Molecular Human Reproduction Vol.11, No.9 pp. 641–647, 2005
Advance Access publication August 26, 2005 doi:10.1093/molehr/gah221

Association between polymorphisms in the progesterone receptor gene and endometriosis

Susan A.Treloar[1], Zhen Zhen Zhao, Trudi Armitage, David L.Duffy, Jacqueline Wicks, Daniel T.O'Connor, Nicholas G.Martin and Grant W.Montgomery

Genetic Epidemiology Laboratory, Queensland Institute of Medical Research, Brisbane, Queensland, Australia

[1]To whom correspondence should be addressed at: Queensland Institute of Medical Research, Brisbane, Queensland 4029, Australia. E-mail: sue.treloar@qimr.edu.au

The progesterone receptor (*PR*) is a candidate gene for the development of endometriosis, a complex disease with strong hormonal features, common in women of reproductive age. We typed the 306 base pair *Alu* insertion (AluIns) polymorphism in intron G of *PR* in 101 individuals, estimated linkage disequilibrium (LD) between five single-nucleotide polymorphisms (SNPs) across the *PR* locus in 980 Australian triads (endometriosis case and two parents) and used transmission disequilibrium testing (TDT) for association with endometriosis. The five SNPs showed strong pairwise LD, and the AluIns was highly correlated with proximal SNPs rs1042839 (Δ^2 = 0.877, D′ = 1.00, *P* < 0.0001) and rs500760 (Δ^2 = 0.438, D′ = 0.942, *P* < 0.0001). TDT showed weak evidence of allelic association between endometriosis and rs500760 (*P* = 0.027) but not in the expected direction. We identified a common susceptibility haplotype *GGGCA* across the five SNPs (*P* = 0.0167) in the whole sample, but likelihood ratio testing of haplotype transmission and non-transmission of the AluIns and flanking SNPs showed no significant pattern. Further, analysis of our results pooled with those from two previous studies suggested that neither the *T2* allele of the AluIns nor the *T1/T2* genotype was associated with endometriosis.

Key words: endometriosis/linkage disequilibrium/polymorphism/progesterone receptor/transmission disequilibrium test

Introduction

Endometriosis is a common and complex gynaecological disorder. Estimates of its true prevalence vary (Zondervan *et al*., 2002), but given a proportion of affected women are asymptomatic, the reported rate of 7% in our large Australian twin study (Treloar *et al*., 1999) suggests a lifetime prevalence of around 10%. Endometriosis is influenced by both environmental and genetic factors (Treloar *et al*., 1999). There is strong evidence for familial aggregation in population based samples (Stefansson *et al*., 2002), hospital samples (Kennedy *et al*., 1995; Simpson and Bischoff, 2002), twin studies (Moen, 1994; Hadfield *et al*., 1997; Treloar *et al*., 1999) and in non-human primates (Zondervan *et al*., 2004). Finding genes predisposing to endometriosis will help to define mechanisms of disease, and efforts are being directed towards association studies of plausible candidate genes, and more recently towards positional cloning approaches (Kennedy, 2003).

In endometriosis, ectopic endometrium-like tissue continues to proliferate and invade in response to the monthly cycles of ovarian steroid hormones (Meresman *et al*., 2002). Endometriosis has characteristics of invasion and metastasis, although pathologically it resembles a benign tumour (Koninckx *et al*., 1999). The complex nature of the disease means variation in many genes may contribute to disease prevalence. Possible candidates include genes involved in endometrial tissue proliferation and invasion including those from steroid hormone pathways. Estrogen is a potent mitogen in the endometrium and endometriosis (Gurates and Bulun, 2003), whereas progesterone counteracts the estrogen-dependent proliferation of endometrial tissue and promotes differentiation (Gurates and Bulun, 2003). Combined oral contraceptives have been shown to down-regulate cell proliferation and to enhance apoptosis in eutopic endometrium from women with endometriosis (Meresman *et al*., 2002). Endometriosis appears less responsive to the antiproliferative and differentiating effects of progesterone (Gurates and Bulun, 2003), but progesterone receptor (PR) modulators are included in the range of promising new medical treatments for endometriosis (D'Hooghe, 2003).

Steroid receptor gene polymorphisms and protein changes have been correlated with endometriosis (Moutsatsou and Sekeris, 2003). The chromosomal region on 11q22-23 encompassing the *PR* gene is a frequent site for loss of heterozygosity in women with endometriosis (Jiang *et al*., 1996). Response to progestins may differ in women with endometriosis and differential *PR* expression has been reported in endometriotic tissue compared with eutopic endometrium (Jiang *et al*., 2002). Variants in *PR* include rs1042838 in exon 4, rs1042839 in exon 5 and an *Alu* insertion (AluIns) polymorphism in intron 7 (Rowe *et al*., 1995). There is strong but not complete linkage disequilibrium (LD) between these variants, and the combination of the *T* allele at rs1042838, *T* allele at rs1042839 and the AluIns has been referred to as PROGINS (Agoulnik *et al*., 1997; Kieback *et al*., 1998). Evidence for the direct involvement of the *PR* gene in endometriosis comes from two small case-control studies that reported a significant association between the *T2* AluIns allele of the intron G (PROGINS) of the *PR* gene (Wieser *et al*., 2002; Lattuada *et al*., 2004). A two-fold increase in risk of endometriosis in women carrying the *T2* allele was found from analysis of the combined data from both studies (Lattuada *et al*., 2004). The same pathophysiology may contribute to progression

of endometriosis and transformation to endometrioid or clear cell ovarian neoplasias (Ness, 2003). The PROGINS AluIns element (*T2* allele) has been associated with sporadic ovarian carcinoma (Rowe *et al*., 1995) and the PROGINS complex has been associated with the prevalence of ovarian cancer in women carrying *BRCA1* and *BRCA2* mutations (Runnebaum *et al*., 2001).

The human *PR* gene encodes several transcripts with two major isoforms, PR-A and PR-B, regulated by distinct promoter regions (Kastner *et al*., 1990). Transcription from separate promoters and translational start sites give rise to the major isoforms PR-A and PR-B, which are identical except for an additional 165 amino acids present only in the N terminus of PR-B (Conneely *et al*., 1987, 1989; Kastner *et al*., 1990). Although PR-A and PR-B share several structural domains, they are distinct transcription factors that mediate their own response genes and physiologic effects with little overlap (Conneely *et al*., 2003). A SNP in the promoter of *PR* (+331G/A; rs10895068) creates a unique start site which increases transcription of the PR-B isoform and has been associated with increased risk for developing endometrial cancer (De Vivo *et al*., 2002). A second polymorphism in the promoter region (+44C/T; rs518162), identified also by De Vivo *et al*. (2002) was not associated with endometrial cancer. Expression of the major isoforms may be altered in endometriotic lesions, but results are inconsistent. In one study a higher ratio of PR-B to PR-A mRNA was found in ectopic compared with eutopic endometrial tissue (Misao *et al*., 1999), whereas in another, PR-B mRNA was absent from endometriotic tissue (Attia *et al*., 2000).

In summary, variation in the *PR* may contribute to the incidence of endometriosis and related ovarian cancers (Del Carmen *et al*., 2003). Therefore we tested association between endometriosis and genetic polymorphism in the *PR* by analysing selected single-nucleotide polymorphisms (SNPs) across the locus in a sample of 980 triads (affected individual and both parents) from our Australian study of endometriosis (Treloar *et al*., 2002). Because typing the AluIns polymorphism is labour intensive, we typed it in only a subsample and used strong LD with one of the flanking SNPs to infer presence/absence of the insertion.

Materials and methods

Participants and sample collection

Samples and genotype data were from Australian participants in the International Endogene Study, recruited between 1996 and 2002 (Treloar *et al*., 2002). The median age at participation of affected daughters was 31 years, with a range of 16–54 years. Median age at diagnosis reported by daughters was 24 years, with a range of 12–50 years. Participants were mostly of European ancestry (Treloar *et al*., 2002). The Australian study recruited 1055 triads of affected women plus two parents, of which 980 families provided adequate DNA. Triads were independent of 931 recruited affected sister-pair families used in our parallel linkage study. All women had been diagnosed by surgical means, with visualisation a minimum requirement for inclusion. Revised American Fertility Society (rAFS) criteria (American Fertility Society, 1985) were used for staging of disease severity (Treloar *et al*., 2002). The project was approved by the Human Research Ethics Committee of the Queensland Institute of Medical Research and all DNA samples were collected with written, informed consent.

Triads included daughters with all rAFS stages of endometriosis. We used two broad categories: stage A (minimal to mild, $N = 673$, 69%) and stage B (moderate to severe, $N = 307$, 31%) for stratified analyses, given that stage was in some cases assigned retrospectively (Treloar *et al*., 2002). Subfertility was measured by reported problems conceiving. In this sample, 388 daughters who had tested their fertility (40%) reported problems conceiving and 215 (22%) reported no problems; 377 cases (38%) had never tried to conceive. In those who had tested their fertility there was a strong positive association ($\chi^2_1 = 24.45$, $P < 0.0001$) between more severe stage and subfertility in this sample, with minimal–mild endometriosis having a relatively protective effect (OR = 0.38, 95% CI = 0.25–0.56).

DNA was extracted from peripheral blood lymphocytes by the salt precipitation method (Miller *et al*., 1988) and was then stored at 4°C at Queensland Institute of Medical Research (QIMR, Brisbane, Australia). Two participants provided a buccal swab sample and DNA was extracted from buccal swabs using Microcon Centrifugal Filter Devices (Amicons) and stored at 4°C. Aliquots of sample genomic DNA were adjusted to final concentrations of 3 ng/ul and 10 ng/ul for Sequenom™ MassARRAY and AluIns PCR, respectively.

SNP selection

We selected four polymorphisms in the *PR* gene on the basis of published reports: the 306 base pair AluIns in intron G, part of the PROGINS complex, which had been reported to be associated with endometriosis (Wieser *et al*., 2002), a functional polymorphism +331G/A (rs10895068) in the promoter region associated with risk of endometrial cancer and two further SNPs, rs518162 and rs1042838 (Val660Leu), identified by De Vivo *et al*. (2002). Five additional SNPs–rs208112, rs3740754 (Ser344Thr), rs1042839 (His770His), rs2020880 (Leu865Ser) and rs500760 (Gln886Gln), the latter two flanking the AluIns–were selected from the NCBI database (NCBI GenBank Database, entry ×51730, human mRNA and promoter DNA for the *PR*). Three SNPs, rs2008112; rs3740754 and rs2020880 were monomorphic in our sample and were excluded from data analyses. Figure 1 shows the positions of the SNPs genotyped in the *PR* gene.

Genotyping

Forward and reverse PCR primers and a primer extension probe were designed using SpectroDESIGNER software (Sequenom™, San Diego, CA, USA) and assembled into a multiplex set. SNPs were typed using the Sequenom™ matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) MassARRAY protocol. The PCR volume of 5 µl contained 15 ng DNA, 10× PCR buffer, 25 mM $MgCl_2$, 25 mM dNTP, forward and reverse primers to a final concentration of 50 nM for all SNPs (excluding rs10895068 which was 75 nM) and 5 U/µl of Taq polymerase. PCR cycling conditions were an initial denaturation step of 15 min at 95°C, followed by 45 cycles of 20 s at 95°C, 30 s at 56°C, 1 min at 72°C and a final extension step of 3 min at 72°C. To neutralize and remove residual unincorporated dNTPs from the PCR, shrimp alkaline phosphatase (1 U/µl) was added and the reaction incubated at 37°C for 20 min and then 85°C for 5 min.

The MassEXTEND PCR mix contained allele-specific extension primers at 300nM each, h-ME (ACT or ACG) extension mix and 32 U/µl of ThermoSequenase. PCR cycling conditions were a denaturation step of 2 min at 94°C, followed by 55 cycles of 5 s at 94°C, 5 s at 52°C and 5 s at 72°C. Purification of h-ME reaction products was performed by the addition of resin. Primer products were spotted onto 384-well chips, precoated with a 3-hydroxypicolinic acid matrix, to help bind and stabilize DNA during ionisation by the MassARRAY. Loading of the products was performed using a nanolitre pipetting system (SpectroCHIP, SpectroJet; Sequenom) and reaction products analysed through MALDI-TOF MassARRAY (Sequenom).

We used a modified protocol to examine the presence/absence of the AluIns polymorphism in a subsample (Spurdle *et al*., 2002). A subset of DNA samples

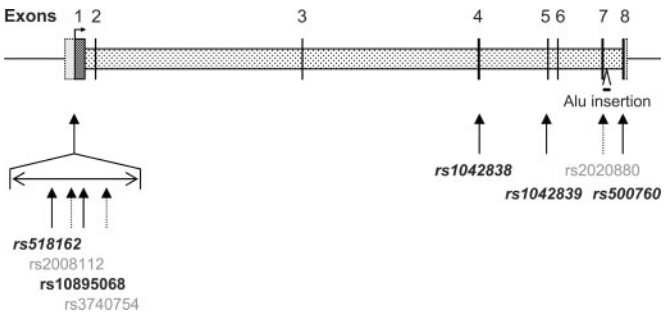

Figure 1. Progesterone receptor (*PR*) gene showing relative positions of single-nucleotide polymorphisms (SNPs) investigated in the study; SNPs that were polymorphic in our sample are indicated by solid arrow lines and bold black font; SNPS that were non-polymorphic in our sample are in grey font and SNPs which appear in HapMap (The International HapMap Consortium, 2003) are in italics.

(n = 101) was typed successfully by PCR using forward and reverse primers 5′-ATACGGTATCCATGACATGAG-3′ and 5′-AAGTATTTTCTTGCTAAATGTCTG-3′, respectively. The reaction volume of 10 ul contained 15 ng of DNA, primer concentrations of 10 pmol each, dNTPs (200 nM), 1× Perkin-Elmer Taq polymerase buffer, 1 unit of Taq polymerase and 1.5 mM $MgCl_2$. Amplification conditions were an initial denaturation step of 10 min at 94°C, 35 cycles of 94°C for 20 s, 51°C for 20 s, 72°C for 60 s and a final extension step of 72°C for 10 min. PCR products of 65 bp for wild type and 385 bp with AluIns were resolved on a 2.5% agarose gel (Agarose 1000 UltraPure, Invitrogen, Carlsbad, CA, USA).

Statistical analysis

To avoid spurious associations arising from population admixture we used a family-based design to study parental alleles/haplotypes transmitted and not transmitted to affected offspring. Our strategy for data analysis was based on (i) typing selected SNPs in the triad set as outlined, (ii) typing the AluIns in a subset of individuals, (iii) estimating LD between the AluIns and flanking SNPs, and between all the SNPs typed and (iv) using the transmission disequilibrium test (TDT) for detecting association between endometriosis and individual SNP and haplotypes. Because inference about association between the AluIns and endometriosis depends on the extent of LD with its proximal SNPs, it was important to estimate LD before proceeding.

Data handling was performed using SAS (SAS Institute Inc., 2001) and Mendelian error checking using SIB-PAIR software (Duffy, 2001). For estimating LD between pairs of variants in the *PR* gene, in addition to Lewontin's D′ linkage disequilibrium statistic, we also estimated Δ^2, effectively providing an estimate of inter-SNP correlation ($r = \sqrt{\Delta^2}$). LD between the SNPs across the gene was estimated using SNPSpD (Nyholt, 2004). LD calculations involving the AluIns were based on data from up to 101 individuals, and for other SNPs using data based on individuals in the 980 triads, numbers ranged between 2889 and 2913.

The single-locus TDT (Spielman *et al*., 1993), a robust test for allelic association in the presence of population stratification, which uses information from heterozygous parents, was performed on SNP data from 980 cases and both parents (2940 individuals) using SIB-PAIR software (Duffy, 2001). This TDT statistic is the Pearson goodness-of-fit based test of symmetry in the square table of transmitted versus non-transmitted alleles to each affected child (Haberman, 1979). The genotypic TDT *P*-value is estimated through gene-dropping based on the genotypes of both typed parents of the proband. The test statistic compares the observed number of each genotype transmitted with the number expected based on the parental genotypes. Multi-locus TDT, using a sliding window approach across the 5 SNP genotypes, was performed using GENEHUNTER 2.0 (Kruglyak *et al*., 1996) to compare the frequencies of transmitted and untransmitted SNP haplotypes, excluding the AluIns. These analyses were also performed for phenotypic strata of endometriosis families, to assess whether any effects were restricted to certain phenotypes. We stratified triad families according to three important phenotypes of cases: age at diagnosis, subfertility and disease stage.

We first estimated haplotype frequencies between AluIns and proximal SNPS rs1042839 and rs500760 in 87 affected probands where complete data were available, using ARLEQUIN version 2.000 (Schneider *et al*., 2000). Because we did not type the full PROGINS complex and typed the AluIns component in a relatively small subsample, we conducted a final check of haplotype transmission. We used TDTPHASE (Dudbridge, 2004) to imput the missing genotypes, enabling estimation of haplotypes for the AluIns and its two flanking SNPs for a much larger sample, and to conduct a full likelihood ratio test in a log-linear model using the expectation maximization (EM) algorithm to obtain maximum-likelihood estimates for haplotype transmission disequilibrium.

Finally, given the strong LD between the AluIns and rs1042839, we used the latter alleles and genotypes to draw comparisons and analyse pooled results from our own and the Austrian (Wieser *et al*., 2002) and Italian (Lattuada *et al*., 2004) studies. Because of the different designs, we used transmitted and untransmitted alleles of rs1042839, which was in strong LD with the AluIns in the Australian data, to combine with case and control data, respectively, from other studies. Given the significant heterogeneity between the results of the three studies, a pooled allelic odds ratio was calculated, under a random effects model, from genotype counts using the methods of DerSimonian and Laird (1986) and Whitehead and Whitehead (1991).

Results

All SNPs were in Hardy–Weinberg equilibrium (probabilities 0.37–0.84, restricted to founders). Minor allele frequencies for the five SNPs in founders ranged from 0.061 to 0.241 (Table I), and for the minor (*T2*) allele of the AluIns was 0.124 (N = 101).

There was evidence for significant allelic association with endometriosis for SNP rs500760 in exon 8 (P = 0.0270), with the ratio of transmission to non-transmission for the A allele of 386:326 (Table II). Genotypic transmission disequilibrium was not significant (empirical P = 0.075). Results were consistent when triads involving identically heterozygous parents were removed from analysis (P = 0.026). There was no evidence for association with SNPs in the promoter region of *PR*.

There was strong LD across the *PR* locus (Table III). Five common haplotypes accounted for more than 99% of chromosomes in the sample (Table IV). There was significant over-transmission of the *GGGCA* haplotype (378:315, P = 0.017), but no significant

Table I. Single-nucleotide polymorphisms (SNPs) at the progesterone receptor (*PR*) locus genotyped in 980 triad families (N = 2940) with endometriosis

SNP	Position	Location	Change (transcribed strand)	Minor allele frequency (founders)
rs518162	–700	5′UTR	G>A	0.068
rs2008112	–569	5′UTR	G>A	Monomorphic
rs10895068	–413	5′UTR	G>A	0.061
rs3740754	168	Exon 1	G>C	Monomorphic
rs1042838	66389	Exon 4	G>T	0.155
rs1042839	77599	Exon 5	C>T	0.154
rs2020880	87073	Exon 7	C>T	Monomorphic
rs500760	89810	Exon 8	A>G	0.241

Table II. Association (TDT) between endometriosis and single-nucleotide polymorphisms (SNPs) at the progesterone receptor (*PR*) locus

SNP	Number of informative probands	χ^2	Over-transmitted allele	*P*
rs518162	225	1.4	G	0.270
rs10895068	210	0.0	G	0.894
rs1042838	438	0.7	G	0.420
rs1042839	435	1.0	C	0.345
rs500760	581	5.1	A	0.027

Table III. Linkage disequilibrium D′ (values denoted with *) and Δ^2 (values denoted with †) measures of association between pairs of variants in the progesterone receptor (*PR*) gene

Variant 1	Variant 2					
	rs518162	rs10895068	rs1042838	Rs1042839	AluIns	rs500760
rs518162	–	**0.998***	**0.995***	**0.995***	0.757*	**0.999***
rs10895068	0.005†	–	**0.996***	**0.995***	1.000*	**0.999***
rs1042838	0.013†	0.012†	–	**0.996***	**0.951***	**0.984***
rs1042839	0.013†	0.012†	0.984†	–	**1.000***	**0.987***
AluIns	0.005†	0.007†	0.826†	0.877†	–	**0.942***
rs500760	0.022†	0.020†	0.570†	0.568†	0.438†	–

Δ is the correlation between two biallelic variants. Significant χ^2 values ($P < 0.0001$) are in bold. All single-nucleotide polymorphism (SNPs) typed on 2867–2881 samples. AluIns typed on 101 individuals including 87 affected probands.

Table IV. Frequency of transmitted and untransmitted progesterone receptor (*PR*) haplotypes (rs518162, rs10895068, rs1042838, rs1042839, rs500760) with frequency greater than one

Haplotype	Transmitted	Untransmitted	χ^2	*P*
AGGCA	99	112	0.80	0.822
GAGCA	101	90	0.63	0.426
GGTCG	3	1	1.00	0.317
GGTTA	4	2	0.67	0.414
GGTTG	173	203	2.39	0.122
GGGCA	378	315	5.73	0.017
GGGCG	101	130	3.64	0.056
GGGTA	0	2	2.00	0.157

under-transmission or 'protective' haplotype. This result stood with the more stringent test after eliminating triads with double-heterozygous parents, although the significance level for the susceptibility haplotype decreased to $P = 0.041$ (Table IV).

We stratified triad families according to phenotype of cases: age at diagnosis, subfertility and disease stage. There were no effects of age at diagnosis. The association with rs500760 was very strong in the subgroup reporting subfertility (problems conceiving), with overtransmission of the *A* allele and consequent undertransmission of *G* allele ($P = 0.007$, Table V). No association was shown in either those who reported no problems conceiving or in those who had never tried conceiving. Interestingly, the evidence for allelic association between endometriosis and rs500760 was also increased in triads with stage A (minimal to mild) disease ($P = 0.006$), despite the negative phenotypic association between subfertility and disease stage. In triads with stage B (moderate to severe) disease, there was no association with rs500760 and no significant association with any other SNP. Over the five SNPs, there was both a significant susceptibility and a protective haplotype in stage A probands (Table V). When triads with double heterozygote parents were removed from analysis, the evidence for significance of the *GGGCA* susceptibility haplotype decreased to $P = 0.049$, and that of the protective *GGTTG* haplotype increased to $P = 0.0280$. In the subset of families with both subfertility and minimal–mild disease overtransmission of the GGGCA haplotype failed to reach significance. The *GGGCG* haplotype appeared protective, although numbers were small.

Inference about association between the AluIns and endometriosis depends on the extent of LD with its proximal SNPs. We estimated Δ^2, effectively providing an estimate of inter-SNP correlation $r = \sqrt{\Delta^2}$. There was very strong LD between rs1042839 and the AluIns ($\Delta^2 = 0.877$, $D' = 1.000$, $P = 0.000$). LD between the AluIns and rs500760 was also strong but not complete ($\Delta^2 = 0.438$, $D' = 0.942$, $P = 0.000$), and similarly between the two flanking SNPs themselves ($\Delta^2 = 0.568$, $D' = 0.987$, $P = 0.000$). Analysis of LD between the AluIns and the two flanking markers in 99 individuals typed for all three markers showed that the *C T1 A* haplotype accounted for 74% of *PR* chromosomes in these families, and the *T T2 G* haplotype accounted for only 13% of the 87 affected probands typed for the AluIns (Table VI). Further, the low frequency of the *T T1 G* haplotype suggests a misclassification rate of no more than 2% if the *T G* haplotype for rs1042839 and rs500760 were used to infer the AluIns genotype.

Because we did not type the full PROGINS complex and typed the AluIns component in a subsample, we assessed haplotype transmission of the AluIns and its flanking SNPs, imputing missing genotypes using TDTPHASE (Dudbridge, 2004). The non-significant likelihood ratio statistic of 6.259, $df = 5$ ($P = 0.2818$), excluding rare haplotypes, suggested no haplotype transmission disequilibrium and provided no reason to modify our conclusion that our data suggested no convincing evidence for association between the AluIns and endometriosis.

Given the strong LD between the AluIns and rs1042839, we used the latter alleles and genotypes to draw comparisons and analyse pooled results from our own and the Austrian (Wieser *et al.*, 2002) and Italian (Lattuada *et al.*, 2004) studies. Because of the different designs, we used transmitted and untransmitted alleles of rs1042839, which was in strong LD with the AluIns in the Australian data, to combine with case and control data, respectively, from other studies. Combined analysis suggests that a significant risk effect of the *T2* allele or of the *T1/T2* genotype for endometriosis is unlikely (Table VII).

Discussion

The AluIns polymorphism in intron G of the *PR* gene was reported to be associated with susceptibility to endometriosis (Wieser *et al.*, 2002; Lattuada *et al.*, 2004). There is strong LD between two SNP variants (rs1042838 in exon 4, rs1042839 in exon 5) and the AluIns polymorphism (Agoulnik *et al.*, 1997; Kieback *et al.*, 1998). We therefore investigated five SNPs in the *PR* gene for association with endometriosis in a large set of Australian families with clinically confirmed endometriosis. We typed the two SNPs reported in LD with the AluIns polymorphism (rs1042838 and rs1042839), one SNP in exon 8

Table V. Significant transmission of progesterone receptor (*PR*) haplotypes (rs518162, rs10895068, rs1042838, rs1042839, rs500760) for phenotypic strata of endometriosis families

Phenotypic stratum	SNP	Allele/haplotype	Transmitted	Untransmitted	*N*	χ^2	*P*
Subfertility							
Subfertility+*	rs500760	A	168	121	231	7.60	0.007
No subfertility	rs500760	A	81	79	135	0.00	0.937
Never tried conceiving	rs500760	A	137	126	215	0.05	0.538
Disease stage							
Stage A	rs500760	A	277	217	404	7.30	0.008
Stage B	rs500760	A	109	109	177	0.00	1.000
Stage A	rs518162–	GGGCA	271	218		5.74	0.017
	rs500760	GGTTG	116	150		4.35	0.037
Subfertility and stage A	rs500760	A	108	71		7.65	0.006
	rs518162–	GGGCA	104	80		3.13	0.077
	rs500760	GGGCG	22	40		5.23	0.022

*Daughters diagnosed with endometriosis reporting problems conceiving.

Table VI. Haplotype frequencies estimation (Schneider *et al*., 2000) between *Alu* insertion (AluIns) and proximal single-nucleotide polymorphisms (SNPs), rs1042839 and rs500760, in 87 affected probands

Number	Haplotype			Frequency	SD
	rs1042839	AluIns	rs100760		
1	C	T1	A	0.744	0.034
2	T	T2	G	0.131	0.026
3	C	T1	G	0.107	0.024
4	T	T1	G	0.018	0.010

flanking the AluIns polymorphism (rs500760) and two SNPs in the promoter region (rs518162 and rs10895068) including +331G/A (rs10895068) reported to be associated with endometrial cancer (De Vivo *et al*., 2002). We found no association between rs1042838 or rs1042839 and endometriosis in our families, and unconvincing evidence of association between endometriosis and the exon 8 SNP rs500760. The confidence bounds of the association were broad considering multiple testing, so we consider the association to be marginal. If multiple comparison adjustment is made using a false discovery rate (Benjamini and Hochnerg, 1995) of 0.05, neither the result in Table II ($P = 0.027$) nor in Table IV ($P = 0.017$) is significant. Further, the rs500760 SNP is located in the third codon position (Gln886Gln) and may not have functional significance.

Our results cannot be interpreted as supporting association between variation in *PR* and susceptibility to endometriosis and do not replicate previous findings. The AluIns has been reported in complete LD with the two exonic SNPs, rs1042838 and 3764C/T (Spurdle *et al*., 2001; De Vivo *et al*., 2002). We genotyped the AluIns allele in a subset of samples. Our data showed very strong, but not complete LD. The frequency of the AluIns mutant *T2* allele in our 87 typed endometriosis probands was 0.128, approximately mid-way between its frequency in cases (0.17) and controls (0.08) in the Austrian study (Wieser *et al*., 2002) and the Italian study (0.17 for cases and 0.11 for controls) (Lattuada *et al*., 2004). Our *T2* frequency was also lower than the frequency of 0.16% reported for an Australian control sample (Spurdle *et al*., 2002). The homozygote AluIns was rare in our sample ($n = 2$, 1.9%), slightly less common than the 3% and 2.52% observed in ovarian cancer cases and controls, respectively (Spurdle *et al*., 2001), although our standard error would be very high. Our heterozygote *T1/T2* genotype was 23%, compared to the 22% and 21% reported for ovarian cases and controls (Spurdle *et al*., 2001).

Unlike Lattuada *et al*.'s study (2004), there was no trend for increased prevalence of the mutant *T2* allele in cases of more severe disease stage. In fact, the *T1/T1* genotype was much more prevalent in cases of stage B disease (87%) compared with stage A disease (66%), and the heterozygous *T1/T2* genotype was more prevalent in stage A than stage B cases (32% compared with 11%) ($P = 0.055$). Given that we found a stronger association with rs500760 and endometriosis in women with minimal to mild disease, and given the very strong LD between the *PR* AluIns polymorphism and proximal SNPs, no association could be inferred for endometriosis with the *T2* allele in our

Table VII. Summary and pooled analysis of *Alu* insertion (AluIns) and endometriosis in three studies

Study	Allele	Cases	Controls	OR	95% CI
Austrian (Wieser *et al*., 2002)	T2	33	17		
	T1	157	197	2.41*	1.31–4.53*
Italian (Lattuada *et al*., 2004)	T2	45	28		
	T1	217	226	1.7*	1.0–2.8*
		Transmitted†	Untransmitted†		
Australian	T	293	305		
	C	1641	1637	0.96	0.81–1.14
Pooled analysis of alleles§	T2/T	371	350		
	T1/C	2015	2060	1.20	0.91–1.48
	Genotype	Cases	Controls	Trend χ^2	*P*
Austrian (Wieser *et al*., 2002)	T2/T2	3	1		
	T2/T1	27	15		
	T1/T1	65	91	7.88	0.005
Italian (Lattuada *et al*., 2004)	T2/T2	3	1		
	T2/T1	39	26		
	T1/T1	89	100	4.16	0.041*
		Transmitted	Untransmitted‡		
Australian†	T/T	25	22		
	T/C	243	261		
	C/C	699	688	0.23	0.633
		Transmitted	Untransmitted†	Pooled OR	95% CI
Pooled analysis of genotypes¶	T2/T2	31	24		
	T2/T1	309	302		
	T1/T1	853	879	1.16	0.92–1.47

CI, confidence interval; OR, odds ratio.
*Published results.
†rs1042839,a flanking SNP in strong linkage disequilibrium (LD) with the AluIns.
‡Frequency in parents minus frequency in affected daughters.
§We pooled major alleles (C) of rs1042839 from the Australian data with the T1 major allele of the AluIns and T minor allele of rs1042839 with the T2 AluIns allele from the other studies. Because of heterogeneity among the three studies, pooled allelic odds ratios were calculated under a random effects model from genotype counts (DerSimonian and Laird, 1986; Whitehead and Whitehead, 1991).
¶We pooled T/T, T/C and C/C from rs1042839 with T2/T2, T2/T1 and T1/T1 genotypes from the AluIns, respectively, in other studies.

data; it was not evident particularly in cases of moderate–severe disease. Therefore we did not replicate findings from the earlier two case control studies. Results of pooling of our data, using a flanking SNP (rs1042839) in strong LD with the AluIns, with the published data from the Austrian (Wieser *et al*., 2002) and Italian (Lattuada *et al*., 2004) studies suggested that association with endometriosis is unlikely.

There was no evidence for association between endometriosis and SNPs in the promoter region of the *PR* gene, which included a functional polymorphism associated with increased endometrial cancer risk (De Vivo *et al*., 2002). We found very strong LD across the *PR* locus. Serum prolactin levels have been implicated as a cause for infertility in patients with minimal to mild endometriosis (Gregoriou *et al*., 1999; Martinez *et al*., 2002). The promoter SNP rs10895068 was associated with prolactin levels in a population sample (Westberg *et al*., 2004); although this SNP did not show individual association with endometriosis in our families, or in the subset with minimal–mild endometriosis and subfertility, it may warrant further investigation.

Like De Vivo *et al*. (2002), we estimated eight haplotypes with a frequency greater than 1%, and five common haplotypes. We typed SNPs in the *PR* promoter (De Vivo *et al*., 2002) and SNPs in LD with the insertion/deletion polymorphism (Wieser *et al*., 2002) implicated in endometriosis. Recent data from large scale genotyping across the human genome typed 44 SNPs across the *PR* locus (http://www.hapmap.org). The data show strong LD in general agreement with results of our study, but greater numbers of SNPs that identify all common alleles would need to be typed to completely exclude association between common variants in *PR* and endometriosis. An in-depth sequence approach to the *PR* gene would be necessary to detect allelic variants that were large in number but individually rare. Nevertheless, caution is again required as transcriptional differences as unmasked by microarray studies may well be the result of epigenetic factors, which are being shown to be increasingly important in influencing gene expression over the lifespan (Fraga *et al*., 2005).

The fact that our results did not support the association between (more severe) endometriosis and the AluIns *T2* allele reported in earlier studies may be a function of the poor predictive value of early studies reporting 'positive' findings (Trikalinos *et al*., 2004); indeed studies are seldom replicated when the initial sample size is low (Ioannidis *et al*., 2001). A more directed approach to targeting potential candidate genes for future association studies may come from different approaches. Using microarray analysis, Arimoto *et al*. (2003) analysed the expression patterns of 23 040 genes of endometrial tissue in various phases throughout the menstrual cycle. Borthwick *et al*. (2003) examined transcript abundance and changes in transcript level for 60 000 gene targets between normal endometrium in the proliferative and secretory phases of the menstrual cycle. These studies have shown altered levels of expression of genes not previously thought to cycle during the different endometrial phases.

In a recent genome-wide scan of microsatellite markers, we found linkage between various chromosomal regions and endometriosis (Treloar *et al*., 2000; Wicks *et al*., 2002). Obviously, these regions warrant further investigation. The genome-wide scan data showed no evidence for a locus for endometriosis on chromosome 11q (Treloar *et al*., 2005). Our findings fail to provide support for association between variants in *PR* and endometriosis and demonstrate the complexity of genotype–phenotype associations in this disease.

Acknowledgements

We thank women with endometriosis and their families for participation in the study, Barbara Haddon for coordinating recruitment, Anjali Henders and Renée Mayne for sample management and genotyping, Lien Le for bioinformatics, David Smyth and Olivia Zheng for data management, Michael James for advice on SNP genotyping methods, and Chris Collet (Queensland University of Technology) for helpful discussions. This study was supported in part by grants from the Cooperative Research Centre for the Discovery of Genes for Common Human Diseases established and supported by the Australian Government's Cooperative Research Centre's Program and National Health and Medical Research Council of Australia (339430). This study was supported by the Cooperative Research Centre for Discovery of Genes for Common Human Diseases, Australia (1997–2004) and the National Health and Medical Research Council of Australia (339430).

References

Agoulnik I, Weigel N, Tong XW, Bingman WE, Estella NM, Blankenberg K, Runnebaum IB, Korner W, Fishman A, Atkinson EN et al. (1997) Functional analysis of mutated progesterone receptor that cosegregates with sporadic ovarian cancer. Proc Am Assoc Cancer Res Annu Meet 38,453.

American Fertility Society (1985) Revised American Fertility Society classification of endometriosis: 1985. Fertil Steril 43,351–352.

Arimoto T, Katagiri T, Oda K, Tsunoda T, Yasugi T, Osuga Y, Yoshikawa H, Nishii O, Yano T, Taketani Y et al. (2003) Genome-wide cDNA microarray analysis of gene-expression profiles involved in ovarian endometriosis. Int J Oncol 22,551–560.

Attia GR, Zeitoun K, Edwards D, Johns A, Carr BR and Bulun SE (2000) Progesterone receptor isoform A but not B is expressed in endometriosis. J Clin Endocrinol Metab 85,2897–2902.

Benjamini Y and Hochnerg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. JR Stat Soc Ser B (Methodology) 57,289–300.

Borthwick JM, Charnock-Jones DS, Tom BD, Hull ML, Teirney R, Phillips SC and Smith SK (2003) Determination of the transcript profile of human endometrium. Mol Hum Reprod 9,19–33.

Conneely OM, Maxwell BL, Toft DO, Schrader WT and O'Malley BW (1987) The A and B forms of the chicken progesterone receptor arise by alternate initiation of translation of a unique mRNA. Biochem Biophys Res Commun 149,493–501.

Conneely OM, Kettelberger DM, Tsai MJ, Schrader WT and O'Malley BW (1989) The chicken progesterone receptor A and B isoforms are products of an alternate translation initiation event. J Biol Chem 264,14062–14064.

Conneely OM, Mulac-Jericevic B and Lydon JP (2003) Progesterone-dependent regulation of female reproductive activity by two distinct progesterone receptor isoforms. Steroids 68,771–778.

D'Hooghe TM (2003) Immunomodulators and aromatase inhibitors: are they the next generation of treatment for endometriosis? Curr Opin Obstet Gynecol 15,243–249.

De Vivo I, Huggins GS, Hankinson SE, Lescault PJ, Boezen M, Colditz GA and Hunter DJ (2002) A functional polymorphism in the promoter of the progesterone receptor gene associated with endometrial cancer risk. Proc Natl Acad Sci USA 99,12263–12268.

Del Carmen MG, Smith Sehdev AE, Fader AN, Zahurak ML, Richardson M, Fruehauf JP, Montz FJ and Bristow RE (2003) Endometriosis-associated ovarian carcinoma. Cancer 98,1658–1663.

DerSimonian R and Laird N (1986) Meta-analysis in clinical trials. Control Clin Trials 7,177–188.

Dudbridge F (2004) TDTPHASE. Hinxton, Cambridge.

Duffy D (2001) SIB-PAIR. Queensland Institute of Medical Research, Brisbane, Australia.

Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, Heine-Suner D, Cigudosa JC, Urioste M, Benitez J et al. (2005) Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci USA 102,10604–10609.

Gregoriou G, Bakas P, Vitoratos N, Papadias K, Goumas K, Chryssicopoulos A and Creatsas G (1999) Evaluation of serum prolactin levels in patients with endometriosis and infertility. Gynecol Obstet Invest 48,48–51.

Gurates B and Bulun SE (2003) Endometriosis: the ultimate hormonal disease. Semin Reprod Med 21,125–134.

Haberman SJ (1979) Analysis of Quantitative Data. New Developments. Academic Press, New York.

Hadfield RM, Mardon HJ, Barlow DH and Kennedy SH (1997) Endometriosis in monozygotic twins. Fertil Steril 68,941–942.

Ioannidis JP, Ntzani EE, Trikalinos TA and Contopoulos-Ioannidis DG (2001) Replication validity of genetic association studies. Nat Genet 29,306–309.

Jiang X, Hitchcock A, Bryan EJ, Watson RH, Englefield P, Thomas EJ and Campbell IG (1996) Microsatellite analysis of endometriosis reveals loss of

heterozygosity at candidate ovarian tumor suppressor gene loci. Cancer Res 56,3534–3539.

Jiang J, Wu R, Wang Z, Sun H, Xu Z and Xiu H (2002) Effect of mifepristone on estrogen and progesterone receptors in human endometrial and endometriotic cells in vitro. Fertil Steril 77,995–1000.

Kastner P, Krust A, Turcotte B, Stropp U, Tora L, Gronemeyer H and Chambon P (1990) Two distinct estrogen-regulated promoters generate transcripts encoding the two functionally different human progesterone receptor forms A and B. EMBO J 9,1603–1614.

Kennedy S (2003) Genetics of endometriosis: a review of the positional cloning approaches. Semin Reprod Med 21,111–118.

Kennedy S, Mardon H and Barlow D (1995) Familial endometriosis. J Assist Reprod Genet 12,32–34.

Kieback DG, Tong XW, Weigel NL and Agoulnik IU (1998) A genetic mutation in the progesterone recptor (PROGINS) leads to an increased risk of non-familial breast and ovarian cancer causing inadequate control of estrogen receptor driven proliferation. J Soc Gynecol Investig 5,40a.

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

Kruglyak L, Daly MJ, Reeve-Daly MP and Lander ES (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet 58,1347–1363.

Lattuada D, Somigliana E, Vigano P, Candiani M, Pardi G and Di Blasio AM (2004) Genetics of endometriosis: a role for the progesterone receptor gene polymorphism PROGINS? Clin Endocrinol (Oxf) 61,190–194.

Martinez LB, Leyva MZ and Romero IC (2002) Prolactin receptor in human endometriotic tissues. Acta Obstet Gynecol Scand 81,5–10.

Meresman GF, Auge L, Baranao RI, Lombardi E, Tesone M and Sueldo C (2002) Oral contraceptives suppress cell proliferation and enhance apoptosis of eutopic endometrial tissue from patients with endometriosis. Fertil Steril 77,1141–1147.

Miller SA, Dykes DD and Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16,1215.

Misao R, Iwagaki S, Fujimoto J, Sun W and Tamaya T (1999) Dominant expression of progesterone receptor form B mRNA in ovarian endometriosis. Horm Res 52,30–34.

Moen MH (1994) Endometriosis in monozygotic twins. Acta Obstet Gynecol Scand 73,59–62.

Moutsatsou P and Sekeris CE (2003) Steroid receptors in the uterus: implications in endometriosis. Ann N Y Acad Sci 997,209–222.

Ness RB (2003) Endometriosis and ovarian cancer: thoughts on shared pathophysiology. Am J Obstet Gynecol 189,280–294.

Nyholt DR (2004) A simple correction for multiple testing for SNPs in linkage disequilibrium with each other. Am J Hum Genet 74,765–769.

Rowe SM, Coughlan SJ, McKenna NJ, Garrett E, Kieback DG, Carney DN and Headon DR (1995) Ovarian carcinoma-associated TaqI restriction fragment length polymorphism in intron G of the progesterone receptor gene is due to an Alu sequence insertion. Cancer Res 55,2743–2745.

Runnebaum IB, Wang-Gohrke S, Vesprini D, Kreienberg R, Lynch H, Moslehi R, Ghadirian P, Weber B, Godwin AK, Risch H et al. (2001) Progesterone receptor variant increases ovarian cancer risk in BRCA1 and BRCA2 mutation carriers who were never exposed to oral contraceptives. Pharmacogenetics 11,635–638.

SAS Institute Inc. (2001) SAS Proprietary Software Release 8.2 1999–2001. SAS Institute Inc., Carey, NC.

Schneider S, Roessli D and Excoffier L (2000) Arlequin ver. 2.000: A Software for Population Genetics Data Analysis. Genetics and Biometry Laboratory, University of Geneva, Geneva, Switzerland.

Simpson JL and Bischoff FZ (2002) Heritability and molecular genetic studies of endometriosis. Ann N Y Acad Sci 955,239–251.

Spielman RS, McGinnis RE and Ewens WJ (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am J Hum Genet 52,506–516.

Spurdle AB, Webb PM, Purdie DM, Chen X, Green A and Chenevix-Trench G (2001) No significant association between progesterone receptor exon 4 Val660Leu G/T polymorphism and risk of ovarian cancer. Carcinogenesis 22,717–721.

Spurdle AB, Hopper JL, Chen X, McCredie MR, Giles GG, Venter DJ, Southey MC and Chenevix-Trench G (2002) The progesterone receptor exon 4 Val660Leu G/T polymorphism and risk of breast cancer in Australian women. Cancer Epidemiol Biomarkers Prev 11,439–443.

Stefansson H, Geirsson RT, Steinthorsdottir V, Jonsson H, Manolescu A, Kong A, Ingadottir G, Gulcher J and Stefansson K (2002) Genetic factors contribute to the risk of developing endometriosis. Hum Reprod 17,555–559.

The International HapMap Consortium (2003) The International HapMap Project. Nature 426,789–796. Available from http://www.hapmap.org

Treloar SA, O'Connor DT, O'Connor VM and Martin NG (1999) Genetic influences on endometriosis in an Australian twin sample. Fertil Steril 71,701–710.

Treloar SA, Bahlo M, Ewen K, O'Connor DT, Duffy DL, Wicking CA, Wainwright BJ, Montgomery GW and Martin NG (2000) Suggestive linkage for endometriosis found in genome-wide scan. Am J Hum Genet 67,1764.

Treloar S, Hadfield R, Montgomery G, Lambert A, Wicks J, Barlow DH, O'Connor DT and Kennedy S (2002) The International Endogene Study: a collection of families for genetic research in endometriosis. Fertil Steril 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. (2005) Genomewide linkage study in 1,176 affected sister pair families identifies a significant susceptibility locus for endometriosis on chromosome 10q26. Am J Hum Genet 77,365–376.

Trikalinos TA, Ntzani EE, Contopoulos-Ioannidis DG and Ioannidis JP (2004) Establishment of genetic associations for complex diseases is independent of early study findings. Eur J Hum Genet 12,762–769.

Westberg L, Ho HP, Baghaei F, Nilsson S, Melke J, Rosmond R, Holm G, Bjorntorp P and Eriksson E (2004) Polymorphisms in oestrogen and progesterone receptor genes: possible influence on prolactin levels in women. Clin Endocrinol (Oxf) 61,216–223.

Whitehead A and Whitehead J (1991) A general parametric approach to the metaanalysis of randomized clinical-trials. Stat Med 10,1665–1677.

Wicks J, Zabaneh D, Treloar SA, Hadfield R, Dawson G, Lambert A, Haddon B, Mackay I, Weeks DE, O'Connor DT et al. (2002) International Endogene Study finds strong evidence of susceptibility loci for endometriosis at two genomic regions. Am J Hum Genet 71, S215.

Wieser F, Schneeberger C, Tong D, Tempfer C, Huber JC and Wenzl R (2002) PROGINS receptor gene polymorphism is associated with endometriosis. Fertil Steril 77,309–312.

Zondervan KT, Cardon LR and Kennedy SH (2002) What makes a good case-control study? Design issues for complex traits such as endometriosis. Hum Reprod 17,1415–1423.

Zondervan KT, Weeks DE, Colman R, Cardon LR, Hadfield R, Schleffler J, Trainor AG, Coe CL, Kemnitz JW and Kennedy SH (2004) Familial aggregation of endometriosis in a large pedigree of rhesus macaques. Hum Reprod 19,448–455.

Submitted on May 9, 2005; accepted on July 27, 2005