

The search for genes contributing to endometriosis risk

Grant W. Montgomery^{1,7}, Dale R. Nyholt¹, Zhen Zhen Zhao¹, Susan A. Treloar¹,
Jodie N. Painter¹, Stacey A. Missmer^{2,3,4}, Stephen H. Kennedy⁵ and Krina T. Zondervan⁶

¹Genetic and Molecular Epidemiology Laboratories, Queensland Institute of Medical Research, Royal Brisbane Hospital, 300 Herston Road, Brisbane, QLD 4029, Australia; ²Channing Laboratory, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA; ³Department of Obstetrics, Gynecology, and Reproductive Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA; ⁴Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA; ⁵Nuffield Department of Obstetrics and Gynaecology, University of Oxford, Oxford, UK; ⁶Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK

⁷Correspondence address. Tel: +61-7-3362-0247; Fax: +61-7-3362-0101; E-mail: grant.montgomery@qimr.edu.au

BACKGROUND: Genetic variation contributes to the risk of developing endometriosis. This review summarizes gene mapping studies in endometriosis and the prospects of finding gene pathways contributing to disease using the latest genome-wide strategies. **METHODS:** To identify candidate-gene association studies of endometriosis, a systematic literature search was conducted in *PubMed* of publications up to 1 April 2008, using the search terms 'endometriosis' plus 'allele' or 'polymorphism' or 'gene'. Papers included were those with information on both case and control selection, showed allelic and/or genotypic results for named germ-line polymorphisms and were published in the English language. **RESULTS:** Genetic variants in 76 genes have been examined for association, but none shows convincing evidence of replication in multiple studies. There is evidence for genetic linkage to chromosomes 7 and 10, but the genes (or variants) in these regions contributing to disease risk have yet to be identified. Genome-wide association is a powerful method that has been successful in locating genetic variants contributing to a range of common diseases. Several groups are planning these studies in endometriosis. For this to be successful, the endometriosis research community must work together to genotype sufficient cases, using clearly defined disease classifications, and conduct the necessary replication studies in several thousands of cases and controls. **CONCLUSIONS:** Genes with convincing evidence for association with endometriosis are likely to be identified in large genome-wide studies. This will provide a starting point for functional and biological studies to develop better diagnosis and treatment for this debilitating disease.

Keywords: endometriosis; genetic variation; candidate gene; linkage; genome-wide association

Introduction

Endometriosis is a complex condition, with tissue resembling endometrium found in extra-uterine sites. Although symptoms vary, they commonly include severe pelvic pain, severe dysmenorrhea (painful periods) and reduced fertility (Giudice and Kao, 2004; Berkley *et al.*, 2005). The main pathological processes associated with the disease are peritoneal inflammation and fibrosis, and the formation of adhesions and ovarian cysts. Different phenotypic classifications have been proposed, the main ones being the 4-stage rAFS classification (based on the total surface size of lesions, presence of adhesions and ovarian lesions) (The American Fertility Society, 1985), ovarian versus peritoneal disease (The American Fertility Society, 1985) and deep infiltrating versus superficial disease (Koninckx *et al.*, 1999). Whether these subphenotypes represent the natural history of one disorder, or are in fact different disease subtypes altogether, is an important

consideration in endometriosis research, but as yet remains unclear. The disorder often recurs and has a major impact on women's health, relationships, productivity and life choices (Mathias *et al.*, 1996; Simoens *et al.*, 2007). The risk of endometriosis increases with age during the reproductive years (Eskenazi and Warner, 1997); the onset can occur from the menarche onwards, but presentation after the menopause is rare, indicating that it is an estrogen-dependent condition (Bulun *et al.*, 2002). The few established risk factors include increased exposure to menstruation (i.e. shorter cycle length, longer duration of flow and nulliparity), positive smoking history (which reduces ovarian production of estrogen and reduces risk of endometriosis) and increased peripheral body fat (which increases estrogen) (Eskenazi and Warner, 1997; Missmer and Cramer, 2003; Berkley *et al.*, 2005).

As the diagnosis of endometriosis is made on visual inspection of the pelvis at laparoscopy, the population prevalence is unknown.

The best estimates suggest endometriosis (all stages) affects 8–10% of women in their reproductive years (Eskenazi and Warner, 1997) and 20–50% of women with infertility (Gao *et al.*, 2006), with an estimated prevalence of moderate–severe endometriosis of up to 2% (Zondervan *et al.*, 2002). Prevalence of all stages of endometriosis in Australia was estimated at 7.2% in a volunteer sample of Australian twins (Treloar *et al.*, 1999a). These figures represent prevalence rates in Caucasian populations; there is some evidence that the prevalence is lower in African-American women (Kyama *et al.*, 2004), but estimates may be biased due to differential access to healthcare (Missmer *et al.*, 2004). The most widely accepted theory to explain the origin of endometriotic tissue is that viable endometrial cells reach the peritoneal cavity through retrograde menstruation along the fallopian tubes (Sampson, 1927). However, menstrual debris is present in the peritoneal cavity of up to 90% of menstruating women. Possible explanations for susceptibility in only some women are summarized in Fig. 1 and include increased exposure to menstrual debris, abnormal eutopic endometrium, altered peritoneal environment, reduced immune surveillance and increased angiogenic capacity (Healy *et al.*, 1998; Vinatier *et al.*, 2001; Treloar *et al.*, 2002; Varma *et al.*, 2004). Biological studies have so far failed to clearly define the mechanisms leading to disease.

Genetic contributions to endometriosis

Endometriosis is commonly regarded as a complex trait, caused by the interplay between genetic and environmental factors.

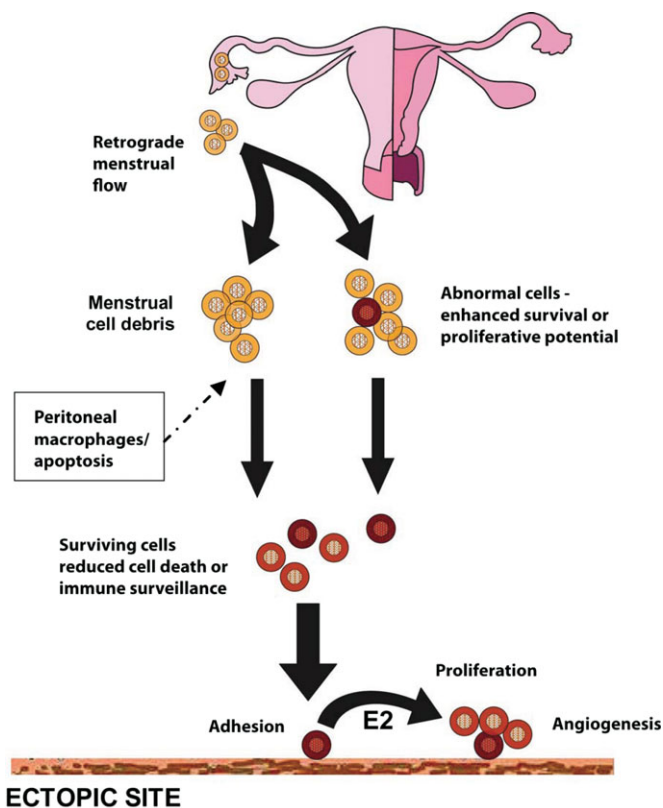


Figure 1: Schematic representation of possible mechanisms contributing to the development of endometriosis.

The contribution of genetic factors in endometriosis susceptibility is supported by a number of different studies (Kennedy, 1998; Zondervan *et al.*, 2001; Simpson and Bischoff, 2002; Stefansson *et al.*, 2002; Treloar *et al.*, 2002; Viganò *et al.*, 2003). Higher rates of endometriosis are found among the relatives of endometriosis cases compared with those of controls in both hospital (Kennedy *et al.*, 1995; Simpson and Bischoff, 2002) and population-based (Stefansson *et al.*, 2002) samples. The relative recurrence risk to sibs, which is the increase in risk of an individual whose sibling is affected compared with the risk in the general population, has been estimated at 2.34 in an Australian sample of twins and their families (Treloar *et al.*, 2002), although estimates from imaging studies on the sisters of women with more severe disease suggest the value may be as high as 15 (Kennedy *et al.*, 1997). It should be noted that, for endometriosis, obtaining an accurate estimate of this recurrence risk is difficult because the population prevalence is unknown and there is inevitable bias in ascertaining endometriosis cases through surgery that will influence the estimated risk to the siblings. Further evidence of genetic risk of endometriosis has been suggested by twin studies, which showed concordance for endometriosis among monozygotic twins (Moen, 1994; Hadfield *et al.*, 1997), and increased concordance in monozygotic compared with dizygotic twins with heritability estimated at 51% (Moen, 1994; Hadfield *et al.*, 1997; Treloar *et al.*, 1999b, 2002). In addition to these studies in humans, familial aggregation has also been shown in non-human primates (Zondervan *et al.*, 2004).

Recently, the genetic contribution to endometriosis has been questioned (Di and Guo, 2007). The authors point out problems with study designs due to small sample size in many studies, ascertainment bias, increased opportunity for diagnosis among family members of cases compared with controls and familial aggregation of confounding risk factors such as early age at menarche (Di and Guo, 2007). These are valid concerns particularly for some published studies, although we should acknowledge the difficulties inherent in endometriosis research because of the lack of a non-invasive diagnostic tool (Zondervan *et al.*, 2002). The strongest evidence for genes influencing endometriosis comes from more recent, large-scale, studies in twins (Treloar *et al.*, 1999b) and in the Icelandic population (Stefansson *et al.*, 2002). A classical twin study in a large sample of Australian twins concluded that genetic factors contributed about half of the variation in endometriosis risk (Treloar *et al.*, 1999b). The potential contribution of increased sharing of environmental risk-factors of endometriosis among identical versus non-identical twins (a potential alternative explanation for the observed results) was investigated and considered unlikely. Analysis of endometriosis cases in the Icelandic population addressed issues of bias in various ways and in each case reached the conclusion that there was evidence for genetic effects on endometriosis risk (Stefansson *et al.*, 2002). Nevertheless, all studies trying to dissect the genetic and non-genetic causes using familial aggregation studies based on phenotypic observations alone must make explicit assumptions about shared environmental influences that are difficult to exclude entirely. Genetic contributions to disease can be tested directly without making these assumptions using genome-wide marker data either through linkage or association studies (Visscher *et al.*, 2006), but very large sample sizes are required for accurate estimates.

Familial aggregation may result from comorbid conditions or traits explaining the observed findings. With regard to endometriosis, Di and Guo (2007) suggested that chromosomal regions linked to endometriosis (Treloar *et al.*, 2005a; Zondervan *et al.*, 2007) may be due to familial aggregation of age at menarche; itself a complex, heritable trait associated with endometriosis risk. Such an explanation—although theoretically possible—is unlikely, since linkage studies specifically designed for one complex trait of interest have low power to detect other complex traits that are loosely associated with it. Indeed, linkage analyses of age at menarche show that linkage signals for endometriosis do not overlap signals for age at menarche (see Linkage Mapping). Therefore, although individual studies carry interpretation difficulties to a greater or lesser extent, current evidence accumulated across a range of studies supports a genetic contribution to endometriosis risk.

Defining pathways to disease is a major goal of research in endometriosis. This knowledge can be used to develop more effective methods of diagnosis and treatment. Genetic studies provide one important approach to define causal pathways influencing endometriosis. The number of gene mapping studies for this disease has increased in recent years as the role of genetic factors has become more widely accepted. In addition, there have been dramatic advances in human genetics in the last few years with many recent papers reporting genetic variants associated with other complex diseases. This review summarizes current studies on genetic variation contributing to endometriosis and the prospects of finding gene pathways contributing to disease using the latest developments in high throughput genotyping and genome-wide strategies.

Gene discovery

Gene mapping methods have been very successful in identifying mutations responsible for monogenic diseases. These mutations generally have large effects and have moderate to high penetrance. Consequently, most mutation carriers present with the disease, and most affected families carry the same mutation. Mutations with large and distinct effects are relatively easy to map and locate. Examples of mutations with large effects influencing reproductive pathways include mutations in KISS1 receptor (*KISS1R*) influencing idiopathic hypogonadotropic hypogonadism (de Roux *et al.*, 2003; Seminara *et al.*, 2003) and mutations in growth differentiation factor 9 (*GDF9*) increasing risk for dizygotic twinning (Palmer *et al.*, 2006).

Finding genetic variants contributing to complex diseases such as endometriosis, diabetes or heart disease is far more difficult because the contribution of individual genes is small, many genes contribute to an individual's risk of developing the disease and disease risk is often modified by environment. However, common diseases present a much greater public health burden than Mendelian diseases, and there are major international efforts to define genetic contributions to these diseases. The basic requirements to map disease genes remain the same (Fig. 2). Many studies are required for both the discovery and replication steps with sufficient power to detect the small effects of any individual variants. In addition, different combinations of variants are likely to be present in cases in different families and cases in the same family may not all carry the same variants. Despite these

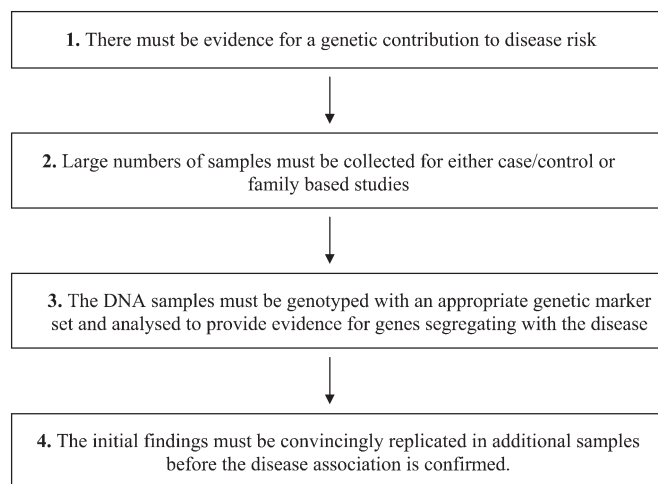


Figure 2: The basic requirements for mapping disease genes.

limitations, gene mapping remains an important approach to understanding pathways involved in complex disease aetiology, including endometriosis.

Candidate gene studies

Most genetic studies of endometriosis to date have followed *a priori* defined biological hypotheses and have analysed a small number of variants in candidate genes (often in small numbers of cases and controls). Candidate genes (candidates) are generally chosen based on biological mechanisms thought to contribute to disease. Variants in these candidate genes are genotyped in samples from cases and controls or in affected families to test for association by statistical analysis of genotype data.

To identify candidate-gene association studies of endometriosis, a systematic literature search was conducted in *PubMed* (<http://www.ncbi.nlm.nih.gov/sites/entrez/>) of publications up to 1 April 2008, using the search terms 'endometriosis' plus 'allele' or 'polymorphism' or 'gene'. Only papers that included information on both case and control selection, and showed allelic and/or genotypic results for named germ-line polymorphisms, were included. To enable the application of these criteria by the authors, the review was limited to publications in the English language. Table I presents an overview of candidate gene studies in a format to allow readers to identify which genes have been tested and whether results have replicated in more than one study.

These studies report results from statistical tests for association with endometriosis susceptibility for 76 genes (summarized in Table I). Many of these studies have been reviewed in detail recently (Guo, 2005, 2006a,b; Falconer *et al.*, 2007). Candidates tested include genes from detoxification pathways, sex steroid pathways and cytokine signalling pathways, adhesion molecules and matrix enzymes, and cell-cycle regulation (Falconer *et al.*, 2007). Glutathione *S*-transferase enzymes involved in the pathway for detoxification of a range of toxic compounds and carcinogens have been studied extensively, in particular because of the suggestion of dioxin exposure being a risk factor (Birnbaum and Cummings, 2002), a finding which was later questioned (Rier *et al.*, 2001; Eskenazi *et al.*, 2002; Guo, 2004). Polymorphisms in glutathione *S*-transferase M1 (*GSTM1*) on chromosome 1p13.3

Table I. A summary of case–control results from genetic association studies with endometriosis.

Candidate gene**	Chr	Number of studies with association*		
		Significant	Non-significant	#Sign/Total
Cytokines/Inflammation				
CCR2	3p21	0	1	0/1
CCR5	3p21	0	1	0/1
CTLA4	2q33	0	1	0/1
IFNG	12q24	1	0	1/1
IL1B	2q13-21	0	2	0/2
IL1R1	2q12	0	1	0/1
IL1RN	2q14	0	1	0/1
IL2RB	22q13	1	0	1/1
IL4	5q23-31	0	2	0/2
IL6	7p21-15	0	3	0/3
IL10	1q31-32	2	1	2/3
IL12RB1	19p13	0	1	0/1
IL18	11q22	0	1	0/1
KIR	19q13	1	0	1/1
MCP1 (CCL2)	17q11-12	0	1	0/1
MPO	17q21-23	0	1	0/1
PTPN22	1q13	1	0	1/0
RANTES (CCL2)	17q11	0	1	0/1
STAT6	12q13	1	0	1/1
TNFA	6p21	3 ^a	5	3/8
TNFR2	1p36	0	1	0/1
Steroid-synthesizing enzymes, detoxifying enzymes and receptors				
AhR	7p15	0	1	0/1
AhRR	5p15	2	0	2/2
ARNT	1q21	0	1	0/1
CYP17A1	10q24	0	6	0/6
CYP19A1	15q21	2	3	2/5
CYP1A1	15q24	1	5	1/6
CYP1B1	2p22	0	3	0/3
GSTM1	1p13	6	6	6/12 ^b
GSTP1	11q13	1	1	1/2
GSTT1	22q11	2	7	2/9 ^b
NAT1	8p23-21	0	1	0/1
NAT2	8p22	2	2	2/4
PPARG	3p25	2	0	2/2
Hormone receptors				
AR	Xq12	1	1	1/2
ESR1	6q24-27	5 ^c	4	5/9 ^d
ESR2	14q21-22	1 (stage IV)	1	1/2
NRIP1	21q11	1	0	1/1
PGR	11q22-23	5 ^c	2	5/7
Estradiol metabolism				
COMT	22q11	0	3 ^f	0/3
HSD17B1	17q11-21	2	0	2/2
Other enzymes and metabolic systems				
ACE	17q23	1	0	1/1
AHSG	3q27	1	0	1/1
EMX2	10q26	0	1	0/1
eNOS (NOS3)	7q36	1	1	1/2
GALT	9p13	1	4	1/5
LH-B	19q13	0	2	0/2
SERPINE (PAI1)	7q21-22	1	0	1/1
TIMP2	17q25	1	0	1/1
XRCC4	5q13-14	1	0	1/1
Growth factor systems				
EGF	4q25	0	1	0/1
EGFR	7p12	1	1	1/2
FGFR2	10q26	0	1	0/1
PTEN	10q23	0	1	0/1

Continued

Table I. Continued

Candidate gene**	Chr	Number of studies with association*		
		Significant	Non-significant	#Sign/Total
TGFB1	19q13	1	1	1/2
VEGFA	6p21-12	5	1	5/6
Adhesion molecules and matrix enzymes				
CDH1	16q22	1	0	1/1
COL18A1	21q22	0	1	0/1
ICAM1	19p13	1	2	1/3
MMP1	11q21-22	1	2	1/3
MMP2	16q12	1 (deep-infil.)	1	1/2
MMP3	11q22	0	3	0/3
MMP7	11q22	1	1	1/2
MMP9	20q11-13	0	1	0/1
MMP12/13	11q22	1	0	1/1
Apoptosis, cell-cycle regulation and oncogenes				
CDKN1A [p21]	6p21	0	1	0/1
FAS	10q24	0	1	0/1
FASL	1q23	0	1	0/1
KRAS	12p12	0	1	0/1
(T)P53	17p13	1	4	1/5
Human leukocyte antigen system and immune components				
HLA-A	6p21	0	1	0/1
HLA-B	6p21	1	0	1/1
HLA-C2	6p21	0	1	0/1
HLA-DPB1	6p21	0	1	0/1
HLA-DQB1	6p21	1	0	1/1
HLA-DRB1	6p21	0	4	0/4

*Only articles published in English were reviewed in preparing this table, with details for individual studies given in Supplementary Table 1. One result for each study sample was counted for each gene; when >1 polymorphism was investigated in a study, the result is indicated as significant if 1 or more of the variants was reported by the authors to be significant.

**Candidate gene groupings are according to those proposed by Falconer *et al.* (2007).

^aTwo of these studies found significant associations with Stage IV only.

^bIn a meta-analysis, Guo (2005) found no evidence for association with GSTM1, and some limited evidence of increased risk with the null genotype of GSTT1 (interpretation hampered by publication bias). The reason for finding GSTT1 not associated in the table above is that in the meta-analysis (Guo, 2005), three or more studies with large significant results were included (non-English manuscripts), two from China (Peng *et al.*, 2003; Ding *et al.*, 2004) and one from Russia (Ivashchenko *et al.*, 2003).

^cOne of these studies found a significant effect for Stage I/II versus controls only, not Stage III/IV, whereas another study found a significant effect for recurrent cases versus controls.

^dCritical meta-analysis by Guo (2006) reported a handful of positive findings that have not been independently replicated.

^eOne of these studies found a significant effect for deep-infiltrating versus controls only.

^fOne of these studies found a significant effect of COMT on risk of adenomyosis.

and glutathione *S*-transferase theta 1 (*GSTT1*) on chromosome 22q11.23 have been evaluated in over twenty studies (Guo, 2005). Pooled odds ratios for both enzymes (*GSTM1*: OR: 1.96, 95% CI: 1.29–2.98; *GSTT1*: OR: 1.77, 95% CI: 1.19–2.63) suggested increased risk of developing endometriosis. However, there is significant heterogeneity between studies for both enzymes and publication bias suggesting the results should be viewed with caution especially for *GSTM1* (Guo, 2005). Meta-analysis for multiple studies for the detoxification enzymes *N*-acetyltransferase 2 (arylamine *N*-acetyltransferase) (*NAT2*) on chromosome 8p22 and cytochrome P450, family 1, subfamily A, polypeptide 1 (*CYP1A1*) on chromosome 15q24.1 found no evidence for association between the *NAT2* acetylation polymorphism (pooled OR: 1.13, 95% CI: 0.70–1.82) and endometriosis (Guo, 2006a). There is some evidence for a small increase in risk for alleles at the *MspI* polymorphism in *CYP1A1* (pooled OR: 1.44, 95% CI: 1.00–2.06), but the evidence is not strong, and further studies are needed to confirm the result (Guo, 2006a).

As noted earlier, endometriosis is an estrogen dependent disease and a number of studies have investigated genes from pathways of sex steroid biosynthesis and signalling. Review of association studies for cytochrome P450, family 17, subfamily A, polypeptide 1 (*CYP17A1*), cytochrome P450, family 19, subfamily A, polypeptide 1 (*CYP19*), androgen receptor (*AR*), progesterone receptor (*PGR*) and estrogen receptors (*ESR1* and *ESR2*) (Guo, 2006b) concluded that many reported positive findings were unsound because of problems with data analysis in the original reports. Meta-analysis of the studies provided some, though limited, support for association between endometriosis and both the *PGR*-PROGINS polymorphism (pooled OR: 1.94, 95% CI: 1.31–2.88) and *ESR1*-PvuII polymorphism (pooled OR: 2.1, 95% CI: 1.20–3.68) (Guo, 2006b). A subsequent study in a large family-based sample failed to support any association between *PGR* and endometriosis (Treloar *et al.*, 2005b).

Convincing evidence for association must include replication. Many replication studies for complex disease associations fail at

the replication step (Ioannidis *et al.*, 2001; Hirschhorn *et al.*, 2002). Less than half of the reported associations with endometriosis have been investigated in a separate sample and many associations are not replicated in subsequent studies (Table I). A number of factors contribute to this failure including low prior odds for association when testing only a few variants in one gene, study power, data analysis, publication bias, population differences, failure to type the same variants and technical issues. Many studies only test small samples and lack the necessary power to detect the small effects expected for most common diseases including endometriosis (Zondervan *et al.*, 2002, 2004). Results that appear significant are more likely to be published (Ioannidis *et al.*, 2001; Hirschhorn *et al.*, 2002). In addition, some studies do not take adequate account of statistical problems of multiple testing (Guo, 2006b). This publication bias, together with problems in experimental designs, suggests many results are false positive associations. Studies in our large Australian sample failed to replicate putative associations for *PGR* and *TNFA* (Treloar *et al.*, 2005b; Zhao *et al.*, 2007). We recently typed the rs2476601 1858T/C (Trp620Arg) variant in *PTPN22* (Ammendola *et al.*, 2007) in the same sample and found no evidence for association with endometriosis ('T' allele frequencies in cases and controls were 0.095 and 0.091, respectively, P -value > 0.5).

For real associations, the strength of the true effect is often over estimated in the initial study (Ioannidis *et al.*, 2001) in an effect referred to as the 'winner's curse' (Zollner and Pritchard, 2007). Consequently, replication studies often need more samples and greater power than the original study to detect the effect and some cases of failure to replicate findings from the original study might be due to the low power of replication studies (Lohmueller *et al.*, 2003; Zollner and Pritchard, 2007). Review of the large number of studies conducted for association with endometriosis (Table I) does not provide support for any gene variants clearly associated with increased risk of endometriosis. Some reported results may represent true associations, but given the small effect sizes expected further studies in very large samples are required to provide convincing evidence. This can be achieved by combining samples from multiple sites. Recent candidate gene studies in breast cancer included up to 18 000 cases and 22 000 controls to identify a common variant in *CASP8* associated with breast cancer risk (Cox *et al.*, 2007), although more modest sample sizes have been used in successful studies for other common diseases.

Linkage mapping

A second gene mapping approach involves a hypothesis-free search for evidence of genomic regions harbouring genetic risk variants for endometriosis—prior to further association or sequencing studies—using linkage mapping across the genome. Over a period of 10 years, laboratories in Australia and the UK recruited families of sisters with surgically confirmed endometriosis for sib-pair linkage analysis. Sib-pair linkage studies use genome-wide analysis of informative polymorphic microsatellite markers to identify regions of significant excess sharing in affected sibs (Risch, 1990a; Kruglyak and Lander, 1995a,b; Lander and Kruglyak, 1995). For common diseases with a high genetic recurrence risk, the most informative relative pairs are distant

ones. For diseases such as endometriosis with lower recurrence risk to sibs, sister pairs constitute the best design (Risch, 1990b). The affected sib method is also more suitable for those conditions where it is difficult to determine 'unaffected' status. In the case of endometriosis, determination of unaffected status would require a laparoscopy.

A genome scan with 1176 affected sister pair families was completed in the combined Australian and the UK families (Treloar *et al.*, 2005a) comprising the International Endogene Consortium (IEC). Power calculations suggested the study sample had 80% power to detect a locus with a recurrence risk to sisters of 1.35 (Treloar *et al.*, 2002). Genetic markers spaced about every 10 cM (approximately every 10 Mb) across the genome were typed in DNA samples from the sisters and other family members. The combined data identified one peak of significant linkage on chromosome 10 with a peak of 3.09 (genome-wide P -value of 0.047). A second peak on chromosome 20 shows suggestive evidence for linkage. The results were consistent for both data sets with evidence for linkage to chromosome 10 in both the Australian and the UK families. Fine mapping with an additional four microsatellite markers on chromosome 10 increased the evidence for linkage slightly. The peak of maximum linkage was located at 148.75 cM (127.92 Mb).

A separate linkage analysis in a subset of families with three or more affected women (Oxford: $n = 52$; Australia: $n = 196$) was conducted to test whether the apparent concentration of cases in these families might reflect the presence of a more 'Mendelian-like' rare genetic variant acting in this subset of families. If present, this situation would be analogous to the discovery of *BRCA1* and *BRCA2* genes causal in a small subset of breast cancer patients with strong familial inheritance patterns (Miki *et al.*, 1994; Wooster *et al.*, 1995). The analysis in the subset of endometriosis families identified an additional peak of linkage on chromosome 7p (Zondervan *et al.*, 2007). The combined analysis identified significant evidence for linkage to the region using a transmission model with a recessive gene conferring a high risk for developing the disease. This suggests there may be a high-penetrance susceptibility locus for endometriosis in this region present in a small subset of families.

It has been suggested that familial aggregation of endometriosis, and linkage to chromosomes 7 and 10, may be due to the association between endometriosis and age at menarche, since age at menarche is known to have a genetic component (Di and Guo, 2007). Specific gene mapping studies of age at menarche suggest that variants influencing this trait do not overlap the regions of chromosome 7 or 10 implicated in our studies of endometriosis (Treloar *et al.*, 2005a; Guo, 2006; Rothenbuhler *et al.*, 2006; Zondervan *et al.*, 2007). Additionally, a large linkage study in 13 697 individuals and 4899 pseudo-independent sister-pairs found no evidence for significant linkage and no suggestive linkage peaks on chromosome 7 or 10 (G.W. Montgomery *et al.*, unpublished observations). Linkage to these regions in endometriosis is unlikely to be related to age at menarche, but it should be noted that common variants influencing both age at menarche and endometriosis would provide valuable information on the causes of endometriosis and have implications for variation in age of menarche.

Subsequent fine mapping and candidate gene resequencing studies have been—and continue to be—conducted for genes

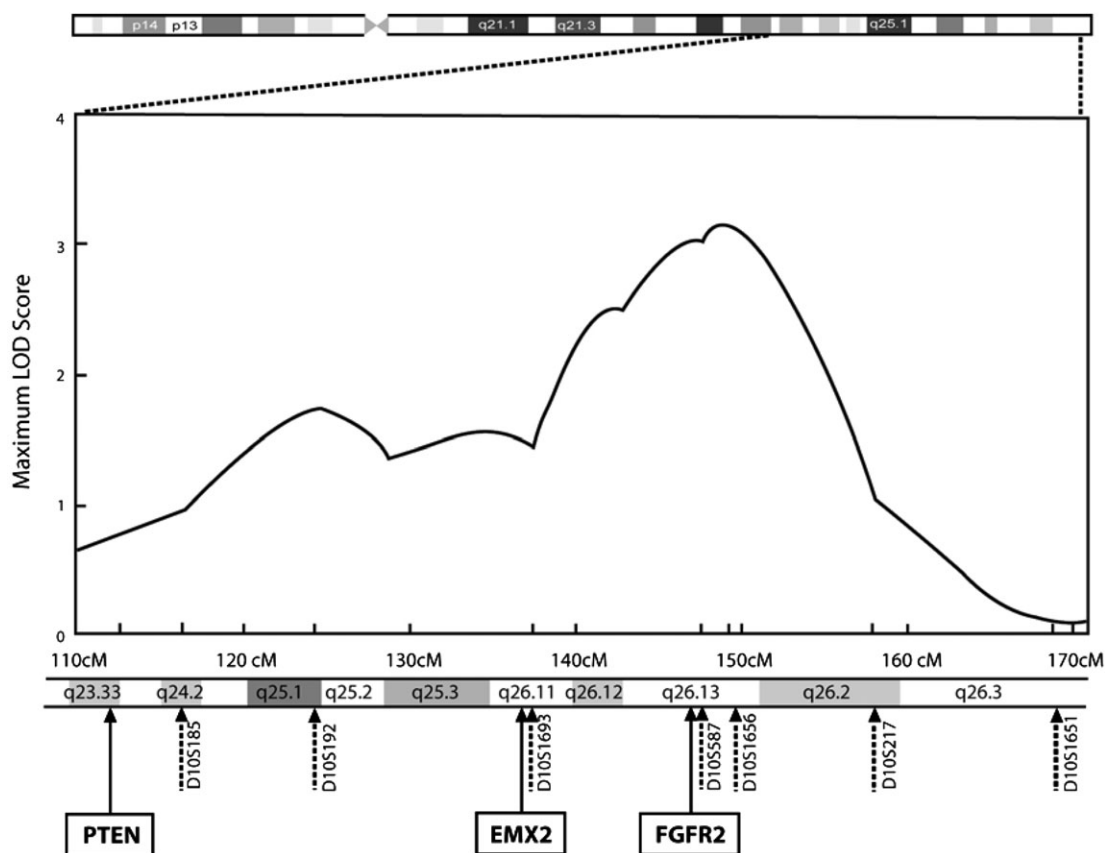


Figure 3: Schematic representation of the distal portion of chromosome 10 showing the region of significant linkage to endometriosis reported in Treloar *et al.* (2005a) and the location of candidate genes in the region screened for association with endometriosis.

under the chromosome 7 and 10 linkage peaks to identify genetic variants contributing to linkage in these regions. Two genes on chromosome 10q (Fig. 3), which were previously implicated in endometriosis and endometrial cancer, have been investigated (Treloar *et al.*, 2007). The genes are Empty Spiracles, Homolog of *Drosophila*, 2 (*EMX2*) and Phosphatase and Tensin Homolog (*PTEN*) (Sato *et al.*, 2000; Kurose *et al.*, 2001, 2002; Fujii *et al.*, 2002; Latta and Chapman, 2002; Martini *et al.*, 2002; Swiersz, 2002; Zhou *et al.*, 2002; Dinulescu *et al.*, 2005). *EMX2* is a transcription factor essential for reproductive tract development also expressed in the adult uterine endometrium, with decreased expression during the luteal phase of the menstrual cycle (Troy *et al.*, 2003; Daftary and Taylor, 2004). *PTEN* promotes cell survival and proliferation and inactivation of *PTEN* is an early event in endometrial hyperplasia and the development of ovarian and endometrial cancers (Maxwell *et al.*, 1998). *PTEN* lies at 89.6 Mb, more centromeric than the region of significant linkage (Fig. 3). However, the linkage peak is broad and there is evidence for linkage and association with endometriosis in Puerto Rican families at marker D10S677 (Flores *et al.*, 2004). D10S677 is located at 113.34 cM (95.95 Mb) close to the *PTEN* locus. Genotyping of a large number of markers across both genes found no evidence of association with endometriosis (Treloar *et al.*, 2007).

More recently, another gene which lies within the region of significant linkage on chromosome 10, the fibroblast growth factor receptor 2 gene (*FGFR2*), was implicated in both endometrial

(Pollock *et al.*, 2007) and breast cancer (Easton *et al.*, 2007; Hunter *et al.*, 2007b). To investigate this gene for involvement in endometriosis, single nucleotide polymorphisms (SNPs) within intron 2 of *FGFR2* including two SNPs (rs2981582 and rs1219648) significantly associated with breast cancer were genotyped. In addition, a dense set of 40 SNPs across 150 kb of the *FGFR2* gene covering common variation within the coding region of the gene was genotyped in a large Australian sample of endometriosis cases and unrelated controls. No evidence for association between common variation in either intron 2 or the entire coding region in the *FGFR2* gene was found, suggesting this gene is not a major contributor to endometriosis susceptibility (Zhao *et al.*, 2008).

Genome-wide association

Experience has shown that, apart from a few notable exceptions, hypothesis-based candidate gene studies and linkage mapping followed by candidate gene targeting have been largely unsuccessful in identifying susceptibility genes for complex diseases. Even with prior linkage information limiting the genomic region of interest, it is difficult to define good candidates because we often lack sufficient knowledge about the biological processes underlying a disease, and—in addition—do not yet know the biological functions of all genes. Indeed, following on from identifying significant linkage on chromosome 10, the IEC has now turned to uncovering the gene(s) involved by utilizing further

hypothesis-free fine mapping association methods using very densely spaced SNPs, chosen to most comprehensively capture the genetic architecture in the region.

We are using recently developed genome-wide association (GWA) methods providing a powerful approach to locating genetic variants contributing to common diseases. These methods have arisen from several genome discovery projects finalized in recent years, combined with staggering developments in increased efficiency and reduced costs of high-throughput genotyping technologies. During 2007, the number of risk alleles identified for common diseases identified using these new methods exceeds the number identified in the previous decade (Petretto *et al.*, 2007). Key developments underpinning GWA studies include the survey of common human variation by the HapMap Consortium (Consortium, 2005), and the development of high-throughput genomics platforms capable of genotyping up to 1 million SNPs in a single experiment. The human HapMap project (<http://www.hapmap.org/>) provided detailed information on population-specific patterns of genomic architecture and genetic variation. The results showed that common SNPs in close proximity to each other are often inherited together within populations, so that the presence of one predicts the presence of another (also termed 'linkage disequilibrium'). The HapMap project also provided a 'map' indicating the dependence between these common SNPs, and this information can be used to greatly reduce the amount of genotyping required in GWA studies to 'cover' the whole genome with regard to common genetic variation.

In parallel with studies exploring genomic architecture, technological advances have led to development of platforms for genotyping many thousands of SNPs in a single experiment. Several high-throughput genotyping platforms can now type up to 1 million SNPs in an individual sample and process many samples per day. Current commercial SNP genotyping chips use sets of SNPs selected based on information from the HapMap data to achieve maximum genome coverage. The chips are being used successfully in GWA studies to discover genes contributing to complex disease.

In a recent series of papers, the Wellcome Trust Case Control Consortium (WTCCC) employed these genome-wide techniques on a common set of 3000 controls with sample collections exceeding 2000 cases and reported associations for a range of complex diseases which were subsequently replicated in further large sample collections including type 1 diabetes (Consortium, 2007; Todd *et al.*, 2007), type 2 diabetes (Consortium, 2007; Saxena *et al.*, 2007; Scott *et al.*, 2007; Sladek *et al.*, 2007), Crohn's disease (Consortium, 2007; Libioulle *et al.*, 2007; Rioux *et al.*, 2007) and coronary heart disease (Consortium, 2007; McPherson *et al.*, 2007). Genome-wide scans have also reported new variants affecting prostate cancer (Gudmundsson *et al.*, 2007; Yeager *et al.*, 2007), and breast cancer (Easton *et al.*, 2007; Hunter *et al.*, 2007a). Many of the papers reported associations to novel loci not previously implicated in the disease. Important features of these studies are they included unprecedented, large sample sizes, they concentrated on the small number of variants with strongest evidence of association and subsequently replicated these associations in additional large numbers of samples. Almost all the new risk alleles have odds ratios (effect sizes) <2.0 and many have odds ratios <1.5 (Petretto *et al.*, 2007). Importantly, ~25% of associated variants are found in regions not coding for

genes and would be missed in gene centric disease discovery programmes.

The GWA methods employed by WTCCC were highly successful in identifying risk alleles in some diseases (such as Crohn's disease), but found very few or no risk alleles in others (such as hypertension and bipolar disorder) (Consortium, 2007; Todd *et al.*, 2007). The reasons for this inconsistency are not immediately apparent. First, they may be related to the phenotype definitions representing a relatively heterogeneous case set (Zondervan and Cardon, 2007). Second, it may be that environmental factors contribute most to their aetiology, and any genetic risk-factors are of such small effect that much larger studies would be required to detect them. Moreover, even for disease where multiple risk alleles are reported, the variants only account for a small proportion of the genetic risk. Further loci are likely to be identified by testing more variants identified in the genome-wide studies, but the effect sizes are all likely to be small.

GWA offers the prospect of making real progress in discovery of genes influencing endometriosis risk and several groups are planning GWA studies in endometriosis. There are major challenges to the successful outcome of these studies. An enduring lesson from all complex disease association studies is the need for very large studies because of the small contribution of each genetic effect, low prior odds for association and the need to adjust for testing many thousands of SNPs (Peltonen, 2007). Studies with sufficient power can only be achieved by collaboration of large consortia. Valid concerns have been expressed frequently in the past regarding the ability to dissect true findings from the large number of 'significant' findings that are produced from GWA studies because of the enormous number of statistical tests conducted when analysing 0.5–1 M SNP genotypes (Zondervan and Cardon, 2004; Di and Guo, 2007). However, the recent GWA and follow-up studies have shown that the combination of GWA study and multiple replication sets does enable the elucidation of common genetic variants underlying a complex trait, as long as both are sufficiently powered (including several thousands of cases and controls).

Although genome-wide studies of gene-environment interaction and epistasis (gene-gene interaction) are likely to require sample sizes much in excess of those currently available, pathway-based approaches can be readily applied to GWA studies of complex disease to yield biological insights that are otherwise undetectable by focusing only on individual genes and/or regions that have the strongest evidence for association. For example, a typical pathway-based approach might rank all genes by their significance of association and then look for whether a particular group of genes is enriched at the significant end of the ranked list more than expected by chance. Application of such pathway-based approaches, where multiple genes in the same pathway contribute to disease aetiology, but common variations in each of the causal genes make modest contributions to disease risk, has enormous potential to both detect novel and confirm hypothesized causal pathways and disease mechanisms (Wang *et al.*, 2007). A practical example where multiple genes in a common pathway influence a complex reproductive trait is twinning (multiple ovulation) frequency, where variation in twinning frequency is influenced by mutations in at least three genes from the intra-ovarian bone morphogenetic signalling pathway (Galloway *et al.*, 2000; Wilson *et al.*, 2001; Hanrahan *et al.*, 2004; Palmer *et al.*, 2006).

Recent experience in the field of breast cancer research provides a good model of how studies might proceed to find the genes contributing to endometriosis (Cox *et al.*, 2007; Easton *et al.*, 2007; Hunter *et al.*, 2007a). Studies must be carefully designed, ensuring case definitions are similar, that large sample sizes are used, and that replication is tested in multiple study populations. No individual groups have sufficient samples to provide convincing evidence for association alone and international consortia must combine data across a range of studies and population samples (Cox *et al.*, 2007; Easton *et al.*, 2007; Hunter *et al.*, 2007a). This combined effort has provided the power to demonstrate convincing evidence for variants with relatively small effects and to replicate findings from the GWA studies (Cox *et al.*, 2007; Easton *et al.*, 2007; Hunter *et al.*, 2007a). Similar strategies are being used in a number of other complex diseases.

It is clear that endometriosis research groups will have to adopt comparable strategies and work closely together to combine data from as many samples as possible to successfully identify genes contributing to this disease. The IEC combining sample sets from Australia, the UK and USA has recently been awarded funding from the Australian National Health and Medical Research Council and The Wellcome Trust in the UK to conduct a GWA study in over 3000 cases and conduct replication studies of the key variants in another 3000 cases and 3000 controls. Results are likely to be similar to other complex diseases and the risk for individual alleles will be small. Despite limitations in results from GWA studies, convincing evidence for variants increasing risk of endometriosis will help to define which pathways contribute to the disease. Subsequent investigations can address whether there is evidence for disease heterogeneity with endometriosis arising through different pathways and also examine interactions between genetic variants and environment. Given the small effect sizes, it is less likely that genetic tests will be used directly to assess individual risk of disease, but knowledge of pathways to disease may help to develop better diagnostic methods.

Summary and conclusions

There is good evidence for a genetic contribution to the risk of developing endometriosis. Gene mapping studies provide an important alternative to biological studies for determining pathways and mechanisms of disease. However, endometriosis is a complex trait. Both theoretical and empirical evidence suggest many genes or variants with small effects are likely to account for the genetic risk. Consequently, powerful, well-designed studies are essential. The tools are now available to find genes predisposing to endometriosis. For this to be successful, we must follow the lead of studies in other diseases. The endometriosis research community must work together to fund the necessary GWA studies and conduct replication studies in many thousands of cases and controls. Convincing evidence for genes associated with endometriosis will provide the starting point for functional and biological studies to develop better diagnosis and treatment for this debilitating disease.

Supplementary Data

Supplementary data are available at <http://humupd.oxfordjournals.org/>

Funding

This study was supported by grants to GWM from the National Institute of Child Health and Human Development (HD050537) and National Health and Medical Research Council of Australia (339430 and 339446).

References

- Ammendola M, Bottini N, Pietropolli A, Saccucci P, Gloria-Bottini F. Association between PTPN22 and endometriosis. *Fertil Steril* 2008;**89**:993–994.
- Berkley KJ, Rapkin AJ, Papka RE. The pains of endometriosis. *Science* 2005;**308**:1587–1589.
- Birnbaum LS, Cummings AM. Dioxins and endometriosis: a plausible hypothesis. *Environ Health Perspect* 2002;**110**:15–21.
- Bulun SE, Gurates B, Fang Z, Tamura M, Sebastian S, Zhou J, Amin S, Yang S. Mechanisms of excessive estrogen formation in endometriosis. *J Reprod Immunol* 2002;**55**:21–33.
- Consortium TIH. A haplotype map of the human genome. *Nature* 2005;**437**:1299–1320.
- Consortium TIH. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;**447**:661–678.
- Cox A, Dunning AM, Garcia-Closas M, Balasubramanian S, Reed MW, Pooley KA, Scollen S, Baynes C, Ponder BA, Chanock S *et al.* A common coding variant in CASP8 is associated with breast cancer risk. *Nat Genet* 2007;**39**:352–358.
- Daftary GS, Taylor HS. EMX2 gene expression in the female reproductive tract and aberrant expression in the endometrium of patients with endometriosis. *J Clin Endocrinol Metab* 2004;**89**:2390–2396.
- de Roux N, Genin E, Carel JC, Matsuda F, Chaussain JL, Milgrom E. Hypogonadotropic hypogonadism due to loss of function of the KiSS1-derived peptide receptor GPR54. *Proc Natl Acad Sci USA* 2003;**100**:10972–10976.
- Di W, Guo SW. The search for genetic variants predisposing women to endometriosis. *Curr Opin Obstet Gynecol* 2007;**19**:395–401.
- Ding Y, Chen ZF, Lin RY, Wang XF, Ding JB, Ai XZ, Wen H. Relationship between endometriosis and glutathione S-transferase M1, T1 genes of the Uyghurs and Hans in Xinjiang. *Zhonghua Fu Chan Ke Za Zhi* 2004;**39**:101–104.
- Dinulescu DM, Ince TA, Quade BJ, Shafer SA, Crowley D, Jacks T. Role of K-ras and Pten in the development of mouse models of endometriosis and endometrioid ovarian cancer. *Nat Med* 2005;**11**:63–70.
- Easton DF, Pooley KA, Dunning AM, Pharoah PD, Thompson D, Ballinger DG, Struwing JP, Morrison J, Field H, Luben R *et al.* Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 2007;**447**:1087–1093.
- Eskenazi B, Warner ML. Epidemiology of endometriosis. *Obstet Gynecol Clin North Am* 1997;**24**:235–258.
- Eskenazi B, Mocarelli P, Warner M, Samuels S, Vercellini P, Olive D, Needham LL, Patterson DG Jr, Brambilla P, Gavoni N *et al.* Serum dioxin concentrations and endometriosis: a cohort study in Seveso, Italy. *Environ Health Perspect* 2002;**110**:629–634.
- Falconer H, D'Hooghe T, Fried G. Endometriosis and genetic polymorphisms. *Obstet Gynecol Surv* 2007;**62**:616–628.
- Flores I, Mandal DM, Bailey-Wilson JE. *Search for Endometriosis Susceptibility Genes in Puerto Rico*. Toronto: American Society of Human Genetics, 2004.
- Fujii H, Matsumoto T, Yoshida M, Furugen Y, Takagaki T, Iwabuchi K, Nakata Y, Takagi Y, Moriya T, Ohtsuji N *et al.* Genetics of synchronous uterine and ovarian endometrioid carcinoma: combined analyses of loss of heterozygosity, PTEN mutation, and microsatellite instability. *Hum Pathol* 2002;**33**:421–428.
- Galloway SM, McNatty KP, Cambridge LM, Laitinen MPE, Juengel JL, Jokiranta S, McLaren RJ, Luiro K, Dodds KG, Montgomery GW *et al.* Mutations in an oocyte-derived growth factor gene (BMP15) cause increased ovulation rate and infertility in a dosage-sensitive manner. *Nat Genet* 2000;**25**:279–283.
- Gao X, Outley J, Botteman M, Spalding J, Simon JA, Pashos CL. Economic burden of endometriosis. *Fertil Steril* 2006;**86**:1561–1572.
- Giudice LC, Kao LC. Endometriosis. *Lancet* 2004;**364**:1789–1799.
- Gudmundsson J, Sulem P, Manolescu A, Amundadottir LT, Gudbjartsson D, Helgason A, Rafnar T, Bergthorsson JT, Agnarsson BA, Baker A *et al.*

- Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. *Nat Genet* 2007;**39**:631–637.
- Guo SW. The link between exposure to dioxin and endometriosis: a critical reappraisal of primate data. *Gynecol Obstet Invest* 2004;**57**:157–173.
- Guo SW. Glutathione S-transferases M1/T1 gene polymorphisms and endometriosis: a meta-analysis of genetic association studies. *Mol Hum Reprod* 2005;**11**:729–743.
- Guo SW. The association of endometriosis risk and genetic polymorphisms involving dioxin detoxification enzymes: a systematic review. *Eur J Obstet Gynecol Reprod Biol* 2006a;**124**:134–143.
- Guo SW. Association of endometriosis risk and genetic polymorphisms involving sex steroid biosynthesis and their receptors: a meta-analysis. *Gynecol Obstet Invest* 2006b;**61**:90–105.
- Guo Y, Shen H, Xiao P, Xiong DH, Yang TL, Guo YF, Long JR, Recker RR, Deng HW. Genomewide linkage scan for quantitative trait loci underlying variation in age at menarche. *J Clin Endocrinol Metab* 2006;**91**:1009–1014.
- Hadfield RM, Mardon HJ, Barlow DH, Kennedy SH. Endometriosis in monozygotic twins. *Fertil Steril* 1997;**68**:941–942.
- Hanrahan JP, Gregan SM, Mulsant P, Mullen M, Davis GH, Powell R, Galloway SM. Mutations in the genes for oocyte-derived GDF9 and BMP15 are associated with both increased ovulation rate and sterility in Cambridge and Belclare sheep (*Ovis aries*). *Biol Reprod* 2004;**70**:900–909.
- Healy DL, Rogers PA, Hii L, Wingfield M. Angiogenesis: a new theory for endometriosis. *Hum Reprod Update* 1998;**4**:736–740.
- Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K. A comprehensive review of genetic association studies. *Genet Med* 2002;**4**:45–61.
- Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, Hankinson SE, Wacholder S, Wang Z, Welch R, Hutchinson A *et al.* A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet* 2007a;**39**:870–874.
- Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, Hankinson SE, Wacholder S, Wang Z, Welch R, Hutchinson A *et al.* A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet* 2007b;**39**:870–874.
- Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. Replication validity of genetic association studies. *Nat Genet* 2001;**29**:306–309.
- Ivashchenko TE, Shved N, Kramareva NA, Ailamazian EK, Baranov VS. Analysis of the polymorphic alleles of genes encoding phase 1 and phase 2 detoxication enzymes in patients with endometriosis. *Genetika* 2003;**39**:525–529.
- Kennedy S. The genetics of endometriosis. *J Reprod Med* 1998;**43**:263–268.
- Kennedy S, Mardon H, Barlow D. Familial endometriosis. *J Assist Reprod Genet* 1995;**12**:32–34.
- Kennedy S, Hadfield R, Barlow D, Weeks DE, Laird E, Golding S. Use of MRI in genetic studies of endometriosis. *Am J Med Genet* 1997;**71**:371–372.
- Koninckx PR, Barlow D, Kennedy S. Implantation versus infiltration: the Sampson versus the endometriotic disease theory. *Gynecol Obstet Invest* 1999;**47**:3–9.
- Kruglyak L, Lander ES. Complete multipoint sib-pair analysis of qualitative and quantitative traits. *Am J Hum Genet* 1995a;**57**:439–454.
- Kruglyak L, Lander ES. High-resolution genetic mapping of complex traits. *Am J Hum Genet* 1995b;**56**:1212–1223.
- Kurose K, Zhou XP, Araki T, Cannistra SA, Maher ER, Eng C. Frequent loss of PTEN expression is linked to elevated phosphorylated Akt levels, but not associated with p27 and cyclin D1 expression, in primary epithelial ovarian carcinomas. *Am J Pathol* 2001;**158**:2097–2106.
- Kurose K, Gilley K, Matsumoto S, Watson PH, Zhou XP, Eng C. Frequent somatic mutations in PTEN and TP53 are mutually exclusive in the stroma of breast carcinomas. *Nat Genet* 2002;**32**:355–357.
- Kyama MC, D'Hooghe TM, Debrock S, Machoki J, Chai DS, Mwenda JM. The prevalence of endometriosis among African-American and African-indigenous women. *Gynecol Obstet Invest* 2004;**57**:40–42.
- Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 1995;**11**:241–247.
- Latta E, Chapman WB. PTEN mutations and evolving concepts in endometrial neoplasia. *Curr Opin Obstet Gynecol* 2002;**14**:59–65.
- Libioulle C, Louis E, Hansoul S, Sandor C, Farnir F, Franchimont D, Vermeire S, Dewit O, de Vos M, Dixon A *et al.* Novel Crohn disease locus identified by genome-wide association maps to a gene desert on 5p13.1 and modulates expression of PTGER4. *PLoS Genet* 2007;**3**:e58.
- Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet* 2003;**33**:177–182.
- Martini M, Ciccarone M, Garganese G, Maggiore C, Evangelista A, Rahimi S, Zannoni G, Vittori G, Larocca LM. Possible involvement of hMLH1, p16(INK4a) and PTEN in the malignant transformation of endometriosis. *Int J Cancer* 2002;**102**:398–406.
- Mathias SD, Kuppermann M, Liberman RF, Lipschutz RC, Steege JF. Chronic pelvic pain: prevalence, health-related quality of life, and economic correlates. *Obstet Gynecol* 1996;**87**:321–327.
- Maxwell GL, Risinger JI, Gumbs C, Shaw H, Bentley RC, Barrett JC, Berchuck A, Futreal PA. Mutation of the PTEN tumor suppressor gene in endometrial hyperplasias. *Cancer Res* 1998;**58**:2500–2503.
- McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR, Hinds DA, Pennacchio LA, Tybjaerg-Hansen A, Folsom AR *et al.* A common allele on chromosome 9 associated with coronary heart disease. *Science* 2007;**316**:1488–1491.
- Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q, Cochran C, Bennett LM, Ding W *et al.* A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* 1994;**266**:66–71.
- Missmer SA, Cramer DW. The epidemiology of endometriosis. *Obstet Gynecol Clin North Am* 2003;**30**:1–19.
- Missmer SA, Hankinson SE, Spiegelman D, Barbieri RL, Marshall LM, Hunter DJ. Incidence of laparoscopically confirmed endometriosis by demographic, anthropometric, and lifestyle factors. *Am J Epidemiol* 2004;**160**:784–796.
- Moen MH. Endometriosis in monozygotic twins. *Acta Obstet Gynecol Scand* 1994;**73**:59–62.
- Palmer JS, Zhao ZZ, Hoekstra C, Hayward NK, Webb PM, Whiteman DC, Martin NG, Boomsma DI, Duffy DL, Montgomery GW. Novel variants in growth differentiation factor 9 in mothers of twins. *J Clin Endocrinol Metab* 2006;**91**:4713–4716.
- Peltonen L. Old suspects found guilty—the first genome profile of multiple sclerosis. *N Engl J Med* 2007;**357**:927–929.
- Peng DX, He YL, Qiu LW, Yang F, Lin JM. Association between glutathione S-transferase M1 gene deletion and genetic susceptibility to endometriosis. *Di Yi Jun Yi Da Xue Xue Bao* 2003;**23**:458–459, 462.
- Petretto E, Liu ET, Aitman TJ. A gene harvest revealing the archeology and complexity of human disease. *Nat Genet* 2007;**39**:1299–1301.
- Pollock PM, Gartside MG, Dejeza LC, Powell MA, Mallon MA, Davies H, Mohammadi M, Futreal PA, Stratton MR, Trent JM *et al.* Frequent activating FGFR2 mutations in endometrial carcinomas parallel germline mutations associated with craniosynostosis and skeletal dysplasia syndromes. *Oncogene* 2007;**26**:7158–7162.
- Rier SE, Turner WE, Martin DC, Morris R, Lucier GW, Clark GC. Serum levels of TCDD and dioxin-like chemicals in Rhesus monkeys chronically exposed to dioxin: correlation of increased serum PCB levels with endometriosis. *Toxicol Sci* 2001;**59**:147–159.
- Rioux JD, Xavier RJ, Taylor KD, Silverberg MS, Goyette P, Huett A, Green T, Kuballa P, Barmada MM, Datta LW *et al.* Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nat Genet* 2007;**39**:596–604.
- Risch N. Linkage strategies for genetically complex traits. I. Multilocus models. *Am J Hum Genet* 1990a;**46**:222–228.
- Risch N. Linkage strategies for genetically complex traits. III. The effect of marker polymorphism on analysis of affected relative pairs. *Am J Hum Genet* 1990b;**46**:242–253.
- Rothenbuhler A, Fradin D, Heath S, Lefevre H, Bouvattier C, Lathrop M, Bougnère P. Weight-adjusted genome scan analysis for mapping quantitative trait Loci for menarchal age. *J Clin Endocrinol Metab* 2006;**91**:3534–3537.
- Sampson JA. Peritoneal endometriosis due to the menstrual dissemination of endometrial tissue to the peritoneal cavity. *Am J Obstet Gynecol* 1927;**14**:422–469.
- Sato N, Tsunoda H, Nishida M, Morishita Y, Takimoto Y, Kubo T, Noguchi M. Loss of heterozygosity on 10q23.3 and mutation of the tumor suppressor gene PTEN in benign endometrial cyst of the ovary: possible sequence progression from benign endometrial cyst to endometrioid carcinoma and clear cell carcinoma of the ovary. *Cancer Res* 2000;**60**:7052–7056.
- Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, Roix JJ, Kathiresan S, Hirschhorn JN, Daly MJ *et al.* Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 2007;**316**:1331–1336.
- Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, Erdos MR, Stringham HM, Chines PS, Jackson AU *et al.* A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* 2007;**316**:1341–1345.
- Seminara SB, Messager S, Chatzidakis EE, Thresher RR, Acierno JS Jr, Shagoury JK, Bo-Abbas Y, Kuohung W, Schwino KM, Hendrick AG

- et al.* The GPR54 gene as a regulator of puberty. *N Engl J Med* 2003;**349**:1614–1627.
- Simoens S, Hummelshoj L, D'Hooghe T. Endometriosis: cost estimates and methodological perspective. *Hum Reprod Update* 2007;**13**:395–404.
- Simpson JL, Bischoff FZ. Heritability and molecular genetic studies of endometriosis. *Ann N Y Acad Sci* 2002;**955**:239–251.
- Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, Boutin P, Vincent D, Belisle A, Hadjadj S *et al.* A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 2007;**445**:881–885.
- Stefansson H, Geirsson RT, Steinthorsdottir V, Jonsson H, Manolescu A, Kong A, Ingadottir G, Gulcher J, Stefansson K. Genetic factors contribute to the risk of developing endometriosis. *Hum Reprod* 2002;**17**:555–559.
- Swiersz LM. Role of endometriosis in cancer and tumor development. *Ann N Y Acad Sci* 2002;**955**:281–292.
- The American Fertility Society. Revised American Fertility Society classification of endometriosis: 1985. *Fertil Steril* 1985;**43**:351–352.
- Todd JA, Walker NM, Cooper JD, Smyth DJ, Downes K, Plagnol V, Bailey R, Nejentsev S, Field SF, Payne F *et al.* Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. *Nat Genet* 2007;**39**:857–864.
- Treloar SA, Do KA, O'Connor VM, O'Connor DT, Yeo MA, Martin NG. Predictors of hysterectomy: an Australian study. *Am J Obstet Gynecol* 1999a;**180**:945–954.
- Treloar SA, O'Connor DT, O'Connor VM, Martin NG. Genetic influences on endometriosis in an Australian twin sample. *Fertil Steril* 1999b;**17**:701–710.
- Treloar S, Hadfield R, Montgomery G, Lambert A, Wicks J, Barlow DH, O'Connor DT, Kennedy S, Group IES. The International Endogene Study: a collection of families for genetic research in endometriosis. *Fertil Steril* 2002;**78**:679–685.
- Treloar SA, Wicks J, Nyholt DR, Montgomery GW, Bahlo M, Smith V, Dawson G, Mackay IJ, Weeks DE, Bennett ST *et al.* Genomewide linkage study in 1,176 affected sister pair families identifies a significant susceptibility locus for endometriosis on chromosome 10q26. *Am J Hum Genet* 2005a;**77**:365–376.
- Treloar SA, Zhen Zhao Z, Armitage T, Duffy DL, Wicks J, O'Connor DT, Martin NG, Montgomery GW. Association between polymorphisms in the progesterone receptor gene and endometriosis. *Mol Hum Reprod* 2005b;**11**:641–647.
- Treloar SA, Zhao ZZ, Le L, Zondervan KT, Martin NG, Kennedy S, Nyholt DR, Montgomery GW. Variants in EMX2 and PTEN do not contribute to risk of endometriosis. *Mol Hum Reprod* 2007;**13**:587–594.
- Troy PJ, Daftary GS, Bagot CN, Taylor HS. Transcriptional repression of peri-implantation EMX2 expression in mammalian reproduction by HOXA10. *Mol Cell Biol* 2003;**23**:1–13.
- Varma R, Rollason T, Gupta JK, Maher ER. Endometriosis and the neoplastic process. *Reproduction* 2004;**127**:293–304.
- Viganò P, Infantino M, Lattuada D, Lauletta R, Ponti E, Somigliana E, Vignali M, DiBlasio AM. Intercellular adhesion molecule-1 (ICAM-1) gene polymorphisms in endometriosis. *Mol Hum Reprod* 2003;**9**:47–52.
- Vinatier D, Orazi G, Cosson M, Dufour P. Theories of endometriosis. *Eur J Obstet Gynecol Reprod Biol* 2001;**96**:21–34.
- Visscher PM, Medland SE, Ferreira MAR, Morley KI, Zhu G, Cornes BK, Montgomery GW, Martin NG. Assumption free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. *PLoS Genet* 2006;**2**:e41.
- Wang K, Li M, Bucan M. Pathway-based approaches for analysis of genomewide association studies. *Am J Hum Genet* 2007;**81**:1278–1283.
- Wilson T, Wu X-Y, Juengel JL, Ross IK, Lumsden JM, Lord EA, Dodds KG, Walling G, McEwan J, O'Connell AR *et al.* Highly prolific Booroola sheep have a mutation in the intracellular kinase domain of bone morphogenetic protein 1B receptor which is expressed in both oocytes and granulosa cells. *Biol Reprod* 2001;**64**:1225–1235.
- Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, Collins N, Gregory S, Gumbs C, Micklem G. Identification of the breast cancer susceptibility gene BRCA2. *Nature* 1995;**378**:789–792.
- Yeager M, Orr N, Hayes RB, Jacobs KB, Kraft P, Wacholder S, Minichiello MJ, Fearnhead P, Yu K, Chatterjee N *et al.* Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. *Nat Genet* 2007;**39**:645–659.
- Zhao ZZ, Nyholt DR, Le L, Thomas S, Engwerda C, Randall L, Treloar SA, Montgomery GW. Genetic variation in tumour necrosis factor and lymphotoxin is not associated with endometriosis in an Australian sample. *Hum Reprod* 2007;**22**:2389–2397.
- Zhao ZZ, Pollock PM, Thomas S, Treloar SA, Nyholt DR, Montgomery GW. Common variation in the fibroblast growth factor receptor 2 gene is not associated with endometriosis risk. *Hum Reprod* 2008; in press.
- Zhou XP, Kuismanen S, Nystrom-Lahti M, Peltomaki P, Eng C. Distinct PTEN mutational spectra in hereditary non-polyposis colon cancer syndrome-related endometrial carcinomas compared to sporadic microsatellite unstable tumors. *Hum Mol Genet* 2002;**11**:445–450.
- Zollner S, Pritchard JK. Overcoming the winner's curse: estimating penetrance parameters from case-control data. *Am J Hum Genet* 2007;**80**:605–615.
- Zondervan KT, Cardon LR. The complex interplay among factors that influence allelic association. *Nat Rev Genet* 2004;**5**:89–100.
- Zondervan KT, Cardon LR. Designing candidate gene and genome-wide case-control association studies. *Nat Protoc* 2007;**2**:2492–2501.
- Zondervan KT, Cardon LR, Kennedy SH. The genetic basis of endometriosis. *Curr Opin Obstet Gynecol* 2001;**13**:309–314.
- Zondervan KT, Cardon LR, Kennedy SH. What makes a good case-control study? Design issues for complex traits such as endometriosis. *Hum Reprod* 2002;**17**:1415–1423.
- Zondervan KT, Weeks DE, Colman R, Cardon L, Hadfield RH, Scheffler J, Trainor AG, Coe CL, Kemnitz JW, Kennedy SH. Familial aggregation of endometriosis in a large pedigree of rhesus macaques. *Hum Reprod* 2004;**19**:448–455.
- Zondervan KT, Treloar SA, Lin J, Weeks DE, Nyholt DR, Mangion J, MacKay IJ, Cardon L, Martin NG, Kennedy SH *et al.* Significant evidence of one or more susceptibility loci for endometriosis with near-Mendelian autosomal inheritance on chromosome 7p13-15. *Hum Reprod* 2007;**22**:717–728.

Submitted on January 22, 2008; resubmitted on April 5, 2008; accepted on April 21, 2008