Published in final edited form as:
Int J Gynaecol Obstet. 2012 January ; 116(1): 61–63. doi:10.1016/j.ijgo.2011.10.003.

A novel gene–environment interaction involved in endometriosis

Catherine A. McCarty[a,*], **Richard L. Berg**[b], **Joseph D. Welter**[c], **Terrie E. Kitchner**[d], and **Joseph W. Kemnitz**[e]

[a]Essentia Institute of Rural Health, Duluth, USA

[b]Biomedical Informatics Research Center, Marshfield Clinic Research Foundation, Marshfield, USA

[c]Department of Obstetrics and Gynecology, Marshfield Clinic, Marshfield, USA

[d]Center for Human Genetics, Marshfield Clinic Research Foundation, Marshfield, USA

[e]Wisconsin National Primate Research Center, University of Wisconsin-Madison, Madison, USA

Abstract

Objective—To establish a well-defined cohort for genetic epidemiology studies of endometriosis and conduct a pilot study to confirm validity using existing data associated with endometriosis.

Methods—Between January and May 2010, a nested cohort within a population-based biobank was established in Marshfield, Wisconsin, USA. The inclusion criteria were women who had laparoscopy or hysterectomy. Fifty-one pleiotropic genetic polymorphisms and other established risk factors, such as smoking status and body mass index, were compared between endometriosis cases and controls.

Results—From the existing biobank, 796 cases and 501 controls were identified, and 259 women with endometriosis were enrolled specifically for the nested cohort within this biobank. A single nucleotide polymorphism in the *MMP1* gene significantly differed between cases and controls only when stratified by smoking status. Minor allele frequency was higher in control women who smoked than in women with endometriosis who smoked (55.5% versus 45.5%, χ^2=8.2, *P*=0.017); the inverse relationship was found in non-smoker control women.

Conclusions—Women with endometriosis were successfully recruited to participate in a general biobank, and a novel gene–environment interaction was identified. The findings suggest that important potential genetic associations may be missed if gene–environment interactions with known epidemiologic risk factors are not considered.

Keywords

Biobank; Endometriosis; Epidemiology; Genetics; Smoking

*Corresponding author: Catherine A. McCarty, Essentia Institute of Rural Health, Maildrop 5AV-2, 502 East Second Street, Duluth, MN 55805, USA. Tel.: +1 218 786 3940; fax: +1 218 722 9952. cmccarty@eirh.org.

Conflict of interest
The authors have no conflicts of interest.

1. Introduction

Endometriosis is a complex disease and a major women's health problem, the causes of which are not entirely understood [1]. Substantiated risk factors include estrogen levels, smoking history (protective), exercise, and peripheral body fat [2]. Endometriosis is an estrogen-dependent inflammatory disease that is characterized by the presence of endometrium-like tissue in sites outside the uterus—mainly, the pelvic peritoneum, ovaries, and rectovaginal septum. Affected women experience severe pelvic pain and reduced fertility, and are at increased risk of developing ovarian cancer and non-Hodgkin lymphoma [2]. Under the assumption of a prevalence rate of 10% among women of reproductive age, the estimated total healthcare cost of treating endometriosis in the United States was US$ 22 billion in 2002 [3].

Although the heritability of endometriosis has been estimated to be 52%, so far genetic polymorphism studies have failed to identify major genetic contributors to endometriosis [4–7]. Familial aggregation of endometriosis has been documented in a large pedigree of rhesus macaques at the Wisconsin National Primate Research Center [8]. Collaboration between the Wisconsin National Primate Research Center and the Marshfield Clinic Research Foundation's Personalized Medicine Research Project (PMRP) [9–12] will facilitate comparative genetic analyses.

The aim of the present study was to establish a well-defined cohort for genetic epidemiology studies of endometriosis and to conduct a pilot study to confirm the validity of the cohort using existing data, some of which has been shown previously to be associated with endometriosis.

2. Materials and methods

In the present pilot study, women with endometriosis attending the Marshfield Clinic, Marshfield, Wisconsin, USA, were identified between January 5 and May 25, 2010, to form a defined cohort for genetic epidemiology studies and to supplement the existing PMRP data. The details of the PMRP have been described previously [9–12]. In brief, adult patients (≥18 years) of the Marshfield Clinic residing in 1 of 19 zip codes surrounding the city of Marshfield in central Wisconsin were asked to provide a blood sample from which DNA was extracted and stored with plasma and serum samples. Ninety-eight percent of participants self-reported being Caucasian and 78% self-reported German ancestry. Participants were asked to allow access to their medical records in order to define phenotype. The informed consent document contained a tick-box that patients could check, indicating their agreement to future contact for research purposes.

For the present study, women with endometriosis attending the Marshfield Clinic but residing outside the original 19 zip codes were recruited; the comparison group comprised control women within the original 19 zip codes who had already participated in the PMRP biobank. The ultimate goal was to assemble a cohort of approximately 1000 women with endometriosis to provide sufficient statistical power for future genome-wide association studies analyses. The protocol was reviewed and approved by the Marshfield Clinic Institutional Review Board, and all women gave written informed consent to participate in the study.

The inclusion criteria for both cases and controls in the present study were women who had undergone laparoscopy or hysterectomy as part of their normal clinical care for reasons other than endometriosis. Endometriosis status was diagnosed by gynecologic visual confirmation on procedure and/or by pathology confirmation. A chart abstraction form was

developed and used to confirm endometriosis status and to document medical and/or surgical treatment and outcomes.

The data analyzed included smoking status collected through a self-administered questionnaire at the time of enrollment into the biobank. Genotyping data on 51 genetic polymorphisms included on a molecular fingerprinting panel [12] available for the whole cohort were evaluated for potential associations with endometriosis (Table 1). Two of the single-nucleotide polymorphisms (SNPs) were located in genes that have been previously shown to be associated with endometriosis [13,14]: the *MMP1* gene (rs1799750) and the gene encoding tumor necrosis factor-α (rs1800629). Data were entered twice and verified. Cases and controls were compared by using χ^2 tests for each SNP across 3 possible categories, and odds ratios for presence of the minor allele were calculated, together with asymptotic 95% confidence intervals (CIs). Analyses were conducted by using SAS version 9.2 (SAS Institute Inc., Cary, NC). Results were considered significant at the 5% level ($P<0.05$) without adjustment for multiple testing; it was noted, however, that maintaining an error rate of 5% over all 51 polymorphisms would require a P value of less than 0.001.

3. Results

By the end of the study period, data for 796 cases and 501 controls had been incorporated in the biobank. Among the cases, 259 women were enrolled in the biobank specifically for the present study (60% of women approached to join the biobank agreed to participate). At enrollment, the affected women ranged from 18 to 86 years (mean 43.9 years) and the control women ranged from 18 to 63 years (40.8 years). Age at diagnosis of endometriosis ranged from 14 to 80 years (35.1 years).

Among the whole cohort, affected women and control women differed significantly in age and body mass index (calculated as weight in kilograms divided by the square of height in meters), but not in smoking status (Table 2)—factors shown previously to be associated with endometriosis that could bias association results [2].

The minor allele or haplotype frequency of 2 of the 51 polymorphisms investigated (IL6 haplotype 3 and MTFHD1) was found to be significantly higher in cases than in controls (Table 3). The allele distribution of an SNP in the *MMP1* gene that has been shown to be associated with endometriosis did not differ significantly between cases and controls except when stratified by smoking status (Table 4). Not only was the association found to be significant only in current smokers, but also the estimated minor allele frequency was higher in control women who smoked than in affected women who smoked (55.5% versus 45.5%, $\chi^2=8.17$, $P=0.017$). By contrast, among those who did not smoke, the frequency was slightly lower in controls than in affected women (47% versus 49.5%), although the difference was not significant.

4. Discussion

In the present pilot study, a cohort of women was assembled for genetic studies of endometriosis. Women with endometriosis were successfully recruited to participate in a general biobank, and the absolute participation rate was 15% higher than that for individuals recruited from the same general population for a biobank to study many potential outcomes in the future [9,11]. This information will be useful for future biobank recruitment targeted to individuals with disease outcomes of immediate research interest. The assembled cohort has been rigorously phenotyped to minimize potential bias caused by the misclassification of case/control status.

Unlike previous studies, current smoking status was not found to be related to endometriosis. One reason for this discrepancy might be temporality. It is possible that women stop smoking after being diagnosed with a chronic condition. This may be especially true for women who are prescribed oral contraceptives to decrease the symptoms of endometriosis. Oral contraceptives are contra-indicated for smokers because of the increased risk of stroke, and women are strongly advised to stop smoking if they are prescribed oral contraceptives. We do not have detailed information about lifetime smoking habits to support or refute this hypothesis. A consideration for future research would be to gather more detailed information on lifetime smoking.

Significant genetic associations were found in the study using existing markers that were not selected specifically for their potential association with endometriosis. One of these markers —an SNP in the *MMP1* gene that has been shown to be related to endometriosis in Chinese women [13]—showed a significant association only in current smokers. The *MMP1* gene encodes an interstitial collagenase. The present findings suggest that a gene–environment interaction involving *MMP1* has a role in endometriosis. As with any association study, the potential for this finding to be a false positive exists, as does the possibility of false negatives for the other polymorphisms investigated that were not found to be associated with endometriosis.

The strengths of the study and its cohort include the rigorous phenotyping, the availability of questionnaire data to quantify personal and environmental exposure, the additional data related to co-morbid conditions, and the availability of treatment history from the medical records. In terms of genetic studies, the relatively high ethnic homogeneity is a limitation to the generalizability of the results. In addition, given the large number of potential associations evaluated, the present results will require validation by further studies.

The first step for successful genetic association studies is to assemble cases and controls that are rigorously phenotyped for the condition of interest. We have shown that women with endometriosis can be successfully recruited into a general biobank and that potential genetic associations may be missed if gene–environment interactions are not considered. Larger sample sizes will be needed to evaluate gene–environment and gene–gene interactions in full.

Acknowledgments

The project was funded in part by an administrative supplement to grant number P51 RR000167 from the National Center for Research Resources (NCRR) to the Wisconsin National Primate Research Center, and was supported by in part by grant 1UL1RR025011 from the Clinical and Translational Science Award program of the NCRR, National Institutes of Health, and grant number U54 RR024377-01 from the NCRR.

References

1. Bulun SE. Endometriosis. N Engl J Med. 2009; 360(3):268–79. [PubMed: 19144942]
2. Cramer DW, Missmer SA. The epidemiology of endometriosis. Ann N Y Acad Sci. 2002; 955:11–22. discussion 34–6,396–406. [PubMed: 11949940]
3. Simoens S, Hummelshoj L, D'Hooghe T. Endometriosis: cost estimates and methodological perspective. Hum Reprod Update. 2007; 13(4):395–404. [PubMed: 17584822]
4. Bischoff FZ, Simpson JL. Heritability and molecular genetic studies of endometriosis. Hum Reprod Update. 2000; 6(1):37–44. [PubMed: 10711828]
5. Falconer H, D'Hooghe T, Fried G. Endometriosis and genetic polymorphisms. Obstet Gynecol Surv. 2007; 62(9):616–28. [PubMed: 17705887]

6. Montgomery GW, Nyholt DR, Zhao ZZ, Treloar SA, Painter JN, Missmer SA, et al. The search for genes contributing to endometriosis risk. Hum Reprod Update. 2008; 14(5):447–57. [PubMed: 18535005]
7. Tempfer CB, Simoni M, Destenaves B, Fauser BC. Functional genetic polymorphisms and female reproductive disorders: part II—endometriosis. Hum Reprod Update. 2009; 15(1):97–118. [PubMed: 18805939]
8. Zondervan KT, Weeks DE, Colman R, Cardon LR, Hadfield R, Schleffler J, et al. Familial aggregation of endometriosis in a large pedigree of rhesus macaques. Hum Reprod. 2004; 19(2): 448–55. [PubMed: 14747196]
9. McCarty, CA.; Wilke, RA.; Giampietro, PF.; Wesbrook, SD.; Caldwell, MD. The Marshfield Clinic Personalized Medicine Research Project (PMRP): design, methods and recruitment for a large population-based biobank. www.marshfieldclinic.org. http://www.marshfieldclinic.org/chg/pages/Proxy.aspx?Content=MCRF-Centers-PMRP-Pubs-McCartyMethodsCHG0522122553.1.pdf. Published 2005
10. McCarty, CA.; Mukesh, BN.; Giampietro, PF.; Wilke, RA. Healthy People 2010 prevalence in the Marshfield Clinic Personalized Medicine Research Project: opportunities for public health genomics research. www.marshfieldclinic.org. http://www.marshfieldclinic.org/chg/pages/proxy.aspx?content=mcrf-centers-pmrp-pubs-mccartypme.1.pdf. Published 2007
11. McCarty, CA.; Peissig, PL.; Caldwell, MD.; Wilke, RA. The Marshfield Clinic Personalized Medicine Research Project: 2008 scientific update and lessons learned in the first 6 years. www.marshfieldclinic.org. http://www.futuremedicine.com/doi/abs/10.2217/17410541.5.5.529. Published 2008
12. Cross DS, Ivacic LC, McCarty CA. Development of a fingerprinting panel using medically relevant polymorphisms. BMC Med Genomics. 2009; 2:17. [PubMed: 19379518]
13. Shan K, Ying W, Jian-Hui Z, Wei G, Na W, Yan L. The function of the SNP in the MMP1 and MMP3 promoter in susceptibility to endometriosis in China. Mol Hum Reprod. 2005; 11(6):423–7. [PubMed: 15879464]
14. Lee GH, Choi YM, Kim SH, Hong MA, Oh ST, Lim YT, et al. Association of tumor necrosis factor-α gene polymorphisms with advanced stage endometriosis. Hum Reprod. 2008; 23(4):977–81. [PubMed: 18272527]

Table 1

Polymorphisms examined in the whole cohort

rs1137101	rs1801253	rs1544410
rs486907	rs2227564	rs16944
rs1042031	rs17999750 (MMP1)	rs1799983
rs231775	rs1063856	rs1800469
rs5186	rs6313	rs1800629
rs6280	rs2236225 (MTHFD1)	rs1800795 (IL6)
rs1799883	rs1800588	rs1800796 (IL6)
rs4961	rs243865	rs1800872
rs1042714	rs4673	rs1801133
rs351855	rs708272	rs268
rs5370	rs4291	rs429358
rs6296	rs4792311	rs4343
rs2227983	rs16430	rs731236
rs213950	rs601338	rs7412
rs7493	rs688	rs7975232
rs328	rs7121	
rs2383206	rs234706	
rs1800861	rs4680	

Abbreviations: IL6, interleukin 6; MMP1, metalloproteinase I; MTHFD1, methylenetetrahydrofolate dehydrogenase 1.

Table 2

Differences between endometriosis cases and controls in epidemiologic factors previously associated with endometriosis[a]

Epidemiologic factor	Controls	Cases	χ^2	*P* value	Odds ratio (95% confidence interval)
Age ≥50 years	93 (18.6)	222 (27.9)	14.5	0.0001	1.70 (1.29–2.23)
Never smoked [b]	266 (67.3)	420 (70.1)	0.9	0.35	1.14 (0.87–1.50)
Upper tertile of BMI distribution	204 (40.9)	270 (34.2)	6.3	0.037	0.75 (0.60–0.95)

Abbreviation: body mass index (BMI, calculated as weight in kilograms divided by the square of height in meters).

[a]Values are given as number (percentage) unless stated otherwise.

[b]Former smokers were not included.

Table 3

Significant univariate associations of endometriosis with genetic markers available for the whole cohort

Genetic marker	Controls	Cases	χ^2	*P* value	Odds ratio (95% confidence interval)
IL6 haplotype H3					
Total number	496	794			
% homozygous or heterozygous	62.3	71.7	12.6	0.0019	1.53 (1.21–1.94)
MTHFD1					
Total number	496	790			
Minor allele frequency, %	49.8	57.1	12.6	0.0018	1.51 (1.15–1.97)

Int J Gynaecol Obstet. Author manuscript; available in PMC 2013 February 26.

Table 4

Association of the *MMP1* SNP with endometriosis among the whole cohort and stratified by smoking status

Study group	Controls	Cases	χ^2	*P* value	Odds ratio (95% confidence interval)
Whole cohort					
Total number	496	791			
Minor allele frequency, %	48.5	48.5	0.84	0.66	0.94 (0.73–1.21)
Never smokers					
Total number	263	416			
Minor allele frequency, %	47.0	49.5	0.90	0.64	1.11 (0.79–1.57)
Current smokers					
Total number	128	178			
Minor allele frequency, %	55.5	45.5	8.17	0.017	0.45 (0.26–0.79)