Immunogenetic Risk and Protective Factors for the Idiopathic Inflammatory Myopathies

Distinct HLA-A, -B, -Cw, -DRB1, and -DQA1 Allelic Profiles Distinguish European American Patients With Different Myositis Autoantibodies

Terrance P. O'Hanlon, PhD, Danielle Mercatante Carrick, PhD, Ira N. Targoff, MD, Frank C. Arnett, MD, John D. Reveille, MD, Mary Carrington, PhD, Xiaojiang Gao, PhD, Chester V. Oddis, MD, Penelope A. Morel, MD, James D. Malley, PhD, Karen Malley, BA, Ejaz A. Shamim, MD, Lisa G. Rider, MD, Stephen J. Chanock, MD, Charles B. Foster, MD, Thomas Bunch, MD, Perry J. Blackshear, MD, DPhil, Paul H. Plotz, MD, Lori A. Love, MD, PhD, and Frederick W. Miller, MD, PhD

Abstract: The idiopathic inflammatory myopathies (IIM) are systemic connective tissue diseases defined by chronic muscle inflammation and weakness associated with autoimmunity. We have performed low to high resolution molecular typing to assess the genetic variability of major histocompatibility complex loci (HLA-A, -B, -Cw, -DRB1, and -DQA1) in a large population of European American patients with IIM (n = 571) representing the major myositis autoantibody groups. We established that alleles of the 8.1 ancestral haplotype (8.1 AH) are important risk factors for the development of IIM in patients producing anti-synthetase/anti-Jo-1, -La, -PM/Scl, and -Ro autoantibodies. Moreover, a random forests classification analysis suggested that 8.1 AH-associated alleles B*0801 and DRB1*0301 are the principal HLA risk markers. In addition, we have identified several novel HLA susceptibility factors associated distinctively with particular myositis-specific (MSA) and myositis-associated autoantibody (MAA) groups of the IIM. IIM patients with anti-PL-7 (anti-threonyl-tRNA synthetase) autoantibodies have a unique HLA Class I risk allele, Cw*0304 (pcorr = 0.046), and lack the 8.1 AH markers associated with other anti-synthetase autoantibodies (for example, anti-Jo-1 and anti-PL-12). In addition, HLA-B*5001 and DQA1*0104 are novel potential risk factors among anti-signal recognition particle autoantibody-positive IIM patients (pcorr = 0.024 and p = 0.010, respectively). Among those patients with MAA, HLA DRB1*11 and DQA1*06 alleles were identified as risk factors for myositis patients with anti-Ku (pcorr = 0.041) and anti-La (pcorr = 0.023) autoantibodies, respectively. Amino acid sequence analysis of the HLA DRB1 third hypervariable region identified a consensus motif, ^{70}D (hydrophilic)/^{71}R (basic)/^{74}A (hydrophobic), conferring protection among patients producing anti-synthetase/anti-Jo-1 and -PM/Scl autoantibodies. Together, these data demonstrate that HLA signatures, comprising both risk and protective alleles or motifs, distinguish IIM patients with different myositis autoantibodies and may have diagnostic and pathogenic implications. Variations in associated polymorphisms for these immune response genes may reflect divergent pathogenic mechanisms and/or responses to unique environmental triggers in different groups of subjects resulting in the heterogeneous syndromes of the IIM.

(*Medicine* 2006;85:111–127)

From National Institute of Environmental Health Sciences (TPO, DMC, EAS, LGR, PJB, FWM), Center for Information Technology (JDM), National Cancer Institute (SJC), and National Institute of Arthritis and Musculoskeletal Disease (PHP) National Institutes of Health, Department of Health and Human Services, Bethesda, Maryland; Veterans Affairs Medical Center (IRT), University of Oklahoma Health Sciences Center, and Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma; University of Texas-Houston Health Science Center (FCA, JDR), Houston, Texas; Basic Research Program (MC, XG), SAIC Frederick National Cancer Institute, Frederick, Maryland; University of Pittsburgh School of Medicine (CVO, PAM), Pittsburgh, Pennsylvania; Malley Research Programming Inc (KM), Rockville, Maryland; Department of Pediatrics (CBF), Johns Hopkins University School of Medicine, Baltimore, Maryland; Mayo Clinic (TB), Rochester, Minnesota; and United States Food and Drug Administration (LAL), Rockville, Maryland.

This work was supported in part by the Center for Biologics Evaluation and Research and the Office of Woman's Health in the United States FDA, and by the Intramural Research programs of the National Institutes of Health, National Institute of Environmental Health Sciences, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Cancer Institute, Center for Cancer Research. The funding sponsors had no role in the design, conduct, or interpretation of the study. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does the mention of trade names, commercial products, or organizations imply endorsement by the United States Government.

Address reprint requests to: Terrance P. O'Hanlon, PhD, National Institute of Environmental Health Sciences, NIH, 9 Memorial Drive, Room 1W101, MSC 0958, Bethesda, MD 20892. Fax: 301-480-4127; e-mail: ohanlont@niehs.nih.gov.

ISSN: 0025-7974/06/8502-0111
DOI: 10.1097/01.md.0000217525.82287.eb

Abbreviations: AH = ancestral haplotype, DM = dermatomyositis, EA = European Americans, HVR3 = third hypervariable region, IBM = inclusion body myositis, IIM = idiopathic inflammatory myopathies, MAA = myositis-associated autoantibodies, MHC = major histocompatibility complex, MSA = myositis-specific autoantibodies, PM = polymyositis, RF = random forests, RSP = restrictive supertype patterns, SRP = signal recognition particle.

INTRODUCTION

The idiopathic inflammatory myopathies (IIM) are a group of heterogeneous systemic autoimmune diseases with a primary feature of muscle inflammation of unknown cause resulting in chronic weakness[30]. These systemic autoimmune syndromes present clinically with symmetric, proximal muscle weakness; elevated muscle enzymes in serum; myopathic changes on electromyography; inflammatory muscle biopsies; and characteristic skin rashes in some cases. Therapy consists of immunosuppressive agents to reduce inflammation and rehabilitation to strengthen weakened muscles.

The IIM syndromes can be divided into multiple clinicopathologic and serologic groups based on distinct clinical signs and symptoms or on the presence of disease-specific autoantibodies[30]. The 2 major clinicopathologic groups of IIM, dermatomyositis (DM) and polymyositis (PM) are distinguished clinically by the presence of photosensitive, pathognomonic rashes in DM. Inclusion body myositis (IBM) shares many of the features of PM but the diagnosis is based on unique histopathologic findings of inclusions in myofibers with surrounding inflammation. Autoantibody subgroups based on the presence of myositis-specific (MSA) and myositis-associated autoantibodies (MAA) have been associated often with different epidemiologic, clinical, prognostic, and immunogenetic features[27].

Autoimmune diseases likely result after chronic immune activation in genetically susceptible individuals following specific environmental exposures. This concept is supported in myositis by familial clustering, immunogenetic associations with IIM, temporal associations of disease onset with drugs and other environmental agents in certain individuals, and seasonal and geographic clustering of myositis onset[39]. Certain polymorphic immune response genes have been associated with myositis; most notable among these are genes of the major histocompatibility complex (MHC) that play an important role in adaptive immunity. Genes encoding human MHC Class I (HLA-A, -B, -Cw) and Class II (HLA-DR, -DQ, -DP) molecules are the most polymorphic in the human genome and among the strongest and most consistently identified genetic factors associated with the development of human autoimmune diseases including IIM[1,9,25,27,33,51].

We and others have documented the association of the northwestern European 8.1 ancestral haplotype (8.1 AH) (containing HLA-A*0101; B*0801; Cw*0701; DRB1*0301; DQA1*0501) with the development of IIM in small cohorts of European American (EA) patients[1,23,25,27,40,43,44]. In addition, we have more recently identified several novel HLA factors associated distinctly with 1 or more clinicopathologic groups of IIM patients[34]. In the present study, we have examined the allelic variability of HLA-A, -B, -Cw, -DRB1, and -DQA1 determinants in a large population of EA myositis patients (n = 571) to assess genetic susceptibility in different autoantibody groups (MSA and/or MAA) of the IIM.

PATIENTS AND METHODS

Study Subjects

EA adult-onset myositis cases and unrelated healthy controls were identified for this study from subjects referred to protocols involving the pathogenesis and treatment of myositis at the National Institutes of Heath Warren Grant Magnuson Clinical Center and the United States Food and Drug Administration between 1983 and 2002 (411 IIM cases and 377 controls). Additional case and control data from this period were provided from collaborators at the University of Texas-Houston Health Science Center, Houston, TX (86 IIM cases and 208 controls); Mayo Clinic, Rochester, MN (41 IIM cases); and the University of Pittsburgh Medical Center, Pittsburgh, PA (65 IIM cases and 196 controls)[22,31]. All subjects were enrolled in investigational review board-approved clinical protocols. Patients were defined as those meeting probable or definite PM or DM[3] or IBM[27] and required the exclusion of inherited, metabolic, or infectious myopathies and other causes of muscle disease. The myositis overlap group was defined when patients met the criteria above and also criteria for another defined connective tissue disease; cancer-associated myositis was defined when cancer was diagnosed within 2years of myositis. Some of these data have been reported in previous studies[1,27].

Laboratory Procedures

Purified genomic DNA was utilized for low to high resolution MHC Class I (HLA-A, -B, -Cw) and Class II (HLA-DRB1 and -DQA1) typing of all presently identified alleles using a combination of laboratory-designed and commercial reagents for PCR-mediated sequence-specific oligonucleotide probe hybridization and sequence-specific priming techniques according to manufacturers' recommendations when applicable (Genovison, West Chester, PA and Dynal Biotech, Lafayette Hill, PA). Complete high-resolution genotyping at the HLA-A, B, Cw, DRB1, and DQA1 loci was obtained based on the availability of sufficient high-quality genomic DNA; additional genotyping results were available from the referral centers. Allele frequencies were determined by the number of allele-positive subjects divided by the total number of subjects for which complete low- or high-resolution HLA data were available at a given locus. HLA allele assignments were consistent with the World Health Organization Nomenclature Committee for Factors of the HLA System (13th International Histocompatibility Workshop, Victoria, Canada).

Hyperdiversity region sequence motifs shared among DRB1 alleles were screened by comparing between cases and controls the combined frequency of all alleles possessing the putative peptide-binding motif. HLA DRB1 restrictive supertype patterns (RSP) were defined as follows: RSP ''A'' (DRB1*0101, *0102, *0401, *0404, *0405, *0408, *1001, *1402); RSP ''R'' (DRB1*0301, *0302); RSP ''E'' (DRB1*0403, *0406, *0407, *0901, *1401); RSP ''Q'' (DRB1*0701); RSP ''Dr'' (DRB1*08, *1101, *1104, *1105, *1106, *12, *1303, *16); RSP ''a'' (DRB1*1501,

*1502); RSP ''De'' (DRB1*0103, *0402, *1102, *1103, *1301, *1302, *1304)[15].

Myositis-specific (anti-synthetase, anti-signal recognition particle [SRP], anti-Mi-2) and myositis-associated (anti-Ku, -La, -Ro, -RNP, -PM/Scl) autoantibodies were identified in serum samples using previously validated methods of protein and RNA immunoprecipitation and double immunodiffusion[49,50].

Statistical Analyses

Analyses were performed using the SAS System for Windows, version 8.02 (SAS Institute, Cary, NC). The Fisher exact test was used to calculate p values for 2×2 tables. To correct for multiple comparisons, the sequential Holm procedure was applied[20]. We defined p values as significant when the adjusted p values were at or below the 0.05 level. The size of the family (k value) over which the Holm procedure was applied varied by the number of testable factors in each group of comparisons. The relative importance that individual HLA alleles confer upon genetic predisposition to disease were estimated using a statistical learning machine[12] using the random forests algorithm developed by Leo Breiman and Adele Cutler (http://stat-www.berkeley.edu/users/breiman/RandomForests/). Random forests (RF) is a prediction and classification tool that generates rank estimates of variable importance and approximates case proximities within clusters. Briefly, open source RF code adapted for the R programming language by Drs. Andy Liaw and Matthew Weiner (Merck Biometrics Research Laboratories) was used for all analyses[45]. All alleles comprising complete high resolution typing data for either HLA Class I (A, B, and Cw) or Class II (DRB1 and DQA1) loci among cases defining different IIM serogroups (MSA or MAA) and controls were classified using RF models each containing 500 independent classification trees. Individual decision trees were constructed from combined, unmatched case and control training datasets using bootstrap sampling with replacement (equal size samples from cases and controls) and random variable selection. RF uses a majority vote across the separate trees and classification was performed on test cases and controls left out of the modeling dataset from each of the respective decision trees. In this fashion, training and test data are randomly re-utilized in the construction of individual decision trees with an ''out-of-bag'' estimate of error rates. All allelic variables in the test population were ranked by their relative importance in discriminating case and control test subjects. We also performed traditional logistic regression analyses as an independent means of corroborating our RF modeling. Thus, to test the hypothesis that specific HLA alleles are associated with IIM, we used SAS E-Guide 2.0 to fit a main effects logistic regression model with controls as the modeled outcome. The global test for the null hypothesis resulted in a likelihood ratio p value of <0.0001 for both HLA Class I and Class II factor analyses (c-statistic = 0.863 and 0.857, respectively), indicating that the models were a good fit and had reasonable predictive power[26]. The estimated logistic regression model for the HLA Class I analysis is the following: $\text{logit}(\pi) = -1.41+(-0.56\_A*0101)+(3.98\_B*0702) + (1.06\_B*0801) + (0.80\_Cw*0701) + (-4.28\_Cw*0702)$. The estimated logistic regression model for the HLA Class II analysis is the following: $\text{logit}(\pi) = -2.62 + (-0.56\_DRB1*0101) + (1.42\_DRB1*0301) + (1.02\_DRB1*0701) + (-1.85\_DQA1*\ 0201) + (-0.07\_DQA1*0501)$.

RESULTS

Overview of the Study Population

MSA and MAA were each detected among approximately one-third of the 603 IIM patients surveyed (Table 1). Female to male ratios among individual MSA and MAA subgroups generally reflected the bias of the total IIM patient population (~2:1). Anti-synthetase autoantibodies were those most frequently detected among all MSA (22%), and anti-Jo-1 comprised 73% of the anti-synthetase subgroup. Anti-EJ (anti-glycyl-tRNA synthetase) autoantibodies had the lowest frequency among all MSA (0.5%). As expected, anti-Mi-2 and -SRP autoantibodies were detected primarily among DM (92%) and PM (95%) patients, respectively. MSA were found in low frequency in cancer-associated myositis and were not detected among the 47 IBM patients surveyed. MAA were detected among 59.4% of the patients in the connective tissue disease-myositis overlap group; a result likely consistent with the broader spectrum of connective tissue diseases represented in the overlap group.

Analyses examining the coincident detection of individual MSA and MAA in IIM patients revealed that no patient had more than 1 MSA, and the frequency of patients coproducing any combination of MSA and MAA serogroups was relatively low (MSA+MAA+ = 6.1%). Other findings included the infrequent codetection of anti-PM/Scl autoantibodies with any MSA serogroup (0.2%, pcorr < 0.0001); a negative association largely attributable to anti-Jo-1 given its prevalence among the MSA. A similar negative association was observed between anti-Jo-1- and anti-RNP-positive patients (0.3%, pcorr = 0.0082) as well as between anti-Mi-2 and any MAA serogroup (0.2%, pcorr = 0.0004).

HLA Associations With IIM MSA and MAA Serogroups

HLA alleles found to be significant risk or protective factors for MSA subgroups of the IIM (pcorr < 0.05 after correction for multiple comparisons) are summarized in Table 2. The identification of HLA A*01, B*08, Cw*0701, DRB1*0301, and DQA1*0501 alleles as risk factors for anti-synthetase/anti-Jo-1-positive IIM is consistent with an extended haplotype in linkage disequilibrium (that is, the 8.1 AH) prevalent among those of northwestern European descent. Our previous observations revealed the importance of genetic factors possibly mapping between and including the HLA-B*08 and DRB1*0301 fragment of the 8.1 AH in IIM and all clinicopathologic groups[34]. While these genetic factors are associated with IIM irrespective of autoantibody status, it is clear that they are even more strongly associated with anti-Jo-1 autoantibodies compared to patients without defined autoantibodies (for example, the DRB1*0301 risk

TABLE 1. Autoantibody Groups of Adult EA Patients With Idiopathic Inflammatory Myopathies (IIM)*

	IIM		Clinical Composition, No. (%)				
	N	Freq (%)	PM (n = 227)	DM (n = 177)	IBM (n = 47)	CTM (n = 101)	CAM (n = 51)
Myositis-Specific Autoantibodies (MSA)†							
All MSA	199	33.0	100 (44.0)	75 (42.4)	0	20 (19.8)	4 (7.8)
Anti-synthetases	133	22.0	74 (32.6)	41 (23.2)	0	17 (16.8)	1 (2.0)
Anti-Jo-1	97	16.1	55 (24.2)	31 (17.5)	0	10 (9.9)	1 (2.0)
Anti-PL-7	15	2.5	7 (3.1)	5 (2.8)	0	3 (3.0)	0
Anti-PL-12	11	1.8	5 (2.2)	4 (2.2)	0	2 (2.0)	0
Anti-OJ	7	1.2	5 (2.2)	0	0	2 (2.0)	0
Anti-EJ	3	0.5	2 (0.9)	1 (0.6)	0	0	0
Anti-Mi-2	38	6.3	3 (1.3)	32 (18.1)	0	1 (1.0)	2 (3.9)
Anti-SRP	22	3.6	20 (8.8)	1 (0.6)	0	0	1 (2.0)
Negative	404	67.0	127 (55.9)	102 (57.6)	47 (100)	81 (80.2)	47 (92.2)
Myositis-Associated Autoantibodies (MAA)							
All MAA	169	28.0	49 (21.6)	47 (26.6)	8 (17.0)	60 (59.4)	5 (9.8)
Anti-PM/Scl	65	10.8	13 (5.7)	19 (10.7)	0	32 (31.7)	1 (2.0)
Anti-Ro	59	9.8	25 (11.0)	17 (9.6)	7 (14.9)	10 (9.9)	0
Anti-La	27	4.5	9 (4.0)	7 (4.0)	3 (6.4)	7 (6.9)	1 (2.0)
Anti-U-RNP	32	5.3	9 (4.0)	8 (4.5)	0	15 (14.8)	0
Anti-Ku	7	1.2	1 (0.4)	1 (0.6)	0	6 (5.9)	0
Negative	434	72.0	178 (78.4)	130 (73.4)	39 (83.0)	41 (40.6)	46 (90.2)

Abbreviations: CTM = connective tissue disease myositis overlap; CAM = cancer-associated myositis. Anti-synthetases: anti-Jo-1 = anti-histidyl-tRNA synthetase; anti-PL-7 = anti-threonyl-tRNA synthetase; anti-PL-12 = anti-alanyl-tRNA synthetase; anti-OJ = anti-isoleucyl-tRNA synthetase; anti-EJ = anti-glycyl-tRNA synthetase.

*Of the 603 total patients (including 32 patients for whom HLA data were unavailable), 410 (68%) were female and 193 (32%) were male.

factors in anti-Jo-1-positive and MSA-negative IIM patients have odds ratios [OR] of 15.5 and 3.6, respectively; see Table 2). Moreover, a significant difference exists in the frequency of DRB1*0301 between anti-Jo-1-positive and MSA-negative IIM patients (82.1% vs. 51.9%; pcorr < 0.0001), suggesting that the anti-Jo-1 marker may define a more homogeneous population of IIM patients.

The linked alleles DRB1*0701 and DQA1*0201 were identified as protective factors for development of anti-synthetase/anti-Jo-1 autoantibodies (see Table 2). In contrast, these same alleles represent risk factors for development of anti-Mi-2 autoantibodies, as observed previously[7,27,29,44]. This finding is consistent with the mutually exclusive production of each of these autoantibodies in IIM patients (that is, each MSA-positive patient almost invariably produces a single MSA). The DQA1*0201 allele was also observed as a protective factor among MSA-negative patients (pcorr = 0.0008). The frequently linked alleles DRB1*01 and DQA1*0101 were also identified as novel protective factors for development of the anti-synthetase/anti-Jo-1 autoantibodies compared to controls. In this instance, it appears the protective effect is dependent on anti-Jo-1 positivity, considering that MSA-negative patients exhibit no such effect.

Additional HLA alleles, Cw*04 and DRB1*1501, were identified as susceptibility factors for anti-synthetase and anti-Jo-1 autoantibodies, respectively. The DRB1*1501 association with anti-synthetase may be explained by inclusion of anti-PL-7 (anti-threonyl tRNA synthetase)-positive patients wherein DRB1*1501 allele frequency is increased compared to MSA-negative patients (53.8% vs. 13.6%, pcorr = 0.039). In addition, HLA-Cw*0304 appears to be a risk factor attributable to the presence of anti-PL-7 autoantibodies compared to both MSA-negative patients and controls. In contrast, DQA1*0501 allele carriage is significantly less frequent among anti-PL-7-positive patients (8.3% vs. 64% MSA-negative patients, pcorr = 0.002). Novel HLA associations were identified among the small number of anti-SRP-positive patients analyzed; HLA-B*5001 and DQA1*0104 alleles were each overrepresented in SRP patients relative to MSA-negative patients and controls.

As summarized in Table 3, alleles comprising the 8.1 AH were also identified as HLA risk factors for different MAA subgroups of IIM patients (that is, anti-La, -PM/Scl, and -Ro). One notable exception was seen in patients producing anti-Ku autoantibodies wherein DRB1*11 alleles were detected as prominent risk factors. In addition, DRB1*15/*16 (DR2) and DQA1*0101 alleles, commonly

TABLE 2. Immunogenetic Differences Among EA IIM Patients With Or Without Myositis-Specific Autoantibodies or Unrelated, Ethnically Matched Controls

LA Allele	IIM MSA+ % (N†)	IIM MSA- % (N†)	CON % (N†)	IIM MSA+ vs. IIM MSA- pcorr (OR, 95% CI)	IIM MSA+ vs. CON IIM pcorr (OR, 95% CI)	MSA- vs. CON pcorr (OR, 95% CI)
	MSA: Anti-synthetase					
HLA-A						
*01	49.3 (34/69)	37.4 (55/147)	27.3 (83/304)	NS	0.015 (2.6, 1.4–4.6)	NS
HLA-B						
*08	78.1 (50/64)	40.1 (55/137)	20.6 (65/316)	<0.0001 (5.3, 2.6–11.4)	<0.0001 (13.8, 6.9–28.5)	0.0008 (2.6, 1.6–4.1)
HLA-C						
*0701	78.8 (52/66)	42.0 (60/143)	27.3 (36/132)	<0.0001 (5.1, 2.5–10.9)	<0.0001 (9.9, 4.7–21.6)	NS
DRB1						
*01	*9.2 (11/120)*	*25.0 (78/312)*	*25.1 (155/618)*	*0.0023 (0.3, 0.1–0.6)*	*0.0009 (0.3, 0.1–0.6)*	*NS*
*0301	72.6 (85/117)	51.9 (152/293)	22.9 (111/485)	0.0041 (2.5, 1.5–4.1)	<0.0001 (9.0, 5.5–14.6)	<0.0001 (3.6, 2.6–5.0)
*0701	*11.5 (13/113)*	*18.4 (53/288)*	*27.2 (134/493)*	*NS*	*0.017 (0.3, 0.2–0.6)*	*NS*
*1501	29.2 (33/113)	13.6 (39/286)	21.6 (103/477)	0.017 (2.6, 1.5–4.6)	NS	NS
DQA1						
*0101	*16.8 (19/113)*	*29.8 (87/292)*	*33.6 (146/435)*	*NS*	*0.0061 (0.4, 0.2–0.7)*	*NS*
*0201	*8.0 (9/113)*	*17.3 (50/289)*	*30.6 (133/435)*	*NS*	*<0.0001 (0.2, 0.1–0.4)*	*0.0008 (0.5, 0.3–0.7)*
*0501	76.5 (88/115)	64.0 (190/297)	50.8 (231/455)	NS	<0.0001 (3.2, 1.9–5.2)	0.0053 (1.7, 1.2–2.4)
	MSA: Anti-Jo-1					
HLA-A						
*01	51.7 (30/58)	37.4 (55/147)	27.3 (83/304)	NS	0.01 (2.8, 1.5–5.3)	NS
HLA-B						
*08	84.9 (45/53)	40.1 (55/137)	20.6 (65/316)	<0.0001 (8.4, 3.5–22.0)	<0.0001 (21.7, (9.4–55.4)	0.0008 (2.6, 1.6–4.1)
HLA-C						
*0701	85.5 (47/55)	42.0 (60/143)	27.3 (36/132)	0.0051 (4.0, 1.7–10.5)	<0.0001 (15.7, 6.4–41.5)	NS
*04	*10.9 (6/55)*	*26.6 (38/143)*	*30.5 (60/197)*	*NS*	*0.036 (0.3, 0.1–0.7)*	*NS*
DRB1						
*01	*9.3 (8/86)*	*25.0 (78/312)*	*25.1 (155/618)*	*NS*	*0.0079 (0.3, 0.1–0.6)*	*NS*
*0301	82.1 (69/84)	51.9 (152/293)	22.9 (111/485)	<0.0001 (4.3, 2.3–8.4)	<0.0001 (15.5, 8.3–30.2)	<0.0001 (3.6, 2.6–5.0)
*0701	*9.9 (8/81)*	*18.4 (53/288)*	*27.2 (134/493)*	*NS*	*0.021 (0.3, 0.1–0.6)*	*NS*
DQA1						
*0101	*16.5 (13/79)*	*29.8 (87/292)*	*33.6 (146/435)*	*NS*	*0.027 (0.4, 0.2–0.7)*	*NS*
*0201	*7.5 (6/80)*	*17.3 (50/289)*	*30.6 (133/435)*	*NS*	*<0.0001 (0.1, 0.1–0.4)*	*0.0008 (0.5, 0.3–0.7)*
*0501	84.1 (69/82)	64.0 (190/297)	50.8 (231/455)	0.0057 (3.0, 1.5–6.1)	<0.0001 (5.1, 2.7–10.4)	0.0053 (1.7, 1.2–2.4)
	MSA: Anti-PL-7					
HLA-C						
*0304	80.0 (4/5)	14.7 (21/143)	13.6 (18/132)	0.05 (23.2, 2.5–218.3)	0.046 (25.3, 2.2–1257.8)	NS
DRB1						
*1501	53.8 (7/13)	13.6 (39/286)	21.6 (103/477)	0.039 (7.4, 2.0–27.8)	NS	NS
DQA1						
*0501	*8.3 (1/12)*	*64.0 (190/297)*	*50.8 (231/455)*	*0.002 (0.05, 0.001–0.4)*	*NS*	*0.0053 (1.7, 1.2–2.4)*
	MSA: Anti-PL-12					
DRB1						
*0301	80.0 (8/10)	51.9 (152/293)	22.9 (111/485)	NS	0.013 (13.5, 2.6–131.2)	<0.0001 (3.6, 2.6–5.0)

continued

TABLE 2. *(continued)*

LA Allele	IIM MSA+ % (N†)	IIM MSA- % (N†)	CON % (N†)	IIM MSA+ vs. IIM MSA- pcorr (OR, 95% CI)	IIM MSA+ vs. CON IIM pcorr (OR, 95% CI)	MSA- vs. CON pcorr (OR, 95% CI)
	MSA: Mi-2					
DRB1						
*0701	64.5 (20/31)	18.4 (53/288)	27.2 (134/493)	<0.0001 (8.0, 3.4–19.6)	0.0021 (4.9, 2.2–11.5)	NS
DQA1						
*0201	59.4 (19/32)	17.3 (50/289)	30.6 (133/435)	<0.0001 (7.0, 3.0–16.4)	0.02 (3.3, 1.5–7.5)	0.0008 (0.5, 0.3–0.7)
	MSA: SRP					
HLA-B						
*5001	50.0 (2/4)	0.7 (1/137)	0 (0/129)	0.045 (136.0, 4.1–8059.6)	0.024 (ND)	NS
DQA1						
*0104	21.4 (3/14)	0.3 (1/288)	2.9 (12/413)	0.0040 (78.3, 5.4–4120.0)	NS	NS

Abbreviations: pcorr = p values (Fisher exact test) corrected for multiple comparisons within each genetic locus; OR = odds ratio; 95% CI = confidence interval; NS = not significant after correction for multiple comparisons; ND = not determined.

†The number (N) of allele-positive subjects/the total number of subjects for whom complete low- or high-resolution HLA data were available at a given locus. Table excludes CAM patients. Alleles identified as potential protective factors for the IIM are listed in italics.

Additional comparisons of connective tissue disease myositis overlap patients (CTM) with their respective non-overlap PM, DM, and IBM groups did not reveal any significant differences in HLA allele frequencies (data not shown). Therefore, myositis overlap patients (CTM-PM, CTM-DM, and CTM-IBM) were included in the respective PM, DM, and IBM clinical groups for the comparisons described above.

found in linkage disequilibrium, were identified as protective factors for the PM/Scl subgroup. The association of the DQA1*0201 protective factor among anti-Ro-positive patients was also observed in MAA-negative patients compared to controls, suggesting that the association is not autoantibody specific but rather attributable to the PM clinical phenotype as described previously[34], since approximately 43% of anti-Ro-positive IIM patients have PM.

In addition to the DRB1*0301 risk and DRB1*0701, DQA1*0201 protective factors, we identified several HLA Class I risk factors associated with patients without myositis autoantibodies (those without either MSA or MAA). These included A*68 (OR, 6.6; pcorr = 0.0025), B*15 (OR, 4.3; pcorr = 0.0020), and C*14 (OR, 14.8; pcorr = 0.023). These HLA Class I alleles, which are infrequent and typically unlinked in EA populations, were previously identified as possible risk factors in PM, DM, and IBM patients analyzed irrespective of autoantibody status[34]. Together, the data suggest that these HLA Class I allele groups may represent independent risk factors in a subset of MSA/MAA-negative IIM patients.

HLA Associations Among Combined Clinical and Serologic Subgroups of IIM Patients

Because of the known HLA associations with different clinicopathologic groups[34], we analyzed associations with MSA and MAA in the context of these clinicopathologic groups (Table 4). Alleles consistent with the 8.1 AH are prominent risk factors associated with PM and DM patients expressing anti-synthetase autoantibodies, and anti-Jo-1 autoantibodies in particular. Among these alleles, DRB1*0301 was the only risk factor detected consistently among MSA-negative PM and DM patients despite having lower OR values (OR, 3.6 and 3.2, respectively) than their anti-synthetase-positive (OR, 9.4 and 8.1, respectively) or anti-Jo-1-positive counterparts (OR, 15.5). The DQA1*0201 allele was identified as a protective factor for anti-synthetase/anti-Jo-1-positive PM and DM patients (see Table 4). The protective association of DQA1*0201 among MSA-negative IIM patients suggests that while this factor may not be MSA dependent, the strength of association increases among anti-synthetase/anti-Jo-1-positive patients (for example, OR = 0.5 and 0.1 in MSA-negative vs. anti-Jo-1-positive IIM, respectively) (see Table 2). Conversely, DQA1*0201 and its frequently linked allele DRB1*0701 were identified as risk factors for DM patients producing anti-Mi-2 autoantibodies; a clinical subgroup of IIM patients in which these autoantibodies are frequently detected (84.2%). Consistent with what was observed among total IIM patients, linked alleles DRB1*01 and DQA1*0101 were identified as protective factors among anti-synthetase-positive PM patients although these same alleles were similarly protective for total DM patients independent of MSA status as reported previously[34]. We note that novel HLA risk factors consistent with the B*15, Cw*0304 haplotype were associated with anti-PL-7-positive PM patients.

Although the number of subjects is more limited for some of the subgroups and make estimates less precise, various risk factors consistent with alleles of the 8.1 AH were also detected among PM, DM, and IBM patients producing anti-La, -PM/Scl, and -Ro MAA as summarized in Table 5. Notable exceptions included increased frequencies of DRB1*1104 detection among anti-Ku-positive PM patients (75% vs. 4.2% and 6.3% of MAA-negative patients and controls, respectively), DQA1*06 alleles among anti-La-positive

DM patients and HLA-B*4501 among anti-PM/Scl-positive PM patients. Although DQA1*01 protective alleles were again detected among PM/Scl-positive patients (PM and DM), the further identification of DQA1*03 alleles as protective was unexpected given our prior definition of DQA1*0301 as a possible risk factor for PM and DM[34]. Together, these data suggest that in addition to shared HLA susceptibility factors, combined clinical and serologic groups of IIM patients have distinct immunogenetic features perhaps consistent with different etiopathogenic pathways of disease development.

Interrelationships of HLA allele associations among IIM MSA and MAA Serogroups

The interrelationships of HLA alleles identified as potential susceptibility factors were compared among different MSA and MAA serogroups of the IIM (Figure 1).

TABLE 3. Immunogenetic Differences Among EA IIM Patients With or Without Myositis-Associated Autoantibodies or Unrelated, Ethnically Matched Controls

HLA Allele	IIM MAA+ % (N†)	IIM MAA- % (N†)	CON % (N†)	IIM MAA+ vs. IIM MAA- pcorr (OR, 95% CI)	IIM MAA+ vs. CON pcorr (OR, 95% CI)	IIM MAA- vs. CON pcorr (OR, 95% CI)
	MAA: Ku					
DRB1						
*11	80.0 (4/5)	16.3 (56/343)	15.8 (97/613)	0.049 (20.5, 2.0–1012.9)	0.041 (21.3, 2.1–1049.8)	NS
	MAA: La					
HLA-B						
*0801	81.8 (9/11)	46.1 (77/167)	22.5 (29/129)	<0.0001 (8.4, 3.5–22.0)	0.0046 (15.5, 2.9–151.7)	0.0014 (3.0, 1.7–5.1)
HLA-C						
*0701	81.8 (9/11)	49.7 (87/175)	27.3 (36/132)	0.0051 (4.0, 1.7–10.5)	0.011 (12.0, 2.3–117.0)	0.0021 (2.6, 1.6–4.4)
DRB1						
*0301	66.7 (14/21)	47.5 (152/320)	22.9 (111/485)	<0.0001 (4.3, 2.3–8.4)	0.0019 (6.7, 2.5–20.1)	<0.0001 (3.0, 2.2–4.2)
DQA1						
*06	8.3 (2/24)	0.0 (0/331)	0.8 (4/514)	0.026 (ND)	NS	NS
	MAA: PM/Scl					
HLA-A						
*0101	80.0 (8/10)	38.8 (69/178)	28.7 (39/136)	NS	0.049 (9.9, 1.8–98.5)	NS
HLA-B						
*08	70.0 (7/10)	46.1 (77/167)	20.6 (65/316)	NS	0.045 (9.0, 2.0–55.0)	<0.0001 (3.3, 2.1–5.1)
DRB1						
*0301	95.8 (46/48)	47.5 (152/320)	22.9 (111/485)	<0.0001 (25.4, 6.4–218.5)	<0.0001 (77.5, 19.6–663.8)	<0.0001 (3.0, 2.2–4.2)
**15/ *16*	*8.2 (4/49)*	*23.8 (82/344)*	*27.0 (167/618)*	*NS*	*0.027 (0.2, 0.1–0.7)*	*NS*
DQA1						
**0101*	*8.5 (4/47)*	*28.5 (90/316)*	*33.6 (146/435)*	*0.025 (0.2, 0.1–0.7)*	*0.0026 (0.2, 0.05–0.5)*	*NS*
*0501	94.0 (47/50)	60.2 (194/322)	50.8 (231/455)	<0.0001 (10.3, 3.2–52.8)	<0.0001 (15.2, 4.8–77.1)	NS
	MAA: Ro					
HLA-B						
*0801	66.7 (18/27)	46.1 (77/167)	22.5 (29/129)	NS	0.0006 (6.9, 2.6–19.1)	0.0014 (3.0, 1.7–5.1)
HLA-C						
*0701	66.7 (18/27)	49.7 (87/175)	27.3 (36/132)	NS	0.0045 (5.3, 2.0–14.6)	0.0021 (2.6, 1.6–4.4)
DRB1						
*0301	72.5 (37/51)	47.5 (152/320)	22.9 (111/485)	0.032 (2.9, 1.5–6.1)	<0.0001 (8.9, 4.5–18.4)	<0.0001 (3.0, 2.2–4.2)
DQA1						
**0201*	*9.1 (5/55)*	*18.3 (58/317)*	*30.6 (133/435)*	*NS*	*0.0054 (0.2, 0.1–0.6)*	*0.0018 (0.5, 0.4–0.7)*
*0501	76.8 (43/56)	60.2 (194/322)	50.8 (231/455)	NS	0.0027 (3.2, 1.6–6.7)	NS

Table excludes CAM patients. Alleles identified as potential protective factors for the IIM are listed in italics.
†The number (N) of allele-positive subjects/the total number of subjects for whom complete low- or high-resolution HLA data were available at a given locus.

TABLE 4. Immunogenetic Differences Among Different Clinicopathologic Groups of IIM Patients With or Without Myositis-Specific Autoantibodies or Unrelated, Ethnically Matched Controls

HLA Allele	IIM MSA+ % (N†)	IIM MSA- % (N†)	CON % (N†)	IIM MSA+ vs. IIM MSA- pcorr (OR, 95% CI)	IIM MSA+ vs. CON pcorr (OR, 95% CI)	IIM MSA- vs. CON pcorr (OR, 95% CI)
MSA: Anti-synthetase/clinicopathologic group: PM						
HLA-B						
*08	80.0 (32/40)	43.1 (28/65)	20.6 (65/316)	0.0054 (5.3, 2.0–15.2)	<0.0001 (15.4, 6.5–40.3)	0.0079 (2.9, 1.6–5.3)
*44	2.5 (1/40)	18.5 (12/65)	24.4 (77/316)	NS	0.026 (0.1, 0.002–0.5)	NS
HLA-C						
*0701	78.0 (32/41)	47.0 (31/66)	27.3 (36/132)	0.043 (4.0, 1.5–11.0)	<0.0001 (9.5, 3.9–24.5)	NS
DRB1						
*01	8.9 (7/79)	24.4 (38/156)	25.1 (155/618)	NS	0.013 (0.3, 0.1–0.6)	NS
*0301	73.7 (56/76)	52.0 (77/148)	22.9 (111/485)	NS	<0.0001 (9.4, 5.3–17.3)	<0.0001 (3.6, 2.4–5.5)
DQA1						
*0101	16.4 (12/73)	28.8 (42/146)	33.6 (146/435)	NS	0.046 (0.4, 0.2–0.8)	NS
*0201	11.0 (8/73)	18.6 (27/145)	30.6 (133/435)	NS	0.0045 (0.3, 0.1–0.6)	NS
*0501	76.0 (57/75)	62.2 (92/148)	50.8 (231/455)	NS	0.0007 (3.1, 1.7–5.7)	NS
MSA: Anti-synthetase/clinicopathologic group: DM						
HLA-A						
*01	65.0 (26/40)	58.2 (89/153)	27.3 (83/304)	NS	0.0084 (4.1, 1.7–10.1)	NS
HLA-B						
*0801	75.0 (18/24)	43.1 (28/65)	22.5 (29/129)	NS	<0.0001 (10.3, 3.5–34.2)	NS
HLA-C						
*0701	80.0 (20/25)	47.0 (31/66)	27.3 (36/132)	0.0085 (7.5, 2.2–29.5)	<0.0001 (10.7, 3.5–38.5)	NS
DRB1						
*0301	70.7 (29/41)	52.0 (77/148)	22.9 (111/485)	NS	<0.0001 (8.1, 3.8–18.0)	<0.0001 (3.2, 2.0–5.1)
*0701	2.6 (1/39)	23.0 (23/100)	27.2 (134/493)	NS	0.0081 (0.1, 0.002–0.4)	NS
DQA1						
*0201	2.5 (1/40)	18.6 (27/145)	30.6 (133/435)	NS	0.0004 (0.1, 0.0001–0.4)	NS
*0501	77.5 (31/40)	62.2 (92/148)	50.8 (231/455)	NS	0.018 (3.3, 1.5–8.1)	NS
MSA: Anti-Jo-1/clinicopathologic group: PM						
HLA-B						
*0801	87.9 (29/33)	43.1 (28/65)	22.5 (29/129)	0.0009 (9.6, 2.8–40.9)	<0.0001 (25.0, 7.7–103.2)	NS
HLA-C						
*0701	85.3 (29/34)	47.0 (31/66)	27.3 (36/132)	0.0040 (6.5, 2.1–23.9)	<0.0001 (15.5, 5.3–54.1)	NS
DRB1						
*01	8.6 (5/58)	25.1 (155/618)	25.1 (155/618)	NS	0.044 (0.3, 0.1–0.7)	NS
*0301	82.1 (46/56)	52.0 (77/148)	22.9 (111/485)	0.0031 (4.2, 1.9–10.1)	<0.0001 (15.5, 7.4–35.4)	<0.0001 (3.6, 2.4–5.5)
DQA1						
*0201	9.4 (5/53)	18.6 (27/145)	30.6 (133/435)	NS	0.0084 (0.2, 0.1–0.6)	NS
*0501	81.8 (45/55)	62.2 (92/148)	50.8 (231/455)	NS	0.0002 (4.4, 2.1–9.9)	NS
MSA: Anti-Jo-1/clinicopathologic group: DM						
HLA-A						
*01	62.5 (15/24)	33.3 (17/51)	27.3 (83/304)	NS	0.013 (4.4, 1.7–11.9)	NS
HLA-B						
*0801	80.0 (16/20)	37.0 (17/46)	22.5 (29/129)	NS	<0.0001 (13.8, 4.0–59.8)	NS
HLA-C						
*0701	85.7 (18/21)	34.7 (1749)	27.3 (36/132)	0.0032 (11.3, 2.6–65.8)	(0.0001 (16.0, 4.2–88.0)	NS

continued

TABLE 4. *(continued)*

HLA Allele	IIM MSA+ % (N†)	IIM MSA- % (N†)	CON % (N†)	IIM MSA+ vs. IIM MSA- pcorr (OR, 95% CI)	IIM MSA+ vs. CON pcorr (OR, 95% CI)	IIM MSA- vs. CON pcorr (OR, 95% CI)
DRB1						
*0301	82.1 (23/28)	48.5 (49/101)	22.9 (111/485)	NS	<0.0001 (15.5, 5.6–53.1)	<0.0001 (3.2, 2.0–5.1)
DQA1						
*0201	*3.7 (1/27)*	*20.4 (20/98)*	*30.6 (133/435)*	*NS*	*0.021 (0.1 (0.002–0.5)*	*NS*
*0501	88.9 (24/27)	63.1 (65/103)	50.8 (231/455)	NS	0.0012 (7.8, 2.3–40.7)	NS
MSA: Anti-PL-7/clinicopathologic group: PM						
HLA-B						
*15	75.0 (3/4)	16.9 (11/65)	5.4 (17/315)	NS	0.027 (52.6, 3.8–2768.5)	NS
HLA-C						
*0304	100.0 (4/4)	10.6 (7/66)	13.6 (18/132)	0.0061 (ND)	0.011 (ND)	NS
DQA1						
*03	75.0 (6/8)	24.2 (37/153)	32.8 (169/516)	0.025 (9.4, 1.6–97.5)	NS	NS
MSA: Mi-2/clinicopathologic group: DM						
DRB1						
*0701	71.4 (20/28)	23.0 (23/100)	27.2 (134/493)	0.0002 (8.4, 3.0–24.6)	0.0001 (6.7, 2.7–17.9)	NS
DQA1						
*0201	65.5 (19/29)	20.4 (20/98)	30.6 (133/435)	0.0002 (7.4, 2.7–20.5)	0.0042 (4.3, 1.8–10.6)	NS

Table excludes CAM patients. Alleles identified as potential protective factors for the IIM are listed in italics.
†The number (N) of allele-positive subjects/the total number of subjects for whom complete low- or high-resolution HLA data were available at a given locus.

As described above, alleles consistent with the 8.1 AH and the DRB1*0301 allele in particular were shared among the anti-synthetase (comprised principally of anti-Jo-1-positive patients), anti-PM/Scl, anti-Ro, and anti-La autoantibody groups. Despite these similarities, other HLA associations were characteristic of 1 or more myositis autoantibody groups. The DQA1*0201 protective factor was identified among anti-synthetase/anti-Jo-1- and anti-Ro-positive patients, while the DQA1*01 allele (commonly linked with DRB1*01) was shared among anti-synthetase/anti-Jo-1 and anti-PM/Scl patients. HLA-DR2 alleles (DRB1*15/*16) were observed as unique protective factors among anti-PM/Scl-positive patients. Other autoantibody groups displayed highly distinctive patterns of allele associations, including HLA-Cw*0304, DRB1*11, and commonly linked alleles DRB1*0701 and DQA1*0201, which were identified as risk factors among the anti-PL-7, -Ku and -Mi-2 autoantibody-positive patients, respectively. In addition, HLA B*5001 and DQA1*0104 alleles were identified as possible risk determinants among anti-SRP-positive patients (pcorr = 0.024 and p = 0.010 vs. controls, respectively) although statistical significance for DQA1*0104 was lost (pcorr > 0.05) after correction for multiple comparisons.

Random Forests Classification Analyses

Among the HLA alleles found associated with the IIM in this study, it is uncertain which factors play a legitimate role in disease predisposition and which are associated as the result of haplotypic linkage disequilibrium. To better define the relative importance of individual HLA susceptibility alleles in discriminating IIM cases and controls, we utilized a random forests (RF) classification algorithm. All HLA Class I (A, B, and Cw) or Class II (DRB1 and DQA1) alleles identified among different myositis autoantibody (MSA and MAA) and clinical groups of disease (IIM, PM, DM, or IBM) and controls were analyzed by RF modeling. As summarized in Table 6, HLA alleles identified as potential risk or protective factors for different MSA serogroups (anti-synthetase, Jo-1, -PL-7, -Mi-2, and -SRP) in univariate analyses are shown ranked by their relative importance in effectively classifying IIM cases and controls. Among the HLA Class I variables, the tightly linked HLA-B*0801 and Cw*0701 alleles of the 8.1 AH consistently ranked highest for anti-synthetase/anti-Jo-1-positive IIM, PM and DM patients. Similarly, the DRB1*0301 allele ranked first among these same groups of patients confirming the importance of the 8.1 Cw-B-DRB1 haplotype fragment in disease susceptibility. The Class II DRB1*0301 linked allele, DQA1*0501, ranked considerably lower than DRB1*0301 in each analysis, again confirming that DRB1*0301 itself or a more closely linked gene(s) is the primary anti-synthetase/anti-Jo-1-associated Class II risk factor. A combined RF analysis of all Class I and Class II alleles, while more restricted by smaller numbers of cases and controls for which complete

TABLE 5. Immunogenetic Differences Among Different Clinicopathologic Groups of IIM Patients With or Without Myositis-Associated Autoantibodies or Unrelated, Ethnically Matched Controls[†]

HLA Allele	IIM MAA+ % (N[†])	IIM MAA- % (N[†])	CON % (N[†])	IIM MAA- vs. IIM MAA+ pcorr (OR, 95% CI)	CON vs. IIM MAA+ pcorr (OR, 95% CI)	CON vs. IIM MAA- pcorr (OR, 95% CI)
MAA: Ku/clinicopathologic group: PM						
DRB1						
*1104	75.0 (3/4)	4.2 (7/165)	6.3 (30/473)	0.018 (67.7, 4.4–3572.9)	NS	NS
MAA: La/clinicopathologic group: PM						
HLA-B						
*0801	100.0 (6/6)	48.8 (39/80)	22.5 (29/129)	NS	0.0073 (ND)	0.0057 (3.3, 1.7–6.3)
HLA-C						
*0701	100.0 (6/6)	53.7 (44/82)	27.3 (36/132)	NS	0.012 (ND)	0.0032 (3.1, 1.7–5.7)
MAA: La/clinicopathologic group: DM						
DQA1						
*06	22.2 (2/9)	0.0 (0/174)	0.8 (4/514)	0.026 (ND)	0.023 (36.4, 2.8–297.4)	NS
MAA: PM/Scl/clinicopathologic group: PM						
HLA-A						
*01	100.0 (5/5)	34.1 (28/82)	27.3 (83/304)	NS	0.029 (ND)	NS
HLA-B						
*4501	40.0 (2/5)	3.8 (3/80)	0.0 (0/129)	NS	0.039 (ND)	NS
DRB1						
*0301	96.0 (24/25)	50.9 (86/169)	22.9 (111/485)	0.0002 (23.2, 3.6–963.7)	<0.0001 (80.9, 12.8–3333.8)	<0.0001 (3.5, 2.4–5.1)
DQA1						
*01	*42.3 (11/26)*	*65.5 (114/174)*	*69.4 (369/532)*	*NS*	*0.04 (0.3, 0.1–0.8)*	*NS*
*03	*7.7 (2/26)*	*27.6 (48/174)*	*32.8 (169/516)*	*NS*	*0.04 (0.1, 0.02–0.7)*	*NS*
*0501	92.3 (24/26)	59.8 (101/169)	50.8 (231/455)	0.0088 (8.1, 1.9–72.2)	0.0003 (11.6, 2.8–102.4)	NS
MAA: PM/Scl/clinicopathologic group: DM						
DRB1						
*0301	95.2 (22/23)	40.9 (47/115)	22.9 (111/485)	<0.0001 (31.8, 4.7–1332.4)	<0.0001 (74.1, 11.6–3067.0)	<0.0001 (3.5, 2.4–5.1)
DQA1						
*0101	*4.5 (1/22)*	*23.5 (27/115)*	*33.6 (146/435)*	*NS*	*0.049 (0.1, 0.002–0.6)*	*NS*
*0501	95.8 (23/24)	57.8 (67/116)	50.8 (231/455)	0.0029 (16.8, 2.5–706.8)	<0.0001 (22.3, 3.5–922.7)	NS
MAA: Ro/clinicopathologic group: PM						
HLA-B						
*0801	75.0 (12/16)	48.8 (39/80)	22.5 (29/129)	NS	0.0018 (10.3, 2.8–46.4)	0.0057 (3.3, 1.7–6.3)
HLA-C						
*0701	75.0 (12/16)	53.7 (44/82)	27.3 (36/132)	NS	0.0059 (8.0, 2.2–35.7)	0.0032 (3.1, 1.7–5.7)
DRB1						
*0301	72.4 (21/29)	50.9 (86/169)	22.9 (111/485)	NS	<0.0001 (8.8, 3.6–23.6)	<0.0001 (3.5, 2.4–5.1)
DQA1						
*0501	83.3 (25/30)	59.8 (101/169)	50.8 (231/455)	NS	0.0070 (4.8, 1.8–16.4)	NS
MAA: Ro/clinicopathologic group: IBM						
DRB1						
*0301	100.0 (7/7)	52.8 (19/36)	22.9 (111/485)	NS	0.0019 (ND)	0.0099 (3.8, 1.8–8.0)

Table excludes CAM patients. Alleles identified as potential protective factors for the IIM are listed in italics.

[†]The number (N) of allele-positive subjects/the total number of subjects for whom complete low- or high-resolution HLA data were available at a given locus.

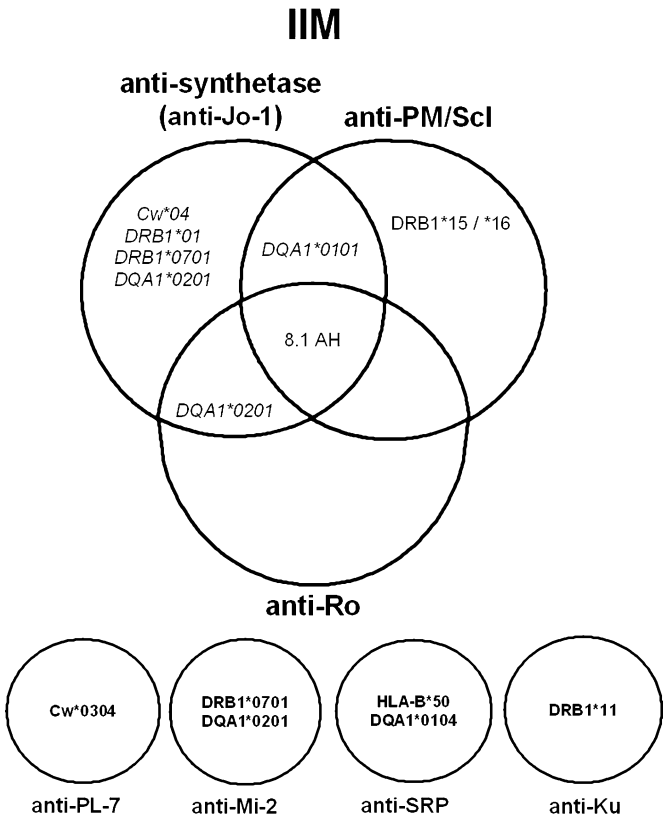

FIGURE 1. Venn diagram depicting the interrelationships of HLA allele associations among different IIM serologic groups (MSA and MAA). HLA alleles identified as either risk or protective factors are shown in boldface type or italics, respectively. Not shown is the association of the 8.1 AH and DQA1*06 risk factors with anti-La-positive IIM patients.

high-resolution data were available across all 5 HLA loci (A, B, Cw, DRB1, and DQA1), again indicated that HLA-B*0801 and DRB1*0301 ranked highest for the accurate classification of anti-Jo-1-positive IIM cases and controls (data not shown). Additionally, a direct comparison of IIM cases and controls that were HLA-B*0801 positive and DRB1*0301 negative, or conversely were HLA-B*0801 negative and DRB1*0301 positive, in an attempt to determine the primary risk allele detected no significant differences between the groups (data not shown). RF analyses of other MSA serotypes ranked HLA risk factors Cw*0304, DRB1*0701, and possibly DQA1*0104 as important classifiers of anti-PL-7-, -Mi-2-, and -SRP-positive PM, DM, and IIM patients, respectively.

Similar to the anti-synthetase/anti-Jo-1 analyses described above, RF classification of MAA serogroups anti-PM/Scl, -Ro, and -La revealed that alleles of the 8.1 AH haploblock Cw*0701-B*0801-DRB1*0301 were the most reliable predictors of IIM cases and controls (Table 7). One notable exception included the high ranking of the 8.1-associated allele HLA-A*0101 as an important classifier of anti-PM/Scl-positive IIM and PM patients, although the number of subjects analyzed was more limited. Nevertheless, these data suggest that HLA-A*0101 or a more closely linked gene(s) may be a risk factor independent of DRB1*0301. Lastly, the DRB1*1104 allele ranked first among all predictors of anti-Ku-positive PM, although only a small number of subjects were available for study.

Traditional logistic regression analyses were used to corroborate these findings independently. For example, a comparative analysis of anti-Jo-1-positive IIM patients ranked HLA-B*0801 ($p = 0.025$; OR, 8.3; 95% confidence interval [CI], 1.3–53.0) and DRB1*0301 ($p < 0.0001$; OR, 17.1; 95% CI, 6.7–43.3) highest among the HLA Class I and Class II alleles discriminating IIM cases and controls, respectively. These data suggest that HLA-B*0801, DRB1*0301, and possibly intervening genes of the 8.1 AH HLA Class III region, are important risk factors for the IIM. The strong and consistent association of the DRB1*0301 allele compared to HLA-B*0801 suggests that additional genetic risk factors associated with the 8.1 AH may map closer to and possibly include determinants of the HLA Class III region.

HLA Peptide-Binding Motifs Among IIM MSA and MAA Serogroups

Given the implied importance of HLA-DRB1 alleles to susceptibility for a host of autoimmune diseases including the IIM, we chose to examine a series of primary amino acid sequence motifs mapping within the third hypervariable region (HVR3) of the DRB1 gene and comprising functional contact residues within pocket 4 of the MHC peptide-binding groove (amino acid positions 70, 71, and 74). As defined in Patients and Methods, these RSP motifs were functionally stratified according to consensus HVR3 amino acid sequences with well-characterized peptide- and/or T-cell receptor-binding properties[36]. As anticipated, the RSP ''R'' motif (Q/K/R) representing the HVR3 domain of DRB1*0301 was a significant risk factor for IIM patients with anti-synthetase, -Jo-1, -PL-12, -PM/Scl, -Ro, and -La serotypes (Table 8). Three different RSP motifs (''A,'' ''Q,'' and ''Dr'') were each detected at significantly lower frequencies among anti-synthetase/anti-Jo-1-positive patients. The putatively protective association of the RSP ''A'' and ''Q'' motifs are likely attributable to the DRB1*01 and DRB1*07 associations described previously. Curiously, the RSP ''Dr'' motif is comprised of a complex mixture of DRB1 alleles (see Patients and Methods), none of which is independently associated with the IIM. Notable among the RSP ''A,'' ''Q,'' and ''Dr'' protective motifs are amino acid residues shared among 2 or more motifs (^{70}D [hydrophilic]/^{71}R [basic]/^{74}A [hydrophobic]). RSP ''A'' and ''Dr'' motifs were also identified as statistically significant protective motifs among anti-PM/Scl-positive patients. The RSP ''a'' motif identified among anti-synthetase- and -PL-7-positive patients appears to be of lesser importance as it is not significantly increased in frequency relative to controls but rather clearly distinguishes anti-synthetase- and -PL-7-positive from MSA-negative patients. As expected, the DRB1*0701 defined RSP ''Q'' motif was identified as a significant risk factor for

TABLE 6. Relative Importance of MSA-Associated Risk and Protective Factors for the IIM as Predicted by Random Forest Classification Models

	SYN (RI%†)	Jo-1 (RI%†)	PL-7 (RI%†)	Mi-2 (RI%†)	SRP (RI%†)
HLA Class I	**IIM (N‡ = 64)**	**IIM (N = 53)**	**IIM (N = 5)**	**NA**	**IIM (N = 4)**
	B*0801 (100)	B*0801 (100)	Cw*0304 (100)		B*5001 (93.2)
	Cw*0701 (81.2)	Cw*0701 (96.4)	**PM (N = 4)**		
	A*0101 (16.0)	A*0101 (14.8)	Cw*0304 (100)		
	PM (N = 40)	*Cw*0401 (11.4)*	B*15 (17.0)		
	B*0801 (100)	**PM (N = 33)**			
	Cw*0701 (79.1)	B*0801 (100)			
	*B*44 (18.5)*	Cw*0701 (78.1)			
	DM (N = 24)	**DM (N = 20)**			
	Cw*0701 (100)	Cw*0701 (100)			
	B*0801 (98.9)	B*0801 (73.3)			
	A*0101 (30.7)	A*0101 (18.2)			
HLA Class II	**IIM (N = 108)**	**IIM (N = 77)**	**IIM (N = 12)**	**IIM (N = 26)**	**IIM (N = 14)**
	DRB1*0301 (100)	DRB1*0301 (100)	*DQA1*0501 (100)*	DRB1*0701 (97.6)	DQA1*0104 (100)
	DQA1*0501 (32.2)	DQA1*0501 (39.0)	DRB1*1501 (64.7)	DQA1*0201 (78.3)	
	*DQA1*0201 (26.2)*	*DQA1*0201 (21.5)*	**PM (N = 7)**	**DM (N = 23)**	
	*DRB1*0701 (15.9)*	*DRB1*0701 (12.2)*	DQA1*03 (92.6)	DRB1*0701 (100)	
	*DQA1*0101 (12.7)*	*DQA1*0101 (7.8)*		DQA1*0201 (71.3)	
	DRB1*1501 (8.6)	*DRB1*01 (7.0)*			
	*DRB1*01 (7.9)*	**PM (N = 50)**			
	PM (N = 69)	DRB1*0301 (100)			
	DRB1*0301 (100)	DQA1*0501 (34.3)			
	DQA1*0501 (27.6)	*DQA1*0201 (19.4)*			
	*DQA1*0201 (20.8)*	*DRB1*01 (6.3)*			
	*DQA1*0101 (11.0)*	**DM (N = 27)**			
	*DRB1*01 (6.8)*	DRB1*0301 (100)			
	DM (N = 39)	DQA1*0501 (39.9)			
	DRB1*0301 (100)	*DQA1*0201 (36.2)*			
	DQA1*0501 (39.9)				
	*DRB1*0701 (36.5)*				
	*DQA1*0201 (36.2)*				

Abbreviations: NA = not applicable (no significant associations of HLA Class I alleles were identified among Mi-2 patients).

Out-of-bag (oob) estimates of error rates (%) for HLA Class I and Class II analyses, respectively: synthetase (IIM, 24.9 and 26.4; PM, 24.2 and 25.2; DM, 25.5 and 24.9), Jo-1 (IIM, 23.0 and 21.3; PM, 22.2 and 23.9; DM, 24.8 and 21.8), PL-7 (IIM, 23.1 and 28.3; PM, 14.7 and 16.8), Mi-2 (IIM, 31.4; DM, 34.9), SRP (IIM, 35.7 and 39.0).

†RI, relative variable importance scores normalized (%) to the highest-ranking factor (GINI score) in a given analysis; GINI scores were calculated using the GINI impurity criterion for individual variables over all classification trees in the forest.

‡N = number of cases for whom complete high-resolution HLA data were available for each locus in the analysis (number of controls = 125 (Class I) and 327 (Class II).

anti-Mi-2-positive patients while serving a protective role among the anti-synthetase, -Jo-1, and -Ro serogroups.

DISCUSSION

To our knowledge, the current study is the largest immunogenetic study of HLA Class I and Class II allelic associations in the IIM to date. We have both confirmed certain previous findings and identified a number of novel genetic risk and protective factors for the development of different MSA and MAA. Among the heterogeneous IIM syndromes, myositis autoantibodies, and particularly those that are disease specific (that is, MSA), have proven invaluable in classifying patients into more homogeneous groups in regard to various clinical, diagnostic, immunopathologic, therapeutic, and prognostic features[30]. Previous studies in our laboratory and by other authors have further demonstrated that myositis autoantibody-defined subgroups of patients share distinct immunogenetic features, observations perhaps consistent with alternative pathways of disease development[1,7,11,27,29,40,44].

TABLE 7. Relative Importance of MAA-Associated Risk and Protective Factors for the IIM as Predicted by Random Forest Classification Models

	Pm/Scl (RI%†)	Ro (RI%†)	La (RI%†)	Ku (RI%; Rank†)
HLA Class I	**IIM (N‡ = 9)**	**IIM (N = 27)**	**IIM (N = 11)**	**NA**
	A*0101 (95.9)	B*0801 (100)	B*0801 (100)	
	B*0801 (79.6)	Cw*0701 (67.4)	Cw*0701 (80.8)	
	PM (N = 5)	**PM (N = 16)**	**PM (N = 6)**	
	A*0101 (100)	B*0801 (100)	Cw*0701 (100)	
	B*4501 (43.4)	Cw*0701 (74.5)	B*0801 (90.7)	
HLA Class II	**IIM (N = 43)**	**IIM (N = 48)**	**IIM (N = 21)**	**IIM (N = 5)**
	DRB1*0301 (100)	DRB1*0301 (100)	DRB1*0301 (100)	DRB1*11 (100)
	DQA1*0501 (37.9)	DQA1*0501 (47.3)	DQA1*06 (8.3)	**PM (N = 4)**
	*DQA1*0101 (13.1)*	*DQA1*0201 (23.6)*	**DM (N = 7)**	DRB1*1104 (100)
	*DRB1*15/16 (4.4)*	**PM (N = 28)**	DQA1*06 (14.3)	
	PM (N = 24)	DRB1*0301 (100)		
	DRB1*0301 (100)	DQA1*0501 (47.6)		
	DQA1*0501 (33.2)	**IBM (N = 5)**		
	*DQA1*01 (8.1)*	DRB1*0301 (100)		
	*DQA1*03 (3.2)*			
	DM (N = 19)			
	DRB1*0301 (100)			
	DQA1*0501 (53.5)			
	*DQA1*0101 (23.4)*			

Abbreviations: NA = not applicable (no significant associations of HLA Class I alleles were identified among Ku patients).

Out-of-bag (oob) estimates of error rates (%) for HLA Class I and Class II analyses, respectively: PM/Scl (IIM, 29.1 and 20.8; PM, 23.8 and 21.6; DM, 22.0), Ro (IIM, 32.9 and 23.7; PM, 28.4 and 20.3; IBM, 31.9), La (IIM, 24.3 and 23.6; PM, 25.2; DM, 23.4), Ku (IIM, 19.6; PM, 19.0).

†RI, relative variable importance scores normalized (%) to the highest-ranking factor (GINI score) in a given analysis; GINI scores were calculated using the GINI impurity criterion for individual variables over all classification trees in the forest.

‡N = number of cases for whom complete high-resolution HLA data were available for each locus in the analysis (number of controls = 125 (Class I) and 327 (Class II).

Among the 603 IIM patients for which serologic data were available, we detected various MSA and MAA serotypes at frequencies consistent with those reported for several distinct European and North American populations[1,4,19,54]. Other studies reporting higher or lower frequencies of myositis autoantibody detection might possibly be explained by methodologic variations, referral bias resulting in different patient populations, and/or smaller sample sizes[14,24]. We also observed that the overall frequency of MSA and MAA codetection among individual patients was low with the possible exception of anti-Jo-1 and anti-Ro coreactivities (detected in approximately 6% of patients). Other studies have reported considerably higher frequencies of anti-Jo-1 and anti-Ro codetection, which may be due to referral bias, patient heterogeneity, and/or variable study methodologies[1,42].

Polymorphic gene variants of the human MHC are among the strongest and most consistent genetic factors associated with autoimmune disease[9,33,51]. The relationship between MHC genetic variability and autoimmune disease susceptibility may relate to the essential role MHC molecules play in T-cell receptor repertoire development, peripheral tolerance to self-antigens, and regulating the types and degree of immune responses to environmental agents[13,28,51]. Since the ability to develop specific lymphocyte-mediated immune responses depends on selective peptide antigen presentation by HLA molecules, the dysregulation of MHC-mediated functions may contribute to a loss of self-tolerance and resultant autoimmune pathology. A direct role for MHC in disease susceptibility has been demonstrated convincingly by the capacity for certain HLA alleles to confer autoimmune pathologies in various transgenic rodent models of human disease[41,46,48].

Previous studies of HLA associations with IIM were often limited by lower resolution serologic typing and/or small numbers of subjects. In our present study, we performed low to high resolution molecular typing (PCR-SSP and -SSOPH) to characterize the allelic variability of HLA-A, -B, -Cw, -DRB1, and -DQA1 determinants in a large population of EA IIM patients (n = 571) representing the major MSA and MAA serogroups. We hypothesized that an adequately powered study of genetic variation among the IIM serogroups might permit the identification of specific HLA alleles or groups of HLA alleles with common amino acid peptide-binding motifs.

To this end, we have corroborated earlier reports describing associations of HLA alleles comprising the

TABLE 8. Immunogenetic differences in DRB1 RSP Functional Motifs Among Caucasian IIM Patients With or Without Myositis-Specific or Myositis-Associated Autoantibodies and Unrelated, Ethnically Matched Controls

DRB1 RSP Pocket 4 Motifs (AA 70/71/74)[†]	IIM MA+[†] N[§] (%)	IIM MA- N[§] (%)	CON N[§] (%)	IIM MA- vs. IIM MA+ pcorr (OR, 95% CI)	CON vs. IIM MA+ pcorr (OR, 95% CI)	CON vs. IIM MA- pcorr (OR, 95% CI)
MSA: Anti-synthetase						
RSP ''A'' (QR/RK/A)	*32/113 (28.3)*	*120/288 (41.7)*	*221/490 (45.1)*	*NS*	*0.0052 (0.5, 0.30–0.76)*	*NS*
RSP ''R'' (Q/K/R)	86/117 (73.5)	154/294 (52.4)	111/485 (22.9)	0.0007 (2.5, 1.54–4.18)	<0.0001 (15.5, 8.33–30.19)	<0.0001 (3.7, 2.68–5.12)
RSP ''Q'' (D/R/Q)	*13/113 (11.5)*	*53/289 (18.3)*	*134/493 (27.2)*	*NS*	*0.0022 (0.35, 0.17–0.65)*	*0.032 (0.6, 0.41–0.87)*
RSP ''Dr'' (D/RK/AL)	*17/113 (15.0)*	*72/288 (25.0)*	*132/479 (27.6)*	*NS*	*0.022 (0.47, 0.25–0.82)*	*NS*
RSP ''a'' (Q/A/A)	34/113 (30.1)	45/287 (15.7)	112/477 (23.5)	0.012 (2.3, 1.34–3.98)	*NS*	0.049 (0.6, 0.40–0.90)
MSA: Anti-Jo-1						
RSP ''A'' (QR/RK/A)	*21/81 (25.9)*	*120/288 (41.7)*	*221/490 (45.1)*	*NS*	*0.0074 (0.4, 0.24–0.74)*	*NS*
RSP ''R'' (Q/K/R)	69/84 (82.1)	154/294 (52.4)	111/485 (22.9)	<0.0001 (4.2, 2.24–8.22)	<0.0001 (15.5, 8.33–30.19)	<0.0001 (3.7, 2.68–5.12)
RSP ''Q'' (D/R/Q)	*8/81 (9.9)*	*53/289 (18.3)*	*134/493 (27.2)*	*NS*	*0.0028 (0.3, 0.12–0.63)*	*0.032 (0.6, 0.41–0.87)*
RSP ''Dr'' (D/RK/AL)	*11/81 (13.6)*	*72/288 (25.0)*	*132/479 (27.6)*	*NS*	*0.034 (0.4, 0.19–0.82)*	*NS*
MSA: Anti-PL-7						
RSP ''a'' (Q/A/A)	7/13 (53.8)	45/287 (15.7)	112/477 (23.5)	0.017 (6.3, 1.70–23.52)	*NS*	0.049 (0.6, 0.40–0.90)
MSA: Anti-PL-12						
RSP ''R'' (Q/K/R)	4/9 (44.4)	45/287 (15.7)	111/485 (22.9)	NS	0.0001 (30.3, 4.09–1332.5)	<0.0001 (3.7, 2.68–5.12)
MSA: Mi-2						
RSP ''Q'' (D/R/Q)	20/31 (64.5)	53/289 (18.3)	134/493 (27.2)	<0.0001 (8.1, 3.43–19.74)	0.0003 (4.9, 2.15–11.53)	0.032 (0.6, 0.41–0.87)
MAA: PM/Scl						
RSP ''A'' (QR/RK/A)	*10/46 (21.7)*	*130/317 (41.0)*	*221/490 (45.1)*	*NS*	*0.016 (0.3, 0.15–0.72)*	*NS*
RSP ''R'' (Q/K/R)	46/48 (95.8)	158/321 (49.2)	111/485 (22.9)	<0.0001 (23.7, 5.99–203.99)	<0.0001 (3.3, 2.38–4.48)	NS
RSP ''Dr'' (D/RK/AL)	*4/47 (8.5)*	*74/315 (23.5)*	*132/479 (27.6)*	*NS*	*0.016 (0.2, 0.06–0.69)*	*NS*
MAA: Ro						
RSP ''R'' (Q/K/R)	38/52 (73.1)	158/321 (49.2)	111/485 (22.9)	0.011 (2.8, 1.14–5.81)	<0.0001 (9.1, 4.62–18.87)	*NS*
RSP ''Q'' (D/R/Q)	*5/51 (9.8)*	*63/321 (19.6)*	*134/493 (27.2)*	*NS*	*0.038 (0.3, 0.09–0.75)*	*NS*
MAA: La						
RSP ''R'' (Q/K/R)	15/22 (68.2)	158/321 (49.2)	111/485 (22.9)	NS	<0.0001 (7.2, 2.68–21.37)	NS

Table excludes CAM patients. Alleles identified as potential protective factors for the IIM are listed in italics.

[†]MA, myositis autoantibody.

[‡]RSP, restrictive supertype pattern defined as amino acid motifs occupying positions 70, 71. and 74 comprising pocket 4 of the HLA DRB1 peptide-binding region.

[§]The number (N) of allele-positive subjects/the total number of subjects for whom complete low- or high-resolution HLA data were available at a given locus.

Caucasian 8.1 AH with particular MSA (anti-Jo-1) and MAA (anti-Ro, -PM/Scl, and -La) serogroups[1,11,27,43,44]. Moreover, we have extended previous findings of genetically linked myositis-associated risk factors (DRB1*0301; DQA1*0501; DQB1*0201) to include high-resolution Class I data (HLA-A*0101; B*0801; Cw*0701), further supporting an important role for the 8.1 AH. The 8.1 AH is the most common haplotype among EA subjects, and its constituent alleles have been identified as risk factors for many autoimmune diseases (for example, systemic lupus erythematosus, Sjögren syndrome, myasthenia gravis, insulin-dependent diabetes mellitus, and Graves disease) and other immune abnormalities[6,47,52]. Increasing evidence suggests that polymorphic variants of multiple immune response genes residing along the 8.1 AH (for example, HLA Class III loci) play important roles in immune regulation and may also contribute to disease susceptibility[6,17,37]. Yet, despite the prevalence of 8.1 AH alleles in the EA population, only a small fraction of these individuals will ever develop an autoimmune pathology. It is conceivable that 8.1 AH gene variants, in combination with other predisposing genetic factors and environmental exposures, may subvert fundamental pathways of immune regulation. These primary deregulating events may comprise a shared pathogenic pathway among multiple immune disorders. Additional genetic factors and environmental exposures along with various stochastic and epigenetic events may ultimately define the tissue-specificity or prevailing clinical phenotype of a particular disorder (for example, IIM vs. systemic lupus erythematosus). These observations highlight the complexity, polygenicity, and multifactorial nature of autoimmune disease susceptibility[9]. It is also likely that genetic and environmental risk and protective factors differ in various subsets or in particular phenotypes of patients as defined by current criteria, as well as in various ethnogeographic groups, further complicating our understanding of the complex pathogenesis of autoimmunity[35,43,44].

In addition to alleles of the 8.1 AH, we also identified several novel genetic markers that are associated uniquely with 1 or more serogroups of EA IIM patients. Most notably, HLA associations with anti-PL-7, an autoantibody targeting threonyl tRNA synthetase, clearly differed from other anti-synthetase associations in lacking 8.1 AH-derived risk factors (for example, DRB1*0301) and having an unusual HLA Class I association (Cw*0304). We have also identified a novel association in a small number of anti-Ku-positive patients with HLA DRB1*1104. A previous study assessing a larger group of anti-Ku-positive patients revealed a strong association with the DQw1 marker; a group of DQB1 alleles from which 2 alleles (DQB1*0502 and *0603) exist in linkage disequilibrium with DRB1*1104[53]. These data are in contrast to a Polish IIM study wherein anti-Ku-positive patients shared the DRB1*0301, DQA1*0501 association[18]. Also, our analyses failed to confirm a previous report of a DR5 (particular alleles of DRB1*11/*12 groups) association with SRP-positive PM patients[27]. Here, we have described novel associations between the anti-SRP serogroup and HLA-B*5001 and possibly DQA1*0104. The relatively small number of patients with anti-SRP autoantibodies, however, limits the power of these findings.

More surprising was the identification of 3 independent HLA Class I associations (HLA-A*68, B*15, and Cw*14) with a subset of IIM patients characterized as MSA/MAA double negatives. In a previous report, we identified each of these alleles as risk factors for IIM irrespective of serologic status[34]. While each of these HLA Class I allele groups is less prevalent in the EA population overall, the MSA/MAA-negative patients had higher OR values relative to their unstratified counterparts compared to controls (data not shown). These data suggest that a proportion of seronegative IIM patients have unique risk factors in addition to those associated with alleles of the 8.1 AH.

Random forests analysis is a powerful and robust prediction and classification tool employed in the study of large and complex datasets (for example, tumor cell microarray, high-throughput structural genomics, and multivariate mapping of complex traits)[5,16,21,55]. Of particular interest is the potential for RF to rank variable importance by testing the effects of randomly chosen and independent variables on prediction accuracy (that is, discriminating cases and controls). Among the linked alleles of the 8.1 AH, multivariate RF analyses consistently ranked the HLA-B*0801/Cw*0701 and DRB1*0301 alleles as the strongest discriminators of IIM cases and controls among anti-synthetase/anti-Jo-1, -PM/Scl, -Ro, and -La autoantibody-positive patients. These RF data, in combination with traditional logistic regression analyses, indicated the importance of DRB1*0301, and possibly linked alleles of the HLA Class III region, as important risk factors associated with the IIM, in contrast to the tightly linked DQA1*0501 allele also found in the extended 8.1 AH. Other HLA alleles, including Cw*0304, DRB1*0701, B*5001/DQA1*0104, and DRB1*1104, were accurate predictors among anti-PL-7-, -Mi-2-, -SRP-, and -Ku-positive IIM patients, respectively. Collectively, these analyses provide considerable utility in distinguishing primary susceptibility factors from other alleles in linkage disequilibrium that are frequently coidentified in univariate analyses.

A higher resolution analysis of DRB1 HVR3 functional domains (RSP motifs) revealed the expected associations between DRB1*0301 (RSP ''R'')- and *0701 (RSP ''Q'')-derived motifs and the anti-synthetase/anti-Jo-1, -PL-12, -PM/Scl, -Ro, -La, and anti-Mi-2 serotypes, respectively. Additionally, we have proposed a consensus sequence (^{70}D/^{71}R/^{74}A) derived from 3 RSP motifs (RSP ''A,'' ''Dr,'' ''Q'') conferring a protective effect among anti-synthetase/anti-Jo-1-positive patients. We also identified a novel DRB1*15 derived motif (RSP ''a'') in the anti-synthetase and anti-PL-7 analyses that distinguished MSA-positive from MSA-negative patients. In fact, the RSP ''a'' motif is a protective factor among MSA-negative patients compared to healthy control subjects. Our discovery of many protective alleles and motifs for certain autoantibody groups is also informative. We use the term ''protective'' here as an operational definition whereby a statistically significant increase in the frequency of a genetic marker is observed in

the control compared with the respective patient group. The implication is that these observations have a physiologic basis although the underlying mechanisms are uncertain. Similar observations have been described in other human autoimmune diseases; most notably among various risk and protective alleles associated with the shared epitope region in rheumatoid arthritis[15]. One might speculate that risk and protective effects associated with specific HLA alleles may represent differential binding of peptides conferring alternatively autoreactive or favorable immunoregulatory properties. In addition, protective HLA markers might also exist in linkage disequilibrium with other presently unidentified genes responsible for the observed protective effects. Nevertheless, these findings may have important implications for identifying immunogenic peptides that bind these susceptibility and protective factors and either initiate, sustain, or block the immunopathology of myositis.

There are several limitations to our case-control, candidate gene study design including diminished statistical power when comparing smaller subsets of patients, incomplete data for all HLA loci in the total subject population, and the inherent genetic, serologic, and clinical heterogeneity among IIM patients overall. Despite these limitations, our data suggest that in addition to being susceptibility markers for IIM, HLA alleles are markers for different MSA and MAA serogroups and possibly relate to divergent pathogenic mechanisms. These variations in associated HLA polymorphisms may reflect responses to different environmental triggers that ultimately result in the tissue pathospecificity and the distinct clinicopathologic syndromes of the IIM. Our findings are consistent with studies of other autoimmune disorders that have identified differing genetic risk factors for a given disease and its varying phenotypes[10,30,32,38].

In summary, our findings support the case that immunogenetic associations seem strongest for subgroups of patients defined by disease-associated immune responses rather than by clinicopathologic features. Additionally, most autoantibody groups have a distinct genetic signature, sometimes defined by the presence of both risk and protective factors. Autoimmune diseases are heterogeneous syndromes and likely comprise multiple entities or elemental disorders, each of which appears to develop from the interaction of the necessary and sufficient genetic and environmental risk and protective factors, which trigger pathologic events that ultimately culminate in a particular sign-symptom-laboratory complex[2,8,30,43]. Distinct immunogenetic features associated with the specific clinical and/or serologic group(s) of the IIM may be due to various gene-environment interactions, which result in different immune responses that are collectively reflected in the many phenotypes of the myositis syndromes.

ACKNOWLEDGMENTS

The authors thank Drs. Derek Middleton and Gabor Illei for their critical reviews of the manuscript.

REFERENCES

1. Arnett FC, Targoff IN, Mimori T, Goldstein R, Warner NB, Reveille JD. Interrelationship of major histocompatibility complex class II alleles and autoantibodies in four ethnic groups with various forms of myositis. *Arthritis Rheum*. 1996;39:1507–1518.
2. Becker KG. The common variants/multiple disease hypothesis of common complex genetic disorders. *Med Hypotheses*. 2004;62:309–317.
3. Bohan A, Peter JB. Polymyositis and dermatomyositis (first of two parts). *N Engl J Med*. 1975;292:344–347.
4. Brouwer R, Hengstman GJ, Vree Egberts W, Ehrfeld H, Bozic B, Ghirardello A, Grondal G, Hietarinta M, Isenberg D, Kalden JR, Lundberg I, Moutsopoulos H, Roux-Lombard P, Vencovsky J, Wikman A, Seelig HP, van Engelen BG, van Venrooij WJ. Autoantibody profiles in the sera of European patients with myositis. *Ann Rheum Dis*. 2001;60:116–123.
5. Bureau A, Dupuis J, Hayward B, Falls K, Van Eerdewegh P. Mapping complex traits using Random Forests. *BMC Genet*. 2003;4(Suppl 1):S64.
6. Candore G, Lio D, Colonna RG, Caruso C. Pathogenesis of autoimmune diseases associated with 8.1 ancestral haplotype: effect of multiple gene interactions. *Autoimmun Rev*. 2002;1:29–35.
7. Chinoy H, Ollier WE, Cooper RG. Have recent immunogenetic investigations increased our understanding of disease mechanisms in the idiopathic inflammatory myopathies? *Curr Opin Rheumatol*. 2004;16:707–713.
8. Cooper GS, Miller FW, Pandey JP. The role of genetic factors in autoimmune disease: implications for environmental research. *Environ Health Perspect*. 1999;107(Suppl 5):693–700.
9. Davidson A, Diamond B. Autoimmune diseases. *N Engl J Med*. 2001;345:340–350.
10. Falkner D, Wilson J, Fertig N, Clawson K, Medsger TA, Jr., Morel PA. Studies of HLA-DR and DQ alleles in systemic sclerosis patients with autoantibodies to RNA polymerases and U3-RNP (fibrillarin). *J Rheumatol*. 2000;27:1196–1202.
11. Franceschini F, Cavazzana I. Anti-Ro/SSA and La/SSB antibodies. *Autoimmunity*. 2005;38:55–63.
12. Friedman JH, Hastie T, Tibshirani R. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer-Verlag; 2001.
13. Gebe JA, Swanson E, Kwok WW. HLA class II peptide-binding and autoimmunity. *Tissue Antigens*. 2002;59:78–87.
14. Ghirardello A, Zampieri S, Iaccarino L, Tarricone E, Tonello M, Bendo R, Rondinone R, Cozzi F, Doria A. Myositis specific and myositis associated autoantibodies in idiopathic inflammatory myopathies: a serologic study of 46 patients. *Reumatismo*. 2005;57:22–28.
15. Gibert M, Balandraud N, Touinssi M, Mercier P, Roudier J, Reviron D. Functional categorization of HLA-DRB1 alleles in rheumatoid arthritis: the protective effect. *Hum Immunol*. 2003;64:930–935.
16. Goh CS, Lan N, Douglas SM, Wu B, Echols N, Smith A, Milburn D, Montelione GT, Zhao H, Gerstein M. Mining the structural genomics pipeline: identification of protein properties that affect high-throughput experimental analysis. *J Mol Biol*. 2004;336:115–130.
17. Hassan AB, Nikitina-Zake L, Sanjeevi CB, Lundberg IE, Padyukov L. Association of the proinflammatory haplotype (MICA5.1/TNF2/TNFa2/DRB1*03) with polymyositis and dermatomyositis. *Arthritis Rheum*. 2004;50:1013–1015.
18. Hausmanowa-Petrusewicz I, Kowalska-Oledzka E, Miller FW, Jarzabek-Chorzelska M, Targoff IN, Blaszczyk-Kostanecka M, Jablonska S. Clinical, serologic, and immunogenetic features in Polish patients with idiopathic inflammatory myopathies. *Arthritis Rheum*. 1997;40:1257–1266.
19. Hengstman GJ, Brouwer R, Egberts WT, Seelig HP, Jongen PJ, van Venrooij WJ, van Engelen BG. Clinical and serological characteristics of 125 Dutch myositis patients. Myositis specific autoantibodies aid in the differential diagnosis of the idiopathic inflammatory myopathies. *J Neurol*. 2002;249:69–75.
20. Holm S. A simple sequentially rejective multiple test procedure. *Scand J Statist*. 1979;6:65–70.
21. Jiang H, Deng Y, Chen HS, Tao L, Sha Q, Chen J, Tsai CJ, Zhang S. Joint analysis of two microarray gene-expression data sets to select lung adenocarcinoma marker genes. *BMC Bioinformatics*. 2004;5:81.
22. Kao AH, Lacomis D, Lucas M, Fertig N, Oddis CV. Anti-signal recognition particle autoantibody in patients with and patients without idiopathic inflammatory myopathy. *Arthritis Rheum*. 2004;50:209–215.
23. Koffman BM, Sivakumar K, Simonis T, Stroncek D, Dalakas MC. HLA

allele distribution distinguishes sporadic inclusion body myositis from hereditary inclusion body myopathies. *J Neuroimmunol*. 1998;84: 139–142.
24. Kowalska-Oledzka E, Stern LZ. Myositis specific antibodies: frequency in different populations. *Folia Neuropathol*. 2000;38:101–103.
25. Lampe JB, Gossrau G, Kempe A, Fussel M, Schwurack K, Schroder R, Krause S, Kohnen R, Walter MC, Reichmann H, Lochmuller H. Analysis of HLA class I and II alleles in sporadic inclusion-body myositis. *J Neurol*. 2003;250:1313–1317.
26. Lemeshow S, Hosmer DW, Jr. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol*. 1982;115:92–106.
27. Love LA, Leff RL, Fraser DD, Targoff IN, Dalakas M, Plotz PH, Miller FW. A new approach to the classification of idiopathic inflammatory myopathy: myositis-specific autoantibodies define useful homogeneous patient groups. *Medicine (Baltimore)*. 1991;70:360–374.
28. Marrack P, Kappler J, Kotzin BL. Autoimmune disease: why and where it occurs. *Nat Med*. 2001;7:899–905.
29. Mierau R, Dick T, Genth E, Bartz-Bazzanella P, Keller E, Albert ED. An update on HLA association of Mi-2 autoantibodies: the association with a tryptophan at position 9 of the HLA-DRbeta chain is strong but not absolute. *Arthritis Rheum*. 1999;42:1552–1553.
30. Miller F. Inflammatory myopathies: polymyositis, dermatomyositis, and related conditions. In: Koopman W, Moreland L, eds. *Arthritis and Allied Conditions, A Textbook of Rheumatology. Vol. 15*. Philadelphia: Lippincott, Williams & Wilkins; 2004:1593–1620.
31. Morel PA, Chang HJ, Wilson JW, Conte C, Saidman SL, Bray JD, Tweardy DJ, Medsger TA, Jr. Severe systemic sclerosis with anti-topoisomerase I antibodies is associated with an HLA-DRw11 allele. *Hum Immunol*. 1994;40:101–110.
32. Nath SK, Kilpatrick J, Harley JB. Genetics of human systemic lupus erythematosus: the emerging picture. *Curr Opin Immunol*. 2004;16: 794–800.
33. Newton JL, Harney SM, Wordsworth BP, Brown MA. A review of the MHC genetics of rheumatoid arthritis. *Genes Immun*. 2004;5:151–157.
34. O'Hanlon TP, Carrick DM, Arnett FC, Reveille JD, Carrington M, Gao X, Oddis CV, Morel PA, Malley JD, Malley K, Dreyfuss J, Shamim EA, Rider LG, Chanock SJ, Foster CB, Bunch T, Plotz PH, Love LA, Miller FW. Immunogenetic risk and protective factors for the idiopathic inflammatory myopathies: distinct HLA-A, -B, -Cw, -DRB1 and -DQA1 allelic profiles and motifs define clinicopathologic groups in caucasians. *Medicine (Baltimore)*. 2005;84:338–349.
35. Okada S, Weatherhead E, Targoff IN, Wesley R, Miller FW. Global surface ultraviolet radiation intensity may modulate the clinical and immunologic expression of autoimmune muscle disease. *Arthritis Rheum*. 2003;48:2285–2293.
36. Ou D, Mitchell LA, Tingle AJ. A new categorization of HLA DR alleles on a functional basis. *Hum Immunol*. 1998;59:665–676.
37. Price P, Witt C, Allcock R, Sayer D, Garlepp M, Kok CC, French M, Mallal S, Christiansen F. The genetic basis for the association of the 8.1 ancestral haplotype (A1, B8, DR3) with multiple immunopathological diseases. *Immunol Rev*. 1999;167:257–274.
38. Quintero-Del-Rio AI, Kelly JA, Kilpatrick J, James JA, Harley JB. The genetics of systemic lupus erythematosus stratified by renal disease: linkage at 10q22.3 (SLEN1), 2q34-35 (SLEN2), and 11p15.6 (SLEN3). *Genes Immun*. 2002;3 Suppl 1:S57–S62.
39. Reed AM, Ytterberg SR. Genetic and environmental risk factors for idiopathic inflammatory myopathies. *Rheum Dis Clin North Am*. 2002; 28:891–916.
40. Rider LG, Shamim E, Okada S, Pandey JP, Targoff IN, O'Hanlon TP, Kim HA, Lim YS, Han H, Song YW, Miller FW. Genetic risk and protective factors for idiopathic inflammatory myopathy in Koreans and American whites: a tale of two loci. *Arthritis Rheum*. 1999;42:1285–1290.
41. Rosloniec EF, Whittington KB, He X, Stuart JM, Kang AH. Collagen-induced arthritis mediated by HLA-DR1 (*0101) and HLA-DR4 (*0401). *Am J Med Sci*. 2004;327:169–179.
42. Rutjes SA, Vree Egberts WT, Jongen P, Van Den Hoogen F, Pruijn GJ, Van Venrooij WJ. Anti-Ro52 antibodies frequently co-occur with anti-Jo-1 antibodies in sera from patients with idiopathic inflammatory myopathy. *Clin Exp Immunol*. 1997;109:32–40.
43. Shamim EA, Rider LG, Miller FW. Update on the genetics of the idiopathic inflammatory myopathies. *Curr Opin Rheumatol*. 2000;12: 482–491.
44. Shamim EA, Rider LG, Pandey JP, O'Hanlon TP, Jara LJ, Samayoa EA, Burgos-Vargas R, Vazquez-Mellado J, Alcocer-Varela J, Salazar-Paramo M, Kutzbach AG, Malley JD, Targoff IN, Garcia-De la Torre I, Miller FW. Differences in idiopathic inflammatory myopathy phenotypes and genotypes between Mesoamerican Mestizos and North American Caucasians: ethnogeographic influences in the genetics and clinical expression of myositis. *Arthritis Rheum*. 2002;46:1885–1893.
45. Svetnik V, Liaw A, Tong C, Culberson JC, Sheridan RP, Feuston BP. Random forest: a classification and regression tool for compound classification and QSAR modeling. *J Chem Inf Comput Sci*. 2003; 43:1947–1958.
46. Szanto S, Bardos T, Szabo Z, David CS, Buzas EI, Mikecz K, Glant TT. Induction of arthritis in HLA-DR4-humanized and HLA-DQ8-humanized mice by human cartilage proteoglycan aggrecan but only in the presence of an appropriate (non-MHC) genetic background. *Arthritis Rheum*. 2004;50:1984–1995.
47. Tan FK, Arnett FC. The genetics of lupus. *Curr Opin Rheumatol*. 1998; 10:399–408.
48. Taneja V, David CS. HLA class II transgenic mice as models of human diseases. *Immunol Rev*. 1999;169:67–79.
49. Targoff IN. Autoantibodies to aminoacyl-transfer RNA synthetases for isoleucine and glycine. Two additional synthetases are antigenic in myositis. *J Immunol*. 1990;144:1737–1743.
50. Targoff IN, Trieu EP, Miller FW. Reaction of anti-OJ autoantibodies with components of the multi-enzyme complex of aminoacyl-tRNA synthetases in addition to isoleucyl-tRNA synthetase. *J Clin Invest*. 1993;91:2556–2564.
51. Thorsby E, Lie BA. HLA associated genetic predisposition to autoimmune diseases: Genes involved and possible mechanisms. *Transpl Immunol*. 2005;14:175–182.
52. Vandiedonck C, Beaurain G, Giraud M, Hue-Beauvais C, Eymard B, Tranchant C, Gajdos P, Dausset J, Garchon HJ. Pleiotropic effects of the 8.1 HLA haplotype in patients with autoimmune myasthenia gravis and thymus hyperplasia. *Proc Natl Acad Sci U S A*. 2004;101:15464–15469.
53. Yaneva M, Arnett FC. Antibodies against Ku protein in sera from patients with autoimmune diseases. *Clin Exp Immunol*. 1989;76:366–372.
54. Zampieri S, Ghirardello A, Iaccarino L, Tarricone E, Gambari PF, Doria A. Anti-Jo-1 antibodies. *Autoimmunity*. 2005;38:73–78.
55. Zhang H, Yu CY, Singer B. Cell and tumor classification using gene expression data: construction of forests. *Proc Natl Acad Sci U S A*. 2003;100:4168–4172.

