ORIGINAL ARTICLE

Co-stimulatory *CD28* and transcription factor *NFKB1* gene variants affect idiopathic recurrent miscarriages

Maneesh Kumar Misra[1], Bharti Singh[2,4], Aditi Mishra[3,4] and Suraksha Agrawal[2]

Co-stimulatory *CD28* and transcription factor *NFKB1* genes are considered as a crucial player in the determination of inflammatory responses; genetic variability in these may modulate the risk for idiopathic recurrent miscarriages (IRM). We investigated the association of functional variants of *CD28* (rs3116496 T/C) and *NFKB1* (rs28362491 ins/del and rs696 A/G) with IRM cases. We recruited 200 IRM women with a history of at least three consecutive pregnancy losses before 20th week of pregnancy and 300 fertile control women. Determination of *CD28* (rs3116496 T/C) and *NFKB1* (rs28362491 ins/del and rs696 A/G) gene variants were based on the polymerase chain reaction pursued by restriction fragment length polymorphism analysis and validated with Sanger sequencing. Single marker analysis and multifactor dimensionality reduction (MDR) model used to predict the IRM risk. We observed nearly three- to twofold increased risk in single marker analysis for minor homozygous genotypes of rs3116496 T/C, rs28362491 ins/del and rs696 A/G tag-SNPs in IRM cases, suggesting the risk association. In MDR analysis, we observed 10.5-fold augmented risk among IRM women in three-SNP model (rs3116496 T/C, rs28362491 ins/del and rs696 A/G). The eQTL mapping analyses was performed to strengthen the results of our study. The eQTL mapping analysis revealed that the variations in *CD28* and *NFKB1* gene content might affect the abundance of transcripts of *CD28* and *Family with sequence similarity 177 member A1* (*FAM177A1*) genes, respectively. These results suggest that *CD28* and *NFKB1* gene variants may be associated with increased risks to IRM.
Journal of Human Genetics (2016) **61,** 1035–1041; doi:10.1038/jhg.2016.100; published online 4 August 2016

INTRODUCTION

Recurrent miscarriages (RM-OMIM: 614389) are considered as post-implantation failures in natural conception; and clinically described as three or more consecutive miscarriages before the 20th week of pregnancy. RM is a heterogeneous disease associated with various maternal and fetal factors.[1] Pregnancy is considered as an unique immunological paradox in which a specific population of T lymphocyte described as CD4+CD25+ regulatory T (Treg) cells protect fetus during gestation by inhibiting maternal allo-immunity reactions in opposition to paternal antigens among fetal cells.[2,3] Treg cells regulate the effector function of T cell, for instance helper T-cells and Th17 cells.[4] Th17 cells secrets interleukin (IL)-17, which is involved in the determination of inflammatory reactions, act together compassionately with helper T-cells and are probably associated with idiopathic recurrent miscarriages (IRM).[5] As IRM women show a considerably decreased frequencies of Treg cells,[6] and decreased inhibitory capability of regulatory T-cells have been reported in peripheral blood along with decidual tissues.[7] CD28-mediated activation of NF-kappaB family of transcription factors initiate Treg cell development.[8] Hence, we hypothesized that the decreased number and/or functional

insufficiency of Treg cells is due to the genetic variations in *CD28* and *NFKB1* may increase the risk to IRM.

Human *CD28* gene is positioned on the long arm q33.2 band of chromosome 2 with a 52.44 Kb nucleotide size, and comprises 4 exons and 3 introns. CD28 is recognized as a co-stimulatory molecule, which is constitutively expressed on naive and activated T-cells. The interaction of CD28 to their CD80 (B7.1)/CD86 (B7.2) ligand-mediated signals encourages differentiation and proliferation of T lymphocyte, and increases production of antibody through B lymphocytes.[9,10] Deficiency of co-stimulatory signal leads to T-lymphocyte tolerance and anergy.[11] *In vitro* and *in vivo* studies suggested the deregulation in CD28 pathways result into the perturbed adaptive immune responses.[12,13] Thus, it may be postulated that the deregulation in *CD28* pathways due to genetic variants may cause the predisposition to IRM.

Human *NFKB1* maps to band q24 of chromosome 4 and constitute a size of ~ 115.97 kb having 25 exons and 24 introns. The NF-κB of mammalian family consists of five proteins namely; NF-κB1, NF-κB2, RelA, RelB and c-Rel.[14] NF-kB has been considered as an essential transcription factor for sustaining normal immune homeostasis; an

[1]Department of Neurology, University of California San Francisco, San Francisco, CA, USA; [2]Department of Medical Genetics, Sanjay Gandhi Post-Graduate Institute of Medical Sciences, Lucknow, India and [3]Department of Computer Science and Engineering, Shri Ramswaroop Memorial University, Barabanki, India
[4]These authors contributed equally to this work.
Correspondence: Dr MK Misra, Department of Neurology, University of California San Francisco, San Francisco, CA 94158, USA.
E-mail: Maneesh.Misra@ucsf.edu
or Dr S Agrawal, Department of Medical Genetics, Sanjay Gandhi Post-graduate Institute of Medical Sciences, Raebareli Road, Lucknow, UP 226014, India.
E-mail: sur_ksha_agrawal@yahoo.co.in

insufficient NF-kB stimulation may mediate inflammation.[15] NF-κB1 suppresses NF-κB2 and act as a main player in the regulation of their activity. *NFKB1* and *NFKB2* genes encode the proteins NF-κB1 and NF-κB2, respectively. Considering the important role of NF-κB signaling pathway in the control of inflammatory responses, the genetic alterations in *NFKB1* gene content may modulate the risk for IRM.

Although the research on CD28 and transcription factor NF-κB family of proteins in IRM are still in its infancy, however, few studies have provided the insight into their potential role in IRM. Interestingly, a study revealed significantly reduced *CD28* expression in IRM women and recommended that a disproportion in the ratio of *CTLA4/CD28* expression on the feto-maternal membrane inhibits the activity of T cell in pregnant women, which may corroborate increased risk to IRM.[16] Further, a study conducted on animal model of early pregnancy has suggested stimulated NF-κB for the duration of uterine receptivity among equally the cyclic (breeding cycle) and pregnant endometrium.[17] Similarly, in human it has been reported that placental NF-κB stimulated ∼ 10-fold in women with pre-eclampsia.[18] Therefore, in this study we made an assumption that the deregulation in the expression and activity of CD28 and NF-κB may be due to the genetic variations or single nucleotide polymorphisms (SNP) in these genes and hence, may be associated with increased risk to IRM.

CD28 gene polymorphisms have been linked with the autoimmunity disorders, like Bechet's disease[19] and rheumatoid arthritis.[20] Recently, *NFKB1* gene variants have been reported in relation to autoimmunity disorders, such as systemic lupus erythematosus (SLE),[21] rheumatoid arthritis[21] and extensive colitis in inflammatory bowel disease.[22] However, *CD28* and *NFKB1* gene polymorphisms have not been investigated previously in IRM cases. Therefore, we conducted this study to fulfill these spaces.

Tag-SNP rs3116496 T/C exists in intronic region of *CD28* gene. *NFKB1* tag-SNPs namely; rs28362491 ins/del and rs696 A/G are located in the promoter and 3′ un-translated region (UTR), respectively. We aimed to investigate the association of functional variants of *CD28* (rs3116496 T/C) and *NFKB1* (rs28362491 ins/del and rs696 A/G) with IRM cases.

MATERIALS AND METHODS
Subjects and recruitment criteria
This study cohort consists of cases that went under assessment for the RM in the out-patients department (OPD) Department of Medical Genetics, Sanjay Gandhi Post-Graduate Institute of Medical Sciences (SGPGIMS), Lucknow. RM cases included in this study were recruited from November, 2011–August, 2015. All the registered RM women were primary aborters with a history of at least three recurrent pregnancy losses. We used a well-structured study proforma to note the detailed clinical information of RM and control women before their recruitment in the present study. All recurrent miscarriages cases included in this study are of clinical intrauterine pregnancies. The exclusion and inclusion criteria for recurrent miscarriages and control women have been provided in the section below.

RM women were assessed for various known causes of pregnancy losses through proper investigations such as karyotypes of the couple; antiphospholipid antibodies including lupus anticoagulant (PLR, 0.8–1.05) and anticardiolipin antibodies (IgG 0–12 IgG anticardiolipin units, immunoglobulin M 0–5 IgM anticardiolipin units) and prothrombotic risk factors including factor V Leiden and prothrombin mutations; day 21 progesterone levels, prolactin level; glycaemic curve, thyroid hormone levels. The uterine anomalies and cervical incompetence were ruled out by history, ultrasonography and appropriate investigations.

Several genetic association studies have been conducted in IRM cases. Some studies have found genetic variants associated with IRM,[23–25] whereas other studies have reported no association for genetic factors with IRM.[26,27] Thus, the field is still muddled and only the genetic studies conducted on IRM, which excluded the cases with abnormal parental karyotype will clarify this situation. Considering the facts that if a fetus was aneuploid, investigating the *CD28* and *NFKB1* tag-SNPs with clinical outcome of IRM would be illogical, hence, we excluded both the couples with aneuploid abortuses as well as couples having abnormal parental karyotypes.

Among initially screened 566 individuals, 35.34% (*n* = 200) had unexplained origin of RM and were recruited for the present study. Our center is a tertiary care referral hospital, where RM women visited after third miscarriages, only with abortus material were included in this study and then the karyotyping was done on the abortus material. In case of women with more than three miscarriages included in this study, the karyotyping was done on the abortus material at the time of their inclusion in the study, but not on their previous miscarriages due to the unavailability of the samples. The following RM cases were excluded, which do not meet the inclusion criteria: secondary aborters (*n* = 44), abnormal parental karyotype (*n* = 38), abortus material having aneuploidy (*n* = 52), maternal thrombophilias (*n* = 56), endocrine defects (*n* = 76), abnormal uterine anatomy (*n* = 52) and plasma lupus anticoagulant (*n* = 14); we have also excluded cases, which were not from the same linguistic group (*n* = 34). Among cases abnormal parental karyotype (*n* = 38) was seen in 14 male and 24 female partners. The 14 male cases revealed structural chromosomal abnormalities like (i) balanced translocations (*n* = 10), (ii) balanced Robertsonian translocation (*n* = 2) and (iii) inversion (*n* = 2). In female cases, 8 had numerical and 16 had structural chromosomal abnormalities. The structural chromosomal abnormalities were either balanced translocations (*n* = 12) or balanced Robertsonian translocation (*n* = 4). The anueploidy of abortus material (*n* = 52) included monosomy X (*n* = 18), triploidy (*n* = 10), trisomy of chromosome 13 (*n* = 4), trisomy of chromosome 15 (*n* = 2), trisomy of chromosome 16 (*n* = 14), trisomy of chromosome 18 (*n* = 2) and trisomy of chromosome 21 (*n* = 2). After excluding all the known causes of recurrent miscarriages among 566 RM women recruited initially for the present study, we remained with only 200 RM women with no known cause of RM for the genetic analysis of *CD28* and *NFKB1* tag-SNPs.

We included 300 control women with age and ethnicity matched to IRM (Supplementary Table S1) with minimum two live births and with no history of miscarriages, pre-eclampsia, ectopic pregnancy or preterm delivery. Controls were randomly chosen from the same population (based on caste system) during the same period, matched by age (±3 years) and residential location with cases. Two independent IRM cases and controls sets were constructed. The first set of cases–controls comprised Hindu women (100 cases and 200 controls), which matched literally 1:2 as they are more frequent in North Indians. The second set of case–control cohort matched 1:1 and comprises North Indian Muslim women (100 cases and 100 controls) as their populations are comparatively lower. The ethnic classification was based on caste system among North Indian populations. The study cohort recruited in the present study belong to the Uttar Pradesh state of North India and fall within the same linguistic group (Indo-Aryan speakers), which indicates that IRM and control women were of the same ethnicity. Proper stratification of IRM and control women was performed using different Alu repeat markers as reported in our earlier study.[28] Institutional Ethics Committee of SGPGIMS, Lucknow granted the permission for the present study and the study was conducted in accordance to the ethical standards laid down by the Declaration of Helsinki for performing medical research in human subjects. All subject provided written informed consent before their inclusion in the study.

Selection of tag-SNPs in *CD28* and *NFKB1* gene regions
We carried out selection of tag-SNPs in *CD28* and *NFKB1* gene regions by using International HapMap database (http://hapmap.ncbi.nlm.nih.gov/index.html.en) and Ensembl (http://www.ensembl.org/index.html) SNP databases. Tagger software was employed for tagging with the pairwise tagging algorithm $r^2 \geqslant 0.8$ with a minor allele frequency (MAF) more than 5% available in each database.[29] We selected a MAF of > 5% to maintain the balance for the factors such as: (a) the existing sample size, (b) a reasonable power of the study

(at least 80%) to detect the risk allele and (c) the number of tag-SNPs to be investigated. Additional criteria for the selection of tag-SNPs was based on published scientific reports to demonstrate their potential role in autoimmunity or inflammatory diseases in *CD28*[19,20] and *NFKB1* gene region.[22,30] Consequently, in this study we selected three-tag SNPs in the gene regions of *CD28* (rs3116496 T/C) and *NFKB1* (rs28362491 ins/del and rs696 A/G) with minor allele frequencies >5% in Asian population.

Laboratory analysis

Three milliliters of venous blood was taken from each IRM and control women in collection vials coated with ethylene diamine tetra acetic acid and preserved at or below −20 °C until use. Extraction of genomic DNA was carried out by QIAGEN genomic DNA extraction kit (Brand GMbII and Co KG, Cat # 51104, Valencia, CA, USA).

Determination of genetic variants of *CD28* (rs3116496 T/C) and *NFKB1* (rs28362491 ins/del and rs696 A/G) gene regions were performed through polymerase chain reaction (PCR) pursued by restriction fragment length polymorphism (RFLP) analysis as mentioned in the earlier published reports.[20,31] The primers and restriction enzymes used for the determination of *CD28* (rs3116496 T/C) and *NFKB1* (rs28362491 ins/del and rs696 A/G) polymorphic markers have been provided in Supplementary Table S2. Detection of genotypes in each IRM and control women has been carried out in absence of prior information about their control or disease status. We constituted two independent set each with 100 cases and 150 controls. Two independent workers have carried out genotyping on each set independently without the knowledge of subject's case or control status. On comparison we found 10% discrepancy in genotyping results of PCR–RFLP assay. These samples were again repeated blindly, and on comparison we found the concordance in the genotyping result. Further, Sanger sequencing was performed on nearly 20% of the randomly selected samples for the validation of PCR–RFLP genotyping analysis. The genotypes obtained from both Sanger sequencing and PCR–RFLP analysis were compared for quality control, which revealed 100% similar genotype call rate.

Statistical analysis

The estimation of sample size for IRM and control groups was carried out by using QUANTO version 1.2 (http://biostats.usc.edu/Quanto.html) program in the supervision of a statistician. We considered a power of 80% with 5% type 1 error ($\alpha = 0.05$) for sample size calculation. Chi-square (χ^2) test was used for estimation of the Hardy–Weinberg equilibrium *P*-value for rs3116496, rs28362491 and rs696 tag-SNPs in both IRM and control cohort. We used logistic regression analysis for the estimation of odds ratios (OR) with 95% confidence intervals (CI) using genotypes as well as alleles of the rs3116496, rs28362491 and rs696 tag-SNPs as the independent factors, and the risk of IRM being the dependent variable. Several confounding variables like ethnicity and age of IRM women at the time of diagnosis were included in the analysis, but removed from the final model because of their minimal affect on disease causation (not altering OR>5%), therefore, unadjusted logistic regression analysis was used.

We address the issue of multiple testing for *CD28* and *NFKB1* tag-SNPs by using the Benjamini–Hochberg approach, which regulates the false discovery rate (FDR).[32] The FDR-adjusted *P*-value has been computed by regulating FDR at 5% level to evaluate the statistical significance of studied *CD28* and *NFKB1* tag-SNPs following correction for multiple testing. After correction for multiple testing *P*-values less than or equal to 0.05 were considered as statistically significant. This study employed bootstrap re-sampling approach for the purpose of internal validation of the presented observations. We produced 500 bootstrapped samples from the genetic data of *CD28* (rs3116496) and *NFKB1* (rs28362491 and rs696) SNPs, and then computed genotype and allele specific *P*-value for three studied SNPs. Statistical tests were carried out by using the R language for statistical computing (https://cran.r-project.org/bin/windows/base/old/3.2.4/).

Multifactor dimensionality reduction (MDR) analysis was carried out by using MDR software version 1.2 (https://cran.r-project.org/web/packages/MDR/index.html). MDR analysis uses a non-parametric test to calculate all the possible one to three way SNP–SNP interactions. MDR approach has been explained in detail elsewhere,[33] and applied in this study among IRM and control cohort. In MDR analysis, a permutation test was used to calculate the multiple tests-adjusted *P*-value.

Expressed quantitative trait locus (eQTL) mapping analysis

We determined the functional affects of *CD28* and *NFKB1* tag-SNPs using eQTL mapping analysis. The eQTL mapping analysis was performed for CD28 tag-SNP by using blood eQTL.[34] Human genetic variation database (http://www.genome.med.kyoto-u.ac.jp/SnpDB/index.html) was used for the eQTL mapping analysis of *NFKB1* tag-SNPs.

RESULTS

Demographic clinical features of IRM and control women

Two hundred IRM cases and three hundred fertile control women were screened for the evaluation of *CD28* (rs3116496 T/C) and *NFKB1* (rs28362491 ins/del and rs696 A/G) polymorphic markers. We mentioned the demographic characteristics of cases and control women in Supplementary Table S1. We have selected age and ethnicity matched IRM (mean age: 28.6 ± 3.2) and control (mean age: 29.2 ± 6.8) women.

Distribution of *CD28* and *NFKB1* gene variants

Distribution of minor allele frequencies of the *CD28* (rs3116496 T/C) and *NFKB1* (rs28362491 ins/del and rs696 A/G) tag-SNPs are shown in Table 1. *CD28* and *NFKB1* tag-SNPs included in this study were observed in Hardy–Weinberg equilibrium among women of IRM and control group in both combined data set and sub-group analysis of Hindu and Muslim populations (Supplementary Table S3). The combined data set consists of both IRM cases and controls from the sub-group of Hindu and Muslim populations (Table 1). In combined data analysis, we observed nearly three- to twofold increased risk in single marker analysis for *CD28* rs3116496 T/C tag-SNP in additive and recessive models, and *NFKB1* tag-SNPs rs696 A/G and rs28362491 ins/del in additive, dominant and recessive models in IRM cases (Table 1). Further, in sub-group analysis of Hindu and Muslim populations the minor allele homozygous genotype carriers of IRM women revealed an increased risk of almost threefold in additive model and twofold in recessive model for both *NFKB1* tag-SNPs rs28362491 ins/del and rs696 A/G (Table 1). However, it is important to note that almost similar frequencies of minor allele homozygous genotypes of rs3116496 T/C (*CD28*) were found among IRM women (25% and 24.5%) and controls (19% and 19%) in sub-group analysis of both Hindu and Muslim populations, however, differences were significant only in cases of Hindu women but not in Muslim women (Table 1), this may be because of smaller sample size in later group. Interestingly, *CD28* (rs3116496 T/C) and *NFKB1* (rs28362491 ins/del and rs696 A/G) tag-SNPs, which revealed significance after corrections for multiple testing in IRM cases were found significant in the bootstrap analysis presenting internal validation to our observations in single marker analysis in both combined data analysis and sub-group analysis.

SNP–SNP interaction analysis

We determined all the possible one to three way SNP–SNP interactions by using MDR model. MDR analysis revealed 10.5-fold increased risk under three-factor model of rs3116496 T/C, rs28362491 ins/del and rs696 A/G tag-SNPs (OR = 10.5, 95% CI = 6.50–16.94, *P* = 1.000E − 04). The three-tag SNPs model shows maximum testing balanced accuracy (TBA), highest (10-fold) cross validation consistency (CVC) and greater significance (Table 2). Meanwhile, in two-factor (rs28362491 ins/del and rs696 A/G) and one-factor

Table 1 Distribution of *CD28* (rs3116496) and *NFKB1* (rs28362491 and rs696) gene variants among IRM cases and controls

Hindu women

Markers	Alleles[a]	MAF[b] (%) IRM, n=100	MAF[b] (%) Controls, n=200	Additive model[c] OR (95% CI)	P-value[f]	P-value[g]	Dominant model[d] OR (95% CI)	P-value[f]	P-value[g]	Recessive model[e] OR (95% CI)	P-value[f]	P-value[g]
rs3116496	T/C	25%	19%	2.85 (1.01–8.03)	4.686E−02[h]	2.395E−02[h]	1.31 (0.80–2.16)	0.2713	0.2794	2.72 (0.98–7.55)	0.0536	2.595E−02[h]
rs28362491	ins/del	57.50%	44.25%	2.70 (1.38–5.30)	3.692E−03[h]	3.900E−03[h]	1.92 (1.08–3.41)	2.482E−02[h]	1.597E−02[h]	2.02 (1.18–3.45)	9.508E−03[h]	7.984E−03[h]
rs696	A/G	51.50%	39.75%	2.59 (1.31–5.09)	5.810E−03[h]	5.988E−03[h]	1.60 (0.94–2.72)	0.0830	0.0958	2.22 (1.25–3.96)	6.566E−03[h]	1.966E−03[h]

Muslim women

Markers	Alleles[a]	MAF[b] (%) IRM, n=100	MAF[b] (%) Controls, n=100	Additive model[c] OR (95% CI)	P-value[f]	P-value[g]	Dominant model[d] OR (95% CI)	P-value[f]	P-value[g]	Recessive model[e] OR (95% CI)	P-value[f]	P-value[g]
rs3116496	T/C	24.50%	19%	2.93 (0.74–11.59)	0.1239	0.0798	1.29 (0.72–2.28)	0.3825	0.3733	2.81 (0.72–10.92)	0.1355	0.0818
rs28362491	ins/del	57.25%	44%	2.61 (1.21–5.65)	1.418E−02[h]	1.198E−02[h]	1.85 (0.98–3.50)	0.0576	2.994E−02[h]	2.02 (1.07–3.81)	2.879E−02[h]	2.196E−02[h]
rs696	A/G	51%	40%	2.41 (1.09–5.31)	2.826E−02[h]	2.395E−02[h]	1.52 (0.83–2.77)	0.1718	0.1557	2.14 (1.07–4.26)	2.960E−02[h]	3.393E−02[h]

Combined

Markers	Alleles[a]	MAF[b] (%) IRM, n=200	MAF[b] (%) Controls, n=300	Additive model[c] OR (95% CI)	P-value[f]	P-value[g]	Dominant model[d] OR (95% CI)	P-value[f]	P-value[g]	Recessive model[e] OR (95% CI)	P-value[f]	P-value[g]
rs3116496	T/C	24.75%	19%	2.82(1.25–6.37)	1.242E−02[h]	1.198E−02[h]	1.31 (0.90–1.89)	0.1534	0.1557	2.69 (1.20–6.01)	1.553E−02[h]	1.597E−02[h]
rs28362491	ins/del	57.25%	44.17%	2.65 (1.61–4.37)	1.240E−04[h]	1.996E−03[h]	1.88 (1.23–2.86)	3.155E−03[h]	7.984E−03[h]	2.02 (1.35–3.02)	5.790E−04[h]	1.996E−03[h]
rs696	A/G	51.25%	39.83%	2.51 (1.51–4.17)	3.540E−04[h]	1.996E−03[h]	1.56 (1.05–2.31)	2.640E−02[h]	3.590E−02[h]	2.19 (1.42–3.40)	3.990E−04[h]	1.996E−03[h]

Abbreviations: CI, confidence interval; MAF, minor allele frequency; OR, odd ratio.
[a]Major/minor alleles.
[b]Minor allele frequencies.
[c]Additive model: comparing mutant homozygous and heterozygous genotypes individually with wild-homozygous genotypes.
[d]Dominant model: mutant homozygous and heterozygous genotype taken together compared with wild-homozygous genotype.
[e]Recessive model: comparing mutant homozygous genotype with wild homozygous and heterozygous genotype taken together.
[f]False discovery rate (FDR)-adjusted P-value (yielding corrected P-value).
[g]On the basis of 500 bootstrap samples.
[h]Statistically risk-associated genotypes/alleles for idiopathic recurrent miscarriages (IRM).

Table 2 Multifactor dimensionality reduction (MDR) models showing association of high-order SNP–SNP interactions with IRM cases

Number of risk factors	Best interaction model	Training bal. acc.	Testing bal. acc.	CVC	P-value	OR (95% CI)
1	rs28362491	0.571	0.555	9/10	5.000E−04*	2.02 (1.35−3.02)
2	rs28362491, rs696	0.732	0.725	10/10	1.000E−04*	9.61 (6.02−15.32)
3	rs3116496, rs28362491, rs696[a]	0.739	0.701	10/10	1.000E−04*	10.5 (6.50−16.94)

Abbreviations: Bal. acc. balanced accuracy; CI, confidence interval; CVC, cross validation consistency; MDR, multifactor dimensionality reduction; OR, odd ratio.
[a]Best model was selected based on maximum testing balanced accuracy, highest CVC and greater significance.
*Significant P-values.
Balanced accuracy is the measurement of arithmetic mean of sensitivity and specificity of model.

Table 3 eQTL mapping analysis for CD28 and NFKB1 tag-SNPs

	SNP				Probe			eQTL	
SNP	Minor allele	Position	P-value	Gene	Gene	Probe	Position	Cis/Trans	Possible effects
rs3116496	C	204302757	1.933E−20	CD28	CD28	4890722	204310303	Cis-eQTL	Changes in abundance of transcript
rs696	G	35871093	4.467E−03	NFKBI	FAM177A1	P322704	35546376	Cis-eQTL	Changes in abundance of transcript

The eQTL mapping analysis was performed for CD28 tag-SNP by using Blood eQTL.[34] Human Genetic Variation Database (http://www.genome.med.kyoto-u.ac.jp/SnpDB/index.html) was used for the eQTL mapping analysis of NFKB1 tag-SNPs.

(rs28362491 ins/del) models the TBA, CVC and significance were reduced, respectively (Table 2).

eQTL mapping analysis for CD28 and NFKB1 tag-SNPs
We performed eQTL mapping analysis to determine the functional affects of bi-allelic tag-SNPs of CD28 rs3116496 T/C, and NFKB1 rs696 A/G for which we observed the significant difference between IRM's and control women at genomic level in this study. However, we did not included NFKB1 rs28362491 marker in eQTL mapping analysis because it is an insertion/deletion polymorphism and may not capture by the same probes used in eQTL mapping analysis for the bi-allelic SNPs. As tag-SNPs are situated in the 5′UTR/3′UTR, it has been reasonable that tag-SNPs may affect the abundance of transcript of the studied polymorphisms. Blood eQTL and Human Genetic Variation Database was used for eQTL mapping analysis for minor variant at 5′UTR of CD28 rs3116496 and 3′UTR of NFKB1 rs696 tag-SNP, respectively, which revealed that these tag-SNPs belong to Cis-eQTL and may be responsible for the changes in abundance of transcripts of CD28 and Family with sequence similarity 177 member A1 (FAM177A1) genes, respectively (Table 3). Interestingly, the eQTL analysis suggested that the genotype of rs696 tag-SNP was significantly associated with the changes in expression level of FAM177A1, but not NFKB1.

DISCUSSION
Co-stimulatory gene CD28 is vital for the production of immune responses. The binding of CD28 to CD80 (B7.1) and CD86 (B7.2) on antigen presenting cells, is arguably the most effective co-stimulatory molecule. Disturbance of CD28/B7 co-stimulation has been suggested to attenuate the severity of inflammation in some autoimmune disease models.[35] The vital controllers of NF-kB stimulation in T-cells are antigenic T cell receptor and CD28 receptor.[36] During inflammatory responses, NF-kB is important for the stimulation of T-cells, differentiation and survival. As CD28 and NFKB1 genes have an important function in the control of inflammatory responses, therefore, the genetic variation in these genes may be associated with etiology of IRM. This study has been designed to evaluate the relationship of CD28 (rs3116496 T/C) and NFKB1 (rs28362491 ins/del and rs696 A/G) gene variants with IRM cases.

This study found considerably increased prevalence of minor homozygous genotypes and minor allele frequencies of rs3116496 T/C, rs28362491 ins/del and rs696 A/G tag-SNPs in IRM cases as compared with controls, suggesting the risk association ranges from three- to twofolds for IRM (Table 1). Similarly, in sub-group analysis of Hindu and Muslim populations the minor allele homozygous genotype carriers of IRM women revealed an increased risk of almost threefold in additive model and twofold in recessive model for both NFKB1 tag-SNPs rs28362491 ins/del and rs696 A/G. Further, it has been worth to mention that almost similar frequencies of minor allele homozygous genotypes of rs3116496 T/C (CD28) were found among IRM women (25% and 24.5%) and controls (19% and 19%) in sub-group analysis of both Hindu and Muslim populations, however, differences were significant only in cases of Hindu women but not in Muslim women (Table 1), this may be because of smaller sample size in later group. Interestingly, rs3116496 T/C, rs28362491 ins/del and rs696 A/G tag-SNPs, which revealed significance after corrections for multiple testing in IRM cases stayed significant in bootstrap analysis presenting internal validation to our observations in both combined data analysis and sub-group analysis. MDR analysis was performed to estimate the possible SNP–SNP interactions in women of IRM and control groups. Interestingly, MDR analysis suggested 10.5-fold enhanced risk in three-factor model (rs3116496 T/C, rs28362491 ins/del and rs696 A/G) for IRM as compared with controls. The three-tag-SNP model revealed maximum TBA, greatest CVC and highest significance for possible prediction of IRM. The eQTL mapping analysis further strengthens our observations and revealed the influence of minor variants at CD28 rs3116496 T/C, and NFKB1 rs696 A/G tag-SNPs on the abundance of transcripts of CD28 and FAM177A1 genes, respectively (Table 3). It is important to consider that the genotype of rs696 tag-SNP was significantly associated with the changes in expression level of FAM177A1, but not NFKB1. We performed eQTL mapping analysis only for bi-allelic SNPs of CD28 rs3116496 T/C, and rs696 A/G but not for rs28362491 ins/del polymorphisms because eQTL for ins/del polymorphisms may require different probe set and may not be capture by the same probes used in eQTL mapping analysis for the bi-allelic SNPs. The earlier report may recommend the influence of CD28 gene variants on human diseases

risk, potentially by altering *CD28* function and/or expression.[20] The findings of this study are in concordance with earlier studies that suggested a risk association for rs3116496 T/C tag-SNP of *CD28* with human diseases like rheumatoid arthritis[20] and Bechet's disease.[19] It has been shown that minor homozygous (C/C) genotype of rs3116496 T/C tag-SNP reduces the sCD28 levels.[20] The down-regulation of sCD28 levels may affect CD28/B7 co-stimulation, and therefore, might be associated with the susceptibility to IRM.

Previous studies have reported that minor homozygous genotypes of *NFKB1* namely; rs28362491 ins/del and rs696 A/G tag-SNPs are linked with augmented risk to endometriosis[37] and extensive colitis in inflammatory bowel disease,[22] respectively, which are consistent with the observations of this study among IRM women. The potential role of tag-SNPs of *NFKB1* namely; rs28362491 ins/del and rs696 A/G in the correlation of disease risk appears to be associated with functional activity and gene expression of *NFKB1*, which have been considered vital in the regulation of crucial cellular mechanisms such as apoptosis and cell death independent of the NF-κB complex.[38] The earlier studies suggested that rs28362491 ins/del tag-SNP in the promoter (5′UTR) element of *NFKB1* has regulatory affect on gene expression of *NFKB1*, and for insertion (ATTG1) allele the functional activity has been found two times more than the deletion (ATTG2) allele.[39] A functional report has proposed that minor alleles of 3′UTR of *NFKB1* gene significantly decrease *NFKB1* mRNA stability.[31] Thus, it may be anticipated that the difference in the expression of *NFKB1* gene due to the variants of 5′UTR and 3′UTR of this gene may modify the risk of IRM.

Limitation of this study is that we have not estimated the *CD28* and *NFKB1* at proteomic levels, hence, future studies need to accomplish this issue for the complete understanding of the function of *CD28* and *NFKB1* in the etiology of IRM. The evaluation of potential link between alleles of risk variant and protein concentrations of *CD28* and *NFKB1* in blood obtained from women with IRM may provide useful information in the identification of a suitable biomarker. The pharmacologic approaches that could modulate these concentrations might be evaluated for the identification of their potential role in the determination of IRM risk.

Taken together, this study suggests an association of *CD28* (rs3116496 T/C) and *NFKB1* (rs28362491 ins/del and rs696 A/G) gene variants with augmented risk to IRM. The eQTL mapping analyses performed in this study strengthen these observations, and suggested that the variation in the *CD28* and *NFKB1* gene content may affect the abundance of transcripts of *CD28* and *FAM177A1* genes, respectively. To the best of our understanding based on the available information, this is the first study, which has examined the role of *CD28* and *NFKB1* tag-SNPs in IRM women. This appears to be an initial report, which revealed significant findings, however, needs further validation at both global and functional level.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

1 Christiansen, O. B., Steffensen, R., Nielsen, H. S. & Varming, K. Multifactorial etiology of recurrent miscarriage and its scientific and clinical implications. *Gynecol. Obstetr. Investig.* **66**, 257–267 (2008).

2 Zenclussen, A. C., Gerlof, K., Zenclussen, M. L., Sollwedel, A., Bertoja, A. Z., Ritter, T. et al. Abnormal T-cell reactivity against paternal antigens in spontaneous abortion: adoptive transfer of pregnancy-induced CD4+CD25+ T regulatory cells prevents fetal rejection in a murine abortion model. *Am. J. Pathol.* **166**, 811–822 (2005).

3 Mold, J. E., Michaelsson, J., Burt, T. D., Muench, M. O., Beckerman, K. P., Busch, M. P. et al. Maternal alloantigens promote the development of tolerogenic fetal regulatory T cells in utero. *Science* **322**, 1562–1565 (2008).

4 Saito, S., Nakashima, A., Shima, T. & Ito, M. Th1/Th2/Th17 and regulatory T-cell paradigm in pregnancy. *Am. J. Reprod. Immunol.* **63**, 601–610 (2010).

5 Lee, S. K., Kim, J. Y., Lee, M., Gilman-Sachs, A. & Kwak-Kim, J. Th17 and regulatory T cells in women with recurrent pregnancy loss. *Am. J. Reprod. Immunol.* **67**, 311–318 (2012).

6 Mei, S., Tan, J., Chen, H., Chen, Y. & Zhang, J. Changes of CD4+CD25 high regulatory T cells and FOXP3 expression in unexplained recurrent spontaneous abortion patients. *Fertil. Steril.* **94**, 2244–2247 (2010).

7 Jiang, G. J., Qiu, L. H. & Lin, Q. D. Regulative effects of regulatory T cells on dendritic cells in peripheral blood and deciduas from unexplained recurrent spontaneous abortion patients. *Zhonghua fu chan ke za zhi* **44**, 257–259 (2009).

8 Vang, K. B., Yang, J., Pagan, A. J., Li, L. X., Wang, J., Green, J. M. et al. Cutting edge: CD28 and c-Rel-dependent pathways initiate regulatory T cell development. *J. Immunol.* **184**, 4074–4077 (2010).

9 Rudd, C. E., Taylor, A. & Schneider, H. CD28 and CTLA-4 coreceptor expression and signal transduction. *Immunol. Rev.* **229**, 12–26 (2009).

10 Collins, M., Ling, V. & Carreno, B. M. The B7 family of immune-regulatory ligands. *Genome Biol.* **6**, 223 (2005).

11 Chen, L. Co-inhibitory molecules of the B7-CD28 family in the control of T-cell immunity. *Nat. Rev. Immunol.* **4**, 336–347 (2004).

12 Hutloff, A., Dittrich, A. M., Beier, K. C., Eljaschewitsch, B., Kraft, R., Anagnostopoulos, I. et al. ICOS is an inducible T-cell co-stimulator structurally and functionally related to CD28. *Nature* **397**, 263–266 (1999).

13 Yoshinaga, S. K., Whoriskey, J. S., Khare, S. D., Sarmiento, U., Guo, J., Horan, T. et al. T-cell co-stimulation through B7RP-1 and ICOS. *Nature* **402**, 827–832 (1999).

14 Nabel, G. J. & Verma, I. M. Proposed NF-kappa B/I kappa B family nomenclature. *Genes Dev.* **7**, 2063 (1993).

15 Abdallah, A., Sato, H., Grutters, J. C., Veeraraghavan, S., Lympany, P. A., Ruven, H. J. et al. Inhibitor kappa B-alpha (IkappaB-alpha) promoter polymorphisms in UK and Dutch sarcoidosis. *Genes Immun.* **4**, 450–454 (2003).

16 Wang, X., Ma, Z., Hong, Y., Lu, P. & Lin, Q. Expression of CD28 and cytotoxic T lymphocyte antigen 4 at the maternal-fetal interface in women with unexplained pregnancy loss. *Int. J. Gynaecol. Obstetr.* **93**, 123–129 (2006).

17 Ross, J. W., Ashworth, M. D., Mathew, D., Reagan, P., Ritchey, J. W., Hayashi, K. et al. Activation of the transcription factor, nuclear factor kappa-B, during the estrous cycle and early pregnancy in the pig. *Reprod. Biol. Endocrinol.* **8**, 39 (2010).

18 Vaughan, J. E. & Walsh, S. W. Activation of NF-kappaB in placentas of women with preeclampsia. *Hypertens. Pregnancy* **31**, 243–251 (2012).

19 Gunesacar, R., Erken, E., Bozkurt, B., Ozer, H. T., Dinkci, S., Erken, E. G. et al. Analysis of CD28 and CTLA-4 gene polymorphisms in Turkish patients with Behcet's disease. *Int. J. Immunogenet.* **34**, 45–49 (2007).

20 Ledezma-Lozano, I. Y., Padilla-Martinez, J. J., Leyva-Torres, S. D., Parra-Rojas, I., Ramirez-Duenas, M. G., Pereira-Suarez, A. L. et al. Association of CD28 IVS3 +17 T/C polymorphism with soluble CD28 in rheumatoid arthritis. *Dis. Markers* **30**, 25–29 (2011).

21 Orozco, G., Sanchez, E., Collado, M. D., Lopez-Nevot, M. A., Paco, L., Garcia, A. et al. Analysis of the functional NFKB1 promoter polymorphism in rheumatoid arthritis and systemic lupus erythematosus. *Tissue Antigens* **65**, 183–186 (2005).

22 Szamosi, T., Lakatos, P. L., Hungarian IBD Study Group, Szilvasi, A., Lakatos, L., Kovacs, A. et al. The 3'UTR NFKBIA variant is associated with extensive colitis in Hungarian IBD patients. *Dig. Dis. Sci.* **54**, 351–359 (2009).

23 Wu, Z., You, Z., Zhang, C., Li, Z., Su, X., Zhang, X. et al. Association between functional polymorphisms of Foxp3 gene and the occurrence of unexplained recurrent spontaneous abortion in a Chinese Han population. *Clin. Immunol.* **2012**, 896458 (2012).

24 Saxena, D., Misra, M. K., Parveen, F., Phadke, S. R. & Agrawal, S. The transcription factor Forkhead Box P3 gene variants affect idiopathic recurrent pregnancy loss. *Placenta* **36**, 226–231 (2015).

25 Siddesh, A., Parveen, F., Misra, M. K., Phadke, S. R. & Agrawal, S. Platelet-specific collagen receptor glycoprotein VI gene variants affect recurrent pregnancy loss. *Fertil. Steril.* **102**, 1078–1084 e1073 (2014).

26 Iversen, A. C., Nguyen, O. T., Tommerdal, L. F., Eide, I. P., Landsem, V. M., Acar, N. et al. The HLA-G 14 bp gene polymorphism and decidual HLA-G 14 bp gene expression in pre-eclamptic and normal pregnancies. *J. Reprod. Immunol.* **78**, 158–165 (2008).

27 Lin, A., Yan, W. H., Dai, M. Z., Chen, X. J., Li, B. L., Chen, B. G. et al. Maternal human leukocyte antigen-G polymorphism is not associated with pre-eclampsia in a Chinese Han population. *Tissue Antigens* **68**, 311–316 (2006).

28 Tripathi, M., Chauhan, U. K., Tripathi, P. & Agrawal, S. Role of Alu element in detecting population diversity. *Int. J. Human Genet.* **8**, 61–74 (2008).

29 de Bakker, P. I., Yelensky, R., Pe'er, I., Gabriel, S. B., Daly, M. J. & Altshuler, D. Efficiency and power in genetic association studies. *Nat. Genet.* **37**, 1217–1223 (2005).

30 Borm, M. E., van Bodegraven, A. A., Mulder, C. J., Kraal, G. & Bouma, G. A NFKB1 promoter polymorphism is involved in susceptibility to ulcerative colitis. *Int. J. Immunogenet.* **32**, 401–405 (2005).

31 Song, S., Chen, D., Lu, J., Liao, J., Luo, Y., Yang, Z. *et al.* NFkappaB1 and NFkappaBIA polymorphisms are associated with increased risk for sporadic colorectal cancer in a southern Chinese population. *PLoS ONE* **6**, e21726 (2011).

32 Benjamini, Y. & Hochberg, Y. On the adaptive control of the false discovery rate in multiple testing with independent statistics. *J. Educ. Behav. Stat.* **25**, 60–83 (2000).

33 Hahn, L. W., Ritchie, M. D. & Moore, J. H. Multifactor dimensionality reduction software for detecting gene–gene and gene–environment interactions. *Bioinformatics* **19**, 376–382 (2003).

34 Westra, H. J., Peters, M. J., Esko, T., Yaghootkar, H., Schurmann, C., Kettunen, J. *et al.* Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat. Genet.* **45**, 1238–1243 (2013).

35 Salomon, B., Lenschow, D. J., Rhee, L., Ashourian, N., Singh, B., Sharpe, A. *et al.* B7/CD28 costimulation is essential for the homeostasis of the CD4+CD25+ immunoregulatory T cells that control autoimmune diabetes. *Immunity* **12**, 431–440 (2000).

36 Dennehy, K. M., Kerstan, A., Bischof, A., Park, J. H., Na, S. Y. & Hunig, T. Mitogenic signals through CD28 activate the protein kinase Ctheta-NF-kappaB pathway in primary peripheral T cells. *Int. Immunol.* **15**, 655–663 (2003).

37 Zhou, B., Rao, L., Peng, Y., Wang, Y., Qie, M., Zhang, Z. *et al.* A functional promoter polymorphism in NFKB1 increases susceptibility to endometriosis. *DNA Cell Biol.* **29**, 235–239 (2010).

38 Yu, Y., Wan, Y. & Huang, C. The biological functions of NF-kappaB1 (p50) and its potential as an anti-cancer target. *Curr. Cancer Drug Targets* **9**, 566–571 (2009).

39 Karban, A. S., Okazaki, T., Panhuysen, C. I., Gallegos, T., Potter, J. J., Bailey-Wilson, J. E. *et al.* Functional annotation of a novel NFKB1 promoter polymorphism that increases risk for ulcerative colitis. *Human Mol. Genet.* **13**, 35–45 (2004).

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