Rheumatology Advance Access published March 24, 2010
RHEUMATOLOGY

Original article

doi:10.1093/rheumatology/kep470

Non-synonymous variant (Gly307Ser) in *CD226* is associated with susceptibility to multiple autoimmune diseases

Amit K. Maiti[1], Xana Kim-Howard[1], Parvathi Viswanathan[1], Laura Guillén[2], Xiaoxia Qian[3], Adriana Rojas-Villarraga[4], Celi Sun[1], Carlos Cañas[5], Gabriel J. Tobón[5], Koichi Matsuda[6], Nan Shen[3], Alejandra C. Cherñavsky[2], Juan-Manuel Anaya[4] and Swapan K. Nath[1]

Abstract

Objectives. Recently, a non-synonymous (Gly307Ser) variant, rs763361, in the *CD226* gene was shown to be associated with multiple autoimmune diseases (ADs) in European Caucasian populations. However, shared autoimmunity with *CD226* has not been evaluated in non-European populations. The aim of the present study is to assess the association of this single nucleotide polymorphism (SNP) with ADs in non-European populations.

Methods. To replicate this association in non-European populations, we evaluated case–control association between rs763361 and coeliac disease (CED) samples from Argentina; SLE, RA, type-1 diabetes (T1D) and primary SS (pSS) from Colombia; and SLE samples from China and Japan. We genotyped rs763361 and evaluated its genetic association with multiple ADs, using χ^2-test. For each association, odds ratio (OR) and 95% CI were calculated.

Results. We show that rs763361 is significantly associated with Argentinean CED ($P = 0.0009$, OR = 1.60). We also observed a trend of possible association with Chinese SLE ($P = 0.01$, OR = 1.19), RA ($P = 0.047$, OR = 1.25), SLE ($P = 0.0899$, OR = 1.24) and pSS ($P = 0.09$, OR = 1.33) in Colombians. Meta-analyses for SLE (using our three populations) and T1D (our population and three published populations) yielded significant association with rs763361, $P = 0.009$ (OR = 1.16) and $P = 1.1.46 \times 10^{-9}$ (OR = 1.14), respectively.

Conclusions. Our results demonstrate that the coding variant rs763361 in *CD226* gene is associated with multiple ADs in non-European populations.

Key words: CD226, Autoimmunity, Latin-America, Asia.

[1]Genetic Epidemiology Unit, Arthritis and Immunology Research Program, Oklahoma Medical Research Foundation, Oklahoma City, OK, USA, [2]Immunogenetic Laboratory, Hospital de Clínicas José de San Martín, Universidad de Buenos Aires, Buenos Aires, Argentina, [3]Shanghai Institute of Rheumatology, Renji Hospital, JiaoTong University School of Medicine, Shanghai, P.R. China, [4]Centre for Autoimmune Diseases Research (CREA), Universidad del Rosario-Corporación para Investigaciones Biológicas, Bogota, [5]Rheumatology Unit, Fundación Valle del Lili, Cali, Colombia and [6]Laboratory of Molecular Medicine, Human Genome Centre, Institute of Medical Science, University of Tokyo, Tokyo, Japan.

Submitted 11 August 2009; revised version accepted 18 December 2009.

Correspondence to: Swapan K. Nath, Genetic Epidemiology Unit, Arthritis and Immunology Research Program, Oklahoma Medical Research Foundation, 825 N.E. 13th Street, Oklahoma City, OK 73104, USA. E-mail: swapan-nath@omrf.org

Introduction

There is a growing understanding that susceptibility to the autoimmune diseases (ADs) is due to a complex interaction of multiple genes and environmental factors, and some of these may be shared among many ADs. Genetic susceptibility of multiple organ-specific ADs often share underlying commonalities [1, 2]. For example, recently accumulated evidence indicates that common single nucleotide polymorphisms (SNPs), such as rs7574865 (*STAT4*), rs6920220 and rs10499194 (*TNFAIP3* region) are associated with multiple ADs, including RA, SLE and type-1 diabetes (T1D) [3–8]. Similarly, shared autoimmunity with multiple autoimmune disorders,

including T1D, RA and SLE, is also observed for the C1858T mutation of *PTPN22* (rs2476601) [9–14] and multiple polymorphisms of *IRF5* [15–17].

The non-synonymous (Gly307Ser) coding variant, rs763361 (C/T polymorphism), in the *CD226* gene, located at 18q22.3, has recently been shown to be associated with T1D, multiple sclerosis (MS), coeliac disease (CED), RA, WG and autoimmune thyroid disease (AITD) in European Caucasian populations [18–20]. A genome-wide association study (GWAS) identified *CD226* association at 18q22.3 with T1D ($P_{overall} = 1.38 \times 10^{-8}$) and thyroid disease [18, 20]. The *CD226* (*DNAM1*) gene encodes a 67 kDa [336 amino acid (aa)] cell surface membrane protein with two immunoglobulin V set domains (aa31–aa125 and aa135–aa240) with an extracellular region. CD226, an immunoglobulin supergene family receptor, is expressed in NKT cells, $CD4^+$ and $CD8^+$ cells, monocytes and in a subset of B cells. rs763361 is located at the C-terminal end of the protein and is presumed to modulate the expression of a signalling molecule in the cytoplasm. In autoimmune encephalomyelitis [21], anti-CD226 treatment delayed disease progression and *CD226* expression deficiency initiates anti-CD95-induced apoptosis in NKT^+ cells in SLE patients [22].

While similar immunogenetic mechanisms may underlie ADs, the genetic factors of ADs could vary depending on ethnicity and geography. Ethnicity-specific association of an SNP in ADs in different population has been documented, such as the rs1143679 in *ITGAM*, which is associated with European American, Hispanic and African-American populations but is monomorphic in many Asian (Korean and Japanese) populations [23, 24]. Strong ethnicity-specific associations are also observed for *PTPN22* (C1858T, rs2476601) [10, 13, 25], *IRF5* (multiple polymorphisms) [16, 26, 27] and the HLA locus [9, 28] in Caucasian populations but not in Asian populations. Although genetic association of rs763361 in *CD226* has been tested in a number of ADs in European Caucasian populations, to our knowledge this association has not been evaluated in non-European populations. Therefore, in the present study, we studied the genetic associations of rs763361 with T1D, SLE, RA and primary SS (pSS) samples from Colombia, CED from Argentina and SLE samples from Asian populations (Chinese and Japanese). We also included published data for meta-analyses to examine disease-specific associations with *CD226*.

Materials and methods

Sample collection

All subjects were enrolled into their respective studies after obtaining their written informed consent and following protocols approved by the appropriate institutional review boards. The samples used in the present study are shown in Table 1.

Argentinean 180 CED patients and 220 controls were recruited through the Hospital de Clínicas José de San Martín, Buenos Aires, Argentina. Diagnosis of CED was based on the presence of clinical features, including characteristic coeliac enteropathy [29], the presence of a positive CED-related serology (disease-specific antibodies) and the response to a strictly monitored gluten-free diet.

Colombian participants consisted of 738 cases and 347 controls, of which 648 cases were enrolled at the Corporación para Investigaciones Biológicas in Medellin and the remaining 90 cases were enrolled at the Fundación Clínica Vale de Lili, in Cali, Colombia. All patients with ADs met the international classification criteria for their respective disease [30–32] and their clinical and immunological characteristics were similar to those previously reported [20, 33, 34]. The common controls for all Colombian analyses included 347 individuals without a history of chronic inflammatory autoimmune or infectious diseases and were unrelated to patients.

Chinese samples included 912 SLE cases and 1080 healthy controls recruited from the Shanghai Renji hospital. Japanese samples, 175 cases and 363 controls, were collected from a Tokyo suburb. All SLE patients met the ACR criteria for the classification of SLE [31, 35] and controls were from healthy blood donors.

Genotyping

All de-identified genomic DNA samples from Argentina and Colombia were genotyped for rs763361 at the Oklahoma Medical Research Foundation using the

Table 1 Allelic association of rs763361 in Argentinian, Colombian, Chinese and Japanese populations

Population	Disease phenotype	Case/control	Male/female	Risk allele frequency		Genotype		χ^2	*P*-value	OR	95% CI		Power[b]
				Case	Control	Case	Control				Lower	Upper	
Argentinian	CED	180/220	156/202	0.558	0.441	61/79/40	43/108/69	10.92	0.00095	1.60	1.21	2.12	0.25
Colombian	RA	323/347[a]	159/511	0.440	0.386	57/170/96	50/168/129	3.95	0.0469	1.25	1.00	1.55	0.38
Colombian	T1D	129/347[a]	183/293	0.438	0.386	26/61/42	50/168/129	2.10	0.1469	1.24	0.93	1.66	0.24
Colombian	pSS	90/347[a]	117/320	0.456	0.386	13/56/21	50/168/129	2.87	0.0905	1.33	0.96	1.85	0.20
Colombian	SLE	196/347[a]	124/419	0.439	0.386	38/96/62	50/168/129	2.88	0.0899	1.24	0.97	1.60	0.30
Japanese	SLE	175/363	311/227	0.443	0.449	35/85/55	72/182/109	0.04	0.8485	0.98	0.76	1.26	0.29
Chinese	SLE	912/1080	424/1397	0.378	0.338	111/344/294	107/486/444	6.18	0.0129	1.19	1.04	1.37	0.88

[a]Colombian controls are the same for Colombian case–control association analysis; [b]Power analysis was performed using OR = 1.2, control minor allele frequency = 40% and the number of cases and controls for each population.

TaqMan assay (Applied Biosystems). Chinese samples were genotyped using the TaqMan assay using ABI7900 (Applied Biosystems) at Dr Shen's laboratory, and Japanese samples were genotyped by the Invader assay combined with multiplex-PCR using ABI 7900 (Applied Biosystems) at Dr Matsuda's laboratory.

Statistical analysis

For each disease phenotype, we estimated the Hardy–Weinberg equilibrium separately for cases and controls. Case–control association studies were analysed by χ^2-test using 2×3 and 2×2 contingency tables of genotype and allele counts, respectively. Odds ratios (ORs) and 95% CIs for allelic tests and genotypic tests under different genetic models were calculated using PLINK 1.05 [36]. To identify a plausible genetic model for each disease phenotype, model-based association was assessed using both *P*-values and Akaike information criterion (AIC) under dominant, recessive and multiplicative models, where at least a marginally significant association was found from the allelic association test. We considered $P < 0.05$ to be significant. However, since we are testing association under many genetic models with many ADs, each association test can represent a separate hypothesis. Therefore, we set a predefined threshold for a *P*-value (0.1–0.01) to be considered as a 'suggestive association'.

Meta-analysis

To see an overall effect of rs763361 with SLE and T1D, meta-analyses were performed using CatMap software [37]. For the SLE meta-analysis, Colombian, Chinese and Japanese SLE patients and associated controls were used. For T1D, we performed meta-analysis with our Colombian T1D population and three previously published populations [18, 38].

Power analysis

Power analyses were performed using CATS software [39]. For power calculations, OR was set at 1.2, minor allele frequencies (MAF) at 40% and the numbers of cases and controls were the sample sets that we used for each population.

Results

The SNP rs763361 was in Hardy–Weinberg equilibrium ($P \geqslant 0.05$) in both cases and controls in all populations. The disease allele (*T*) frequency varies from population to population, with the highest and lowest MAF estimated to be 45 and 34% in Japanese and Chinese controls, respectively. The results of our power analysis were performed retrospectively, the approximate values for the MAF from controls (40%), effect size (OR = 1.2) and the number of cases and controls for each population. The MAF (46%) and effect size (OR = 1.16) were not too different from the control's MAF and effect size reported in Caucasian populations [18].

The summary of the allelic association test and relevant information are given in Table 1. The strongest allelic association was observed between Argentinean CED cases and healthy controls [$P = 0.00095$, OR (95% CI) = 1.60 (1.21, 2.12)]. Suggestive associations were also observed between Colombian RA cases and controls [$P = 0.047$, OR (95% CI) = 1.25 (1.003, 1.55)], as well as between Chinese SLE cases and controls [$P = 0.013$, OR (95% CI) = 1.19 (1.04, 1.37)]. No allelic association was found between Colombian T1D, Colombian pSS or Japanese SLE cases and their respective controls. To identify the best fitted model, *P*-value and AIC estimate under dominant, recessive and multiplicative models were compared. While we observed association with different models for some of the populations, we, however, did not find a single model that was significantly better than the other models.

To see the association between rs763361 and different autoimmune phenotypes (T1D, SLE, pSS and RA) from Colombian populations, we performed a combined analysis using all four autoimmune phenotypes *vs* controls. We have observed that rs763361 showed 'suggestive' association [$P = 0.0158$, OR (95% CI) = 1.25 (1.04, 1.51)] (data not shown).

To see an overall effect with SLE, a meta-analysis was performed on Colombian, Japanese and Chinese SLE cases and controls. Since the test of heterogeneity was not significant ($P_{het} = 0.33$), the combined *P*-value was calculated under a fixed effects model (Table 2). The overall association was significant [$P_{meta} = 0.0086$, OR (95% CI) = 1.16 (1.04, 1.29)]. Additionally, we also performed a meta-analysis of T1D using our Colombian data and three published data sets (Estonians [38], UK case–controls and trio [18]). We found a significantly robust association with rs763361 [$P_{meta} = 1.46 \times 10^{-9}$, OR (95% CI) = 1.14 (1.09, 1.18); $P_{het} = 0.13$ under the fixed effects model].

Discussion

The genetic association study is a powerful tool to identify associated SNPs influencing disease susceptibility. Recent GWAS, as well as candidate gene analyses, indicate that rs763361 is significantly associated with multiple ADs, including T1D, CED, MS, AITD, WG and RA in European Caucasian populations [19, 20, 38, 40, 41]. However, this association had not been evaluated in any non-European samples.

To evaluate the impact of rs763361 on non-European populations, we genotyped relatively homogeneous samples with ethnically matched controls in multiple ADs in two geographical regions of Latin America (Colombia and Argentina) and two geographical populations from Asia (China and Japan). We found evidence of significant association of this SNP in Argentinian CED. We also observed a trend of possible association or 'suggestive associations' (*P*-value between 0.01 and 0.1) with SLE, RA and pSS in Colombian samples and Chinese SLE sample. These results suggest that rs763361 is associated not only in European-derived samples but also with other phenotypes in South American populations. Model-based

Table 2 Meta-analyses with rs763361 and SLE and T1D

Population	Reference	Case/control	Risk allele frequency		OR	95% CI		*P*-value	Heterogeneity	
			Case	Control		Upper	Lower		χ^2	*P*-value
SLE										
Pooled			–	–	1.16	1.04	1.29	0.0086	2.19	0.3344
Colombian	Present	196/347	0.439	0.386	1.24	0.97	1.60	0.0899	–	–
Chinese	Present	912/1080	0.378	0.338	1.19	1.04	1.37	0.0129	–	–
Japanese	Present	175/363	0.443	0.449	0.98	0.76	1.26	0.8485	–	–
T1D										
Pooled			–	–	1.14	1.09	1.18	1.46×10^{-9}	5.68	0.1282
Estonian	[38]	154/230	0.510	0.413	1.48	1.11	1.98	0.0084	–	–
Colombian	Present	129/347	0.438	0.386	1.24	0.93	1.66	0.1469	–	–
UK	[18]	6021/6088	0.504	0.469	1.15	1.09	1.21	2.82×10^{-8}	–	–
UK trio	[18]	2997 trio	1396[a]	1297[b]	1.08	0.998	1.16	0.0281	–	–

[a]Transmitted alleles; [b]non-transmitted alleles.

(dominant, recessive and multiplicative) association was assessed for all populations to find a best fitted genetic model. However, none of the models clearly and consistently fit the association between rs763361 and ADs.

The results of our meta-analyses show that the association between *CD226* SNP rs763361 and multiple ADs is consistent and robust ($P = 0.0056$ for SLE, $P = 1.46 \times 10^{-9}$). Our T1D and SLE meta-analyses also demonstrated that the ORs from different populations were not significantly different (i.e. homogeneous) for any of the disease-specific analyses.

One potential problem for association studies is the presence of population substructure in the samples, which raises the potential for confounding and spurious results, especially the admixed population. It is now well known that populations of both Argentina and Colombia have varying degrees of admixture with European, Native American and African ancestral populations. Therefore, if the samples come from several ancestral subpopulations with different allele frequencies, and if the proportions of cases and controls sampled from each subpopulation differ, differences in allele frequencies between cases and controls will appear, mimicking a statistical signal of association leading to false-positive results. It would be ideal to check thoroughly for hidden population substructure in both cases and controls for population stratification while performing this ancestral study. Unfortunately, we do not have enough ancestry informative markers available for this analysis. However, most of our Antioquia, Colombian samples were collected from the Paisa community from Colombia, which is historically considered an ethnically homogeneous population [42]. In this association study, for all four disease phenotypes (pSS, T1D, RA and SLE) from the Colombian population we have used common controls. The use of common controls is a valid general strategy and has been previously applied successfully in disease-association studies conducted by the Wellcome Trust [43] and others, including Jimenez-Morales *et al*. [44].

We did not obtain significant association, allelic or model-based, in the association with T1D, which has been reported earlier [8, 20]. The limited sample size might explain this lack of association, as our power analysis demonstrates that both these association tests were under-powered. The small cohort size in Colombian T1D and Japanese SLE may be a reason for not obtaining significant associations. Based on our power analysis using population-specific allele frequencies and ORs, it is also reflected that except Chinese SLE, all other groups were under-powered. It is also possible that the admixture in the Latin-American samples reduced the power to detect significant associations. However, despite suggestive associations, our meta-analyses, including published data, yielded significant association for SLE and T1D, and suggest that the disease-specific association with rs763361 is significant for many ADs across diverse ethnic populations.

Given the high degree of overlap between ADs, in terms of symptoms and the degree of multiple ADs within certain individuals, for example, SLE and pSS, or SLE and RA, it would be ideal to perform these analyses with individuals without overlapping syndromes, and to perform subgroup analyses with specific disease sub-phenotypes. However, while we have verified diagnosis for all AD patients in this study, we do not have clinical data with regard to comorbid conditions or for individual disease sub-phenotypes. The study of comorbid autoimmune diseases would be an important follow-up study to develop an understanding of the underlying functional mechanisms of this genetic association.

It has been hypothesized that ADs may be related through an underlying cellular mechanism involved in immune response cells for humoral and cell-mediated immunity. *CD226* is also an important cell surface receptor molecule involved in T-cell adhesion and activation. The exon-7 variant, rs763361/(Gly307Ser) in the cytoplasmic tail can affect the signalling function in T and/or other cells, and the underlying cellular mechanisms may lead to different phenotypes in multiple ADs. Another compelling hypothesis [18] is that this variant could disrupt a splice site enhancer or silencer, leading to alternative RNA splicing. The residue (serine) is conserved (data not shown) in higher mammals such as humans, chimpanzees and

gorillas, although isoleucine in place of serine is present in mice and rats. However, it is still to be explored as to how this single amino acid variation modulates signalling pathway genes or proteins, and gives rise to different clinical phenotypes that lead to various autoimmune phenotypes.

Rheumatology key messages

- Functional variant, rs763361, within *CD226* is associated with multiple autoimmune diseases.
- Together with earlier studies, this association is replicated in both European and non-European populations.

Acknowledgements

We wish to thank the patients and their families for their co-operation. We also thank anonymous reviewers for their critical comments to help improve the quality of this manuscript.

Funding: This study was made possible by funding from National Institutes of Health (R01AI063622), grants from Colciencias (2213-04-16484); School of Medicine, Rosario University and the Colombian Association of Rheumatology, Bogota, Colombia.

Disclosure statement: The authors have declared no conflicts of interest.

References

1 Deshmukh H, Kim-Howard X, Nath SK. Replication of recently identified associated SNPs from 6 autoimmune diseases in GAW16 RA data. BMC Proc 2009; Suppl. 7:831.

2 Lettre G, Rioux JD. Autoimmune diseases: insights from genome-wide association studies. Hum Mol Genet 2008; 17:R116–21.

3 Plenge RM, Cotsapas C, Davies L *et al*. Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. Nat Genet 2007;39:1477–82.

4 Remmers EF, Plenge RM, Lee AT *et al*. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N Engl J Med 2007;357:977–86.

5 Thomson W, Barton A, Ke X *et al*. Rheumatoid arthritis association at 6q23. Nat Genet 2007;39:1431–3.

6 Graham RR, Cotsapas C, Davies L *et al*. Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus. Nat Genet 2008;40:1059–61.

7 Musone SL, Taylor KE, Lu TT *et al*. Multiple polymorphisms in the TNFAIP3 region are independently associated with systemic lupus erythematosus. Nat Genet 2008;40: 1062–4.

8 Fung EY, Smyth DJ, Howson JM *et al*. Analysis of 17 autoimmune disease-associated variants in type 1 diabetes identifies 6q23/TNFAIP3 as a susceptibility locus. Genes Immun 2009;10:188–91.

9 Plenge RM. Recent progress in rheumatoid arthritis genetics: one step towards improved patient care. Curr Opin Rheumatol 2009;21:262–71.

10 Lee YH, Rho YH, Choi SJ *et al*. The PTPN22 C1858T functional polymorphism and autoimmune diseases–a meta-analysis. Rheumatology 2007;46:49–56.

11 Plenge RM, Padyukov L, Remmers EF *et al*. Replication of putative candidate-gene associations with rheumatoid arthritis in >4000 samples from North America and Sweden: association of susceptibility with PTPN22, CTLA4, and PADI4. Am J Hum Genet 2005;77:1044–60.

12 Begovich AB, Carlton VE, Honigberg LA *et al*. A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. Am J Hum Genet 2004;75:330–7.

13 Kyogoku C, Langefeld CD, Ortmann WA *et al*. Genetic association of the R620W polymorphism of protein tyrosine phosphatase PTPN22 with human SLE. Am J Hum Genet 2004;75:504–7.

14 Ladner MB, Bottini N, Valdes AM, Noble JA. Association of the single nucleotide polymorphism C1858T of the PTPN22 gene with type 1 diabetes. Hum Immunol 2005; 66:60–4.

15 Sigurdsson S, Nordmark G, Goring HH *et al*. Polymorphisms in the tyrosine kinase 2 and interferon regulatory factor 5 genes are associated with systemic lupus erythematosus. Am J Hum Genet 2005;76:528–37.

16 Graham RR, Kozyrev SV, Baechler EC *et al*. A common haplotype of interferon regulatory factor 5 (IRF5) regulates splicing and expression and is associated with increased risk of systemic lupus erythematosus. Nat Genet 2006;38: 550–5.

17 Sigurdsson S, Padyukov L, Kurreeman FA *et al*. Association of a haplotype in the promoter region of the interferon regulatory factor 5 gene with rheumatoid arthritis. Arthritis Rheum 2007;56:2202–10.

18 Todd JA, Walker NM, Cooper JD *et al*. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. Nat Genet 2007;39:857–64.

19 Smyth DJ, Plagnol V, Walker NM *et al*. Shared and distinct genetic variants in type 1 diabetes and celiac disease. N Engl J Med 2008;359:2767–77.

20 Hafler JP, Maier LM, Cooper JD *et al*. CD226 Gly307Ser association with multiple autoimmune diseases. Genes Immun 2009;10:5–10.

21 Dardalhon V, Schubart AS, Reddy J *et al*. CD226 is specifically expressed on the surface of Th1 cells and regulates their expansion and effector functions. J Immunol 2005;175:1558–65.

22 Tao D, Shangwu L, Qun W *et al*. CD226 expression deficiency causes high sensitivity to apoptosis in NK T cells from patients with systemic lupus erythematosus. J Immunol 2005;174:1281–90.

23 Nath SK, Han S, Kim-Howard X *et al*. A nonsynonymous functional variant in integrin-alpha(M) (encoded by ITGAM) is associated with systemic lupus erythematosus. Nat Genet 2008;40:152–4.

24 Han S, Kim-Howard X, Deshmukh H *et al*. Evaluation of imputation-based association in and around the integrin-alpha-M (ITGAM) gene and replication of robust association between a non-synonymous functional variant within itgam and systemic lupus erythematosus (SLE). Hum Mol Genet 2009;18:1171–80.

25 Ikari K, Momohara S, Inoue E *et al*. Haplotype analysis revealed no association between the PTPN22 gene and

RA in a Japanese population. Rheumatology 2006;45: 1345–8.

26 Shin HD, Kim I, Choi CB, Lee SO, Lee HW, Bae SC. Different genetic effects of interferon regulatory factor 5 (IRF5) polymorphisms on systemic lupus erythematosus in a Korean population. J Rheumatol 2008;35:2148–51.

27 Ito I, Kawaguchi Y, Kawasaki A *et al*. Association of a functional polymorphism in the IRF5 region with systemic sclerosis in a Japanese population. Arthritis Rheum 2009; 60:1845–50.

28 Fernando MM, Stevens CR, Walsh EC *et al*. Defining the role of the MHC in autoimmunity: a review and pooled analysis. PLoS Genet 2008;4:e1000024.

29 Marsh MN. Gluten, major histocompatibility complex, and the small intestine. A molecular and immunobiologic approach to the spectrum of gluten sensitivity ('celiac sprue'). Gastroenterology 1992;102:330–54.

30 Arnett FC, Edworthy SM, Bloch DA *et al*. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988; 31:315–24.

31 Tan EM, Cohen AS, Fries JF *et al*. The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1982;25:1271–7.

32 Vitali C, Bombardieri S, Jonsson R *et al*. Classification criteria for Sjogren's syndrome: a revised version of the European criteria proposed by the American-European consensus group. Ann Rheum Dis 2002;61:554–8.

33 Delgado-Vega AM, Castiblanco J, Gomez LM, Diaz-Gallo LM, Rojas-Villarraga A, Anaya JM. BCL2 antagonist killer 1 (BAK1) polymorphisms influence the risk of developing autoimmune rheumatic diseases in women. Ann Rheum Dis, Advance Access published March 11, 2009, doi:10.1136/ard.2008.100818.

34 Anaya JM, Castiblanco J, Tobon GJ *et al*. Familial clustering of autoimmune diseases in patients with type 1 diabetes mellitus. J Autoimmun 2006;26:208–14.

35 Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1997;40: 1725.

36 Purcell S, Neale B, Todd-Brown K *et al*. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–75.

37 Nicodemus KK. Catmap: case-control and TDT meta-analysis package. BMC Bioinformatics 2008;9:130.

38 Douroudis K, Nemvalts V, Rajasalu T, Kisand K, Uibo R. The CD226 gene in susceptibility of type 1 diabetes. Tissue Antigens 2009;74:417–9.

39 Skol AD, Scott LJ, Abecasis GR, Boehnke M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. Nat Genet 2006;38:209–13.

40 The expanding genetic overlap between multiple sclerosis and type I diabetes. Genes Immun 2009;10:11–4.

41 Wieczorek S, Hoffjan S, Chan A *et al*. Novel association of the CD226 (DNAM-1) Gly307Ser polymorphism in Wegener's granulomatosis and confirmation for multiple sclerosis in German patients. Genes Immun 2009;10: 591–5.

42 Bravo ML, Valenzuela CY, Arcos-Burgos OM. Polymorphisms and phyletic relationships of the Paisa community from Antioquia (Colombia). Gene Geogr 1996; 10:11–7.

43 Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of sev-en common diseases and 3,000 shared controls. Nature 2007;447:661–78.

44 Jimenez-Morales S, Velazquez-Cruz R, Ramirez-Bello J *et al*. Tumor necrosis factor-alpha is a common genetic risk factor for asthma, juvenile rheumatoid arthritis, and systemic lupus erythematosus in a Mexican pediatric population. Hum Immunol 2009;70:251–6.