Chapter 4

The contribution of genetic risk factors other than the HLA shared epitope alleles to the genetic variance of rheumatoid arthritis

Diane van der Woude, Jeanine J. Houwing-Duistermaat, Ann Morgan, Jane Worthington, Tom W.J. Huizinga, Bart O. Roep, René E.M. Toes and René R.P. de Vries

Manuscript in preparation

ABSTRACT

Objective Genetic factors explain 66% of susceptibility to rheumatoid arthritis, but it is unclear how much of this genetic risk can be explained by the genetic risk factors known to date. In addition to the contribution of the HLA shared epitope (SE) alleles which we have quantified previously, we now investigated the contribution of the protective HLA-DRB1*13 alleles and of 12 replicated single nucleotide polymorphisms (SNPs), to the genetic variance of total, anti-citrullinated protein antibody (ACPA)-positive and ACPA-negative RA.

Methods A cohort of 148 twin pairs, of which at least one twin had RA, were tested for ACPA. The contribution of the HLA-DRB1 alleles and SNPs to the genetic variance was assessed by a logistic regression model with genotype-specific parameters and a random effect model representing the contribution of unobserved genetic factors.

Results In the ACPA-positive subset, the HLA SE alleles contributed 19.2% to the genetic variance, while the contribution of the protective HLA-DRB1*13 alleles was very small: 0.13%. Both SE and DRB1*13 alleles hardly contributed to ACPA-negative RA (1.52%). The contributions of the separate SNPs to the different subsets of RA were small with an average of below 1%.

Conclusion A modest amount of the genetic variance of ACPA-positive, but not ACPA-negative RA can be explained by the HLA SE alleles, while the contribution of the protective HLA-DRB1*13 alleles and non-HLA loci is very small. These results indicate that the largest part of the genetic contribution to RA remains unexplained.

INTRODUCTION

Rheumatoid arthritis (RA) affects approximately 0.5-1% of the population in western civilizations [1] and investigations of factors which determine a person's risk of developing RA have been ongoing since decades. Several risk factors for RA are now known, and these can be broadly classified as genetic or environmental risk factors. The relative importance of genetic and environmental risk factors can be investigated by estimating the heritability: the amount of disease variance which can be explained by genetic factors. The heritability of RA has consistently been reported to be approximately 66% [2, 3].

The hypotheses concerning how risk factors determine the development of RA have changed dramatically since the discovery of anti-citrullinated protein antibodies (ACPA) [4]. Many of the risk factors for RA have now been found to be specific for either ACPA-positive or ACPA-negative RA. For example, the genetic risk factors which confer the highest risk of disease, the human leucocyte antigen (HLA) DRB1 shared epitope (SE) alleles, have been described to specifically predispose to autoantibody-positive, and most prominently ACPA-positive, disease [5, 6]. This predilection for ACPA-positive RA has also been reported for the majority of the other genetic risk factors known to date [7]. It was therefore surprising to discover that the heritability of ACPA-positive and ACPA-negative disease does not differ, and is 66% for both subsets of disease, as we have recently described [3]. This suggests that many genetic risk factors for ACPA-negative RA remain to be discovered. Furthermore, the effect of the HLA SE alleles explained a relatively modest 18% of the genetic variance of ACPA-positive RA, raising the question which other genetic risk factors could account for the "missing" 82% of the heritability.

A possible answer to this question may lie in genes which confer protection from developing RA. In a recent meta-analysis of HLA-DRB1 associations in 2800 RA patients and 3000 controls, the presence of HLA-DRB1*13 alleles was associated with a markedly lower risk of ACPA-positive disease [8]. This protective effect was present in both SE-positive and SE-negative individuals, demonstrating that protection was not just due to the absence of predisposing effects, but was an independent phenomenon. Furthermore, the DRB1*13 alleles were the only alleles with a significant protective effect, and the protection which had in the past been ascribed to alleles harboring a DERAA or D70, could be explained solely by the protective effect of the DRB1*13 alleles. Within the DRB1*13 alleles, especially DRB1*1301 was associated with protection and in light of the strength of this protective effect (OR= 0.24 - 0.50), it seems plausible that these protective alleles may contribute considerably to the heritability.

Another factor which may play a role in explaining the remaining genetic variance of RA, is the effect of non-HLA genes associated with single nucleotide polymorphisms (SNPs). As a result of recent genome-wide association studies (GWAS), several new SNPs have been described to be associated with RA [9] and although the effects of these SNPs by

themselves tend to be limited, the combined effect of several SNPs may nonetheless be substantial with regard to the genetic variance of the disease.

For these reasons, we wished to investigate to what extent protective HLA effects, in the form of DRB1*13, and the effect of 12 well-replicated SNPs associated with RA, contribute to the genetic variance of RA. Using genotype frequencies reported in large case-control studies, we applied quantitative genetic methods to data from a relatively large twin cohort to assess the effect of these genetic risk factors for RA in general, as well as for ACPA-positive and ACPA-negative disease.

METHODS

Study population

A nationwide study of twin pairs of which one or both had RA was performed in the United Kingdom in 1989 [10]. Twin pairs were eligible for inclusion if at least one twin fulfilled the 1987 revised ACR criteria [11]. Blood samples were collected and zygosity status was verified by DNA fingerprinting on all same-sexed pairs using non-radioactive Southern blotting with a multilocus probe [12]. SNP genotyping was not available for this twin cohort.

ACPA (Anti-CCP2) assays

Total IgG anti-CCP2 was measured in serum samples by enzyme-linked immunosorbent assay (ELISA) (Immunoscan RA Mark 2; Eurodiagnostica, Arnhem, The Netherlands). Samples with a value above 25 units/ml were considered positive according to the manufacturer's instructions.

Statistical analysis

The contribution of the HLA genotypes and the 12 SNPs to the total genetic variance of RA, ACPA-positive and ACPA-negative RA was computed from the genotype frequencies available for Dutch and UK cohorts and the total genetic variance (τ^2) estimated previously [3].

When investigating the genetic contribution of DRB1*13, it is important to take both DBR1*13 and SE alleles into account because these alleles are mutually exclusive at the HLA-DRB1 locus. All six possible genotype combinations of DRB1*13 (denoted as DR13) and SE therefore need to be taken into consideration (SE/SE, SE/x, x/x, DR13/x, DR13/DR13 and SE/DR13, where x denotes genotypes which are neither SE nor DR13). In order to quantify the effect of SE and DRB1*13 alleles separately, the following dummy variables were defined:

X_1 to measure the effect of 1 SE allele (assuming the value 1 for genotypes SE/x and SE/DR13 and 0 for the other four genotypes),

X_2 measuring the effect of 2 SE alleles (value=1 for genotype SE/SE and 0 for the other five genotypes),

X_3 measuring the effect of 1 DRB1*13 allele (value=1 for genotypes DR13/x and SE/DR13 and 0 for the other four genotypes),

X_4 measuring the effect of 2 DRB1*13 alleles (value=1 for genotype DR13/DR13 and 0 for the other five genotypes).

We assumed the following logistic regression model for the relationship between the HLA genotypes and the risk for RA:

$$P(Y=1|\text{genotypes},u)=\frac{\exp(\alpha^* +\beta_1^* X_1 +\beta_2^* X_2 +\beta_3^* X_3 +\beta_4^* X_4 +u_r)}{1+\exp(\alpha^* +\beta_1^* X_1 +\beta_2^* X_2 +\beta_3^* X_3 +\beta_4^* X_4 +u_r)} \quad (1)$$

With Y = 1 for RA patients and Y=0 for controls, and u_r is a Gausian random effect with zero mean and variance τ^2_r which represents unmeasured genetic factors.

Previously we estimated the total genetic variance τ^2 by using model (1) without the genotypes (assuming $\beta 1, \beta 2, \beta 3, \beta 4 = 0$[3]). τ^2 equals the sum of the genetic variance explained by HLA (var(G)) and the genetic variance explained by other genetic risk factors (τ^2_r):

$$\tau^2 = \text{var(G)} + \tau_r^2 \quad (2)$$

The contribution of HLA (var(G)) to the total genetic variance τ^2 can thus be measured by

$$\frac{\text{var(G)}}{\text{var(G)}+\tau_r^2}=\frac{\text{var(G)}}{\tau^2}$$

For the genetic effect of HLA we considered the effect of the six genotypes on RA as described above. Var(G) can be expressed as

$$\text{var}(G)=\text{var}(\beta_1^* X_1 +\beta_2^* X_2 +\beta_3^* X_3 +\beta_4^* X_4)=$$
$$\beta_1^{*2}\,\text{var}(X_1)+\beta_2^{*2}\,\text{var}(X_2)+\beta_3^{*2}\,\text{var}(X_3)+\beta_4^{*2}\,\text{var}(X_4)+2\beta_1^*\beta_2^*\,\text{cov}(X_1X_2)+2\beta_1^*\beta_3^*\,\text{cov}(X_1X_3) \quad (3)$$
$$+2\beta_1^*\beta_4^*\,\text{cov}(X_1X_4)+2\beta_2^*\beta_3^*\,\text{cov}(X_2X_3)+2\beta_2^*\beta_4^*\,\text{cov}(X_2X_4)+2\beta_3^*\beta_4^*\,\text{cov}(X_3X_4)$$

The variance terms in this equation can be computed from the genotype frequencies. The β^* are the parameters of the random effects model (1) and they are related to the β of a standard logistic regression model without random effects (marginal model) by

$$\beta \approx (0.346\tau_r^2+1)^{-1/2}\beta^* \quad (4)$$

[13]. The β can be computed from the genotype frequencies in cases and controls and from the prevalence of RA.

Given β and τ^2 and using the two equations given by formula (2) and (4), we can compute $\tau^2_{r,}$ β* and hence var(G). Note that var(G) in formula (2) is a function of β* as given by formula (3).

The contribution of SE alleles was defined as

$$\text{var}(G) = \text{var}(\beta_1^* X_1 + \beta_2^* X_2) = \beta_1^{*2}\,\text{var}(X_1) + \beta_2^{*2}\,\text{var}(X_2) + 2\beta_1^*\beta_2^*\,\text{cov}(X_1 X_2) \quad (5)$$

Analogously the contribution of DRB1*13 alleles was defined as

$$\text{var}(G) = \text{var}(\beta_3^* X_3 + \beta_4^* X_4) = \beta_3^{*2}\,\text{var}(X_3) + \beta_4^{*2}\,\text{var}(X_4) + 2\beta_3^*\beta_4^*\,\text{cov}(X_3 X_4) \quad (6)$$

The calculation of the genetic contribution of the separate SNPs was performed in a similar manner with the logistic regression formula (1) containing only 1 β and 1 X. Now X assumed the value 1 for genotypes containing 1 or 2 minor alleles.

All computations were performed in the freely available software environment for statistical computing R.

RESULTS

Contribution of HLA-DRB1*13 and *1301 to the genetic variance

The study cohort consisted of 64 monozygotic and 84 dizygotic twin pairs from the United Kingdom (UK). The heritability modeling using twin pairs with at least one affected requires that the prevalence for each genotype be fixed to the specific population prevalence, in this case from the UK. Genotype data from 2914 RA patients, (1470 ACPA-positive, 723 ACPA-negative RA patients and 721 RA patients without ACPA-status) and 1348 controls were used to calculate the HLA population prevalences listed in Table 1 (unpublished data received from Ann Morgan). Since four-digit HLA typing was not available for British controls, the effect of the HLA-DRB1*1301 alleles, which have been shown to be associated with the strongest protective effect, could not be assessed. The analysis was therefore

Table 1. HLA population prevalences from the United Kingdom used for modeling.

Prevalence of	Population			
	All RA patients	ACPA-positive	ACPA-negative	Controls
ACPA	67%	100%	0%	1%
HLA-DRB1				
SE/SE	26.7%	32.6%	12.7%	8.6%
SE/x	44.0%	46.5%	38.5%	34.1%
x/x	18.2%	12.5%	31.5%	42.0%
*13/x	5.2%	3.0%	10.0%	9.6%
*13/*13	0.5%	0.3%	1.0%	0.7%
SE/*13	5.3%	5.0%	6.4%	5.0%

performed for the HLA-DRB1*13 alleles, which in meta-analysis were also associated with a significant protective effect.

Before performing the heritability analysis, we first investigated whether the protective effect of HLA-DRB1*13 on the risk of ACPA-positive RA was also present in the UK, based on the prevalences shown in Table 1. In order to correct for skewing due to the association between ACPA-positive RA and the SE alleles, the analysis was stratified for the SE alleles as described previously [8]. There was no effect of DRB1*13 on the development of ACPA-positive RA in the SE-negative stratum (OR: 1.08, 95% CI: 0.74-1.58), while in the SE-positive stratum there was a trend towards a protective effect (OR: 0.72, 95% CI: 0.50-1.04). As expected, there was no effect of DRB1*13 on the risk of developing ACPA-negative RA (OR: 1.15, 95% CI: 0.90-1.47).

Although the effect of HLA-DRB1*13 in the UK was smaller than observed in a large European meta-analysis, this did not exclude a contribution of the HLA-DRB1*13 alleles to the genetic variance of RA. Using the genotype frequencies from the UK listed in Table 1, we estimated the contribution of HLA-DRB1*13 to the genetic variance, while also taking into account the effect of the SE alleles. As shown in Table 2, the contribution of the HLA-DRB1*13 alleles to the genetic variance of total, ACPA-positive and ACPA-negative RA was very limited.

Table 2. Contribution of HLA SE alleles and HLA-DRB1*13 alleles to the genetic variance.

	SE	DRB1*13	Combined SE and DRB1*13
Total RA	12.9%	0.1%	12.8%
ACPA-positive RA	19.2%	0.1%	19.6%
ACPA-negative RA	1.4%	0.4%	1.5%

The fact that the combined contribution of SE and DRB1*13 is not equal to the sum of their separate contributions can be explained by the existence of covariance between the various genotypes, as described by the covariance term in formulas (4) and (5) in the Methods section. The figures for the contribution of the SE alleles differed slightly from the previously reported estimates [3], which is most likely due several factors: the statistical methods differed as described in the methods section, the current approach took both heterozygote and homozygote effects into account, used another reference category (no carriers of SE or DRB1*13 alleles) and was based on slightly different population prevalences.

The observation that the DRB1*13 alleles were not associated with a significant protective effect with regard to the risk of ACPA-positive RA in the UK, raised the question whether the contribution of the DRB1*13 alleles might have been larger in a population in which there was a strong effect of these alleles. To test this hypothesis, the contribution of DRB1*13 to the genetic variance of ACPA-positive RA was also assessed using population

prevalences from the Netherlands where there is a strong protective effect of DRB1*13. Genotype prevalences shown in Table 3 were extracted from the Leiden Early Arthritis Clinic (EAC) and the BeSt trial, which together consisted of 844 RA patients, (492 ACPA-positive and 352 ACPA-negative RA patients), as well as from a cohort of 1213 random healthy controls [8].

Table 3. Population prevalences from the Netherlands (ref. 8) used for modeling.

Prevalence of	Population			
	All RA patients	ACPA-positive	ACPA-negative	Controls
ACPA	58%	100%	0%	1%
HLA-DRB1				
SE/SE	18.7%	25.2%	7.6%	6.0%
SE/x	42.1%	46.8%	34.9%	30.8%
x/x	22.7%	14.9%	35.2%	36.2%
***13/x**	8.6%	6.3%	12.3%	17.5%
***13/*13**	0.6%	0%	1.5%	2.6%
SE/*13	7.4%	6.7%	8.5%	6.9%

This however did not result in a larger contribution of DRB1*13 to the genetic variance. As shown in Table 4, the DRB1*13 alleles contributed merely 1.46% to the genetic variance of ACPA-positive RA. The increase in the combined estimate of the contribution of SE and DRB1*13 over the two separate contributions is due to the positive covariance inherently present between a predisposing (SE) effect and protective (DRB1*13) effect. For ACPA-negative RA, the amount of explained genetic variance by the HLA alleles was very small and hardly increased by taking the effect of DRB1*13 into account. Since DRB1*1301 is the allele which is mainly responsible for the protective effect of DRB1*13, the analysis was repeated using the population prevalences of DRB1*1301 from the Netherlands. The resulting estimates were similar (data not shown), demonstrating that neither the DRB1*13 nor the DRB1*1301 alleles explain a large part of the genetic variance of RA.

Table 4. Contribution of HLA SE alleles and HLA-DRB1*13 alleles to the genetic variance assuming population prevalences from the Netherlands.

	SE	DRB1*13	Combined SE and DRB1*13
Total RA	8.9%	0.9%	11.1%
ACPA-positive RA	13.5%	1.5%	17.0%
ACPA-negative RA	0.5%	0.3%	0.9%

Contribution of SNPs to the genetic variance

Recent GWAS and meta-analyses of GWAS have revealed numerous SNPs to be associated with RA [7]. For the current analysis, we assessed the effect of 12 SNPs which have been

well-replicated, and of which data on the prevalence in the UK in RA patients and controls was available. The prevalence figures in the population of the United Kingdom which were used for the logistic regression model are listed in Table 5. If the prevalence figures were not directly reported in the literature, they were determined based on the reported odds ratio's and patient numbers for the different RA subsets. For some of the SNPs (e.g. TRAF1-C5, STAT4, IL2RB), the prevalence in the total RA patient population was not completely in line with the prevalence in the ACPA-positive and ACPA-negative subgroups. This is most likely due to the fact that the patient population used to determine the prevalence in all RA patients was much larger than the groups of ACPA-positive and ACPA-negative RA patients combined, thereby causing small discrepancies in the SNP prevalences.

Table 5. Minor allele frequencies reported in studies from the United Kingdom used for modeling.

Gene region of SNP (UK reference)	Population			
	All RA patients	ACPA-positive	ACPA-negative	Controls
PTPN22 (T allele) [23]	15.0%	16.7%	12.1%	10.3%
TRAF1-C5 [24]	45.6%	45.5%	45.0%	43.8%
STAT4 [24]	24.2%	24.7%	25.0%	21.7%
TNFAIP3-OLIG3 [25]	25.1%	26.6%	23.2%	21.5%
CTLA4 [26]	43%	-	-	43.8%
CD40 [27]	22.6%	21.5%	24.5%	25.2%
MMEL1-TNFRSF14 [17]	29.4%	-	-	33.2%
KIF5A [28]	34.2%	34.4%	33.9%	37.1%
PRKCQ [28]	17.1%	16.6%	17.7%	19.5%
CCL21 [27]	36.5%	37.0%	33.6%	34.8%
IL2RA [29]	24.4%	-	-	28.6%
IL2RB [28]	28.4%	29.2%	29.1%	26.1%

Table 6 shows the calculated contribution of the SNPs to the genetic variance. The contributions associated with the separate SNPs were small and, with the exception of PTPN22, were all well below 1%. If the effects of these SNPs were to be completely independent without interaction, then the addition of the contribution of the separate SNPs would give an impression of the total contribution of these SNPs. Assuming this to be the case, the 12 SNPs would together explain 3.4% of the genetic variance of total RA. Although this calculation is not possible for the ACPA-positive and ACPA-negative subset due to the lack of subset-specific prevalences for some of the SNPs, it appears unlikely, based on Table 3, that these estimates would exceed 5% of the explained genetic variance of total RA.

Table 6. Contribution of SNPs to the genetic variance.

Prevalence of	Population		
	All RA patients	ACPA-positive	ACPA-negative
PTPN22 (T allele)	1.1%	1.5%	0.2%
TRAF1-C5	0.1%	0.1%	0.0%
STAT4	0.2%	0.2%	0.3%
TNFAIP3-OLIG3	0.4%	0.7%	0.1%
CTLA4	0.0%	-	-
CD40	0.2%	0.4%	0.0%
MMEL1-TNFRSF14	0.4%	-	-
KIF5A	0.2%	0.1%	0.2%
PRKCQ	0.2%	0.3%	0.1%
CCL21	0.1%	0.1%	0.0%
IL2RA	0.5%	-	-
IL2RB	0.2%	0.2%	0.2%

DISCUSSION

In the present study, we investigated to what extent the effects of protective HLA alleles and non-HLA SNPs contribute to the, thus far largely unexplained, genetic variance of RA. The DRB1*13 and *1301 alleles had only a small contribution to ACPA-positive RA, and together with the HLA SE alleles explained 19.6% of the genetic variance of this subset of disease. Neither the HLA SE alleles nor the DRB1*13 alleles contributed substantially to the genetic variance of ACPA-negative RA, which was to be expected based on the fact that these alleles have not been found to confer a strong predisposing or protective effect for ACPA-negative RA [8]. The effects of 12 well-replicated SNPs on the genetic variance of RA were very modest. Most SNPs explained far less than 1% of the genetic variance, and assuming independent effects, all SNPs combined would explain less than 5%.

The HLA-DRB1*13 alleles were not significantly associated with protection from ACPA-positive RA in the UK, in contrast to the findings of a recent meta-analysis of case-control cohorts from Norway, Sweden, the Netherlands and Spain. While the allele frequency of DRB1*13 was similar in RA patients in the UK and the four populations contributing to the meta-analysis, the prevalence of these alleles in the UK controls used in the current report was 8%, compared to at least 14% in the controls from the other cohorts. The lack of effect in the UK may thus be due to a difference in population structure and distribution of HLA alleles in the general British population.

In light of the non-significant effect of the HLA-DRB1*13 alleles in the UK, it was not surprising that this did not result in a large contribution of DRB1*13 to the genetic vari-

ance. To investigate whether the DRB1*13 alleles might explain a larger part of the genetic variance of RA if they were associated with a stronger protective effect, the analysis was repeated in a hypothetical fashion assuming the Dutch prevalences for DRB1*13. Unexpectedly, this also did not result in a marked contribution of the HLA-DRB1*13 alleles to the genetic variance of ACPA-positive RA.

The most likely explanation for the limited contribution of HLA-DRB1*13 is the low prevalence of these alleles. Despite the fact that the HLA-DRB1*13 alleles are more prevalent in the Netherlands than in the UK, an allele frequency of 14.8% in the healthy controls used in this study still places a considerable limit on the maximally achievable effect. The explanatory value of a genetic risk factor is ultimately dependent on both the associated effect size and the prevalence in the general population. In case of the other type of genetic risk factor investigated in the current study: the SNPs, the contribution to the genetic variance is most likely constrained by the modest effect sizes with which they are associated.

It is remarkable that there is a historic estimate of 37% contribution of HLA genes to the overall inherited risk of developing RA, as described by Deighton et al [14]. Due to the haplotype sharing approach which was used in the study by Deighton, protective HLA effects were also taken into account and the resulting 37% reflects far more than the contribution of solely the SE alleles. A possible reason for the fact that this estimate is much larger than our figure could be that, besides the effect of SE and DRB1*1301 alleles, other effects may remain within the MHC region [15, 16], which were, perhaps unconsciously, also taken into account by the haplotype sharing approach.

The observation that 12 well-replicated SNPs had a relatively minor contribution to the genetic variance is in line with the results of a previous study which estimated that all non-MHC genetic variants (known at the end of 2008) would explain just 3.6% of the total disease variance [17]. These results justify the conclusion that SNPs associated with moderate odds ratio's probably have a very limited contribution to the genetic variance.

In order to calculate the contribution of the 12 SNPs to the total genetic variance, we used the genotypic specific population prevalences reported in the largest number of total RA, ACPA-positive and ACPA-negative patients. This resulted in slightly deviant estimates for STAT4 and IL2RB, since the reported prevalences for ACPA-positive and ACPA-negative RA patients were higher than for the total RA population. This inconsistency can be resolved as soon as data on larger cohorts become available, but the contribution of these SNPs to the genetic variance will nonetheless remain small.

The results of the current study improve our understanding of the relative importance of distinct genetic factors in developing RA. However, most of the genetic variance remains unexplained, and RA therefore still suffers from "missing heritability", as do many other diseases [18]. Where should one look for this missing heritability? It appears unlikely that other common variants (SNPs) with increasingly smaller odds ratio's will have a large con-

tribution, although one cannot exclude that the true causal genetic variants, which may be incompletely tagged by the current SNPs, may have a much stronger effect. Another possibility is that rare variants, with a low minor allele frequency precluding identification by the GWAS, could play an important role [19, 20]. Such rare variants could very well have considerable odds ratio's (> 2), and if in addition some of these variants were found to be inherited in a Mendelian fashion, a small number of families could thereby explain a large part of the genetic component of RA [21].

Structural genomic variation, such as copy number variants (CNV), may contribute to the heritability, although recent data indicate that the SNP genotyping arrays used for the GWAS have captured most of the variability due to this trait [22]. Another possibility is that interaction between different genetic factors may play an important role, although the latest reports indicate that there is no gene-gene interaction between the known genetic risk factors and the most recently identified SNPs [17]. Similarly, gene-environment interaction could be a contributing factor to the heritability, although it is unclear to what extent this would be reflected in the current heritability estimates. All of the factors listed above may play a part in determining an individual's risk for RA. Future studies including large numbers of patients and controls, and using increasingly advanced techniques to study genetic variation will be required to shed more light on the causes of missing heritability.

In conclusion, the current data show that the HLA-DRB1*13 alleles do not substantially contribute to the genetic variance of this RA subset. Similarly, the effect of 12 well-replicated SNPs on the genetic variance of RA is small. Despite a substantial contribution of HLA, the largest part of the genetic risk for both ACPA-positive and ACPA-negative RA remains unexplained.

REFERENCES

1. Silman AJ, Pearson JE. Epidemiology and genetics of rheumatoid arthritis. Arthritis Res 2002;4 Suppl 3:S265-S272.
2. Macgregor AJ, Snieder H, Rigby AS et al. Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. Arthritis Rheum 2000;43(1):30-37.
3. van der Woude D, Houwing-Duistermaat JJ, Toes RE et al. Quantitative heritability of anti-citrullinated protein antibody-positive and anti-citrullinated protein antibody-negative rheumatoid arthritis. Arthritis Rheum 2009;60(4):916-923.
4. Schellekens GA, de Jong BA, van den Hoogen FH et al. Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies. J Clin Invest 1998;101(1):273-281.
5. Huizinga TW, Amos CI, van der Helm-van Mil AH et al. Refining the complex rheumatoid arthritis phenotype based on specificity of the HLA-DRB1 shared epitope for antibodies to citrullinated proteins. Arthritis Rheum 2005;52(11):3433-3438.
6. Morgan AW, Haroon-Rashid L, Martin SG et al. The shared epitope hypothesis in rheumatoid arthritis: evaluation of alternative classification criteria in a large UK Caucasian cohort. Arthritis Rheum 2008;58(5):1275-1283.
7. Delgado-Vega A, Sanchez E, Lofgren S et al. Recent findings on genetics of systemic autoimmune diseases. Curr Opin Immunol 2010;22(6):698-705.
8. van der Woude D, Lie BA, Lundstrom E et al. Protection against anti-citrullinated protein antibody-positive rheumatoid arthritis is predominantly associated with HLA-DRB1*1301: a meta-analysis of HLA-DRB1 associations with anti-citrullinated protein antibody-positive and anti-citrullinated protein antibody-negative rheumatoid arthritis in four European populations. Arthritis Rheum 2010;62(5):1236-1245.
9. Raychaudhuri S, Thomson BP, Remmers EF et al. Genetic variants at CD28, PRDM1 and CD2/CD58 are associated with rheumatoid arthritis risk. Nat Genet 2009;41(12):1313-1318.
10. Silman AJ, Macgregor AJ, Thomson W et al. Twin concordance rates for rheumatoid arthritis: results from a nationwide study. Br J Rheumatol 1993;32(10):903-907.
11. 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(3):315-324.
12. Jeffreys AJ, Wilson V, Thein SL. Individual-specific 'fingerprints' of human DNA. Nature 1985;316(6023):76-79.
13. Diggle PJ, Liang K-Y, Zeger SL. Analysis of longitudinal data (Book). P. 142. 1994. Oxford Science Publications.
14. Deighton CM, Walker DJ, Griffiths ID et al. The contribution of HLA to rheumatoid arthritis. Clin Genet 1989;36(3):178-182.
15. Okada Y, Yamada R, Suzuki A et al. Contribution of a haplotype in the HLA region to anti-cyclic citrullinated peptide antibody positivity in rheumatoid arthritis, independently of HLA-DRB1. Arthritis Rheum 2009;60(12):3582-3590.
16. Rioux JD, Goyette P, Vyse TJ et al. Mapping of multiple susceptibility variants within the MHC region for 7 immune-mediated diseases. Proc Natl Acad Sci U S A 2009;106(44):18680-18685.
17. Raychaudhuri S, Remmers EF, Lee AT et al. Common variants at CD40 and other loci confer risk of rheumatoid arthritis. Nat Genet 2008;40(10):1216-1223.

18. Manolio TA, Collins FS, Cox NJ et al. Finding the missing heritability of complex diseases. Nature 2009;461(7265):747-753.
19. McCarthy MI, Hirschhorn JN. Genome-wide association studies: potential next steps on a genetic journey. Hum Mol Genet 2008;17(R2):R156-R165.
20. Pritchard JK. Are rare variants responsible for susceptibility to complex diseases? Am J Hum Genet 2001;69(1):124-137.
21. Casanova JL, Abel L. Revisiting Crohn's disease as a primary immunodeficiency of macrophages. J Exp Med 2009;206(9):1839-1843.
22. Conrad DF, Pinto D, Redon R et al. Origins and functional impact of copy number variation in the human genome. Nature 2010;464(7289):704-712.
23. Morgan AW, Thomson W, Martin SG et al. Reevaluation of the interaction between HLA-DRB1 shared epitope alleles, PTPN22, and smoking in determining susceptibility to autoantibody-positive and autoantibody-negative rheumatoid arthritis in a large UK Caucasian population. Arthritis Rheum 2009;60(9):2565-2576.
24. Barton A, Thomson W, Ke X et al. Re-evaluation of putative rheumatoid arthritis susceptibility genes in the post-genome wide association study era and hypothesis of a key pathway underlying susceptibility. Hum Mol Genet 2008;17(15):2274-2279.
25. Thomson W, Barton A, Ke X et al. Rheumatoid arthritis association at 6q23. Nat Genet 2007;39(12):1431-1433.
26. Barton A, Jury F, Eyre S et al. Haplotype analysis in simplex families and novel analytic approaches in a case-control cohort reveal no evidence of association of the CTLA-4 gene with rheumatoid arthritis. Arthritis Rheum 2004;50(3):748-752.
27. Orozco G, Eyre S, Hinks A et al. Association of CD40 with rheumatoid arthritis confirmed in a large UK case-control study. Ann Rheum Dis 2010;69(5):813-816.
28. Barton A, Thomson W, Ke X et al. Rheumatoid arthritis susceptibility loci at chromosomes 10p15, 12q13 and 22q13. Nat Genet 2008;40(10):1156-1159.
29. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447(7145):661-678.