

# Major histocompatibility complex (MHC) class II alleles, haplotypes and epitopes which confer susceptibility or protection in systemic sclerosis: analyses in 1300 Caucasian, African-American and Hispanic cases and 1000 controls

Frank C Arnett,<sup>1</sup> Pravitt Gourh,<sup>1</sup> Sanjay Shete,<sup>2</sup> Chul W Ahn,<sup>1</sup> Robert E Honey,<sup>1</sup> Sandeep K Agarwal,<sup>1</sup> Filemon K Tan,<sup>1</sup> Terry McNearney,<sup>3</sup> Michael Fischbach,<sup>4</sup> Marvin J Fritzler,<sup>5</sup> Maureen D Mayes,<sup>1</sup> John D Reveille<sup>1</sup>

► Additional data are published online only at <http://ard.bmj.com/content/vol69/issue5>

<sup>1</sup>Division of Rheumatology and Clinical Immunogenetics, Department of Internal Medicine, The University of Texas Health Science Center at Houston (UTH), Houston, Texas, USA

<sup>2</sup>Department of Epidemiology, MD Anderson Cancer Center, Houston, Texas, USA

<sup>3</sup>Division of Rheumatology, Department of Internal Medicine, The University of Texas Medical Branch (UTMB), Galveston, Texas, USA

<sup>4</sup>Division of Rheumatology, Department of Internal Medicine, The University of Texas at San Antonio Health Sciences Center (UT-SA), San Antonio, Texas, USA

<sup>5</sup>Department of Medicine, University of Calgary, Calgary, Alberta, Canada

## Correspondence to

Dr Frank C Arnett, Division of Rheumatology, The University of Texas Health Science Center at Houston (UTHSC-H), 6431 Fannin Street, MSB 5.270, Houston, TX 77030, USA; [Frank.C.Arnett@uth.tmc.edu](mailto:Frank.C.Arnett@uth.tmc.edu)

Accepted 27 June 2009  
Published Online First  
12 July 2009



This paper is freely available online under the BMJ Journals unlocked scheme, see <http://ard.bmj.com/info/unlocked.dtl>

## ABSTRACT

**Objective** To determine human leucocyte antigen-class II (HLA-class II) (DRB1, DQB1, DQA1 and DPB1) alleles, haplotypes and shared epitopes associated with scleroderma (systemic sclerosis (SSc)) and its subphenotypes in a large multi-ethnic US cohort by a case-control association study.

**Patients and methods** 1300 SSc cases (961 white, 178 black and 161 Hispanic subjects) characterised for clinical skin forms (limited vs diffuse), SSc-specific autoantibodies (anticentromere (ACA), anti-topoisomerase I (ATA), anti-RNA polymerase III (ARA), anti-U3 ribonucleoprotein (fibrillarin)) and others were studied using molecular genotyping. Statistical analyses in SSc itself by ethnicity, gender, skin type and autoantibodies were performed using exact logistic regression modelling for dominant, additive and recessive effects from HLA.

**Results** The strongest positive class II associations with SSc in white and Hispanic subjects were the *DRB1\*1104*, *DQA1\*0501*, *DQB1\*0301* haplotype and *DQB1* alleles encoding a non-leucine residue at position 26 (*DQB1 26 epi*), while the *DRB1\*0701*, *DQA1\*0201*, *DQB1\*0202* haplotype and *DRB1\*1501* haplotype were negatively correlated and possibly protective in dominant and recessive models, respectively. These associations did not discriminate between limited and diffuse SSc. SSc in black subjects was associated with *DRB1\*0804*, *DQA1\*0501*, *DQB1\*0301* alleles. *DPB1\*1301* showed the highest odds ratio for ATA (OR = 14). Moreover, it showed no linkage disequilibrium or gene interaction with DR/DQ. ACA was best explained by *DQB1\*0501* and *DQB1\*26 epi* alleles and ARA by *DRB1\*0404*, *DRB1\*11* and *DQB1\*03* alleles in white and Hispanic subjects but *DRB1\*08* in black subjects.

**Conclusion** These data indicate unique and multiple HLA-class II effects in SSc, especially on autoantibody markers of different subphenotypes.

## INTRODUCTION

Scleroderma (systemic sclerosis, SSc) is a chronic complex autoimmune disease characterised by (a) organ fibrosis involving the skin, lungs, gastrointestinal tract and/or heart; (b) a proliferative vasculopathy primarily affecting small blood vessels and capillaries; (c) immune activation with production of disease-specific autoantibodies.<sup>1–4</sup> The

disease is further classified into limited and diffuse forms based on extent and distribution of cutaneous thickening.

The most common SSc-specific autoantibodies are directed against centromeric proteins anticentromere (ACA) (CENP B and A), anti-topoisomerase I (ATA) (also termed Scl-70) and anti-RNA polymerase III (ARA); however, a variety of less common specificities, typically antinucleolar, can be found, which include anti-U3 ribonucleoprotein (fibrillarin) AFA, anti-Th/To, anti-Pm-Scl, anti-RNA polymerase I and anti-U1-ribonucleoprotein (RNP).<sup>5</sup> Importantly, each patient with SSc typically produces only one of these autoantibodies and each one currently serves as a biomarker for different patterns of skin and visceral involvement, as well as prognosis. In addition, certain SSc-specific antibodies occur in different frequencies among different ethnic groups.

Scleroderma is thought to be a complex genetic disease, influenced by multiple genes, with a substantial environmental or non-germline component based on twin studies.<sup>6</sup> African-American and Hispanic patients with scleroderma tend to have more severe disease than their Caucasian counterparts and disease in African-American subjects begins at an earlier age.<sup>7</sup> The Choctaw Indians of southeastern Oklahoma have a nearly 10-fold prevalence of the disease compared with other ethnicities.<sup>8</sup> It appears likely that there are different combinations of genes, the interacting effects of which influence disease susceptibility and severity. Recently, we have found that the same polymorphism in the *PTPN22* gene associated with rheumatoid arthritis, systemic lupus erythematosus and other autoimmune diseases is also associated with SSc, especially in those patients having ATA or ACA antibodies.<sup>9–11</sup> Additional 'autoimmunity' genes now reported to be associated with SSc include allograft inflammatory factor, *IL1A*, *IRF5*, *STAT4* and *FAS*.<sup>12–16</sup>

Major histocompatibility complex (MHC) or human leucocyte antigen-class II (HLA-class II) allelic associations with SSc have been reported in European and North American Caucasian subjects (*DRB1\*0301*, *DRB1\*11*, *DRB1\*07*) and Japanese and Koreans (*DRB1\*1502*) over the past two decades but have been relatively weak.<sup>5 17–22</sup>

Much stronger correlations, however, have been demonstrated between certain HLA-class II alleles and each of the SSc-specific autoantibodies.<sup>21–23–27</sup> Different HLA-*DRB1*, *DQB1*, *DQA1* and/or *DPB1* alleles, or combinations thereof, are associated with expression of ACA (*DQB1*\*0501 and other *DQB1* alleles encoding non-leucine residues at position 26 in the peptide binding groove),<sup>21–25–27</sup> ATA (*DRB1*\*11, especially the *DRB1*\*1104, *DQB1*\*0301 haplotype) in Caucasian subjects and *DRB1*\*1502, *DQB1*\*0601 in Japanese and Korean subjects,<sup>21–23</sup> AFA antibody (the *DRB1*\*1302, *DQB1*\*0604 haplotype),<sup>24</sup> and anti-PM-Scl (*DRB1*\*0301).<sup>26</sup> No consistent HLA correlations heretofore have been made with ARA.<sup>28–29–38–40</sup> Only a few studies have examined HLA-*DPB1* alleles in SSc; however, one such allele (*DPB1*\*1301) has been associated with ATA.<sup>30–32</sup> It has been unclear whether this allele is an independent disease correlate or the result of linkage disequilibrium (LD) with HLA-*DRB1* and *DQB1* haplotypes. Although genetic influences are thought to have an important role in susceptibility to SSc, genetic studies of scleroderma have included relatively small numbers of patients, especially black and Hispanic subjects. HLA associations with specific autoantibodies may be clinically relevant because each of the autoantibody subsets of scleroderma is associated with certain disease features and different prognostic implications.

Thus, the overall aims of this study were to determine specific HLA-class II alleles, haplotypes and epitopes influencing susceptibility to, and/or expression of, SSc itself, its limited or diffuse forms, or its various autoantibody subsets across ethnic lines in the largest cohort yet of American patients.

## METHODS

### Selection of SSc cases and normal controls

A case-control association study was performed (total SSc cases 1300; total controls 1000) along with subanalyses by ethnicity, gender, clinical subsets and specific autoantibody profiles. Patients were included from the NIH/NIAMS Scleroderma Family Registry and DNA Repository,<sup>33</sup> the Genetics versus Environment in Scleroderma Outcomes Study (GENISOS) cohort followed up in the NIH/NIAMS Center of Research Translation in Scleroderma,<sup>6</sup> and the UT-Houston Rheumatology Division. Patients with SSc included 961 Caucasian, 178 African-American and 161 Hispanic subjects; normal local control totals were 539 Caucasian, 263 African-American and 198 Hispanic subjects. Only Hispanic patients with SSc and controls of Mexican or Central American ancestry were included.

HLA-*DPB1* alleles were determined in 705 Caucasian patients with SSc and 287 Caucasian controls in the Scleroderma Registry. Only 82 African-American and Hispanic cases were studied for HLA-*DPB1* alleles but were added to the Caucasian subjects for an overall comparison. All patients fulfilled the preliminary American College of Rheumatology criteria for the diagnosis of SSc<sup>34</sup> or had three of the five clinical features of the CREST syndrome (calcinosis, Raynaud's phenomenon, esophageal dysfunction, sclerodactyly or telangiectasia). Patients having overlapping Sjögren's syndrome, myositis or features of rheumatoid arthritis or systemic lupus were not excluded. Results of HLA allele frequencies in some of these patients have been reported previously<sup>23–25–30</sup> and are included here. Because of the cross-sectional nature of two of these three cohorts, reliable clinical information on specific organ involvement, such as pulmonary fibrosis, pulmonary hypertension or renal crisis, was not available to assess HLA associations. HLA registry controls were primarily spouse or friend controls, and GENISOS/Division controls were healthy medical centre personnel or blood bank

donors from the local Houston area. All controls were screened for a history of any autoimmune diseases and excluded if positive. All study subjects provided written informed consent and the study was approved by the UTH Committee for the Protection of Human Subjects.

### Autoantibody identification

Antinuclear antibodies and ACA were determined in all patients with SSc by indirect immunofluorescence on HEp-2 cells (Antibodies, Davis, California, USA). Immunodiffusion against calf thymus extract was used to determine the presence of ATA (Scl-70), anti-Ro/SS-A, anti-La/SS-B, anti-Smith (Sm) and anti-RNP autoantibodies (Inova Diagnostics, San Diego, California, USA). ARA were determined using a commercially available enzyme-linked immunoassay (EIA) kit (MBL, Nagoya, Japan). AFA were determined only in those Division and GENISOS patients with SSc with a nucleolar antinuclear antibody pattern using immunoprecipitation.

### HLA-class II genotyping

HLA-*DQA1*, *-DQB1* and *-DPB1* alleles were oligotyped and *DRB1* alleles directly sequenced from extracted genomic DNA as previously described.

### Statistical analyses

We used statistical analysis software SAS 9.1.3 and the SAS Genetic Package (SAS Institute, Cary, North Carolina, USA).  $\chi^2$  Tests or Fisher's exact tests were used to compare HLA-class II allele frequencies between the normal control and SSc groups. Mantel-Haenszel tests were performed for the analysis of HLA-class II allelic frequencies to control for the confounding effects of ethnicity. In the online supplementary tables, we only reported the odds ratio (OR) and the corresponding 95% CI for p values <0.017, which accounts for the Bonferroni corrections by the number of comparisons for each HLA genotype (0.017 = 0.05/3). LD coefficients and the corresponding p values were computed to examine the allelic associations that occur between alleles at different loci using SAS Genetics software.

These  $\chi^2$  data are presented in the supplementary tables. We then performed exact logistic regression for each of the HLA variants using dominant (D), additive (A) and recessive (R) modelling. To account for the multiple comparisons, we used a false discovery rate approach.<sup>35</sup> We used an experimental threshold of  $\alpha = 0.05$  and accounted for 1000 potential multiple comparisons that we performed. Using these values for  $\alpha$  and number of tests performed, a point significance of 0.004 is considered statistically significant. Nonetheless, p values of <0.05 also are selectively shown for alleles of interest and those which have been associated with SSc or its subgroups in other reports. Use of the 0.004 as the threshold to declare significance leads to a rough false discovery rate of 0.025. For LD analyses we used the procedure *haplotype* in the SAS Genetic Package. In addition, we examined frequencies of the presence of non-leucine residues at position 26 in the DQ  $\beta$  chain based on our previous studies implicating this 'epitope' (*DQB1*\*26 *epi*) in susceptibility to the ACA in SSc.<sup>23–25</sup>

## RESULTS

### HLA-class II alleles in SSc

HLA-class II (*DRB1*, *DQA1* and *DQB1*) genotyping was completed in 1300 SSc cases and 1000 normal controls and HLA-*DPB1* in 705 white SSc cases and 287 white controls (tables 1–3 and supplementary tables 1 and 2).

For all white patients with SSc versus ethnically matched controls, a strong association was found with HLA-*DRB1\*11*, but specifically with *DRB1\*1104* and not with *DRB1\*1101*, *DQA1\*0501* and *DQB1\*0301* also were significantly increased. These three alleles constitute a haplotype with most *DRB1\*11* alleles including both *DRB1\*1101* and *DRB1\*1104*.<sup>36</sup> Given the highest odds ratio (OR = 2.48) being conferred by *DRB1\*1104* rather than *DQA1\*0501* (OR = 2.29) or *DQB1\*0301* (OR = 1.50), the primary associated allele appeared likely to be *DRB1\*1104* in a dominant model. Also, the shared *DQB1\*26* epitope (absence of leucine in position 26) was increased (OR = 1.59). HLA-*DRB1\*0404* also was significantly increased (OR = 2.33); however, the strongest HLA-class II association with SSc was with HLA-*DPB1\*1301* (OR = 3.18). Testing for LD between HLA-*DPB1\*1301* and *DRB1\*1104* showed an  $r^2$  value of 0.0001 and for *DQB1\*0301* 0.009, thus demonstrating no significant LD of this DR/DQ haplotype with *DPB1\*1301*.

The HLA-*DRB1\*0701*, *DQA1\*0201*, *DQB1\*0202* haplotype was negatively associated with SSc in an additive model (HLA-*DRB1\*0701*, *DQA1\*0201*) and in dominant model (HLA-*DQB1\*0202*) suggesting that it conferred a 'protective' effect (table 1). Similarly, the HLA-*DRB1\*1501*, *DQA1\*0102*, *DQB1\*0602* haplotypes were significantly decreased but in a recessive model.

**Table 1** HLA-class II allelic frequencies significantly associated with Caucasian patients with systemic sclerosis (SSc) compared with Caucasian normal controls using logistic regression

	SSc (%)	Controls (%)	Model (D/R/A)	p Value	OR	95% CI
*0404	4.3	1.9	D	0.003	2.33	1.3 to 4.1
*0701	8.0	14.2	A	<0.0001	0.39	0.1 to 1.1
					0.43	0.3 to 0.6
*11	13.5	8.2	D	0.0002	1.81	1.3 to 2.4
*1104	7.1	3.0	D	<0.0001	2.48	1.6 to 3.9
*1501	11.0	14.3	R	0.0003	0.22	0.1 to 0.5
DQA1†						
*0102	18.1	21.3	R	0.0002	0.36	0.2 to 0.6
*0201	7.6	14.4	A	<0.0001	0.20	0.1 to 0.6
					0.43	0.3 to 0.6
*0501	31.6	24.0	A	<0.0001	2.29	1.4 to 3.7
					1.66	1.3 to 2.2
DQB1†						
*0202	5.4	10.3	D	<0.0001	0.46	0.3 to 0.7
*0301	23.0	16.8	D	0.002	1.50	1.2 to 1.9
*0602	11.5	14.6	R	0.003	0.29	0.1 to 0.7
*DQB1	64.4	53.8	–	0.0003	1.59	1.2 to 2.0
26 epi						
DPB1‡						
*1301	3.7	1.4	D	0.004	3.18	1.4 to 7.3

†SSc cases (n = 961) versus normal controls (n = 539); ‡SSc cases (n = 705) versus normal controls (n = 287) typed for HLA-DPB1.

**Table 2** HLA-class II allelic frequencies significantly associated with African-American patients with systemic sclerosis (SSc) compared with normal African-American controls using logistic regression

	SSc (%)	Controls (%)	Model (D/R/A)	p Value	OR	95% CI
DRB1†						
*0804	10.2	3.6	D	0.0003	3.31	1.7 to 6.5
DQA1†						
*0501	34.2	24.1	A	0.003	3.55	1.6 to 8.1
					1.63	1.1 to 2.5
DQB1†						
*0301	33.0	21.9	D	<0.0001	2.44	1.6 to 3.7

†SSc cases (n = 178) versus normal controls (n = 263).

**Table 3** HLA-class II allelic frequencies significantly associated with Hispanic patients with systemic sclerosis (SSc) compared with normal Hispanic controls using logistic regression

	SSc (%)	Controls (%)	Model (D/R/A)	p Value	OR	95% CI
DRB1†						
*1104	8.8	2.0	D	0.0003	4.99	2.0 to 12.4
*0701	3.8	7.6	D	0.01	0.39	0.2 to 0.8
DQA1†						
*0501	29.9	21.7	R	0.002	4.09	1.6 to 10.4
*0201	3.8	7.8	D	0.01	0.39	0.2 to 0.8
DQB1†						
*0301	24.2	14.9	D	0.003	2.13	1.3 to 3.5
*DQB1	72.7	61.7	–	0.019	1.81	1.1 to 3.0
26 epi						

†SSc cases (n = 161) versus normal controls (n = 198).

In African-Americans, the HLA-*DRB1\*0804/DQA1\*0501/DQB1\*0301* haplotype was significantly increased, while *DRB1\*1104* was not associated with SSc (table 2).

Significant associations with SSc in Hispanic subjects were found for *DRB1\*1104* (OR = 4.99), *DQA1\*0501* (OR = 4.09) and *DQB1\*0301* (OR = 2.13), along with the *DQB1\*26* epi (OR = 1.81) (table 3). Again, *DRB1\*0701* and *DQA1\*0201* were decreased in frequency.

HLA-*DQA1\*0501* was significantly more frequent in male (101/165 or 61%) than in female (583/1135 or 51%) patients with SSc (p = 0.02, OR = 1.49, 95% CI 1.06 to 2.12), thus confirming one earlier report.

#### HLA alleles in limited versus diffuse SSc

HLA associations with both limited and diffuse forms of SSc largely mirrored the findings in the disease group as a whole, except in black patients with limited disease whose numbers were small (n = 54) (supplementary tables 3a and 4a by logistic regression and supplementary tables 3, 4 and 5 by  $\chi^2$ ).

#### HLA allele frequencies in autoantibody subsets of SSc

Frequencies and ethnic differences in SSc autoantibodies are shown in online supplementary table 5. ACA were found to be associated with several class II haplotypes in white subjects, including HLA-*DRB1\*0101*, *DQA1\*0101* and *DQB1\*0501*, *DRB1\*04*, especially *DRB1\*0404*, but weakly *DRB1\*0401* which carried *DQB1\*0301* alleles and weakly with *DRB1\*08* (*\*0801*) and *DQA1\*0401* (table 4, supplementary table 6). The strongest associated alleles were *DRB1\*0401* (OR = 8.98) in a recessive model, *DRB1\*0404* (OR = 4.18) in a dominant model, *DQB1\*0301* in a recessive model (OR = 3.41), *DQB1\*0501* (dominant OR = 2.56) and the *DQB1* 26 epitope (OR = 2.93). No HLA-*DPB1* alleles were associated with ACA.

No HLA alleles were associated with ACA in black subjects and only HLA-*DRB1\*0407* in Hispanic subjects.

ATA in white subjects were strongly associated with the HLA-*DRB1\*1104*, *DQA1\*0501*, *DQB1\*0301* haplotype (OR = 6.93) and even more so with HLA-*DPB1\*1301* (OR = 14.02), both showing dominant models (table 5) (supplementary table 7).

The *DRB1\*1104* association also was present in the black subjects and Hispanic subjects, as was the *DQB1* 26 epitope (table 5, supplementary table 7). In addition, HLA-*DRB1\*08* alleles, *DRB1\*0804* in black subjects (OR = 3.42) and *DRB1\*0802* in Hispanic subjects (OR = 1.91) also were increased in these groups showing both additive and dominant effects. There appeared to be no negative or 'protective' effect from the *DRB1\*0701* haplotype. HLA-*DPB1\*1301* was highly significantly associated with



**Table 4** HLA-class II allelic frequencies significantly associated with anticentromere antibody positive Caucasian and Hispanic patients with systemic sclerosis (SSc) compared with ethnically matched normal controls using logistic regression

	SSc (%)	Controls (%)	Model (D/R/A)	p Value	OR	95% CI
Caucasian†						
<b>DRB1</b>						
*01	21.0	12.5	D	<0.0001	2.19	1.5 to 3.2
*0101	16.6	8.2	D	<0.0001	2.72	1.8 to 4.1
*04	24.9	11.7	D	0.0004	1.90	1.3 to 2.7
*0401	11.0	8.3	R	0.03	8.98	1.4 to 57.2
*0404	5.5	1.9	D	<0.0001	4.18	2.0 to 8.6
*0701	2.9	14.2	A	<0.0001	0.10	0.01 to 0.8
					0.14	0.1 to 0.2
*08	6.1	3.4	D	0.04	1.90	1.1 to 3.4
*0801	4.8	2.0	D	0.03	3.17	1.5 to 6.6
*11	11.9	8.2	D	0.02	1.67	1.1 to 2.6
*1104	5.0	3.0		NS		
*1501	8.9	14.3	R	0.006	0.67	0.4 to 1.02
					0.16	0.03 to 0.7
<b>DQA1</b>						
*0101	23.1	13.4	D	<0.0001	2.57	1.8 to 3.7
*0102	14.4	21.3	R	0.001	0.25	0.1 to 0.6
*0201	3.0	14.4	A	<0.0001	0.07	0.01 to 0.6
					0.15	0.1 to 0.2
<b>DQB1</b>						
*0202	5.3	2.8	D	<0.0001	0.14	0.07 to 0.3
*0301	23.5	16.8	R	0.004	3.41	1.4 to 8.0
*0302	14.6	9.4	D	0.007	2.04	1.2 to 3.4
*0501	22.8	12.6	D	<0.0001	2.56	2.0 to 3.7
*DQB1 26	75.1	53.8	–	<0.0001	2.93	2.0 to 4.2
epi						
<b>DPB1</b>						
*1301	1.5	1.4		NS		
Hispanic‡						
<b>DRB1</b>						
*0407	26.0	6.8	D	0.0009	5.07	2.0 to 12.9

†SSc cases (n = 281) versus normal controls (n = 539) in Caucasian subjects; ‡SSc cases (n = 25) versus normal controls (n = 198) in Hispanic subjects; NS=not significant for any of the three models

ATA when all ethnic groups were combined ( $p < 0.0001$ , OR = 9.96) (supplementary table 7). Using logistic regression, there was no evidence of gene–gene interaction between *DPB1*\*1301 and *DRB1*\*1104 ( $p = 0.9863$ ) or between *DPB1*\*1301 and *DQB1*\*0301 ( $p = 0.9999$ ) in the ATA positive group or in the total SSc group.

ARA were found to be most strongly associated in white subjects with HLA-*DRB1*\*0404 (OR = 5.13) and *DRB1*\*11 alleles (OR = 1.55 for one copy, 6.78 for two copies), especially *DRB1*\*1104 in additive and dominant models, respectively (table 6) (supplementary table 8). In addition, a positive recessive effect was seen from *DQB1*\*03 alleles (OR = 2.38), especially *DQB1*\*0301 (OR = 1.50 for one copy, 3.77 for two copies). In black patients, the strongest associations were with *DRB1*\*08 alleles (OR = 3.92), primarily *DRB1*\*0804 (OR = 2.98), along with *DQA1*\*0501 (OR = 3.10 for one copy, 6.03 for two copies) in an additive model and *DQB1*\*0301 (OR = 3.60) showing a dominant effect. Among Hispanic subjects, *DRB1*\*11, *DQA1*\*0501 and the *DQB1* 26 epitope showed possible weak positive correlations, but the strongest association was with *DQB1*\*0301 (OR = 4.07).

#### Other autoantibodies and HLA

The HLA-*DRB1*\*1302, *DQB1*\*0604 haplotype was most strongly associated with AFA in white subjects ( $p = 0.0003$ , OR=6.87), along with the *DQB1* 26 epitope ( $p = 0.0009$ , OR=1.60), while in black subjects *DRB1*\*08 alleles ( $p = 0.0003$ , OR=5.76), especially *DQB1*\*0804 ( $p = 0.002$ , OR=5.70), showed the strongest associations (supplementary tables 9 and 9a).

**Table 5** HLA-class II allelic frequencies significantly associated with antitopoisomerase I antibody positive Caucasian, African-American and Hispanic patients with systemic sclerosis (SSc) compared with ethnically matched normal controls using logistic regression

	SSc (%)	Controls (%)	Model (D/R/A)	p Value	OR	95% CI
Caucasian†						
<b>DRB1</b>						
*0401	3.6	8.3	D	0.003	0.35	0.2 to 0.7
*1101	6.2	3.8	D	0.03	2.13	1.1 to 4.1
*1104	16.6	3.0	D	<0.0001	6.93	3.9 to 12.2
<b>DQA1</b>						
*0501	34.7	24.0	A	<0.0001	2.01	0.9 to 4.7
					2.80	1.8 to 4.2
<b>DQB1</b>						
*0202	5.2	10.3	D	0.006	0.44	0.2 to 0.8
*03	45.1	30.1	A	0.0004	3.30	1.8 to 6.2
					1.76	1.1 to 2.7
*0301	31.2	16.8	D	<0.0001	3.20	2.1 to 4.8
*DQB1	72.2	53.8	–	<0.0001	2.47	1.6 to 3.8
26 epi						
<b>DPB1</b>						
*1301	12.2	1.4	D	<0.0001	14.02	6.5 to 30.3
African-American‡						
<b>DRB1</b>						
*08	12.5	5.5	D	0.04	2.64	1.1 to 6.1
*0804	11.3	3.6	D	0.01	3.42	1.3 to 9.0
*11	28.8	13.4	D	0.0005	3.55	1.8 to 7.2
*1101	12.5	5.7	D	0.13	1.98	0.8 to 4.9
*1104	8.8	1.9	D	0.007	4.66	1.6 to 13.4
<b>DQB1</b>						
*0301	40.0	14.9	D	0.0002	3.90	1.9 to 8.1
*DQB1	87.5	63.0	–	0.003	3.91	1.5 to 10.4
26 epi						
Hispanic§						
<b>DRB1</b>						
*08	23.6	9.9	D	0.08	2.00	0.9 to 4.4
*0802	20.8	8.3	D	0.13	1.91	0.8 to 4.3
*1104	22.2	2.0	D	<0.0001	15.96	5.3 to 48.0
<b>DQA1</b>						
*0401	26.4	9.1	D	0.008	3.05	1.4 to 6.7
*0301	24.3	14.9	D	0.01	2.88	1.3 to 6.3
<b>DQB1</b>						
*0402	25.7	8.1	D	0.003	3.46	1.6 to 7.7
*DQB1	78.4	61.7	–	0.03	2.59	1.1 to 6.1
26 epi						

†SSc cases (n = 154) versus normal controls (n = 539) in Caucasian subjects; ‡SSc cases (n = 40) versus normal controls (n = 263) in African-American subjects; §SSc cases (n = 37) versus normal controls (n = 198) in Hispanic subjects.

Anti-RNP, anti-Smith (Sm), anti-Ro/SSA and anti-La/SSB antibodies showed no significant HLA associations.

#### DISCUSSION

new to this study was the finding that the African HLA-*DRB1*\*0804 allele, along with *DQA1*\*0501 and *DQB1*\*0301, showed the most significant associations in black subjects and recurred in those with diffuse disease, ATA, ARA and AFA.

Perhaps equally important was the susceptibility effect of other *DQB1* alleles (besides \*0301 and \*0501) encoding polar amino acids at position 26 of the DQ  $\beta$  chain (*DQB1*\*26 epi). Previously, we reported this shared epitope in the antigen binding cleft to be most important in the ACA autoimmune response<sup>25</sup>; however, this current larger study suggests that it occurs in the majority of patients with SSc regardless of ethnicity or autoantibody.

A second unreported and potentially 'protective' haplotype in white subjects, HLA-*DRB1*\*1501, *DQA1*\*0102, *DQB1*\*0602, was significantly decreased in a recessive model in Caucasian

**Table 6** HLA-class II allelic frequencies significantly associated with anti-RNA polymerase III positive antibody positive Caucasian, African-American and Hispanic patients with systemic sclerosis (SSc) compared with ethnically matched normal controls using logistic regression.

	SSc (%)	Controls (%)	Model (D/R/A)	p Value	OR	95% CI
<b>Caucasian†</b>						
<i>DRB1</i>						
*0404	7.5	1.9	D	<0.0001	5.13	2.4 to 11.0
*11	12.4	8.2	A	0.02	6.78	1.2 to 37.7
					1.55	0.9 to 2.6
*1104	6.1	3.0	D	0.02	2.34	1.2 to 4.6
<i>DQB1</i>						
*03	40.6	30.1	R	0.004	2.38	1.3 to 4.2
*0301	22.9	16.8	A	0.01	3.77	1.3 to 11.0
					1.50	1.002 to 2.26
<b>African-American‡</b>						
<i>DRB1</i>						
*08	13.6	5.5	D	0.003	3.92	1.7 to 9.2
*0804	9.1	3.6	D	0.05	2.98	1.04 to 8.5
*1104	3.0	1.9		NS		
<i>DQA1</i>						
*0501	42.2	24.1	A	0.003	6.03	1.5 to 23.8
					3.10	1.3 to 7.2
<i>DQB1</i>						
*0301	37.5	21.9	D	0.002	3.60	1.6 to 8.0
<b>Hispanic§</b>						
<i>DRB1</i>						
*11	12.5	5.6	D	0.10	2.17	2.0 to 11.6
*1104	3.8	2.0	D	0.005	10.50	2.3 to 46.8
<i>DQA1</i>						
*0501	28.8	21.7	R	0.047	3.65	1.1 to 12.6
<i>DQB1</i>						
*03	50.0	34.1	D	0.04	2.33	1.1 to 5.1
*0301	32.5	14.9	D	0.0002	4.07	2.0 to 8.5
*DQB1	78.0	61.7	–	0.04	2.39	1.1 to 5.4

26 epi

†SSc cases (n = 182) versus normal controls (n = 539) in Caucasian subjects; ‡SSc cases (n = 33) versus normal controls (n = 263) in African-American subjects; §SSc cases (n = 40) versus normal controls (n = 198) in Hispanic subjects; NS=not significant for any of the three models

subjects with SSc overall and in both limited and diffuse subgroups, as well as in those patients with ACA.

An HLA-*DPB1* allele, *DPB1*\*1301 in a dominant model, showed the highest odds ratios in SSc (OR = 3.18) and in both the limited (OR = 4.20) and diffuse (OR = 4.38) forms in Caucasian subjects; however, these associations were completely explained by its strong correlation with ATA (OR = 14.02). Testing for LD of this *DPB1* allele with the *DRB1*\*1104 and *DQB1*\*0301 alleles showed that these were independently associated class II effects. Interestingly, the ATA response is strongly associated with pulmonary fibrosis in SSc and certain HLA-DP alleles have been clearly shown to promote susceptibility to occupationally acquired berylliosis, another fibrosing lung disease.<sup>40</sup>

Finally, ARA characteristically are markers of rapidly progressive skin thickening and renal crises and this study is the first to demonstrate that they occur in similar prevalences in each

of these ethnic groups and are associated with certain class II MHC alleles.

In previous studies of HLA and ARA, there have been contradictory findings, perhaps because small samples have been studied.<sup>28 38 39</sup>

**Funding** NIH.

**Competing interests** None.

**Ethics approval** This study was conducted with the approval of the University of Texas Internal Review Board (IRB).

**Provenance and peer review** Not commissioned; externally peer reviewed.

## REFERENCES

- Charles C, Clements P, Furst DE. Systemic sclerosis: hypothesis-driven treatment strategies. *Lancet* 2006;**367**:1683–91.
- Jimenez SA, Derk CT. Following the molecular pathways toward an understanding of the pathogenesis of systemic sclerosis. *Ann Intern Med* 2004;**140**:37–50.
- LeRoy EC, Black C, Fleischmajer R, et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. *J Rheumatol* 1988;**15**:202–5.
- Mayes MD. Scleroderma epidemiology. *Rheum Dis Clin North Am* 2003;**29**:239–54.
- Arnett FC. HLA and autoimmunity in scleroderma (systemic sclerosis). *Int Rev Immunol* 1995;**12**:107–28.
- Feghali-Bostwick C, Medsger TA Jr, Wright TM. Analysis of systemic sclerosis in twins reveals low concordance for disease and high concordance for the presence of antinuclear antibodies. *Arthritis Rheum* 2003;**48**:1956–63.
- Reveille JD, Fischbach M, McNamee T, et al. Systemic sclerosis in 3 US ethnic groups: a comparison of clinical, sociodemographic, serologic, and immunogenetic determinants. *Semin Arthritis Rheum* 2001;**30**:332–46.
- Arnett FC, Howard RF, Tan F, et al. Increased prevalence of systemic sclerosis in a Native American tribe in Oklahoma. Association with an Amerindian HLA haplotype. *Arthritis Rheum* 1996;**39**:1362–70.
- Gourh P, Tan FK, Assassi S, et al. Association of the PTPN22 R620W polymorphism with anti-topoisomerase I- and anticentromere antibody-positive systemic sclerosis. *Arthritis Rheum* 2006;**54**:3945–53.
- Orozco G, Sánchez E, González-Gay MA, et al. Association of a functional single-nucleotide polymorphism of PTPN22, encoding lymphoid protein phosphatase, with rheumatoid arthritis and systemic lupus erythematosus. *Arthritis Rheum* 2005;**52**:219–24.
- Plenge RM, Padyukov L, Remmers EF, et al. Replication of putative candidate-gene associations with rheumatoid arthritis in >4,000 samples from North America and Sweden: association of susceptibility with PTPN22, CTLA4, and PADI4. *Am J Hum Genet* 2005;**77**:1044–60.
- Alkassab F, Gourh P, Tan FK, et al. An allograft inflammatory factor 1 (AIF1) single nucleotide polymorphism (SNP) is associated with antinuclear antibody positive systemic sclerosis. *Rheumatology (Oxford)* 2007;**46**:1248–51.
- Dieudé P, Guedj M, Wipff J, et al. Association between the IRF5 rs2004640 functional polymorphism and systemic sclerosis: a new perspective for pulmonary fibrosis. *Arthritis Rheum* 2009;**60**:225–33.
- Kawaguchi Y, Tochimoto A, Ichikawa N, et al. Association of IL1A gene polymorphisms with susceptibility to and severity of systemic sclerosis in the Japanese population. *Arthritis Rheum* 2003;**48**:186–92.
- Rueda B, Broen J, Simeon C, et al. The STAT4 gene influences the genetic predisposition to systemic sclerosis phenotype. *Hum Mol Genet* 2009;**18**:2071–7.
- Liakouli V, Manetti M, Pacini A, et al. The -670G>A polymorphism in the FAS gene promoter region influences the susceptibility to systemic sclerosis. *Ann Rheum Dis* 2009;**68**:584–90.
- Gladman DD, Keystone EC, Baron M, et al. Increased frequency of HLA-DR5 in scleroderma. *Arthritis Rheum* 1981;**24**:854–6.
- Briggs D, Welsh KI. Major histocompatibility complex class II genes and systemic sclerosis. *Ann Rheum Dis* 1991;**50**(Suppl 4):862–5.
- Gladman DD, Kung TN, Siannis F, et al. HLA markers for susceptibility and expression in scleroderma. *J Rheumatol* 2005;**32**:1481–7.
- Kuwana M, Inoko H, Kameda H, et al. Association of human leukocyte antigen class II genes with autoantibody profiles, but not with disease susceptibility in Japanese patients with systemic sclerosis. *Intern Med* 1999;**38**:336–44.
- Kuwana M, Okano Y, Kaburaki J, et al. HLA class II genes associated with antinuclear antibody in Japanese patients with systemic sclerosis (scleroderma). *Ann Rheum Dis* 1995;**54**:983–7.
- Kang SH, Park MH, Song EY, et al. Association of HLA class II genes with systemic sclerosis in Koreans. *J Rheumatol* 2001;**28**:1577–83.
- Reveille JD, Durban E, MacLeod-St Clair MJ, et al. Association of amino acid sequences in the HLA-DQB1 first domain with antitopoisomerase I autoantibody response in scleroderma (progressive systemic sclerosis). *J Clin Invest* 1992;**90**:973–80.

24. **Arnett FC**, Reveille JD, Goldstein R, *et al*. Autoantibodies to fibrillarin in systemic sclerosis (scleroderma). An immunogenetic, serologic, and clinical analysis. *Arthritis Rheum* 1996;**39**:1151–60.
25. **Reveille JD**, Owerbach D, Goldstein R, *et al*. Association of polar amino acids at position 26 of the HLA-DQB1 first domain with the antientromere autoantibody response in systemic sclerosis (scleroderma). *J Clin Invest* 1992;**89**:1208–13.
26. **Marguerie C**, Bunn CC, Copier J, *et al*. The clinical and immunogenetic features of patients with autoantibodies to the nucleolar antigen PM-Scl. *Medicine (Baltimore)* 1992;**71**:327–36.
27. **Morel PA**, Chang HJ, Wilson JW, *et al*. Severe systemic sclerosis with anti-topoisomerase I antibodies is associated with an HLA-DRw11 allele. *Hum Immunol* 1994;**40**:101–10.
28. **Falkner D**, Wilson J, Fertig N, *et al*. Studies of HLA-DR and DQ alleles in systemic sclerosis patients with autoantibodies to RNA polymerases and U3-RNP (fibrillarin). *J Rheumatol* 2000;**27**:1196–202.
29. **Kuwana M**, Kaburaki J, Arnett FC, *et al*. Influence of ethnic background on clinical and serologic features in patients with systemic sclerosis and anti-DNA topoisomerase I antibody. *Arthritis Rheum* 1999;**42**:465–74.
30. **Reveille JD**, Brady J, MacLeod-St Clair M, *et al*. HLA-DPB1 alleles and autoantibody subsets in systemic lupus erythematosus, Sjögren's syndrome and progressive systemic sclerosis: a question of disease relevance. *Tissue Antigens* 1992;**40**:45–8.
31. **Gilchrist FC**, Bunn C, Foley PJ, *et al*. Class II HLA associations with autoantibodies in scleroderma: a highly significant role for HLA-DP. *Genes Immun* 2001;**2**:76–81.
32. **Stephens CO**, Briggs DC, Vaughan RW, *et al*. The HLA-DP locus in systemic sclerosis—no primary association. *Tissue Antigens* 1993;**42**:144–5.
33. **Mayes MD**. The establishment and utility of a population-based registry to understand the epidemiology of systemic sclerosis. *Curr Rheumatol Rep* 2000;**2**:512–16.
34. **Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee**. Preliminary criteria for the classification of systemic sclerosis (scleroderma). *Arthritis Rheum* 1980;**23**:581–90.
35. **Benjamini Y**, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Stat Soc* 1995;**57**:289–300.
36. **Schreuder GM**, Hurley CK, Marsh SG, *et al*. HLA dictionary 2004: summary of HLA-A, -B, -C, -DRB1/3/4/5, -DQB1 alleles and their association with serologically defined HLA-A, -B, -C, -DR, and -DQ antigens. *Hum Immunol* 2005;**66**:170–210.
37. **Lambert NC**, Distler O, Müller-Ladner U, *et al*. HLA-DQA1\*0501 is associated with diffuse systemic sclerosis in Caucasian men. *Arthritis Rheum* 2000;**43**:2005–10.
38. **Fanning GC**, Welsh KI, Bunn C, *et al*. HLA associations in three mutually exclusive autoantibody subgroups in UK systemic sclerosis patients. *Br J Rheumatol* 1998;**37**:201–7.
39. **Kuwana M**, Pandey JP, Silver RM, *et al*. HLA class II alleles in systemic sclerosis patients with anti-RNA polymerase I/III antibody: associations with subunit reactivities. *J Rheumatol* 2003;**30**:2392–7.
40. **Saltini C**, Amicosante M, Franchi A, *et al*. Immunogenetic basis of environmental lung disease: lessons from the berylliosis model. *Eur Respir J* 1998;**12**:1463–75.