). The aim of this analysis was to evaluate the aggregate effects of many SNPs of small effect and assess whether subsets of SNPs selected in this manner from one disease/phenotype GWAS predicted disease/phenotype status in another, thus providing a measure of a common polygenic component with concordant directions of effect underlying the traits. Briefly, subsets of SNPs were selected from the WHRadjBMI GWAS data based on their association with WHRadjBMI using increasingly liberal thresholds, that is, \(P < 0.01\), \(P < 0.05\), \(P < 0.1\), \(P < 0.2\), \(P < 0.3\), \(P < 0.4\), \(P < 0.5\) and \(P < 0.75\). Using these thresholds, we defined sets of allele-specific scores in the WHRadjBMI dataset to generate risk profile scores for individuals in the endometriosis dataset. For each individual in the endometriosis dataset, we calculated the number of score alleles they possessed, each weighted by their effect size \((\beta\)-value\) of association in the WHRadjBMI dataset. To assess whether the aggregate scores were associated with endometriosis risk, we tested for a higher mean score in cases compared with controls. Logistic regression was used to assess the relationship between endometriosis disease status and aggregate risk score.

**Expression analyses**

**MGEx-Udb**

The mammalian gene expression uterus database (MGEx-Udb) is a manually curated uterus-specific database created using a meta-analysis approach from published papers \((28)\) that provides lists of transcribed and dormant genes for various normal, pathological (e.g. endometriosis, cervical cancer and endometrial cancer) and experimental (e.g. treatment and
knockout) conditions. Each gene’s expression status is indicated by a reliability score, derived based on the consensus across multiple samples and studies which highly variable (http://resource.ibab.ac.in/ MGEx-Udb/).

MuTHER
The MuTHER resource includes LCLs, skin and adipose tissue-derived simultaneously from a subset of well-phenotyped healthy female twins (29). Whole-genome expression profiling of the samples, each with two or three technical replicates, was performed using the Illumina Human HT-12 V3 BeadChips (Illumina, Inc.) according to the protocol supplied by the manufacturer. Log2 transformed expression signals were normalized separately per tissue as follows: quantile normalization was performed across technical replicates of each individual followed by quantile normalization across all individuals.

Genotyping was conducted using a combination of Illumina arrays (HumanHap300, HumanHap610Q, 1M-Duo and 1.2MDuo 1 M). Untyped HapMap2 SNPs were imputed using the IMPUTE software package (v2). In total, there were 776 samples with genotypes and expression values in adipose tissue. Association between all SNPs (MAF > 5%, IMPUTE info score > 0.8) within a gene or within 1 Mb of the gene transcription start or end site, and normalized expression values, were performed with the GenABEL/ProbABEL packages (42) using polygenic linear models incorporating a kinship matrix (GenABEL) followed by the mm score test with imputed genotypes (ProbABEL). Age and experimental batch were included as cofactors in the analysis. Benjamini Hochberg corrected \(P\)-values are reported.

MolOBB
We performed differential cis-eQTL analysis to compare the expression levels in gluteal and abdominal fat tissue from 49 individuals in the MolOBB dataset (24 with and 25 without metabolic syndrome—MetSyn) (30). We first checked for the presence of the SNP in the MolOBB genotype data and, in the case of absence, selected any proxies \(r^2 > 0.8\) available. We then searched for nearby genes (±500 kb) covered by the expression data using the bioconductor R package GenomiCRanges (43) and tested for association at each pair using a linear model with the expression level as the outcome and the SNP allelic dosage as a predictor, adjusting for age, gender and MetSyn case–control status. This analysis was carried out for both abdominal and gluteal subcutaneous adipose tissue. To investigate whether genes were differentially expressed between the two tissues, we applied a linear mixed model with tissue, MetSyn case–control status, gender and plate as fixed effects, and subject as a random effect using MAANOVA (44), as previously described in Min et al. (30). We report the uncorrected and genome-wide FDR corrected \(F\)s test \(P\)-values (30).

Biological pathway analysis

**PANTHER**
We conducted analyses using the PANTHER 8.1 database containing pathway information on 20 000 genes ( Homo sapiens) (32). We selected independent SNPs, which had association \(P\)-values < 1 \(\times\) 10\(^{-3}\) in the endometriosis datasets and an association \(P\)-value of <0.05 in the WHRadjBMI dataset, resulting in (i) 91 SNPs for all endometriosis and WHRadjBMI and (ii) 108 SNPs for Stage B endometriosis and WHRadjBMI. Each SNP was mapped to the closest gene within 1 Mb; 88 of 91 and 103 of 108 genes were present in the PANTHER database, and these subsets were tested for correlation with 241 biological processes and 176 pathways classified in the database (32). For each biological process/pathway, the difference between the observed fraction of genes in that pathway and the number expected by chance was tested using Fisher exact test. A Bonferroni correction was used as a conservative method for adjusting for the maximum number of biological processes \((n = 278); P = 1.80 \times 10^{-4}\) and pathways \((n = 78); P = 6.41 \times 10^{-4}\) tested.

SUPPLEMENTARY MATERIAL
Supplementary Material is available at HMG online.

ACKNOWLEDGEMENTS
We acknowledge with appreciation all the women who participated in the QIMR and Oxford endometriosis studies, and the many hospital directors and staff, gynecologists, general practitioners and pathology services in Australia and the UK who provided assistance with confirmation of diagnoses, and the many research assistants and interviewers for assistance with the studies.

Conflict of Interest statement. K.T.Z. has been a member of scientific advisory boards for AbbVie, Inc., Bayer Pharma AG and Roche Diagnostics.

FUNDING
The endometriosis GWAS was supported by a grant from the Wellcome Trust (WT084766/Z/08/Z) and makes use of WTCCC2 control data generated by the Wellcome Trust Case-Control Consortium. A full list of the investigators who contributed to the generation of these data is available from http://www.wtccc.org.uk. Funding for the WTCCC project was provided by the Wellcome Trust under awards 076113 and 085475. The QIMR study was supported by grants from the National Health and Medical Research Council (NHMRC) of Australia (241944, 339462, 389927, 389875, 389891, 389892, 389938, 443036, 442915, 442981, 496610, 496739, 552485 and 552498), the Cooperative Research Centre for Discovery of Genes for Common Human Diseases (CRC), Cerylid Biosciences (Melbourne) and donations from N. Hawkins and S. Hawkins. S.M. was supported by NHMRC Career Development Awards (496674, 613705). D.R.N. was supported by the NHMRC Fellowship (339462 and 613674) and the ARC Future Fellowship (FT0991022) schemes. A.P.M. was supported by a Wellcome Trust Senior Research Fellowship. G.W.M. was supported by the NHMRC Fellowships Scheme (339446, 619667). K.T.Z. was supported by a Wellcome Trust Research Career Development Fellowship (WT085235/Z/08/Z). C.M.L. was supported by a Wellcome Trust Research Career Development Fellow (086596/Z/08/Z). N.R. was supported by an MRC grant (MR/K011480/1). Funding to pay the
Open Access publication charges for this article were funded by the Wellcome Trust.

REFERENCES


APPENDIX

The International Endogene Consortium

Carl A. Anderson1,2, Scott D. Gordon3, Qun Guo4, Anjali K. Henders3, Ann Lambert2, Sang Hong Lee2, Peter Kraft1, Stephen H. Kennedy5, Kristina Eisinger3, Buechler31, Kristina Eisinger31, DIAGRAM Consortium, Amelie Bonne-...