Insights into Assessing the Genetics of Endometriosis

Nilufer Rahmioglu · Stacey A. Missmer · Grant W. Montgomery · Krina T. Zondervan

Abstract Endometriosis is a complex disease arising from the interplay between multiple genetic and environmental factors. The genetic variants potentially underlying the hereditary component of endometriosis have been widely investigated through hypothesis-driven candidate gene studies, an approach that generally has proven to be inherently difficult and problematic for a number of reasons. Recently, through major collaborative efforts in the endometriosis research field, hypothesis-free genome-wide approaches have started to provide new insights into potential pathways leading to development of endometriosis, as well as highlighting the phenotypic heterogeneity of the condition. This review summarizes the most recent studies investigating the genetic variation contributing to endometriosis, with a particular focus on genome-wide approaches, and discusses promising future directions of genetic research.

Keywords Endometriosis · Genetics · Heritability · Genetic epidemiology · Candidate gene · Genome-wide association study · Linkage study

N. Rahmioglu · K. T. Zondervan (✉)
Wellcome Trust Centre for Human Genetics, University of Oxford,
Oxford, UK
e-mail: krinaz@well.ox.ac.uk

S. A. Missmer
Department of Obstetrics, Gynecology, and Reproductive Biology,
Brigham and Women's Hospital and Harvard Medical School,
Boston, MA, USA

S. A. Missmer
Channing Laboratory, Department of Medicine,
Brigham and Women's Hospital and Harvard Medical School,
Boston, MA, USA

S. A. Missmer
Department of Epidemiology, Harvard School of Public Health,
Boston, MA, USA

G. W. Montgomery
Molecular Epidemiology,
Queensland Institute for Medical Research,
Brisbane, Australia

K. T. Zondervan
Nuffield Department of Obstetrics and Gynaecology,
University of Oxford,
Oxford, UK

Introduction

Endometriosis is a common estrogen (E)-dependent gynecological disorder, characterized by presence of tissue resembling endometrium tissue outside the uterine cavity. Typical locations for the ectopic deposits are on the pelvic peritoneum, the ovaries, uterosacral ligaments, pouch of Douglas, and the rectovaginal septum. Endometriosis has been associated with chronic pelvic pain, reduced fertility, and severe dysmenorrhea [1, 2]. Definitive diagnosis can only be established through surgery, and, though other classifications exist, the disease is most commonly staged using the revised American Fertility Society (rAFS) classification, based on the total surface size of the lesions, presence of adhesions, and ovarian lesions [3]. Because endometriosis is E-dependent, it occurs almost exclusively in women of reproductive age [4]. Because of the need for surgical diagnosis, the exact population prevalence is unknown. Based on community prevalence rates of pelvic pain and infertility, the prevalence of endometriosis is estimated to be about 5–10 % in premenopausal women [5], increasing to 35–50 % in women undergoing laparoscopy for pelvic pain and infertility [1, 6].

The etiology of endometriosis remains unclear. The most widely accepted mechanism for development of the peritoneal endometriotic lesions is via retrograde menstruation [7]. However, menstrual debris is present in the peritoneal cavity of up to 90 % of the menstruating women [8]. Possible explanations for adhesion and growth of endometriotic lesions in only some

women include increased exposure to menstrual debris (e.g., increased menstrual flow, shorter cycle length), abnormal eutopic endometrium, altered peritoneal environment, reduced immune surveillance, and increased angiogenic capacity [9–11]. It is possible that several, if not all, of these factors play a role to an extent, and/or that specific subtypes of disease are due to specific underlying biological pathways; research continues to elucidate these mechanisms.

One approach to study the underlying biological pathways leading to a complex disease such as endometriosis (the development of which is determined by multiple genetic and environmental factors) is to study the effect of genetic variants on disease causation. The involvement of genetic factors in the development of endometriosis is supported by different studies [12–14]. Twin studies have shown increased concordance in monozygotic twins when compared to dizygotic twins [15, 16]; the largest such study carried out to date among Australian 3,096 female twins concluded that about 51 % of the variation in endometriosis risk is heritable [16]. In addition to human studies, familial aggregation of spontaneous (i.e., non-induced) endometriosis also has been shown in nonhuman primates such as the rhesus macaque [17].

Study designs to search for genes underlying endometriosis can be separated into hypothesis-based and hypothesis-free approaches. Hypothesis-based "candidate gene" studies typically rely on prior biological hypotheses/knowledge of the disease or of the approximate genomic location of a disease-predisposing variant derived from a hypothesis-free study. Hypothesis-free studies search the entire genome to identify disease-predisposing variants, without knowledge of their functional relevance. This review summarizes the evidence from the most recent studies investigating the genetic variation contributing to endometriosis. Of particular focus will be the results from large-scale collaborative genome-wide approaches, which have started to provide new insights into potential pathways leading to endometriosis.

Hypothesis-Based Research: Candidate Gene Association Studies

Historically, the search of genes contributing to susceptibility of many complex diseases such as endometriosis began with hypothesis-driven "candidate gene association" studies (CGASs). Although hypothesis-free approaches have since taken off, candidate gene studies based on a biological "hunch" alone are still commonplace. CGASs are based on the a priori selection of genes with inferred biological function and association of variants in these genes with disease risk. With limited budgets and faced with over 20,000 genes in the human genome, the approach seems an attractive one. However, CGASs are inherently limited by current knowledge of biological mechanisms underlying the studied phenotype,

which is attenuated in diseases where the underlying mechanisms are not very well understood, such as in endometriosis. Furthermore, to fully test the involvement of a biological pathway, one would need to investigate all genes making up the pathway (as well as factors regulating their expression), in a large sample of cases and controls; a costly approach that has, to our knowledge, never been adopted. Therefore, the probability of success of a CGAS as commonly conducted, based on the testing of a few, selected, variants in a handful of genes, is extremely low. This probability may be greatly improved if evidence of the likely genomic location of a disease-predisposing variant, derived from a whole genome linkage study for example, is included, although this approach has its own drawbacks (described later in this article).

Association studies these days typically focus on the most abundant genetic variants in the genome: single nucleotide polymorphisms (SNPs), which comprise over 90 % of all common genetic variation in the human genome [18]. SNPs are changes to the DNA at one nucleotide base pair, and can have two identities, or "alleles." Association studies compare the frequency of the alleles of a genetic variant between cases and controls, and can be of direct and indirect design. Direct association studies aim to test specific genetic variants by genotyping these, whereas indirect association studies type a set of common (population allele frequency > 0.05) SNPs carefully selected on the basis of local genomic "linkage disequilibrium" (LD). LD is the nonrandom association of genetic variants in a population. This phenomenon means, in simple terms, that an SNP can predict the status of a genetic variant nearby because of common ancestry of that particular genomic segment. The International HapMap Project [19] has mapped LD patterns across the human genome in a number of ethnic populations, which has enabled the investigation of whole sections of the genome for their association with many common complex diseases; by testing those SNPs that are highly predictive of the status of other SNPs in the region (termed "tagSNPs"), all common genetic variation present in the population can be tested with very high coverage. Genome-wide association studies (GWASs) are of indirect design; they utilize LD by typically genotyping ~500,000–1,000,000 SNPs that tag a large proportion of the roughly 10 million common SNPs present in the human genome. Most CGASs of endometriosis have been of direct design (i.e., a limited number of specific genetic variants were genotyped and tested for association, although indirect tag approaches have also been used [20]). For a detailed review of the design of CGASs and GWASs and their underlying principles, see Zondervan & Cardon [21].

We previously published a review of CGASs published up to April 1, 2008 [20]. For this paper, we identified CGASs of endometriosis published from April 1, 2008 to April 1, 2012, conducting a systematic literature search in

PubMed for English language publications, using the terms "endometriosis" with "genetics" or "genes" or "polymorphisms." An overview of these recent CGASs is presented in Table 1. Variants from 73 candidate genes were tested (Table 1), involved in sex steroid biosynthesis and signaling pathways, adhesion molecules and matrix enzymes, immunological mechanisms, inflammatory pathways, estradiol metabolism, growth factor systems, cell cycle regulation, oncogenes, apoptosis, and angiogenic factors [22].

For evidence of association to be credible, it is paramount that the observed associations should be replicated in an independent sample from the same underlying ethnic population [21, 23]. Generally, many reported associations do not show significant association in independent studies. This can be due to many factors, such as differences in disease definition, inadequate control selection, low coverage of the candidate genes, small sample size of the initial dataset (creating a false-positive result) or of the replication dataset (creating a false-negative result), or population differences between the discovery and the replication datasets. Moreover, to avoid failure to replicate due to Winner's curse (the statistical phenomenon in which the effect size [the odds ratio] of the association in the discovery dataset is larger than in any subsequent replication dataset), the replication studies need to be larger in sample size and have greater power than the discovery study to detect the effect of the putative association [24].

As can be seen from the results in Table 1, no candidate gene has been robustly associated with endometriosis across studies. One of the most frequently investigated and plausible genes, encoding for vascular endothelial growth factor (*VEGF*), probably shows the most suggestive evidence of association. A literature-based meta-analysis of five polymorphisms from 11 studies including 1785 cases and 1879 control patients of Chinese, Indian, Korean, Japanese, Spanish, Turkish, Estonian, and Australian origin suggested one significantly associated polymorphism (+936T/C; TT+TC vs CC: $P=0.02$) [25•], although no such association was found in a large study of 958 Australian cases and 959 control patients ($P=0.31$) [26•]. Similarly, our previous review of 76 studies published until April 1, 2008 concluded that none provided clear support for any genetic variant to be robustly associated with endometriosis, and that some reported results may represent true associations, but that given the small effect sizes expected, large replication datasets or meta-analyses are required. In a large GWAS of 3,194 cases and 7,060 control patients of European ancestry, all candidate genes from the 2008 Montgomery et al. review [20] were investigated for nominal evidence of association [27•]; the only gene with a nominal $P<10^{-3}$ for SNPs in the GWAS data was the gene encoding the progesterone receptor (*PGR*) on chromosome 11, but the result for the SNP in this gene was not significant in the replication stage. This does not necessarily mean that none of these candidate genes are involved in the causation of (subtypes) of endometriosis; it also could mean that even in a large study of all types of endometriosis, power is lacking to detect their effect.

As can be seen in Table 1, the recent CGASs were conducted in ethnically diverse populations. Because different populations may have different genetic contributions to a disease, and LD patterns may vary between ethnic populations, a genetic variant identified as causal in one population may not be acting as a susceptibility variant in another. Therefore, care should be taking in interpreting and comparing findings in terms of "replication."

Hypothesis-Free Research: Whole Genome Linkage Studies and Their Follow-up

One of the two hypothesis-free approaches to investigate underlying genetic etiology of a disease is linkage mapping, which adopts genetic investigation of families with multiple affected individuals. This methodology became popular in the 1980s after the discovery of highly variable genetic markers ("microsatellites"), and works by 1) genotyping about 400 microsatellites across the genome; 2) comparing whether the allelic status of these variants are shared between affected people in a family; 3) summing this evidence of sharing across multiple families; 4) calculating whether the amount of sharing is more than expected by chance based on simple Mendelian laws of inheritance; and 5) identifying those genomic regions where excess sharing is statistically significant. The linkage approach was highly successful in identifying genetic variants responsible for rare, monogenic disorders. Such disorders show clear Mendelian segregation patterns through families [28, 29], and the presence of the causal genetic variant determines whether or not the disease in question will develop. The genetic causes of many of these monogenic diseases could be 'mapped' by investigating a handful of extended families with many affected members.

Mendelian disorders are very different from complex diseases such as endometriosis, in which one genetic variant only carries susceptibility to disease (not 100 % risk). Many more families with multiple affected women are required to conduct a linkage study of endometriosis. Linkage studies are fundamentally different from association studies; the two approaches are complementary. In contrast to association studies, which focus on SNPs common in the general population, linkage studies are designed to detect disease-causing variants responsible for disease in a family, but that are otherwise rare in the general population. Moreover, because linkage studies use recombination events in families, the resolution of the approach is very large compared to association studies: a significantly linked region typically

Table 1 Candidate gene association studies of endometriosis published from April 2008 to April 2012, and total evidence for these from all studies published until 2012

Candidate gene[**]	Chr	Number of studies with association 2008–2012[*]		
		Significant (n cases, n controls)	Nonsignificant (n cases, n controls)	#Sign/Total up to 2012, n/n[***]
Adhesion molecules and matrix enzymes				
CDH1	16q22	Indian (715, 500) [73] Japanese (511, 498) [74]	–	2/2
ICAM1	19p13	Iranian (84, 120) [75]	Korean (390, 351) [76]	2/5
MMP2	16q12	Estonian (150, 199) [77]	–	2/3
MMP9	20q11-13	Estonian (150, 199) [77]	–	1/2
MUC2	11p15	Taiwanese (195, 196) [78]	–	1/1
MUC4	3q29	Taiwanese (140, 150) [79]	–	1/1
Apoptosis, cell-cycle, DNA repair, and oncogenes				
AKT1	14q32	–	Indian (30, 30) [80]	0/1
APE1	14q11-12	–	Turkish (52, 101) [81]	0/1
CDKN1A	6p21	–	Taiwanese (180, 330) [82]	0/2
CDKN1B	12p13	Brazilian (104, 109) [83]	–	1/1
HOGG1	3p26	–	Turkish (52, 101) [81]	0/1
PI3KCA	3q26	–	Indian (30, 30) [80]	0/1
TP53	17p13	Mexican (151, 204) [84] Caucasian, Asian (749, 857)[a] [85•], Taiwainese (180, 330) [82]	Brazilian (19, 19) [86]	4/9
XPD	19q13	–	Turkish (52, 101) [81]	0/1
XPG	13q33	–	Turkish (52, 101) [81]	0/1
XRCC1	19q13	–	Turkish (52, 101) [81]	0/1
XRCC3	14q32	–	Turkish (52, 101) [81]	0/1
Cytokines/inflammation				
C3	19p13	Puerto Rican (214, 147) [87]	–	1/1
CTLA4	2q33	–	Brazilian (177, 172) [88]	0/2
DR4/5	8p21	–	Korean (138, 214) [89]	0/1
FCRL3	1q21-22	Brazilian (167, 167) [90]	–	0/1
IFNG	12q24	Korean (622, 442) [91] Indian (302, 324) [92]	–	3/3
IL10	1q31-32	Danish (100, 358) [93] Taiwanese (196, 397) [94] Chinese (214, 160) [95]	–	5/6
IL16	15q26	Chinese (230, 203) [96]	–	1/1
IL18	11q22	Turkish (135, 84) [97]	–	1/2
IL1B	2q13-21	–	Turkish (118, 78) [98] Taiwainese (196, 397) [94]	0/4
IL2RB	22q13	–	Korean (237, 164) [99]	1/2
IL6	7p21-15	–	Taiwainese (196, 397) [94] Korean (390, 351) [76]	0/5
LOXL4	10q24	Puerto Rican (214, 147) [87]	–	1/1
NFKB1	4q24	Chinese (206, 365) [100]	–	1/1
PTGS2	1q25	–	Taiwainese (196, 397) [94]	0/1
PTPN22	1q13	Brazilian (140,180) [101] Polish (171, 310) [102]	Caucasian (789, 351) [103]	3/4
TCRB	7q34	–	Italian (70, 120) [104]	0/1
TNFA	6p21	Indian (245, 85) [105]	Korean (105, 101) [106]	1/2
TNFR1	12p13	–	Korean (105, 101) [106]	0/1

Table 1 (continued)

Candidate gene[**]	Chr	Number of studies with association 2008–2012[*]		#Sign/Total up to 2012, n/n[***]
		Significant (n cases, n controls)	Nonsignificant (n cases, n controls)	
TNFR2	1p36	–	Korean (105, 101) [106]	0/2
TNFRSF11B	8q24	–	Korean (138, 214) [89]	0/1
TNFSF13B	13q32-34	Dutch (87, 107) [107]	–	1/1
TRAIL	8p22-21	–	Korean (138, 214) [89]	0/1
Estradiol metabolism				
COMT	22q11	–	Brazilian (198, 168) [108] Caucasian (256, 567) [109]	0/5
HSD17B1	17q11-21	Estonian (150, 199) [110]	Caucasian (256, 567) [109]	3/4
HSD17B2	16q24	–	Caucasian (256, 567) [109]	0/1
Growth factor systems				
BNC2	9p22.2	–	Caucasian (789, 351) [111]	0/1
FGF1	5q31	–	Chinese (421, 421) [112]	0/1
FGF2	4q26	Chinese (421, 421) [112]	–	1/1
IGF1	12q23	–	Korean (128, 108) [113]	0/1
IGF1R	15q26	–	Korean (128, 108) [113]	0/1
IGF2	11p15	Korean (128, 108) [113]	–	1/1
TGFB1	19q13	–	Korean (485, 352) [114] Korean (131, 107) [115]	1/4
VEGF	6p21-12	Turkish (98, 94) [116] Chinese (344, 360) [117] Caucasian (186, 180) [118] Estonian (150, 199) [110] Japanese, Australian, Spanish, Chinese, Estonian (1785, 1879)[b] [25•]	Australian (958, 959)[c] [26•]	10/12
Hormone receptors				
ESR1	6q24-27	Chinese (214, 160) [95] Estonian (150, 199) [110]	Caucasian (256, 567) [109]	7/12
ESR2	14q21-22	Brazilian (108, 210) [119] Brazilian (136, 209) [120] Estonian (150, 199) [110]	Caucasian (256, 567) [109]	4/6
PGR	11q22-23	Caucasian (348, 5820)[d] [121•]	Caucasian (256, 567) [109]	6/9
Human leukocyte antigen system and immune components				
CCL21	9p13	Caucasian (789, 351) [103]	–	1/1
CD40	20q12-13	–	Caucasian (789, 351) [103]	0/1
HLA region	6p21	–	Japanese (89, 136) [122]	0/1
HLA-DRB1	6p21	Caucasian (789, 351) [103]	–	1/2
IRF5	7q32	–	Caucasian (789, 351) [103]	0/1
STAT4	2q32	–	Caucasian (789, 351) [103]	0/1
TRAF1-C5	9q33-34	–	Caucasian (789, 351) [103]	0/1
VDR	12q13	–	Brazilian (132, 195) [123]	0/1
Other enzymes and metabolic systems				
eNOS	7q36	Korean (299, 459) [124]	Indian (232, 210) [125]	2/4
IRS2	13q34	Turkish (135, 135) [126]	–	1/1
LHB	19q13	Brazilian (110, 209) [127]	–	1/3
SERPINE	7q21-22	Brazilian (140, 148) [128]	Italian (368, 329) [129]	2/3
Steroid-synthesizing enzymes, detoxifying enzymes, and receptors				
CYP17A1	10q24	Turkish (46, 39) [130]	Caucasian (256, 567) [109]	1/9

Table 1 (continued)

Candidate gene[**]	Chr	Number of studies with association 2008–2012[*]		#Sign/Total up to 2012, n/n[***]
		Significant (n cases, n controls)	Nonsignificant (n cases, n controls)	
			Italian (104, 86) [131]	
CYP19A1	15q21	Caucasian (256, 567) [109] Italian (104, 86) [131]	–	4/6
CYP1A1	15q24	–	Caucasian (256, 567) [109]	1/7
CYP1A2	15q24	–	Caucasian (256, 567) [109]	0/1
CYP2C19	10q24	Turkish (50, 50) [132]	Turkish (46, 39) [130]	1/2
GSTM1	1p13	Iranian (120, 200) [133]	–	7/13
GSTP1	11q13	–	Korean (260, 164) [134]	1/3
PPARG	3p25	Korean (446, 427) [135]	–	3/3

Chr chromosome

[*] When more than one polymorphism was investigated in a study, the result is indicated as significant if one or more of the variants were reported as significant by the authors. [**] Candidate gene categories are as suggested by Falconer et al. [22]. [***] Total number of candidate studies is summed together with data from our review in 2008 [20]

[a] Literature-based meta-analysis of *P53* codon 72 polymorphism from six published studies of Asian and Caucasian ancestry. The study showed significant association of the polymorphism with endometriosis only in the Asian patients [85•]

[b] Literature-based meta-analysis of five polymorphisms in the *VEGF* gene from 11 studies from Chinese, Indian, Korean, Japanese, Spanish, Turkish, Estonian, and Australian patients. The study identified one significant polymorphism with endometriosis from the overall meta-analysis [25•]

[c] Genotyped and tested the association of four *VEGF* polymorphisms with endometriosis in a large Australian sample. Then, performed a meta-analysis with four additional published studies and showed no evidence for association neither in the genotyped association analysis they have performed, nor in the meta-analysis for any of the four *VEGF* polymorphisms [26•]

[d] An international ovarian cancer consortium, including eight ovarian cancer case-controls studies from Australia, Europe, and the United States, investigated the association of polymorphisms in *PGR* with endometriosis in 348 endometriosis patients and 5812 control patients [121•]

extends across 10–50 Mb and contains hundreds of genes, in contrast to the high resolution of association studies which (depending on the local LD structure) are able to pinpoint a signal to within 10–500 Kb.

The largest genome linkage scan to date was conducted by the International Endogene Study (consisting of 1,176 affected sib-pairs from Australian and United Kingdom families), which identified a region of significant linkage on chromosome 10q26 [30]. In a subsequent analysis involving a subset of 248 families with 3 or more affected members, a second region was found on chromosome 7p13-15 likely due to one or more rare variants following near-Mendelian patterns of inheritance [31]. In a recent study, the linkage peak on chromosome 10q26 region was extensively genotyped using 11,984 SNPs in 1,144 familial cases and 1,190 control patients. The study identified three independent signals in the region, at 96.59 Mb (rs11592737), 105.63 Mb (rs1253130), and 124.25 Mb (rs2250804), with nominally significant evidence of association. However, only rs11592737 in the cytochrome P450 subfamily C (*CYP2C19*) gene replicated ($P=0.04$) in an independent sample of 2,079 cases and 7,060 control patients [32•]. *CYP2C19* is a plausible candidate for endometriosis because it is involved in the metabolism of drugs and E including conversion of E2 to estrone (E1), and the production of E1 and E2 2a- and 16a-hydroxylation metabolites

[33, 34]. Future studies should follow up by investigating novel rare genetic variants in this region and conducting gene expression analyses to further understand the role of *CYP2C19* in endometriosis.

As a first approach to follow-up the significant linkage peak on chromosome 7 region, the coding regions and upstream regulatory regions of three promising candidate genes (*INHBA*, *SFRP4*, and *HOXA10* with known roles in endometrial development) located within or close to the linkage signal were sequenced to identify potential causal polymorphisms. Sequencing, rather than genotyping, was chosen because the signal was likely due to one or more novel rare variants not present on any genotyping array. Sequencing was conducted in 47 cases from the 15 families contributing most to the linkage signal, and 11 variants were found. The minor allele frequencies (MAFs) of observed variants were compared with MAFs from two publicly available reference populations of European ancestry: 60 individuals in HapMap and 150 individuals in the 1000 Genomes Project. Five of the 11 variants were common (MAF>0.05); the remaining six variants were rare and unlikely to be individually or cumulatively responsible for the linkage signal. These results indicate that the coding regions of these three genes do not harbor rare mutations responsible for linkage to endometriosis in these families [35]. However, this does not exclude the possibility that

variants responsible for the linkage signal exist in noncoding regions of these three genes, or that they are located in other genes in this region.

Hypothesis-Free Research: Genome-wide Association Studies

GWASs are based on the premise that common diseases such as endometriosis are caused by genetic variants that are common themselves (Common disease-common variant hypothesis [CDCV]). The first and seminal papers using GWAS methodology successfully were published in 2005–2007 [36–38], and since then the method has taken off, resulting in the detection of many common genetic variants associated with complex diseases. The NHGRI (National Human Genome Research Institute) GWA Catalog (www.genome.gov/GWAStudies) [39] counted the number of published significant genome-wide associations ($P \leq 5 \times 10^{-8}$) up to June 2011 at 1,449 for 237 traits. Key developments that enabled GWASs were (1) the documentation of hundreds of thousands of common SNPs, and the LD patterns between them in different populations in the human genome by the HapMap Consortium [19]; and 2) the development of high-throughput genomics platforms capable of genotyping over 1 million SNPs in one assay at ever decreasing cost.

The first GWASs of endometriosis were published in 2010 and 2011: two in women of Japanese ancestry [40•, 41] and one in women of European ancestry [27•]. Only two [27•, 40•] reported genome-wide significant signals, which were replicated (Table 2). The first Japanese endometriosis GWAS included 1,423 cases and 1,318 control patients in the discovery sample, with cases a mixture of surgically confirmed and clinically diagnosed women. After the application of SNP quality control criteria, 460,945 SNPs were included in the discovery analysis. A replication analysis was conducted in an independent set of 484 cases and 3,974 control patients, in which the top 100 most significant SNPs from the discovery set were genotyped. A significant association for one SNP was found, rs10965235, in *CDKN2BAS* on chromosome 9p21 ($P=6.79 \times 10^{-6}$, odds ratio (OR)=1.56 [1.29–1.89], $P=4.89 \times 10^{-4}$). Combining the discovery and replication samples provided a genome-wide significant result for rs10965235 ($P=5.57 \times 10^{-12}$, OR=1.44 [1.30–1.59]). Rs10965235 is located in intron 6 of *CDKN2BAS,* which encodes for the cyclin-dependent kinase inhibitor 2B antisense RNA (discussed further in this article). Uno et al. [40•] observed a second interesting potential association with rs16826658, only providing suggestive evidence, on chromosome 1p36 in a region close to *WNT4* (combined datasets: $P=1.66 \times 10^{-6}$, OR=1.20 [1.11–1.29]). *WNT4* encodes for wingless-type MMTV integration site family, member 4. Both these variants on chromosome 9p21 and 1p36 are novel susceptibility loci for endometriosis in the Japanese population [40•], and are discussed further in this article.

A second, independent, Japanese GWAS was published recently, comprising a meta-analysis of two GWAS on two case–control datasets. After quality control, 282,828 SNPs were tested for association in 696 endometriosis cases (not all surgically confirmed) and 825 control patients. Limited by their sample size, their aim was to detect potential common susceptibility loci with large effect on the disease. The study did not reveal any significant susceptibility loci for endometriosis, which is likely to be due to the small sample size of their dataset [41].

Table 2 Summary of significant genetic variants discovered by GWASs of endometriosis

Study	Novel loci	Top associated SNP	Position	Chr	Population	Sample size (Ca:Co), *n:n*	Rep	Biological explanation
Uno et al. [40•]	*CDKN2BAS*	rs10965235	Genic	9	Japanese	1,423: 1,318	Yes	Silencing of tumor suppressor gene, *CDKN2A*, by *CDKN2BAS* may have an important role in the development of endometriosis
Painter et al. [27•] and Uno et al. [40•]	*WNT4*	rs16826658	Genic	1	Japanese Caucasian	1,423: 1,318 3,194: 7,060	Yes	Plays a role in development of the female reproductive tract, ovarian follicle development and steroidogenesis
Painter et al. [27•]	*mir-148a, NFE2L3, HOXA11*	rs12700667	Intergenic	7	Caucasian	3,194: 7,060	Yes	Variant associated with WHR as a measure of fat distribution. Fat distribution is also hormonally regulated; suggestive pleiotropic locus

GWAS genome-wide association studies; *SNP* single nucleotide polymorphism; *Ch* chromosome; *Ca* cases; *Co* controls; *Rep* replication; *WHR* Waist to hip ratio; *Ref* reference

The largest GWAS on endometriosis to date was performed in women of European ancestry by the International EndoGene Consortium (IEC) [27•]. The study involved 3,194 surgically confirmed endometriosis cases and 7,060 control patients from Australia and the United Kingdom. Disease severity 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; $n=1,686$, 52.7 %) or stage B (stage III or IV disease; $n=1,364$, 42.7 %), or unknown ($n=144$, 4.6 %). After quality control, analyses were performed using 504,723 SNPs. Analyzing all SNPs combined, Painter et al. [27•] showed a significantly increased genetic loading among 1,364 cases with stage B endometriosis compared to 1,666 with stage A disease (proportion of endometriosis variation explained by common SNPs: 0.34 [SD: 0.04] vs 0.15 [SD: 0.15] respectively; $P=1.8\times10^{-3}$). Because of this result, two GWA analyses were performed, using (1) 3,194 "all" endometriosis cases and (2) 1,364 stage B cases. For "all" endometriosis, the strongest signal observed was rs12700667 in an intergenic region on chromosome 7p15.2 ($P=2.6\times10^{-7}$, OR=1.22 [1.13–1.32]), which was considerably stronger when limiting cases to those with stage B endometriosis ($P=1.5\times10^{-9}$, OR=1.38 [1.24–1.53]). A second strong association was found for rs1250248 (2q35) within *FN1* ($P=3.2\times10^{-8}$). In the replication phase, 70 SNPs that produced nominal evidence of association with "all" or stage B were genotyped in an independent dataset comprising 2,392 self-reported surgically confirmed cases and 2,271 control patients from the Nurses' Health Study I and II in the USA. The association on 7p15.2 with rs12700667 was replicated ($P=1.2\times10^{-3}$, OR=1.17 [1.06–1.28]). However, there was no evidence for replication of rs12540248 (*FN1*) or association with the remaining SNPs. Combined analysis of all 5,586 cases and 9,331 control patients from Australian, UK and US datasets further confirmed association between "all" endometriosis and 7p15.2 (rs12700667, $P=1.4\times10^{-9}$, OR=1.20 [1.13–1.27]). Of note is that the estimated percentage of "all" endometriosis variance explained by rs12700667 was only 0.69 % of the estimated 51 % heritability. Rs12700667 is located in a roughly 924-kb intergenic region containing at least one noncoding RNA (AK057379), predicted transcripts and regulatory elements, and a micro RNA (miRNA [hsa-mir-148a]) about 88 kb upstream (discussed further in this article).

Painter et al. [27•] also investigated the associations reported by Uno et al. [40•] in Japanese women. They found no evidence for association with rs10965235 on chromosome 9p21 (rs10965235 is monomorphic in individuals of European descent, reflecting the different genetic (ancestral) backgrounds between the studies), nor with any SNPs in LD with rs10965235. However, there was evidence for replication of rs7521902 on 1p36, close to *WNT4* gene, with the strongest signal for stage B endometriosis ($P=7.5\times10^{-6}$, OR=1.25,

95 % CI 1.13–1.38); meta-analysis of the evidence from "any" endometriosis (as severity of disease was not assessed in the Japanese GWAS) combining the three GWAS datasets resulted in a genome-wide significant P value of 4.2×10^{-8} (OR=1.19, 95 % CI 1.12–1.27).

GWASs for endometriosis, to date, have been conducted on samples of relatively modest sizes compared to other conditions, but these first studies have shown that there are no common variants with large effects. ORs for replicated, genome-wide significant signals are all lower than 1.5, an observation that mirrors results from the many GWAS performed on complex diseases to date. The results from Painter et al. [27•] suggest that future larger studies enriched for surgically confirmed rAFS stage III/IV cases will be better powered to identify risk loci of endometriosis, but further work, requiring larger samples, needs to be conducted to assess whether specific severe subtypes such as deep infiltrating endometriosis [42] or rectovaginal disease [43] are genetically heterogeneous from each other, or indeed if rAFS stage III/IV is genetically heterogenous with respect to presence of endometriomas or scarring/adhesions. Furthermore, the implicated variants from GWASs are likely to be in LD with the actual disease-causing variants, and further in-depth work to explore genetic variation in the regions, as well as their influence on gene expression and downstream biological pathways, now needs to be conducted. The variants themselves are of small effect and explain only a very small proportion of disease prevalence; hence, they are unsuitable individually to serve as diagnostic markers. However, they provide important data to define the underlying pathways contributing to the disease, can be used to define genetically heterogeneous subtypes that are likely to respond differently to treatments, and thus, through the identification of differential pathways, help in the development of diagnostic tests and future treatments.

Genes Implicated by GWAS of Endometriosis to Date

Although there was no evidence of a signal in women of European ancestry [27], the GWAS by Uno et al. [40•] clearly implicated *CDKN2BAS* as potentially involved in endometriosis in Japanese women (Table 2). *CDKN2BAS* is an interesting candidate because it is expressed in the uterus and regulates the expression of *CDKN2B* (p15), *CDKN2A* (p16), and *ARF* (p14), which are tumor suppressor genes [44–46]. *CDKN2A* is a cell-cycle–dependent kinase inhibitor and acts as a negative cell-cycle regulator [47, 48]. Inactivation of *CDKN2A* has been reported in a subset of endometrial carcinomas [49–52]. Moreover, hypermethylation of the *CDKN2A* promoter region has been observed in endometriosis [52], and loss of heterozygosity on the *CDKN2A* locus has been found in endometriosis, suggesting that *CDKN2A* might play a role in the regulation of endometrial cell growth [53]. The evidence from the literature

suggests that silencing of tumor suppressor genes such as *CDKN2A* by *CDKN2BAS* might have an important role in the development of endometriosis [53].

The second novel genetic variant, which was supported by evidence from GWASs in both women of Japanese and European ancestry, is located near *WNT4*, with evidence from the latter mainly limited to rAFS stage III/IV disease. *WNT4* is known to play an important role in the development of the female genital tract from the Müllerian duct that develops into the fallopian tubes and uterus; the loss of *WNT4* in knockout mice was shown to lead to complete absence of the Müllerian duct [54, 55]. Moreover, *WNT4* is expressed in normal peritoneum, suggesting that endometriosis may arise through metaplasia using developmental pathways involved in the development of the female genital tract [54]. Furthermore, it is possible that genetic variants in *WNT4* might contribute to endometriosis susceptibility through abnormal cell growth in female genital tract. With all the evidence from the literature, *WNT4*, through its prominent role in development of the female reproductive tract [55], ovarian follicle development, and steroidogenesis [56], might play a critical role in development of endometriosis.

The third novel genetic variant found only in the GWASs of women of European ancestry [27•], rs12700667, is located in an intergenic region on chromosome 7p15.2. This locus was not reported among the top 100 signals in the Japanese GWAS by Uno et al. [40•]; however, given their sample size they would have had only 13 % power to detect it [27•]. Chromosome 7p15.2 contains multiple genomic features that may be involved in endometriosis development. The region contains a number of expressed sequence tags (ESTs) and a miRNA (*has-mir-148a*), suggesting a role in gene expression regulation. Further downstream, there are several genes including *NFE2L3* (highly expressed in placenta) and two endometriosis candidate genes, *HOXA10* and *HOXA11,* which are transcription factors that play a role in uterine development. Another interesting aspect about this variant comes from a large GWAS ($n=77{,}167$) investigating the genetic variants associated with fat distribution as measured by waist-to-hip ratio adjusted for the effect of body mass index (WHRadjBMI) [57]. Interestingly, rs12700667 is positioned at the same locus as one of 13 genome-wide significant loci they reported as associated with reduced WHRadjBMI, a result that is currently further being explored. As endometriosis and fat distribution are both hormonally dependent, pleiotropic genetic loci acting on both traits may well exist.

Challenges and Future Directions

The first GWASs of endometriosis have identified three genetic loci highly likely to be involved in the pathogenesis of endometriosis in women of Japanese and European ancestry. Although they have identified the loci, they have not pinpointed the actual genetic variants that are causal to endometriosis. Furthermore, they explain only a fraction of the heritability of endometriosis. Therefore, directions for further investigation include the following:

(1) The GWASs of endometriosis to date are relatively crude, in that they generally only describe associations with endometriosis, without regard for subtypes or differences in symptomatology (pain and infertility). Only the IEC GWAS [27•] considered stage B versus stage A endometriosis, showing that the two are genetically different in etiology. Future studies need to include cases who are phenotyped in much greater detail to enable distinction between endometriosis subtypes. For these studies to have sufficient power, however, even larger sample sizes are required than used to date.

(2) Of the 1,449 novel loci identified by GWASs up to June 2011 [39], the number of loci identified per complex trait varies greatly. Visscher et al. [58] showed that the number of discovered variants is strongly correlated with the sample size of different studies. The sample sizes of the endometriosis GWASs are at the lower end of those that have been successful in identifying loci involved in the etiology of a complex trait. This means that increasing the discovery sample size will increase the number of discovered variants, which also was suggested by prediction analyses based on all genotyped SNPs in the GWAS of the IEC [27•]. When individual causal variants only explain a small amount of variation, then the power to detect them is low in studies with small sample sizes. Therefore, there is a need for larger endometriosis GWASs in different populations, and meta-analyses of GWASs, to detect additional common causal variants with modest effect on risk. As mentioned before, the power of these studies can be increased by focusing on more severe subtypes, and in general by ensuring detailed and accurate clinical phenotyping allowing the investigation of such subtypes.

(3) To maximize information gained from GWASs, dense genotyping platforms containing at least 500,000 SNPs selected on the basis of LD should be utilized in GWASs, and imputation approaches should be used to estimate non-genotyped common variants (eg, using 1000 Genomes project data [http://www.1000genomes.org]). To follow up on initial GWA scans and to further improve the coverage, custom-made gene arrays could be designed including dense SNP sets focused on implicated genes (although, if budget allows, sequencing of these may be preferable); these could be supplemented by genes robustly implicated in hypothesized related common disease such as gynecological cancers [59, 60], migraine [61] or complex traits such as age of menarche [62] and body mass index (BMI) [63]. The use of whole exome

genotyping arrays, including carefully selected sets of common and low-frequency variants in exons, also could be a promising approach to identify additional signals.

(4) Although GWASs are unbiased by prior biological knowledge or genomic location, they are limited with regard to their ability to only assess the effect of common SNPs (MAF>~0.05) in the general population. It is possible that some of the unexplained genetic variation may be due to rarer variants (MAF<0.5 %), either single-site or structural, that are not captured by current GWA genotyping arrays [64, 65]. Indeed, the effects of structural variation are generally missed out in GWAS (although some of the GWA arrays contain selected probes to test for common copy number variation [CNV]). Rare and structural variants can be investigated through exome or whole-genome sequencing; however, these studies are still challenging because of their cost, with bioinformatic methodology to accurately call structural variants in particular still under development.

(5) Preliminary studies have suggested the role of gene–gene and gene–environment interactions in the development of endometriosis [66, 67]. There is no doubt that such interactions play a role in the development of many complex traits, but to robustly ascertain them without strong prior hypotheses, even larger sample sizes are generally required than advocated for GWAS. A biological route through which environmental factors could impact on the influence of genes on disease is through epigenetic modification of DNA (mainly methylation). There is considerable interest in epigenetic variants as contributors to the unexplained genetic variation of endometriosis [68]. One method through which the role of epigenetic variants in the development of endometriosis could be tested is by using a twin cohort where age-specific endometriosis concordance rates are obtained from monozygotic and dizygotic twins. If the monozygotic twin concordance rates for endometriosis increase with age, this could provide evidence for the potential role of epigenetic factors [69]. However, a major issue for endometriosis could be age-specific confounding by patterns of diagnosis. Platforms to detect genome-wide DNA methylation are being developed, and although currently still based on selected loci, are increasingly providing exciting new avenues for exploring the effect of epigenetic changes on complex disease such as endometriosis.

(6) Although the study of the genetics of endometriosis is starting to reap fruits, it is still faced with many challenges inherent to the complexities of the disease. This includes the general lack of correlation between disease stage and pain severity; much greater phenotypic detail of cases included in genetic studies is needed to start distinguishing between potential subtypes of disease and

symptomatologies. A particular problem is the need for a surgical diagnosis, which hampers population-based research. A complementary avenue may be to study the disease in animal models, in particular nonhuman primates such as the rhesus macaque [70] or the baboon [71, 72], with the rhesus macaque a promising model for the study of heritable, spontaneously occurring endometriosis [17]. Work on the genetics of endometriosis in the rhesus macaque, comparing human and rhesus macaque genomes, is currently underway.

Conclusions

There is mounting evidence for genetic variants contributing to endometriosis susceptibility. To date, candidate gene studies have not provided robust, replicated genetic variants associated with endometriosis and they have low a-priori chance of success when based solely on biological hypotheses. Linkage studies in pedigrees have found some regions of interest, which are likely to harbor variants implicated in familial endometriosis; however, due to the methodology, the regions identified are large and contain many genes of potential interest. These regions require further investigations to elucidate the susceptibility variants. The much-anticipated initial GWAS results identified three regions showing robust association with endometriosis in women of Japanese and European ancestry. Although they are by far the largest studies of endometriosis to date, they were of relatively small sample size compared to GWASs of other traits, and generally lacked detailed clinical information that would have allowed investigation of more etiologically homogeneous subphenotypes of endometriosis. Such information needs to be collected systematically for future analyses, and well-designed and sufficiently powered GWASs of thousands of cases and controls need to be conducted to identify novel genetic variants associated with endometriosis. These larger studies also will allow for better investigation of potential gene–gene and gene–environment interactions. Moreover, there are upcoming opportunities to study the impact of environment on genetic factors influencing endometriosis, through epigenetic studies, with the development of platforms to detect epigenetic changes genome-wide. Identification of robust, replicated genetic/epigenetic variants of endometriosis will pave the way to functional studies to better understand the underlying mechanisms of endometriosis, leading to better diagnosis and treatment of the disease. Complimentary research in animal models, particularly nonhuman primate models such as the rhesus macaque and the baboon, which can develop endometriosis spontaneously, should help to further elucidate the genetics of this complex condition.

References

Papers of particular interest, published recently, have been highlighted as:

• Of importance

1. Giudice LC, Kao LC. Endometriosis. Lancet. 2004;364:1789–99.
2. Berkley KJ, Rapkin AJ, Papka RE. The pains of endometriosis. Science. 2005;308:1587–9.
3. Society TAF. Revised American Fertility Society classification of endometriosis: 1985. Fertil Steril. 1985;43:351–2.
4. Bulun SE, Gurates B, Fang Z, et al. Mechanisms of excessive estrogen formation in endometriosis. J Reprod Immunol. 2002;55:21–33.
5. Zondervan KT, Cardon LR, Kennedy SH. What makes a good case-control study? Design issues for complex traits such as endometriosis. Hum Reprod. 2002;17:1415–23.
6. Eskenazi B, Warner ML. Epidemiology of endometriosis. Obstet Gynecol Clin North Am. 1997;24:235–58.
7. Sampson JA. Metastatic or embolic endometriosis, due to the menstrual dissemination of endometrial tissue into the venous circulation. Am J Pathol. 1927;3:93–110. 43.
8. Blumenkrantz MJ, Gallagher N, Bashore RA, et al. Retrograde menstruation in women undergoing chronic peritoneal dialysis. Obstet Gynecol. 1981;57:667–70.
9. Healy DL, Rogers PA, Hii L, et al. Angiogenesis: a new theory for endometriosis. Hum Reprod Update. 1998;4:736–40.
10. Vinatier D, Orazi G, Cosson M, et al. Theories of endometriosis. Eur J Obstet Gynecol Reprod Biol. 2001;96:21–34.
11. Varma R, Rollason T, Gupta JK, et al. Endometriosis and the neoplastic process. Reproduction. 2004;127:293–304.
12. Kennedy S. The genetics of endometriosis. J Reprod Med. 1998;43:263–8.
13. Zondervan KT, Cardon LR, Kennedy SH. The genetic basis of endometriosis. Curr Opin Obstet Gynecol. 2001;13:309–14.
14. Stefansson H, Geirsson RT, Steinthorsdottir V, et al. Genetic factors contribute to the risk of developing endometriosis. Hum Reprod. 2002;17:555–9.
15. Hadfield RM, Mardon HJ, Barlow DH, et al. Endometriosis in monozygotic twins. Fertil Steril. 1997;68:941–2.
16. Treloar SA, O'Connor DT, O'Connor VM, et al. Genetic influences on endometriosis in an Australian twin sample. sueT@qimr.edu.au. Fertil Steril. 1999;71:701–10.
17. Zondervan KT, Weeks DE, Colman R, et al. Familial aggregation of endometriosis in a large pedigree of rhesus macaques. Hum Reprod. 2004;19:448–55.
18. Wang DG, Fan JB, Siao CJ, et al. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science. 1998;280:1077–82.
19. The International HapMap C. A haplotype map of the human genome. Nature. 2005;437:1299–320.
20. Montgomery GW, Nyholt DR, Zhao ZZ, et al. The search for genes contributing to endometriosis risk. Hum Reprod Update. 2008;14:447–57.
21. Zondervan KT, Cardon LR. Designing candidate gene and genome-wide case-control association studies. Nat Protoc. 2007;2:2492–501.
22. Falconer H, D'Hooghe T, Fried G. Endometriosis and genetic polymorphisms. Obstet Gynecol Surv. 2007;62:616–28.
23. Hirschhorn JN, Lohmueller K, Byrne E, et al. A comprehensive review of genetic association studies. Genet Med. 2002;4:45–61.
24. Zollner S, Pritchard JK. Overcoming the winner's curse: estimating penetrance parameters from case-control data. Am J Hum Genet. 2007;80:605–15.
25. • Liang S, Huang Y, Fan Y. Vascular endothelial growth factor gene polymorphisms and endometriosis risk: a meta-analysis. Arch Gynecol Obstet 2012. *This is a large meta-analysis of VEGF gene polymorphisms from 11 studies from Chinese, Indian, Korean, Japanese, Spanish, Turkish, Estonian, and Australian patients.*
26. • Zhao ZZ, Nyholt DR, Thomas S, et al. Polymorphisms in the vascular endothelial growth factor gene and the risk of familial endometriosis. Mol Hum Reprod. 2008;14:531–8. *This large association study investigated the VEGF polymorphisms in a large Australian sample and further performed a large literature-based meta-analysis.*
27. • Painter JN, Anderson CA, Nyholt DR, et al. Genome-wide association study identifies a locus at 7p15.2 associated with the development of endometriosis. 2010. *This study, the first endometriosis GWAS performed in women of European ancestry, identified a novel locus on chromosome 7p15.2.*
28. Kerem B, Rommens JM, Buchanan JA, et al. Identification of the cystic fibrosis gene: genetic analysis. Science. 1989;245:1073–80.
29. Gilliam TC, Tanzi RE, Haines JL, et al. Localization of the Huntington's disease gene to a small segment of chromosome 4 flanked by D4S10 and the telomere. Cell. 1987;50:565–71.
30. Treloar SA, Wicks J, Nyholt DR, 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. 2005;77:365–76.
31. Zondervan KT, Treloar SA, Lin J, et al. Significant evidence of one or more susceptibility loci for endometriosis with near-Mendelian inheritance on chromosome 7p13-15. Hum Reprod. 2007;22:717–28.
32. • Painter JN, Nyholt DR, Morris A, et al. High-density fine-mapping of a chromosome 10q26 linkage peak suggests association between endometriosis and variants close to CYP2C19. Fertil Steril. 2011;95:2236–40. *This follow-up of the linkage peak on chromosome 10q26 region, through extensive genotyping of 11,984 SNPs, identified a novel signal in the CYP2C19 gene associated with endometriosis.*
33. Lee AJ, Cai MX, Thomas PE, et al. Characterization of the oxidative metabolites of 17beta-estradiol and estrone formed by 15 selectively expressed human cytochrome p450 isoforms. Endocrinology. 2003;144:3382–98.
34. Cribb AE, Knight MJ, Dryer D, et al. Role of polymorphic human cytochrome P450 enzymes in estrone oxidation. Cancer Epidemiol Biomarkers Prev. 2006;15:551–8.
35. Lin J, Zong L, Kennedy SH, et al. Coding regions of INHBA, SFRP4 and HOXA10 are not implicated in familial endometriosis linked to chromosome 7p13-15. Mol Hum Reprod. 2011;17:605–11.

36. Klein RJ, Zeiss C, Chew EY, et al. Complement factor H polymorphism in age-related macular degeneration. Science. 2005;308:385–9.

37. Duerr RH, Taylor KD, Brant SR, et al. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. Science. 2006;314:1461–3.

38. WTCCC. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007;447:661–78.

39. Hindorff LA, Sethupathy P, Junkins HA, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci USA. 2009;106:9362–7.

40. • Uno S, Zembutsu H, Hirasawa A, et al. A genome-wide association study identifies genetic variants in the CDKN2BAS locus associated with endometriosis in Japanese. Nat Genet. 2010;42:707–10. *This was the first endometriosis GWAS in women of Japanese ancestry, and identified a novel locus on chromosome 9p21.*

41. Adachi S, Tajima A, Quan J, et al. Meta-analysis of genome-wide association scans for genetic susceptibility to endometriosis in Japanese population. J Hum Genet 2010.

42. Koninckx PR, Barlow D, Kennedy S. Implantation versus infiltration: the Sampson versus the endometriotic disease theory. Gynecol Obstet Invest. 1999;47 Suppl 1:3–9. discussion 9-10.

43. Nisolle M, Donnez J. Peritoneal endometriosis, ovarian endometriosis, and adenomyotic nodules of the rectovaginal septum are three different entities. Fertil Steril. 1997;68:585–96.

44. Jarinova O, Stewart AF, Roberts R, et al. Functional analysis of the chromosome 9p21.3 coronary artery disease risk locus. Arterioscler Thromb Vasc Biol. 2009;29:1671–7.

45. Liu Y, Sanoff HK, Cho H, et al. INK4/ARF transcript expression is associated with chromosome 9p21 variants linked to atherosclerosis. PLoS One. 2009;4:e5027.

46. Pasmant E, Laurendeau I, Heron D, et al. Characterization of a germ-line deletion, including the entire INK4/ARF locus, in a melanoma-neural system tumor family: identification of ANRIL, an antisense noncoding RNA whose expression coclusters with ARF. Cancer Res. 2007;67:3963–9.

47. Kool J, Uren AG, Martins CP, et al. Insertional mutagenesis in mice deficient for p15Ink4b, p16Ink4a, p21Cip1, and p27Kip1 reveals cancer gene interactions and correlations with tumor phenotypes. Cancer Res. 2010;70:520–31.

48. Okuda T, Hirai H, Valentine VA, et al. Molecular cloning, expression pattern, and chromosomal localization of human CDKN2D/INK4d, an inhibitor of cyclin D-dependent kinases. Genomics. 1995;29:623–30.

49. Nakashima R, Fujita M, Enomoto T, et al. Alteration of p16 and p15 genes in human uterine tumours. Br J Cancer. 1999;80:458–67.

50. Yanokura M, Banno K, Susumu N, et al. Hypermethylation in the p16 promoter region in the carcinogenesis of endometrial cancer in Japanese patients. Anticancer Res. 2006;26:851–6.

51. Ignatov A, Bischoff J, Schwarzenau C, et al. P16 alterations increase the metastatic potential of endometrial carcinoma. Gynecol Oncol. 2008;111:365–71.

52. Martini M, Ciccarone M, Garganese G, et al. Possible involvement of hMLH1, p16(INK4a) and PTEN in the malignant transformation of endometriosis. Int J Cancer. 2002;102:398–406.

53. Goumenou AG, Arvanitis DA, Matalliotakis IM, et al. Loss of heterozygosity in adenomyosis on hMSH2, hMLH1, p16Ink4 and GALT loci. Int J Mol Med. 2000;6:667–71.

54. Gaetje R, Holtrich U, Engels K, et al. Endometriosis may be generated by mimicking the ontogenetic development of the female genital tract. Fertil Steril. 2007;87:651–6.

55. Vainio S, Heikkila M, Kispert A, et al. Female development in mammals is regulated by Wnt-4 signalling. Nature. 1999;397:405–9.

56. Boyer A, Lapointe E, Zheng X, et al. WNT4 is required for normal ovarian follicle development and female fertility. FASEB J. 2010;24:3010–25.

57. Heid IM, Jackson AU, Randall JC, et al.: Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. Nat Genet 2010.

58. Visscher PM, Brown MA, McCarthy MI, et al. Five years of GWAS discovery. Am J Hum Genet. 2012;90:7–24.

59. Munksgaard PS, Blaakaer J. The association between endometriosis and gynecological cancers and breast cancer: a review of epidemiological data. Gynecol Oncol. 2011;123:157–63.

60. Munksgaard PS, Blaakaer J. The association between endometriosis and ovarian cancer: a review of histological, genetic and molecular alterations. Gynecol Oncol. 2012;124:164–9.

61. Nyholt DR, Gillespie NG, Merikangas KR, et al. Common genetic influences underlie comorbidity of migraine and endometriosis. Genet Epidemiol. 2009;33:105–13.

62. Elks CE, Perry JR, Sulem P, et al. Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. Nat Genet. 2010;42:1077–85.

63. Missmer SA, Hankinson SE, Spiegelman D, et al. Incidence of laparoscopically confirmed endometriosis by demographic, anthropometric, and lifestyle factors. Am J Epidemiol. 2004;160:784–96.

64. Goldstein DB. Common genetic variation and human traits. N Engl J Med. 2009;360:1696–8.

65. Pritchard JK. Are rare variants responsible for susceptibility to complex diseases? Am J Hum Genet. 2001;69:124–37.

66. Barbosa CP, Teles JS, Lerner TG, et al. Genetic association study of polymorphisms FOXP3 and FCRL3 in women with endometriosis. Fertil Steril 2012.

67. McCarty CA, Berg RL, Welter JD, et al. A novel gene-environment interaction involved in endometriosis. Int J Gynaecol Obstet. 2012;116:61–3.

68. Guo SW. Epigenetics of endometriosis. Mol Hum Reprod. 2009;15:587–607.

69. Fraga MF, Ballestar E, Paz MF, et al. Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci USA. 2005;102:10604–9.

70. Zondervan K, Cardon L, Desrosiers R, et al. The genetic epidemiology of spontaneous endometriosis in the rhesus monkey. Ann N Y Acad Sci. 2002;955:233–8. discussion 293–5, 396–406.

71. Braundmeier AG, Fazleabas AT. The non-human primate model of endometriosis: research and implications for fecundity. Mol Hum Reprod. 2009;15:577–86.

72. D'Hooghe TM, Kyama CM, Chai D, et al. Nonhuman primate models for translational research in endometriosis. Reprod Sci. 2009;16:152–61.

73. Govatati S, Tangudu NK, Deenadayal M, et al. Association of E-cadherin single nucleotide polymorphisms with the increased risk of Endometriosis in Indian women. Mol Hum Reprod 2011.

74. Yoshida K, Yoshihara K, Adachi S, et al. Possible involvement of the E-cadherin gene in genetic susceptibility to endometriosis. Hum Reprod 2012.

75. Aghajanpour L, Mashayekhi F, Rajaei F. Intercellular adhesion molecule-1 (ICAM-1) gene polymorphism and endometriosis in northern Iran. Arch Gynecol Obstet. 2011;283:1035–9.

76. Chae SJ, Lee GH, Choi YM, et al. Intercellular adhesion molecule-1 and interleukin-6 gene polymorphisms in patients with advanced-stage endometriosis. Gynecol Obstet Invest. 2010;70:34–9.

77. Saare M, Lamp M, Kaart T, et al. Polymorphisms in MMP-2 and MMP-9 promoter regions are associated with endometriosis. Fertil Steril. 2010;94:1560–3.

78. Chang CY, Chen Y, Lin WC, et al. MUC2 polymorphisms are associated with endometriosis development and infertility: a case-control study. BMC Med Genet. 2012;13:15.

79. Chang CY, Chang HW, Chen CM, et al. MUC4 gene polymorphisms associate with endometriosis development and endometriosis-related infertility. BMC Med. 2011;9:19.

80. Rai P, Deenadayal M, Shivaji S. Absence of activating somatic mutations of PI3KCA and AKT1 genes in South Indian women with endometriosis. Eur J Obstet Gynecol Reprod Biol. 2010;152:78–82.

81. Attar R, Cacina C, Sozen S, et al. DNA repair genes in endometriosis. Genet Mol Res. 2010;9:629–36.

82. Ying TH, Tseng CJ, Tsai SJ, et al. Association of p53 and CDKN1A genotypes with endometriosis. Anticancer Res. 2011;31:4301–6.

83. Camargo-Kosugi CM, da Silva ID, Sato H, et al. The V109G polymorphism in the p27 gene is associated with endometriosis. Eur J Obstet Gynecol Reprod Biol. 2009;145:180–3.

84. Gallegos-Arreola MP, Valencia-Rodriguez LE, Puebla-Perez AM, et al. The TP53 16-bp duplication polymorphism is enriched in endometriosis patients. Gynecol Obstet Invest. 2012;73:118–23.

85. • Jia S, Xu L, Chan Y, et al. p53 codon 72 polymorphism and endometriosis: a meta-analysis. Arch Gynecol Obstet 2012. *This is a literature-based meta-analysis of P53 codon 72 polymorphism from six studies of Asian and Caucasian ancestry.*

86. Ribeiro Junior CL, Arruda JT, Silva CT, et al. Analysis of p53 codon 72 gene polymorphism in Brazilian patients with endometriosis. Genet Mol Res. 2009;8:494–9.

87. Ruiz LA, Dutil J, Ruiz A, et al. Single-nucleotide polymorphisms in the lysyl oxidase-like protein 4 and complement component 3 genes are associated with increased risk for endometriosis and endometriosis-associated infertility. Fertil Steril. 2011;96:512–5.

88. Lerner TG, Bianco B, Teles JS, et al. Analysis of CTLA4 gene variant in infertile Brazilian women with and without endometriosis. Int J Immunogenet. 2011;38:259–62.

89. Kim H, Ku SY, Suh CS, et al. Association between endometriosis and polymorphisms in tumor necrosis factor-related apoptosis-inducing ligand (TRAIL), TRAIL receptor and osteoprotegerin genes and their serum levels. Arch Gynecol Obstet 2012.

90. Teles JS, Bianco B, Vilarino FL, et al. Association of FCRL3 C-169T promoter single-nucleotide polymorphism with idiopathic infertility and infertility-related endometriosis. J Reprod Immunol. 2011;89:212–5.

91. Kim JJ, Choi YM, Hwang SS, et al. Association of the interferon-gamma gene (CA)n repeat polymorphism with endometriosis. BJOG. 2011;118:1061–6.

92. Rozati R, Vanaja MC, Nasaruddin K. Genetic contribution of the interferon gamma dinucleotide-repeat polymorphism in South Indian women with endometriosis. J Obstet Gynaecol Res. 2010;36:825–31.

93. Riiskjaer M, Nielsen K, Steffensen R, et al. Association of interleukin–10 promoter polymorphism and endometriosis. Am J Reprod Immunol. 2011;65:13–9.

94. Juo SH, Wu R, Lin CS, et al. A functional promoter polymorphism in interleukin-10 gene influences susceptibility to endometriosis. Fertil Steril. 2009;92:1228–33.

95. Xie J, Wang S, He B, et al. Association of estrogen receptor alpha and interleukin-10 gene polymorphisms with endometriosis in a Chinese population. Fertil Steril. 2009;92:54–60.

96. Gan XL, Lin YH, Zhang Y, et al. Association of an interleukin-16 gene polymorphism with the risk and pain phenotype of endometriosis. DNA Cell Biol. 2010;29:663–7.

97. Ayaz L, Celik SK, Cayan F, et al. Functional association of interleukin-18 gene -607C/A promoter polymorphisms with endometriosis. Fertil Steril. 2011;95:298–300.

98. Attar R, Agachan B, Kucukhuseyin O, et al. Association of interleukin 1beta gene (+3953) polymorphism and severity of endometriosis in Turkish women. Mol Biol Rep. 2010;37:369–74.

99. Lee GH, Choi YM, Kim SH, et al. Interleukin-2 receptor beta gene C627T polymorphism in Korean women with endometriosis: a case-control study. Hum Reprod. 2009;24:2596–9.

100. Zhou B, Rao L, Peng Y, et al. A functional promoter polymorphism in NFKB1 increases susceptibility to endometriosis. DNA Cell Biol. 2010;29:235–9.

101. Gomes FM, Bianco B, Teles JS, et al. PTPN22 C1858T polymorphism in women with endometriosis. Am J Reprod Immunol. 2010;63:227–32.

102. Ploski R, Dziunycz P, Kostrzewa G, et al. PTPN22/LYP 1858C>T gene polymorphism and susceptibility to endometriosis in a Polish population. J Reprod Immunol. 2009;79:196–200.

103. Sundqvist J, Falconer H, Seddighzadeh M, et al. Endometriosis and autoimmune disease: association of susceptibility to moderate/severe endometriosis with CCL21 and HLA-DRB1. Fertil Steril. 2011;95:437–40.

104. Tripputi P, Bignotto M, Cigognini D, et al. T-cell receptor beta polymorphism is not associated with endometriosis. J Obstet Gynaecol Res. 2011;37:1405–8.

105. Lakshmi KV, Shetty P, Vottam K, et al. Tumor necrosis factor alpha -C850T polymorphism is significantly associated with endometriosis in Asian Indian women. Fertil Steril. 2010;94:453–6.

106. Chae SJ, Kim H, Jee BC, et al. Tumor necrosis factor (TNF)-TNF receptor gene polymorphisms and their serum levels in Korean women with endometriosis. Am J Reprod Immunol. 2008;60: 432–9.

107. de Graaff AA, Dunselman GA, Delvoux B, et al. B lymphocyte stimulator -817C>T promoter polymorphism and the predisposition for the development of deep infiltrating endometriosis. Fertil Steril. 2010;94:1108–10.

108. Christofolini DM, Teles JS, Vilarino FL, et al. COMT polymorphism and the risk of endometriosis-related infertility. Gynecol Endocrinol. 2011;27:1099–102.

109. Trabert B, Schwartz SM, Peters U, et al. Genetic variation in the sex hormone metabolic pathway and endometriosis risk: an evaluation of candidate genes. Fertil Steril. 2011;96:1401–6. e3.

110. Lamp M, Peters M, Reinmaa E, et al. Polymorphisms in ESR1, ESR2 and HSD17B1 genes are associated with fertility status in endometriosis. Gynecol Endocrinol. 2011;27:425–33.

111. Sundqvist J, Falconer H, Seddighzadeh M, et al. Ovarian cancer-associated polymorphisms in the BNC2 gene among women with endometriosis. Hum Reprod. 2011;26:2253–7.

112. Kang S, Li SZ, Wang N, et al. Association between genetic polymorphisms in fibroblast growth factor (FGF)1 and FGF2 and risk of endometriosis and adenomyosis in Chinese women. Hum Reprod. 2010;25:1806–11.

113. Kim H, Ku SY, Kim SH, et al. Association between endometriosis and polymorphisms in insulin-like growth factor binding protein genes in Korean women. Eur J Obstet Gynecol Reprod Biol 2012.

114. Kim JJ, Choi YM, Choung SH, et al. Analysis of the transforming growth factor beta1 gene -509C/T polymorphism in patients with advanced-stage endometriosis. Fertil Steril. 2010;93:2121–4.

115. Lee HJ, Kim H, Ku SY, et al. Transforming growth factor-beta1 gene polymorphisms in Korean women with endometriosis. Am J Reprod Immunol. 2011;66:428–34.

116. Altinkaya SO, Ugur M, Ceylaner G, et al. Vascular endothelial growth factor +405C/G polymorphism is highly associated with an increased risk of endometriosis in Turkish women. Arch Gynecol Obstet. 2011;283:267–72.

117. Liu Q, Li Y, Zhao J, et al. Association of single nucleotide polymorphisms in VEGF gene with the risk of endometriosis and adenomyosis. Zhonghua Yi Xue Yi Chuan Xue Za Zhi. 2009;26:165–9.

118. Cosin R, Gilabert-Estelles J, Ramon LA, et al. Vascular endothelial growth factor polymorphisms (-460C/T, +405G/C, and 936C/T)

and endometriosis: their influence on vascular endothelial growth factor expression. Fertil Steril. 2009;92:1214–20.

119. Bianco B, Christofolini DM, Mafra FA, et al. +1730G/A polymorphism of the estrogen receptor beta gene (ERbeta) may be an important genetic factor predisposing to endometriosis. Acta Obstet Gynecol Scand. 2009;88:1397–401.

120. Zulli K, Bianco B, Mafra FA, et al. Polymorphism of the estrogen receptor beta gene is related to infertility and infertility-associated endometriosis. Arq Bras Endocrinol Metabol. 2010;54:567–71.

121. • Near AM, Wu AH, Templeman C, et al. Progesterone receptor gene polymorphisms and risk of endometriosis: results from an international collaborative effort. Fertil Steril. 2011;95:40–5. *This international ovarian cancer consortium, including eight ovarian cancer case-controls studies from Australia, Europe, and the United States, investigated the association of polymorphisms in PGR with endometriosis.*

122. Matsuzaka Y, Kikuti YY, Izumi S, et al. Mapping of susceptibility locus for endometriosis within the HLA region using microsatellite markers in Japanese women. Tissue Antigens. 2010;75:65–7.

123. Vilarino FL, Bianco B, Lerner TG, et al. Analysis of vitamin D receptor gene polymorphisms in women with and without endometriosis. Hum Immunol. 2011;72:359–63.

124. Kim H, Ku SY, Kim SH, et al. Endothelial nitric oxide synthase gene Glu298Asp polymorphism is associated with advanced stage endometriosis. Hum Reprod. 2009;24:2656–9.

125. Bhanoori M, Kameshwari DB, Zondervan KT, et al. The endothelial nitric oxide synthase Glu298Asp polymorphism is not a risk factor for endometriosis in south Indian women. Eur J Obstet Gynecol Reprod Biol. 2008;139:53–8.

126. Cayan F, Ertunc D, Aras-Ates N, et al. Association of G1057D variant of insulin receptor substrate-2 with endometriosis. Fertil Steril. 2010;94:1622–6.

127. Mafra FA, Bianco B, Christofolini DM, et al. Luteinizing hormone beta-subunit gene (LHbeta) polymorphism in infertility and endometriosis-associated infertility. Eur J Obstet Gynecol Reprod Biol. 2010;151:66–9.

128. Goncalves-Filho RP, Brandes A, Christofolini DM, et al. Plasminogen activator inhibitor-1 4G/5G polymorphism in infertile women with and without endometriosis. Acta Obstet Gynecol Scand. 2011;90:473–7.

129. Gentilini D, Vigano P, Castaldi D, et al. Plasminogen activator inhibitor-1 4G/5G polymorphism and susceptibility to endometriosis in the Italian population. Eur J Obstet Gynecol Reprod Biol. 2009;146:219–21.

130. Bozdag G, Alp A, Saribas Z, et al. CYP17 and CYP2C19 gene polymorphisms in patients with endometriosis. Reprod Biomed Online. 2010;20:286–90.

131. Vietri MT, Cioffi M, Sessa M, et al. CYP17 and CYP19 gene polymorphisms in women affected with endometriosis. Fertil Steril. 2009;92:1532–5.

132. Cayan F, Ayaz L, Aban M, et al. Role of CYP2C19 polymorphisms in patients with endometriosis. Gynecol Endocrinol. 2009;25:530–5.

133. Hosseinzadeh Z, Mashayekhi F, Sorouri ZZ. Association between GSTM1 gene polymorphism in Iranian patients with endometriosis. Gynecol Endocrinol. 2011;27:185–9.

134. Jeon MJ, Choi YM, Hong MA, et al. No association between the GSTP1 exon 5 polymorphism and susceptibility to advanced stage endometriosis in the Korean population. Am J Reprod Immunol. 2010;63:222–6.

135. Hwang KR, Choi YM, Kim JM, et al. Association of peroxisome proliferator-activated receptor-gamma 2 Pro12Ala polymorphism with advanced-stage endometriosis. Am J Reprod Immunol. 2010;64:333–8.