

ORIGINAL ARTICLE

A microRNA-520 mirSNP at the *MMP2* gene influences susceptibility to endometriosis in Chinese women

Eing-Mei Tsai^{1,2}, Yung-Song Wang³, Chang-Shen Lin², Wen-Yin Lin³, Edward Hsi^{2,4}, Ming-Tsang Wu^{5,6} and Suh-Hang H Juo^{4,5}

The *MMP2* gene has been implicated in the pathogenesis of endometriosis. We investigated the role and function of single-nucleotide polymorphisms (SNP) of *MMP2* in relation to endometriosis. First a case–control study was conducted and 17 SNPs were examined in 211 patients and 344 controls. Regression analysis was used to evaluate the genetic effect. We used reporter assay to validate the functional consequences of the significant SNP. Two SNPs (rs243832 and rs7201) had *P*-values <0.05 and they are in strong linkage disequilibrium ($D' = 0.96$ and $r^2 = 0.47$). Further analysis showed that rs7201 but not rs246832 was an independent risk factor and the risk C allele of rs7201 had an odds ratio (OR) of 1.88 ($P = 0.004$). SNP rs7201 is located at the 3'-untranslated region and is predicted to be within the microRNA-520g binding site. The reporter assay for rs7201 showed that the risk C allele had a higher expression level than the A allele ($P = 0.027$). Using microRNA-520g mimic and inhibitor, the results indicated that the A allele but not the risk C allele can be regulated by microRNA-520g. The C allele of SNP rs7201 increases a risk for endometriosis because of out of regulation by microRNA-520g.

Journal of Human Genetics (2013) 58, 202–209; doi:10.1038/jhg.2013.1; published online 31 January 2013

Keywords: endometriosis; microRNA-520; *MMP2*; polymorphism

INTRODUCTION

Endometriosis is a common gynecological disorder. The disease is characterized by the presence of endometrial glands and stroma outside their normal locations. Symptoms associated with endometriosis include dysmenorrhea, dyspareunia, non-cyclic pelvic pain and infertility. A growing body of evidence indicates that genetic factors influence susceptibility to the disease^{1,2} and several genetic polymorphisms have been implicated for the risk of this disease.³ Several steps have been implicated for the development of endometriosis, which include retrograde menstruation, attachment of endometrial fragments to the epithelium of the peritoneum, invasion of the epithelium, establishment of a blood supply and generation of a suboptimum immune response that does not adequately clear the implants.⁴

Proteolytic activity of cell–surface proteins and extracellular matrix has a pivotal role in the biological process of menstruation and pathological process in the development of endometriosis. The process of cell migration or invasion is associated with the degradation and reconstruction of the surrounding microenvironment that consists of extracellular matrix components. The matrix

metalloproteinases (MMPs) are a broad family of zinc-binding endopeptidases that are responsible for the degradation of extracellular matrix components in several organs/systems including female reproductive system.⁵ It was also shown that endometrial stromal cells secrete several MMPs, including *MMP2*.^{6,7} The uterine endometrium from women with endometriosis expresses a higher level of *MMP2*.^{6,7} Accumulating evidence suggests that *MMP2* is involved in the formation of endometriosis.^{6–10}

This study aimed to test for an association between polymorphisms of the *MMP2* gene and to elucidate the underlying molecular mechanism of genetic variation in relation to endometriosis. We first used tagging single-nucleotide polymorphisms (tSNPs) plus a known functional SNP to evaluate systematically the association between the *MMP2* gene and the disease. As genetic association studies are subject to a type I error, we further conducted a series of functional studies to support the statistical result from the association study. This study illustrated that microRNA-520g (miR-520g) and SNP rs7201 has a pivotal role in the regulation of *MMP2* expression, which in turn confers risk for endometriosis.

¹Department of Obstetrics and Gynecology, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan; ²Graduate Institute of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan; ³Department of Genome Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan; ⁴Department of Medical Research, Kaohsiung Medical University, Kaohsiung, Taiwan; ⁵Department of Family Medicine, Kaohsiung Medical University Hospital, Taiwan and ⁶Department of Public Health, Kaohsiung Medical University, Kaohsiung, Taiwan

Correspondence: Dr S-H Juo, Department of Genome Medicine, Kaohsiung Medical University, 100 TzYou First Road, Kaohsiung City 807, Taiwan.

E-mail: hjuo@kmu.edu.tw

Received 26 September 2012; revised 4 December 2012; accepted 4 January 2013; published online 31 January 2013

MATERIALS AND METHODS

Genetic association study

Subjects. The study was approved by the Kaohsiung Medical University Hospital (KMUH) Institutional Review Board. All participants gave the informed consent. For the endometriosis cases, we only included the patients who underwent laparotomy or laparoscopy at KMUH and had pathological confirmation of the disease. All the cases were recruited from the Department of Obstetrics and Gynecology at the KMUH in Taiwan.

We compared the cases with two different control groups: sources 1 and 2. Source 1 included the normal controls who had regular menstruation, did not have any evidence of endometriosis by sonography and had no history of gynecological tumors upon enrollment. The controls from source 2 were recruited from communities. The details of the ascertainment scheme were described elsewhere.¹¹ In brief, the subjects were the participants in community-based Pap smear screening projects.^{12,13} An eligible control in this study did not have any of the following conditions: infertility, dysmenorrhea, hypermenorrhea, irregular menstruation, surgical history for any obstetrical diseases, previously diagnosed endometriosis or adenomyosis. Owing to insufficient DNA samples, some of the controls used in our previous studies^{11,14} were not included in this study. All cases and controls are Chinese descents.

SNP selection and genotyping. One commonly studied functional polymorphism (-1306 C/T, rs243865)¹⁵ and 16 tSNPs selected from the HapMap Project (<http://www.hapmap.org/>) were genotyped in this study. All tSNPs have the minor allele frequency $\geq 10\%$ in the Han Chinese population based on the HapMap database (www.hapmap.org). Some tSNPs indicated by the HapMap Project were replaced by our previously genotyped SNPs¹⁶ that were in strong linkage disequilibrium ($r^2 \geq 0.8$) with the indicated tSNPs. Genotyping was carried out by using the TaqMan technology (Applied Biosystems, Foster City, CA, USA).

Statistical analysis. A goodness-of-fit χ^2 test was used to test for Hardy-Weinberg equilibrium (HWE) in the controls. We compared the mean of age, body mass index and parity between cases and controls. Because these three variables may also be associated with disease risk, they were included in the multivariate regression models. To account for multiple testing, we also performed family-wise permutation to obtain an empirical *P*-value. In this study, we used the minP permutation procedure to calculate the empirical *P*-values. This is a distribution-free statistic, which uses random label swapping to generate the null distribution in our population. The null distribution was obtained from the following steps: (1) swapping refractive errors randomly among all study subjects; (2) performing χ^2 test (or Fisher's test if a cell contains four or fewer samples) on the permuted data to obtain the minimal *P*-value of each run; and (3) repeating steps 1 and 2 to obtain 10 000 permutations. The empirical *P*-values for each genotype were calculated by the following formula: $(R + 1)/(N + 1)$. *R* is the number of runs in which the minimal *P*-value is less than the observed; *N* is the number of permutations.

Odds ratio (OR) was calculated to assess the genetic effect. All quantitative data are presented as mean \pm standard deviation (s.d.). Statistical analyses were performed with the SPSS statistical software (version 12.0, SPSS Inc., Chicago, IL, USA). The Hap-Clustering program¹⁷ was used for haplotype analyses with adjustment for the three covariates.

Functional study

Reporter assay to assess the differential allelic effect on gene expression. A genomic DNA segment of 1004 bp that includes rs7201 was amplified using polymerase chain reaction (PCR) with the forward primer of 5'-AAAA AGCTTGAAGCATCAAATCCGACTG-3' and reverse primer of 5'-AAAGGA TCCGTTACTCTCCCTGAGGTTCTC-3'. The DNA segment was subcloned into the downstream (*Hind*III/*Bam*HI sites) of the green fluorescent protein (GFP)-expressing vector, pEGFP-C1 (Clontech, Mountain View, CA, USA). Two reporter constructs were created. One construct carries the A allele and was designated as pEGFP-A, and the other carries C allele and was designated as pEGFP-C. The constructs were confirmed by DNA sequencing. Transient cotransfection was performed with 1 μ g of each construct and 1 μ g of pDsRed

(Promega, Madison, WI, USA) using Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA). The reporter assay was tested in the three endometrial cancer cell lines (RL95-2, HEC-1A and HEC1B) and one lung cancer cell line (H1299). pDsRed was used to calibrate the transfection efficiency. After 48 h, the GFP and DsRed signals were determined using a fluorometer (BioTek FLx800, Winooski, VT, USA). The GFP signal was detected as green (G) fluorescence and DsRed red (R) fluorescence, and we compared the ratio of G/R from each of the two reporter constructs. Each experiment was independently performed three times and each sample was studied in triplicates in each experiment. Statistical analysis was performed with the Mann-Whitney *U*-test with a two-side *P*-value $\leq 5\%$ as significance.

Immunohistochemical stain. We used the immunohistochemical (IHC) stain for pathological specimens from 24 cases with severe endometriosis (stages III and IV) according to the revised American Society for Reproductive Medicine (Re-ASRM). The technician who conducted the IHC experiment was blind to the clinical and genetic information. Each tissue specimen was routinely embedded in paraffin wax after 10% formalin fixation, and cut into 4- μ m-thick section for MMP2 immunostaining. In brief, sections were deparaffinized and rehydrated. Endogenous peroxidase in the section was soaked in absolute methanol containing 3% hydrogen peroxide. After washing with Tris buffer solution, sections were incubated with the anti-MMP2 primary antibody. Sections were subsequently incubated with biotin-conjugated goat anti-mouse immunoglobulin G and avidin-biotin-peroxidase complex. Finally, the sections were incubated with 0.05% 3,3'-diaminobenzidine tetrahydrochloride (DAKO Cytoimaion liquid DAB + Substrate chromogen System, DAKO Cyto- mation Inc., Carpinteria, CA, USA), followed by counterstained with Mayer's hematoxylin and mounted with Histomount (Invitrogen). Antigen retrieval was carried out. Negative controls were obtained by replacing the primary antibody with non-immune serum, and the positive control sample was breast carcinoma.

The MMP2 IHC stain has species specificity, which means it only reacts with human MMP2 but does not react with mouse and rat.

The MMP2 IHC stain was evaluated by the semiquantitative score system: 0 indicating that the sample was not immunoreactive in the endometriotic cells; 1, 2, 3 and 4 indicating 1–25%, 26–50%, 51–75% and >75% immunoreactivity in the endometriotic cells, respectively. A positive stain was defined as a score ≥ 3 .

Measure endogenous miRNA in the cell lines. Three cell lines (H1299, RL95-2 and HEC-1A) were used to investigate miR-520g effect. First of all, we used the quantitative real-time PCR to measure the endogenous miR-520g levels. The microRNA Isolation Kit (mirVana; Applied Biosystems) was used to extract total RNA that also contains small RNAs. The TaqMan MicroRNA Reverse Transcription Kit (Applied Biosystems) was used to convert microRNA to cDNA. Quantitative real-time PCR was performed on the ABI PRISM 7500 sequence detection system (Applied Biosystems). The typical amplification program was carried out in a 20 μ l reaction and real-time PCR results were analyzed using the ABI PRISM 7500 SDS software (Applied Biosystems). Each experiment was independently performed three times and each sample was studied in triplicates in each experiment. The Mann-Whitney *U*-test was used for statistical analysis with a two-side *P*-value $\leq 5\%$ as significance.

pEGFP-rs7201 reporter assay for differential knockdown effect by miR-520g. In all, 50 and 100 nM of miR-520g mimic (pre-miR miRNA precursor molecule; Ambion-Life Technologies, Carlsbad, CA, USA) were transfected along with 1 μ g of reporter construct and 1 μ g of pDsRed to the HEC-1A cells using the Lipofectamine 2000 reagent (Invitrogen). Similarly, the HEC-1A cells were transiently transfected with 50 and 100 nM of antagomiR-520g (Anti-miR miRNA Inhibitor; Ambion) and a negative control RNA (Ambion), respectively. miR-520g mimic is a synthetic RNA duplex and it acts as a mimic of naturally occurring miR-520g upon delivering into cells. AntagomiR-520g is a chemically engineered oligonucleotide to silence endogenous miR-520g. After 48 h, the GFP and DsRed signals were detected to determine the effects of miR-520g mimic and antagomiR-520g on the two reporter constructs. The results were analyzed using the Mann-Whitney *U*-test.

MMP2 gene is silenced by miR-520g. The effects of miR-520g mimic on *MMP2* mRNA and protein expression were also evaluated in the primary endometrial stromal cells. The primary stromal cells were obtained from an endometriosis patient who carries the AA genotype at SNP rs7201. The stromal cells were grown in the Dulbecco's modified Eagle's medium/F12 medium supplemented with 10% fetal bovine serum. Cells between passages 3 and 6 were applied for the experiments. By using 5 μ l siPort NeoF_x (Ambion) reagent, concentrations of 30, 60 and 120 nM of miR-520g mimic were transfected into the stromal cells for 48 h. Quantification of *MMP2* mRNA was conducted using High Capacity cDNA Reverse Transcription kit (Applied Biosystems) with random primers, 0.5 μ g total RNA and the MultiScribe Reverse Transcriptase according to the protocol. Gene-specific PCR primers and TaqMan probes for human *MMP2* (ABI ID number: Hs01548731_m1) were from the Applied Biosystems. *MMP2* mRNA expression was normalized to glyceraldehyde-3-phosphate dehydrogenase as an internal control. The protein extracts from the cultured stromal cells were determined by the western blot analysis. Equal amounts of protein were subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis. A standard western blot analysis was conducted using *MMP2* propeptide antibody (1:500 dilution; Santa Cruz Biotechnology, Santa Cruz, CA, USA). β -Actin antibody (1:5000 dilution; Cell Signaling, Beverly, MA, USA) was used as a loading control.

Function prediction

Microcosm Targets version 5 database (<http://www.ebi.ac.uk/enright-srv/microcosm>, accessed in March 2010) was used to predict target genes of miR-520g. GeneOntology enrichment analysis was carried out using the MetaCore Pathways Analysis (GeneGo, Encinitas, CA, USA) by uploading the data set of predicted target genes to the MetaCore server. Target genes that were not expressed in the uterus were excluded from the MetaCore analysis. Enrichment analysis identified the most significant biological processes based on the uploaded target genes.

RESULTS

Demographic data

This study included a total of 211 women with endometriosis, and 126 hospital-based and 218 community-based controls. According to Re-ASRM, 179 patients with endometriosis were in stages III and IV and 32 in stages I and II.

SNP and endometriosis

A total of 17 SNPs were selected for genotyping. Two SNPs are in exons (rs1053605 synonymous SNP, and rs243849 synonymous SNP), one SNP is in the 3'-untranslated region (UTR) (rs7201) and the rest SNPs are in the non-coding regions. All SNPs were in HWE except SNP rs9302671 (HW P -value = 0.00035), and this SNP was not included for further genetic analyses. Based on χ^2 or Fisher's exact tests, only two SNPs had P -values < 0.05: rs243832 (P = 0.045 for the overall genetic effect and P = 0.017 for the comparison between CC and GG genotypes) and rs7201 (P = 0.014 for the overall genetic effect and the OR for AC was 1.88 with the P = 0.004 for the comparison between AC and AA genotypes) (Table 1). The family-wise permutation was conducted to adjust for multiple testing and the empirical P -value for the AC genotype of SNP rs7201 was 0.047. Rs243832 and rs7201 SNPs are located in the same block with the D' = 0.96 and r^2 = 0.47 (Figure 1). When both SNPs were included in the regression model, only SNP rs7201 remained significant. We also used these two SNPs (rs243832 and rs7201) for the haplotype analysis and the result yielded a global P -value of 0.045.

As *MMP2* has been reported to be more related to severe endometriosis,¹⁰ we also analyzed 179 severe cases who were in Re-ASRM stage III or IV. The results for SNP rs243832 and rs7201 were less significant than those from all the cases (data not shown).

Therefore, the association between *MMP2* SNPs and endometriosis is not only restricted to the severe cases.

SNP rs7201 influences gene expression

Our statistical results suggest that SNP rs7201 was significantly associated with endometriosis. Because this SNP is located at 3'-UTR, we first conducted a reporter gene assay to test whether this SNP can influence gene expression levels. If so, we would further test whether this SNP could influence a microRNA binding. First, we compared GFP expression between two reporter constructs (pEGFP-A and pEGFP-C). GFP expression was found to be significantly higher for the pEGFP-C construct (G/R value was 2.30 for pEGFP-C and 1.62 for pEGFP-A, P = 0.027 by Mann-Whitney U -test) transfected to the H1299 cell line. That is, the C allele yields a higher protein production than the A allele by 1.42-fold. Similarly, higher GFP expression for the pEGFP-C than pEGFP-A construct was shown in other three endometrial cancer cell lines: G/R ratio in the RL95-2 cell line was 1.99 for pEGFP-C and 1.78 for pEGFP-A; in the HEC-1A cell line was 1.89 for pEGFP-C and 1.73 for pEGFP-A; and in the HEC1B cell line was 1.90 for pEGFP-C and 1.33 for pEGFP-A.

MMP2 expression in different genotypes

We measured *MMP2* expression in the endometrial samples from 24 patients with severe endometriosis. The number of subjects in the AA, AC and CC genotypes of SNP rs7201 was 16, 6 and 2 CC, respectively (Table 2). The P -value from the Mann-Whitney U -test was 0.027 for comparison between positive and negative IHC stains among three genotypes (Table 2). The CC subjects had higher mean IHC score and positive IHC stain, followed by the AC subjects and then AA subjects.

Endogenous miRNA-520g in the cell lines

Given that SNP rs7201 could influence gene expression, we searched for putative microRNAs that were predicted to bind to this locus. Using the bioinformatics database (<http://mirmap.mbc.nctu.edu.tw/> and <http://www.microrna.org/microrna/home.do>), both miR-520g and miR-520h were predicted to bind the 3'-UTR including the location of SNP rs7201 (Figure 2). In addition, both miR-520g and miR-520h are highly expressed in the placenta (<http://mirmap.mbc.nctu.edu.tw/>). As miR-520g and miR-520h have a similar structure, we only investigated the miR-520g effect on the *MMP2* gene. We measured the expression level of endogenous miR-520g in three cell lines (H1299, RL95-2 and HEC-1A) and primary stromal cells obtained from a single endometriosis patient. The endogenous miR-520g level was much higher (P < 0.0001) in the stromal cells than three cancer cells (Supplementary Figure). Among the three cancer cell lines, the endogenous miR-520g levels were also different (P = 0.007 for the comparison between RL95-2 and HEC-1A, P = 0.04 between RL95-2 and H1299 and P = 0.2 between HEC-1A and H1299).

SNP rs7201 influences gene expression via miRNA-520g

We further verified the effect of miR-520g on GFP expression by the overexpression and knockdown strategies. The experiments were conducted in three groups for each reporter construct: Group I—cotransfection of reporter constructs with negative control oligo (Invitrogen; Cat. No. AM17110), Group II—cotransfection of reporter constructs with miR-520g mimic and Group III—cotransfection of reporter constructs with the antagomiR-520g. After transfecting miR-520g mimic to the HEC-1A cells for 24 h, we detected a striking increase (P = 0.0003) of intracellular miR-520g levels by more than 500-fold (Figure 3a). On the contrary, transfecting antagomiR-520g

Table 1 *MMP2* gene polymorphisms and endometriosis

	Case, n (%)	Hospital controls				Community controls				All controls		
		n (%)	OR	95% CI	P-value	n (%)	OR	95% CI	P-value	OR	95% CI	P-value
<i>rs243865</i>												
GG	156 (74.6)	99 (78.6)	1			163 (76.5)	1			1		
AG	49 (23.4)	26 (20.6)	1.20	0.70–2.05	0.514	46 (21.6)	1.11	0.70–1.76	0.647	1.14	0.76–1.73	0.526
AA	4 (1.9)	1 (0.8)	2.54	0.28–23.04	0.652 ^a	4 (1.9)	1.04	0.26–4.25	1.000	1.34	0.36–5.08	0.733 ^a
<i>rs857403</i>												
AA	101 (48.1)	43 (34.1)	1			98 (46.9)	1			0.78	0.54–1.11	0.170
AT	90 (42.9)	70 (55.6)	0.55	0.34–0.88	0.012	92 (44.0)	0.95	0.63–1.42	0.799	0.83	0.44–1.54	0.554
TT	19 (9.0)	13 (10.3)	0.62	0.28–1.37	0.237	19 (9.1)	0.97	0.48–1.94	0.932			
<i>rs1030868</i>												
GG	121 (57.9)	74 (58.7)	1			132 (60.8)	1			1.08	0.75–1.56	0.673
AG	75 (35.9)	47 (37.3)	0.98	0.61–1.55	0.918	71 (32.7)	1.15	0.77–1.73	0.495	1.16	0.56–2.44	0.686
AA	13 (6.2)	5 (4.0)	1.59	0.54–4.64	0.393	14 (6.5)	1.01	0.46–2.24	0.975			
<i>rs1477017</i>												
AA	124 (59.3)	74 (58.7)	1			130 (61.9)	1			1.05	0.73–1.53	0.788
AG	71 (34.0)	46 (36.5)	0.92	0.58–1.47	0.732	65 (31.0)	1.15	0.75–1.74	0.524	1.10	0.54–2.24	0.799
GG	14 (6.7)	6 (4.8)	1.39	0.51–3.78	0.514	15 (7.1)	0.98	0.45–2.11	0.956			
<i>rs1053605</i>												
CC	107 (69.5)	94 (74.6)	1			157 (73.4)	1			1.26	0.82–1.94	0.294
CT	44 (28.6)	30 (23.8)	1.29	0.75–2.21	0.358	52 (24.3)	1.24	0.78–1.99	0.367	1.01	0.26–3.96	1.000 ^a
TT	3 (1.9)	2 (1.6)	1.32	0.22–8.06	1.000 ^a	5 (2.3)	0.88	0.21–3.76	1.000			
<i>rs2241145</i>												
GG	59 (28.4)	48 (38.1)	1			70 (32.4)	1			1.45	0.98–2.16	0.065
CG	111 (53.4)	52 (41.3)	1.74	1.05–2.87	0.031	101 (46.8)	1.30	0.84–2.02	0.236	1.07	0.65–1.77	0.791
CC	38 (18.3)	26 (20.6)	1.19	0.63–2.23	0.589	45 (20.8)	1.00	0.58–1.74	0.995			
<i>rs2241146</i>												
GG	142 (67.9)	88 (69.8)	1			157 (72.7)	1			1.23	0.84–1.81	0.293
AG	62 (29.7)	34 (27.0)	1.13	0.69–1.86	0.629	53 (24.5)	1.29	0.84–1.99	0.242	0.86	0.29–2.57	0.791
AA	5 (2.4)	4 (3.2)	0.77	0.20–2.96	0.736 ^a	6 (2.8)	0.92	0.28–3.08	1.000			
<i>rs243849</i>												
CC	129 (61.4)	64 (50.8)	1			129 (60.0)	1			0.82	0.57–1.18	0.278
CT	71 (33.8)	53 (42.1)	0.66	0.42–1.06	0.085	77 (35.8)	0.92	0.62–1.38	0.694	0.83	0.37–1.86	0.652
TT	10 (4.8)	9 (7.1)	0.55	0.21–1.42	0.214	9 (4.2)	1.11	0.44–2.82	0.825			
<i>rs12599775</i>												
GG	172 (81.5)	107 (84.9)	1			189 (88.7)	1			1.55	0.96–2.52	0.072
CG	37 (17.5)	19 (15.1)	1.21	0.66–2.22	0.533	22 (10.3)	1.85	1.05–3.26	0.032	1.72	0.24–12.33	0.628 ^a
CC	2 (0.9)	0 (0.0)	NA	NA	NA	2 (0.9)	1.10	0.15–7.89	0.925			
<i>rs243847</i>												
TT	71 (34.1)	51 (40.5)	1			68 (32.2)	1			1.04	0.71–1.53	0.835
CT	105 (50.5)	58 (46.0)	1.30	0.80–2.11	0.285	111 (52.6)	0.91	0.59–1.39	0.650	1.09	0.64–1.87	0.740
CC	32 (15.4)	17 (13.5)	1.35	0.68–2.69	0.390	32 (15.2)	0.96	0.53–1.73	0.886			
<i>rs243844</i>												
GG	89 (47.8)	66 (53.2)	1			95 (49.0)	1			1.18	0.81–1.72	0.378
AG	87 (46.8)	50 (40.3)	1.29	0.81–2.07	0.289	83 (42.8)	1.12	0.74–1.70	0.598	0.75	0.34–1.65	0.477
AA	10 (5.4)	8 (6.5)	0.93	0.35–2.48	0.880	16 (8.2)	0.67	0.29–1.55	0.343			
<i>rs243840</i>												
AA	69 (36.9)	32 (25.6)	1			76 (37.8)	1			0.88	0.59–1.30	0.515
AG	89 (47.6)	64 (51.2)	0.64	0.38–1.09	0.106	95 (47.3)	1.03	0.67–1.60	0.888	0.77	0.45–1.32	0.338
GG	29 (15.5)	29 (23.2)	0.46	0.24–0.90	0.022	30 (14.9)	1.06	0.58–1.95	0.839			
<i>rs2287076</i>												
TT	96 (48.0)	65 (51.6)	1			102 (48.8)	1			1.01	0.69–1.46	0.973
CT	81 (40.5)	50 (39.7)	1.10	0.68–1.76	0.701	90 (43.1)	0.96	0.63–1.44	0.830	1.43	0.78–2.62	0.247
CC	23 (11.5)	11 (8.7)	1.42	0.65–3.10	0.384	17 (8.1)	1.44	0.72–2.85	0.298			
<i>rs11639960</i>												
AA	122 (57.8)	67 (53.2)	1			141 (65.0)	1			1.17	0.81–1.67	0.406
AG	80 (37.9)	54 (42.9)	0.81	0.52–1.28	0.375	63 (29.0)	1.47	0.97–2.21	0.066	0.85	0.37–1.96	0.706
GG	9 (4.3)	5 (4.0)	0.99	0.32–3.07	0.984	13 (6.0)	0.80	0.33–1.94	0.620			
<i>rs243832</i>												
GG	78 (37.7)	63 (50.8)	1			95 (44.2)	1			1.33	0.91–1.93	0.136
CG	97 (46.9)	51 (41.1)	1.54	0.96–2.47	0.076	97 (45.1)	1.22	0.81–1.84	0.347	1.96	1.13–3.43	0.017
CC	32 (15.5)	10 (8.1)	2.58	1.18–5.66	0.015	23 (10.7)	1.69	0.92–3.13	0.090			
<i>rs7201</i>												
AA	115 (55.3)	82 (65.0)	1			147 (68.4)	1			1.72	1.19–2.49	0.004
AC	84 (40.4)	40 (31.8)	1.50	0.93–2.40	0.092	57 (26.5)	1.88	1.24–2.85	0.003	1.19	0.51–2.81	0.683
CC	9 (4.3)	4 (3.2)	1.60	0.48–5.39	0.566 ^a	11 (5.1)	1.05	0.42–2.61	0.923			

Abbreviations: CI, confidence interval; OR, odds ratio; MMP, matrix metalloproteinase; NA, not applicable.
^aP-value were adjusted by Fisher's exact test.

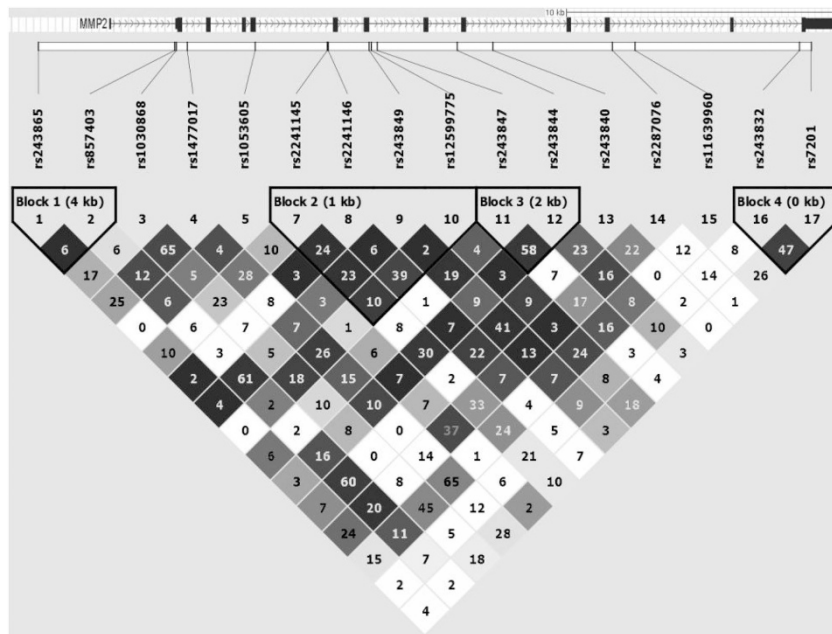


Figure 1 In all, 17 *MMP2* SNPs form four haplotype blocks. The r^2 (%) between any two SNPs is displayed in each cell. The cells in dark gray indicate strong linkage disequilibrium and log of odds (LOD) of 2 or greater and in white indicate uninformative with LOD lower than 2. The bar on the top is the *MMP2* gene structure and location of each SNP; kb indicates kilobase. SNPs rs1053605 and rs243849 are synonymous SNPs, SNP rs7201 is in the 3'-UTR and the res SNPs are not in the coding regions.

led to a significant decrease ($P=0.0003$) of endogenous miR-520g (Figure 3a). Similarly, miR-520g mimic and antagomiR-520g also caused comparable changes in the RL95-2 cells (data not shown). The above experiments indicated that mimic and antagomiR could significantly alter intracellular miR-520 levels.

The G/R ratio for the pEGFP-A construct was significantly suppressed by miR-520g mimic ($P<0.005$ for 50 vs 0 nM, $P<0.005$ for 100 vs 0 nM; Figure 3b). Using trend t -test, we found a significant dose-dependent manner ($P<0.0001$) for miR-520g silencing effect. Similarly, there was a significant increase ($P<0.005$ for 50 vs 0 nM, $P<0.005$ for 100 vs 0 nM; Figure 3b) of G/R ratio by antagomiR-520g for the pEGFP-A construct, and the increase also had a dose-dependent manner ($P<0.0001$). On the contrary, the G/R ratio was not or barely influenced by miR-520g mimic ($P=0.4$ for 50 vs 0 nM, $P=0.4$ for 100 vs 0 nM; Figure 3b) or antagomiR-520g ($P=0.3$ for 50 vs 0 nM, $P=0.04$ for 100 vs 0 nM; Figure 3b) for the pEGFP-C construct. Our results indicate that the risk C allele is less responsive to the miR-520s silencing effect because it causes imperfect complement to miR-520 seed region.

miR-520g directly represses *MMP2* expression

To further confirm that the miR-520g can downregulate the *MMP2* gene, we directly examined the influence of miR-520g on *MMP2* expression in the primary endometrial stromal cells of the AA genotype at SNP rs7201. The result showed that both *MMP2* mRNA and propeptide expressions were reduced in parallel with the concentrations of miR-520g mimic (Figure 4). Transfecting miR-520g mimic led to a decrease of *MMP2* mRNA levels (compared with the baseline status, $P<0.05$ for 30 nM and all P -values <0.01 for 60, 90 and 120 nM; Figure 4a). The reduction of *MMP2* mRNA expression also had a significant trend ($P<0.0001$) with the dose of miR-520g mimic. For *MMP2* propeptide, miR-520 mimic reduced propeptide synthesis (compared with the baseline, $P<0.05$ for 30 nM and all

Table 2 Comparison of *MMP2* expression measured by the IHC stain in three genotypes at SNP rs7201

	Genotype		
	AA (n = 16)	AC (n = 6)	CC (n = 2)
<i>IHC stain</i>			
Mean score	2.375	2.833	3.5
Number of positive IHC	6	5	2
Number of negative IHC	10	1	0
Ratio of positive IHC	37.5% (6/16)	83% (5/6)	100% (2/2)

Abbreviations: IHC, immunohistochemical; MMP, matrix metalloproteinase; SNP, single-nucleotide polymorphism.

The IHC score: 1 indicates not immunoreactive and 2, 3 and 4 indicating weak, moderate and intense immunoreactivity, respectively. A positive stain was defined as a score ≥ 3 . The Mann-Whitney U -test showed the P -value was 0.027 for the distributions between positive and negative IHC stain among three genotypes.

P -values <0.01 for 60 or 120 nM; Figure 4b), which also showed a significant dose-dependent manner ($P<0.0001$). When 120 nM of miR-520g mimic was transfected, the propeptide production could be reduced by 45%.

DISCUSSION

In this study, we systematically investigated the relationship between 17 SNPs of the *MMP2* gene and endometriosis. The previously known functional promoter polymorphism (rs243865) is not associated with endometriosis but the 3'-UTR SNP rs7201 is significant for endometriosis. Our IHC experiment also suggests differential expression of *MMP2* in three different genotypes of SNP rs7201, and the risk C allele had a higher expression level. Both *in silico* and experimental studies indicate that rs7201 is located in the miR-520g binding site. We demonstrate that miR-520g mimic can downregulate the *MMP2* expression. We also show that the risk C allele of SNP rs7201 acts as a

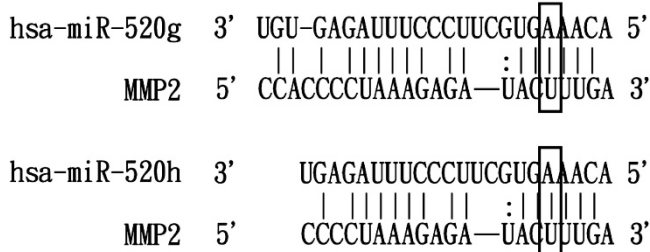


Figure 2 Bioinformatic analysis predicted that miR-520g and miR-520h can bind to the 3'-UTR of MMP2 transcript. The site of SNP rs7201 was indicated.

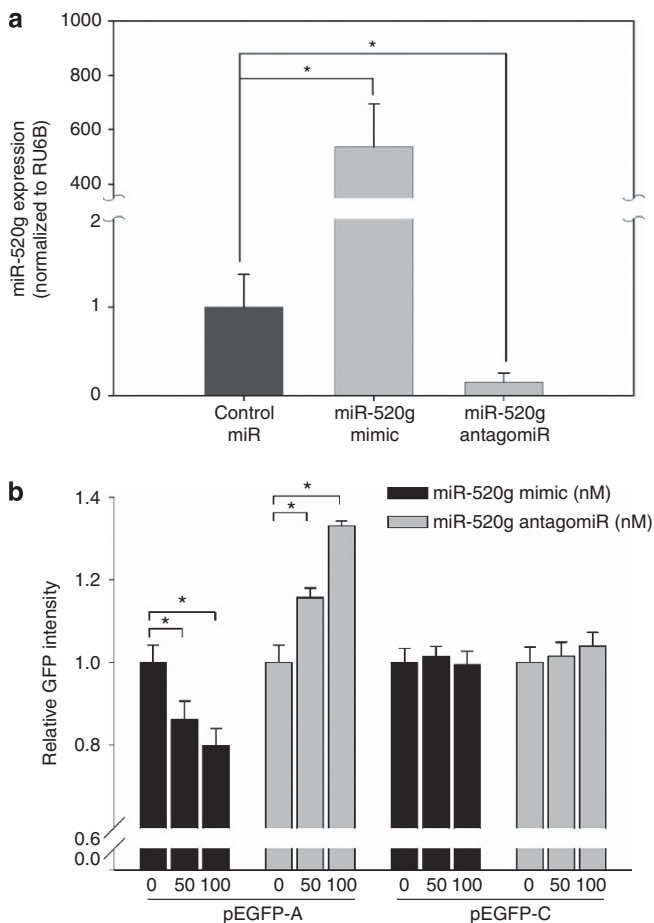


Figure 3 (a) Transfecting miR-520g mimic and antagomiR-520g to the HEC-1A cells significantly changed the intracellular miR-520g levels ($*P$ -value=0.0003). (b) The G/R ratio for the pEGFP-A construct was significantly reduced by miR-520g mimic in the HEC-1A cells. AntagomiR-520g increased the G/R ratio for the pEGFP-A construct. The G/R ratio for pEGFP-C was not influenced by miR-520g mimic. AntagomiR-520g barely increased the G/R ratio ($P=0.3$ for 50 vs 0 nM, $P=0.04$ for 100 vs 0 nM) for the pEGFP-C construct. $*P<0.005$. Each experiment was independently performed three times and each sample was studied in triplicates in each experiment.

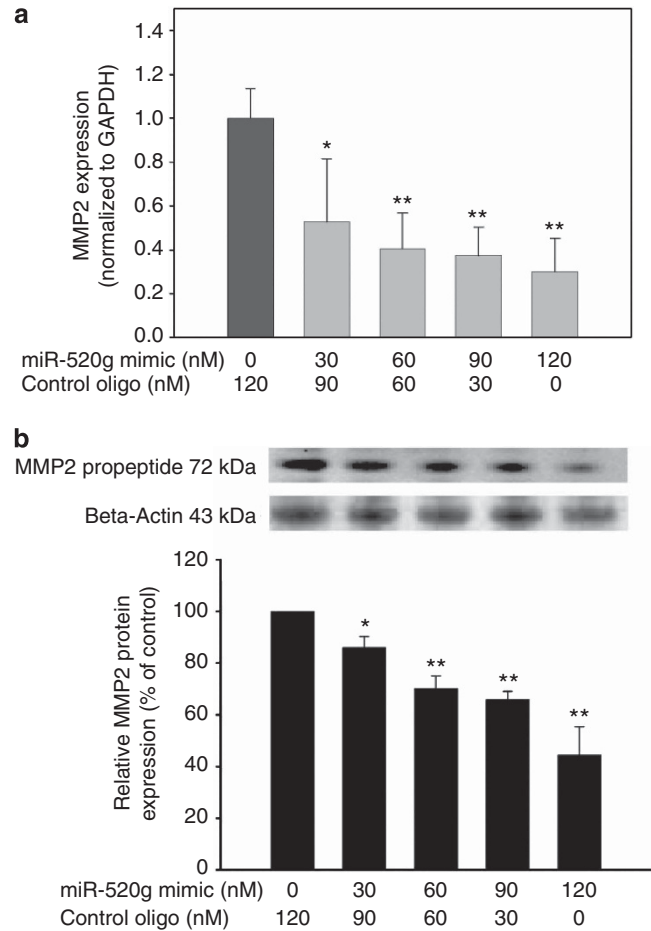


Figure 4 (a) miR-520g mimic significantly downregulated MMP2 mRNA expression in the endometrial stromal cells, and the miR-520 effect also showed a dose-dependent manner. $*P<0.05$; $**P<0.01$ (b) miR-520g suppressed MMP2 propeptide expression in the endometrial stromal cells and the miR-520 effect on propeptide also had a dose-dependent manner ($P<0.0001$). $*P<0.05$; $**P<0.01$. Each experiment was independently performed three times and each sample was studied in triplicates in each experiment.

loss-of-function mutation, because neither miR-520g mimic nor antagomiR-520g can influence the MMP2 expression in the presence of this allele. This loss-of-function mutation results in dysregulation of MMP2 expression and a higher risk for endometriosis. As dysregulation of MMP2 expression has been implicated as a pathogenesis of endometriosis,^{6,8} our study provides an additional insight into the mechanism of endometriosis.

MMP2 is a highly expressed protein in endometrium, especially in ectopic endometrium.¹⁸ Women with endometriosis also have a higher level of MMP2 in the serum and peritoneal fluid than women without this disease.⁸ These observations are compatible with our finding in which the risk C allele causes a higher expression level and causes out of regulation by miR-520g. The *MMP2* gene polymorphisms have been implicated in different diseases but most of previous studies only focused on functional promoter SNPs.^{19–21} Only a limited number of studies investigated endometriosis in relation to MMP2 polymorphisms.^{21–23} Analyzing the two functional promoter SNPs (rs2285053 also known as -735 C/T and rs243865 also known as -1306 C/T) in the Chinese patients with endometriosis, both Kang *et al.*²² and Zhao *et al.*²³ did not find any

significant association. Although Saare *et al.*²¹ reported that -735 C/T SNP was significant ($P=0.02$) for endometriosis, the significant result was primarily from 92 cases with Re-ASRM stages I and II but not from patients with Re-ASRM stages III and IV. On the other hand, we use the tagging SNP approach to investigate comprehensively the overall gene effect of MMP2. Although we do not disclose a relationship between the promoter polymorphism -1306 C/T and endometriosis, our study identifies another functional SNP rs7201 at the 3'-UTR conferring risk for endometriosis. Recently, the C allele of SNP rs7201 was also found to be an independent risk factor for stroke.²⁴

SNP rs7201 is located in the complementary DNA region to the seed region of miR-520g. The MMP2 3'-UTR spans 1249bp and only two validated SNPs are in the vicinity of SNP rs7201. The distance between rs7201 and rs17860019 is 101 base pairs (bp), and between rs7210 and rs17860021 is 140 bp. However, these two neighboring SNPs are too rare (rs17860019 is monomorphic in Asians; and the minor allele frequency of rs17860021 is about 1%) based on the NCBI dbSNP data set (131 version).

Identification the role of microRNAs in endometriosis opens a new avenue for biomarkers and therapeutic tools. Using global microRNA expression profiles in relation to endometriosis, recent studies did not find miR-520g or miR-520h to be differentially expressed between ectopic and eutopic endometrial tissues.^{25–28} Notably, the expression of miR-520g and miR-520h is highly tissue specific. Both microRNAs are barely detectable in other tissues except for placenta (<http://mirnamap.mbc.nctu.edu.tw/>). Unfortunately, there are no available expression data for miR-520g or miR-520h in the uterus. The functional role of miR-520g is hardly mentioned in the literatures. From GeneOntology enrichment analysis, miR-520g target genes are enriched in cGMP-mediated signaling pathway, vasodilation and neural transmitters. The cGMP-mediated signaling pathway was reported to be associated with neointima thickness.²⁹ Accordingly, miR-520g could have an important role in blood vessel function and/or development.

A SNP located at or near a microRNA binding site has been named as mirSNP, which represents a novel class of functional SNPs. Mishra *et al.*³⁰ reported that a mirSNP located 14bp downstream of the miR-24 binding site of the *dihydrofolate reductase* gene can affect dihydrofolate reductase expression. The mutant allele causes over-expression of dihydrofolate reductase protein, which is similar to this study. As dihydrofolate reductase is a target of methotrexate, patients with mutant allele at this mirSNP are resistant to methotrexate treatment. The application of mirSNP may lead to a different type of personalized medicine called microRNA pharmacogenomics.³¹ Recently, mirSNPs located in the *WHSC1*, *SLC22A23*³² and *KRAS*³³ genes have been reported to influence the risk for endometriosis. As MMP2 has been implicated in several common diseases including cancer and cardiovascular disease,^{19–21,34} using RNA interference for the subjects carrying the risk C allele of SNP rs7201 could be one of microRNA pharmacogenomic applications in the future.

There are several limitations in this study. Although adjusted (that is, empirical) P -value for the rs7201 AC genotype is 0.047 by the family-wise permutation test, readers still need to take the statistical results with caution because our sample size is still too moderate to avoid a type I error. However, functional studies in conjunction with SNP association studies can reduce such errors. The biosamples of the rare CC genotype at SNP rs7201 is hard to obtain. We only had two CC subjects for IHC stain and we do not have any stromal cells of the CC genotype. Therefore, the data obtained from subjects of CC genotype need to be taken with caution. The miRNA-520 levels

shown in the Supplementary Figure were from a single patient of endometriosis. It is not clear whether all the endometriosis patients have such a high miR-520 expression level. Further studies to validate our finding are warranted.

In conclusion, SNP rs7201 at the 3'-UTR is at the miRNA-520g binding site. The risk C allele can reduce miRNA-520s silencing effect leading to increase MMP2 expression and endometriosis risk. From a microRNA pharmacogenomic point of view, a subject carrying the risk C allele may be more beneficial for the treatment by an MMP inhibitor.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported by the National Science Council, Taiwan (Grant numbers NSC 97-2314-B-037-010-MY3, 99-2628-B-037-009-MY3); Kaohsiung Medical University Research Foundation (KMUR-004) and the Kaohsiung Medical University Hospital Research Fund (Grant Numbers KMH98-8111, KMH99-9104 and KMH99-9R30).

- Hadfield, R. M., Mardon, H. J., Barlow, D. H. & Kennedy, S. H. Endometriosis in monozygotic twins. *Fertil. Steril.* **68**, 941–942 (1997).
- Simpson, J. L. & Bischoff, F. Z. Heritability and molecular genetic studies of endometriosis. *Ann. NY Acad. Sci.* **955**, 239–251 (2002).
- Tempfer, C. B., Simoni, M., Destenaves, B. & Fauser, B. C. Functional genetic polymorphisms and female reproductive disorders: part II—endometriosis. *Hum. Reprod. Update* **15**, 97–118 (2009).
- Giudice, L. C. & Kao, L. C. Endometriosis. *Lancet* **364**, 1789–1799 (2004).
- Curry, Jr T. E. & Osteen, K. G. The matrix metalloproteinase system: changes, regulation, and impact throughout the ovarian and uterine reproductive cycle. *Endocr. Rev.* **24**, 428–465 (2003).
- Chung, H. W., Lee, J. Y., Moon, H. S., Hur, S. E., Park, M. H., Wen, Y. *et al.* Matrix metalloproteinase-2, membranous type 1 matrix metalloproteinase, and tissue inhibitor of metalloproteinase-2 expression in ectopic and eutopic endometrium. *Fertil. Steril.* **78**, 787–795 (2002).
- Rawdanowicz, T. J., Hampton, A. L., Nagase, H., Woolley, D. E. & Salamonson, L. A. Matrix metalloproteinase production by cultured human endometrial stromal cells: identification of interstitial collagenase, gelatinase-A, gelatinase-B, and stromelysin-1 and their differential regulation by interleukin-1 alpha and tumor necrosis factor-alpha. *J. Clin. Endocrinol. Metab.* **79**, 530–536 (1994).
- Huang, H. F., Hong, L. H., Tan, Y. & Sheng, J. Z. Matrix metalloproteinase 2 is associated with changes in steroid hormones in the sera and peritoneal fluid of patients with endometriosis. *Fertil. Steril.* **81**, 1235–1239 (2004).
- Inagaki, N., Ung, L., Otani, T., Wilkinson, D. & Lopata, A. Uterine cavity matrix metalloproteinases and cytokines in patients with leiomyoma, adenomyosis or endometrial polyp. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **111**, 197–203 (2003).
- Uzan, C., Cortez, A., Dufournet, C., Fauvet, R., Siffroi, J. P. & Darai, E. Eutopic endometrium and peritoneal, ovarian and bowel endometriotic tissues express a different profile of matrix metalloproteinases-2, -3 and -11, and of tissue inhibitor metalloproteinases-1 and -2. *Virchows Arch.* **445**, 603–609 (2004).
- Juo, S. H., Wang, T. N., Lee, J. N., Wu, M. T., Long, C. Y. & Tsai, E. M. CYP17, CYP11A1 and COMT polymorphisms and the risk of adenomyosis and endometriosis in Taiwanese women. *Hum. Reprod.* **21**, 1498–1502 (2006).
- Tsai, H. T., Wu, C. H., Lai, H. L., Li, R. N., Tung, Y. C., Chuang, H. Y. *et al.* Association between quantitative high-risk human papillomavirus DNA load and cervical intraepithelial neoplasia risk. *Cancer Epidemiol. Biomarkers Prev.* **14**, 1–6 (2005).
- Wu, M. T., Lee, L. H., Ho, C. K., Liu, C. L., Wu, T. N., Wu, S. C. *et al.* Lifetime exposure to environmental tobacco smoke and cervical intraepithelial neoplasms among nonsmoking Taiwanese women. *Arch. Environ. Health* **58**, 353–359 (2003).
- Juo, S. H., Wu, R., Lin, C. S., Wu, M. T., Lee, J. N. & Tsai, E. M. A functional promoter polymorphism in interleukin-10 gene influences susceptibility to endometriosis. *Fertil. Steril.* **92**, 1228–1233 (2009).
- Price, S. J., Greaves, D. R. & Watkins, H. Identification of novel, functional genetic variants in the human matrix metalloproteinase-2 gene: role of Sp1 in allele-specific transcriptional regulation. *J. Biol. Chem.* **276**, 7549–7558 (2001).
- Wang, L. F., Chien, C. Y., Kuo, W. R., Tai, C. F. & Juo, S. H. Matrix metalloproteinase-2 gene polymorphisms in nasal polyps. *Arch. Otolaryngol. Head Neck Surg.* **134**, 852–856 (2008).
- Tzeng, J. Y., Wang, C. H., Kao, J. T. & Hsiao, C. K. Regression-based association analysis with clustered haplotypes through use of genotypes. *Am. J. Hum. Genet.* **78**, 231–242 (2006).

- 18 Ueda, M., Yamashita, Y., Takehara, M., Terai, Y., Kumagai, K., Ueki, K. *et al.* Gene expression of adhesion molecules and matrix metalloproteinases in endometriosis. *Gynecol. Endocrinol.* **16**, 391–402 (2002).
- 19 Lin, S. C., Lo, S. S., Liu, C. J., Chung, M. Y., Huang, J. W. & Chang, K. W. Functional genotype in matrix metalloproteinases-2 promoter is a risk factor for oral carcinogenesis. *J. Oral Pathol. Med.* **33**, 405–409 (2004).
- 20 Ye, S. Influence of matrix metalloproteinase genotype on cardiovascular disease susceptibility and outcome. *Cardiovasc. Res.* **69**, 636–645 (2006).
- 21 Saare, M., Lamp, M., Kaart, T., Karro, H., Kadastik, U., Metspalu, A. *et al.* Polymorphisms in MMP-2 and MMP-9 promoter regions are associated with endometriosis. *Fertil. Steril.* **94**, 1560–1563 (2010).
- 22 Kang, S., Zhao, X. W., Wang, N., Chen, S. C., Zhou, R. M. & Li, Y. Association of polymorphisms of the MMP-2 and TIMP-2 genes with the risk of endometriosis in North Chinese women. *Fertil. Steril.* **90**, 2023–2029 (2008).
- 23 Zhao, X. W., Li, Y., Wang, N., Zhao, J., Li, X. L., Liu, Q. *et al.* Study on the association of SNPs of MMP-2 and TIMP-2 genes with the risk of endometriosis and adenomyosis. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* **25**, 280–283 (2008).
- 24 Fatar, M., Stroick, M., Steffens, M., Senn, E., Reuter, B., Bukow, S. *et al.* Single-nucleotide polymorphisms of MMP-2 gene in stroke subtypes. *Cerebrovasc. Dis.* **26**, 113–119 (2008).
- 25 Pan, Q., Luo, X., Toloubeydokhti, T. & Chegini, N. The expression profile of micro-RNA in endometrium and endometriosis and the influence of ovarian steroids on their expression. *Mol. Hum. Reprod.* **13**, 797–806 (2007).
- 26 Ohlsson Teague, E. M., Van der Hoek, K. H., Van der Hoek, M. B., Perry, N., Wagaraachchi, P., Robertson, S. A. *et al.* MicroRNA-regulated pathways associated with endometriosis. *Mol. Endocrinol.* **23**, 265–275 (2009).
- 27 Filigheddu, N., Gregnanin, I., Porporato, P. E., Surico, D., Perego, B., Galli, L. *et al.* Differential expression of microRNAs between eutopic and ectopic endometrium in ovarian endometriosis. *J. Biomed. Biotechnol.* **2010**, 369549 (2010).
- 28 Hawkins, S. M., Creighton, C. J., Han, D. Y., Zariff, A., Anderson, M. L., Gunaratne, P. H. *et al.* Functional microRNA involved in endometriosis. *Mol. Endocrinol.* **25**, 821–832 (2011).
- 29 Melichar, V. O., Behr-Roussel, D., Zabel, U., Uttenthal, L. O., Rodrigo, J., Rupin, A. *et al.* Reduced cGMP signaling associated with neointimal proliferation and vascular dysfunction in late-stage atherosclerosis. *Proc. Natl Acad. Sci. USA* **101**, 16671–16676 (2004).
- 30 Mishra, P. J., Humeniuk, R., Longo-Sorbello, G. S., Banerjee, D. & Bertino, J. R. A miR-24 microRNA binding-site polymorphism in dihydrofolate reductase gene leads to methotrexate resistance. *Proc. Natl Acad. Sci. USA* **104**, 13513–13518 (2007).
- 31 Mishra, P. J., Banerjee, D. & Bertino, J. R. MiRSNPs or MiR-polymorphisms, new players in microRNA mediated regulation of the cell: introducing microRNA pharmacogenomics. *Cell Cycle* **7**, 853–858 (2008).
- 32 Zhao, Z. Z., Croft, L., Nyholt, D. R., Chapman, B., Treloar, S. A., Hull, M. L. *et al.* Evaluation of polymorphisms in predicted target sites for micro RNAs differentially expressed in endometriosis. *Mol. Hum. Reprod.* **17**, 92–103 (2011).
- 33 Grechukhina, O., Petracco, R., Popkhadze, S., Massasa, E., Paranjape, T., Chan, E. *et al.* A polymorphism in a let-7 microRNA binding site of KRAS in women with endometriosis. *EMBO Mol. Med.* **4**, 206–217 (2012).
- 34 Yu, C., Pan, K., Xing, D., Liang, G., Tan, W., Zhang, L. *et al.* Correlation between a single nucleotide polymorphism in the matrix metalloproteinase-2 promoter and risk of lung cancer. *Cancer Res.* **62**, 6430–6433 (2002).

Supplementary Information accompanies the paper on Journal of Human Genetics website (<http://www.nature.com/jhg>)