Low copy numbers of complement C4 and C4A deficiency are risk factors for myositis, its subgroups and autoantibodies

Danlei Zhou,1,2 Emily H King,1,2 Simon Rothwell,3,4 Olga Krystufkova,5 Antonella Notarnicola,6 Samantha Coss,1,2 Rabheb Abdul-Aziz,2,7 Katherine E Miller,1,2 Amanda Dang,1 G Richard Yu,1 Joanne Drew,2 Emeli Lundström,6 Lauren M Pachman,8 Gulnara Mamyrova,9 Rodolfo V Curiel,9 Boel De Paepe,10 Jan L De Bleecker,10 Antony Payton,11 William Ollier,12 Terrance P O’Hanlon,13 Ira N Targoff,14 Willy A Flegel,15 Vidya Sivaraman,2 Edward Oberle,2 Shoghik Akoghlanian,2 Kyla Driest,2 Charles H Spencer,16 Yee Ling Wu,2,17 Haikady N Nagaraja,18 Stacy P Ardoin,2 Hector Chinoy,2,3,4 Lisa G Rider,13 Frederick W Miller,13 Ingrid E Lundberg,6 Leonid Padyukov,6 Jiří Vencovský,5 Janine A Lamb,19 Chuck-Yung Yu,1,2 for MYOGEN Investigators

ABSTRACT

Background Idiopathic inflammatory myopathies (IIM) are a group of autoimmune diseases characterised by myositis-related autoantibodies plus infiltration of leucocytes into muscles and/or the skin, leading to the destruction of blood vessels and muscle fibres, chronic weakness and fatigue. While complement-mediated destruction of capillary endothelia is implicated in paediatric and adult dermatomyositis, the complex diversity of complement C4 in IIM pathology was unknown.

Methods We elucidated the gene copy number (GCN) variations of total C4, C4A and C4B, long and short genes in 1644 Caucasian patients with IIM, plus 3526 matched healthy controls using real-time PCR or Southern blot analyses. Plasma complement levels were determined by single radial immunodiffusion.

Results The large study populations helped establish the distribution patterns of various C4 GCN groups. Low GCNs of C4T(C4T=2+3) and C4A deficiency (C4A=0+1) were strongly correlated with increased risk of IIM with OR equalled to 2.58 (2.28–2.91), p=5.0×10−53 for C4T, and 2.82 (2.48–3.21), p=7.0×10−57 for C4A deficiency. Contingency and regression analyses showed that among patients with C4A deficiency, the presence of HLA-DR3 became insignificant as a risk factor in IIM except for inclusion body myositis (IBM), by which 98.2% had HLA-DR3 with an OR of 11.02 (1.44–84.4). Intragroup analyses of patients with IIM for C4 protein levels and IIM-related autoantibodies showed that those with anti-Jo-1 or with anti-Pm/Scl had significantly lower C4 plasma concentrations than those without these autoantibodies.

Conclusions C4A deficiency is relevant in dermatomyositis, HLA-DRB1*03 is important in IBM and both C4A deficiency and HLA-DRB1*03 contribute interactively to risk of polymyositis.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Complement activation causes damage to the skin and muscles in myositis through the formation of membrane attack complexes on capillaries and activation products that stimulate inflammation. How genetic diversity of complement contributes to differential susceptibility among human patients with myositis was unknown.

WHAT THIS STUDY ADDS

⇒ We deciphered gene copy number (GCN) variations for complement total C4 (C4T), acidic C4A, basic C4B, long genes (C4L) and short genes (C4S) in >1600 patients with IIM and >3500 healthy subjects of European ancestry. Low GCNs of C4T, C4A and C4L strongly correlated with elevated risk of juvenile dermatomyositis, adult-onset dermatomyositis and polymyositis. The presence of HLA-DR3 with deficiencies for C4A or C4B was the predominant genetic factor for inclusion body myositis. Lower plasma protein levels of C4 and C3 were present among patients with IIM with anti-Jo1 and myositis-associated autoantibodies.

INTRODUCTION

Idiopathic inflammatory myopathies (IIM) are a group of autoimmune diseases characterised by chronic muscle weakness and fatigue.1,3 Pathology in IIM includes the generation of myositis-related autoantibodies and infiltration of leucocytes into muscles and/or the skin leading to inflammation with high levels of muscle enzymes in the circulation.1,4 Four major subgroups of IIM include juvenile dermatomyositis (JDM),1 adult-onset dermatomyositis (DM), polymyositis (PM) and inclusion body myositis (IBM).1,2
myositis (IBM). Immune-mediated necrotising myositis and anti-synthetase syndrome are recently defined categories.

JDM is the most common form of myositis in children that has a mean age of diagnosis between 7 and 8 years. Patients with JDM have similar muscle and skin manifestations as in adult-onset DM but do not have the increased risk of interstitial lung disease (ILD) and malignancy that are more common among adult patients. Specific patterns of rash involving the eyelids, face, shoulders and body areas frequently exposed to sunlight are prevalent among JDM and DM. Muscle weakness is symmetric and proximal to the body axis. In pathognomonic muscle biopsies, there is remarkable complement-mediated destruction of perivascular endothelium leading to perifascicular ischaemia and degeneration of muscle fibres. However, triggers for complement activation and whether complement genetic diversity is engaged in the breakdown of immune tolerance have not been investigated. PM is more common in women over the age of 30.

Among subjects of European ancestry, the presence of HLA-DRB1*03:01 or HLA-DRB3 tends to strongly associate with IIM and short C4B genes (C4B<26) in disease susceptibility for IIM and its four major subtypes. The relative roles of HLA-DRB1*03 and C4A deficiency on genetic risk of IIM, and how the C4 GCN variations and complement protein levels correlated with the presence of myositis-related autoantibodies were also examined.

Isolation of genomic DNA, EDTA-plasma and Southern blot analyses

For subjects recruited in Ohio, preparation of genomic DNA from peripheral blood samples, performance of TaqI, PstI-PvuII restriction fragment length polymorphisms and Pmel pulsed-field gel electrophoresis to elucidate RP-C4-CYP21-TNX (RCCX) modular structures were as described.

Copy numbers and sizes of C4A and C4B genes by real-time PCR

When quantities of genomic DNA were limiting, copy numbers of C4 genes were determined by TaqMan-based quantitative real-time PCR with internal control using cosmid DNA with both test and control amplions. Five independent test amplions specific for total C4 (C4T), C4A, C4B, long genes and short genes were performed. Verification was achieved when GCN of C4T=GCNs of C4A+C4B and/or GCNs of C4L+C4S.
Protein concentrations and polymorphic variants
Complement C4 and C3 protein concentrations were measured by single radial immunodiffusion (RID) using EDTA-plasma and an RID kit from the Binding Site (UK). C4A and C4B protein allotypes in plasma samples were resolved by high-voltage agarose gel electrophoresis, followed by immunofixation using antiserum against human C4.40

Genotyping of HLA-DRB1
Genotyping for HLA-DRB1 alleles for samples from the USA and Sweden was performed at low resolution using the sequence-specific primer-PCR methods (eg, DR low-resolution kit: Olerup SSP, Saltsjobaden, Sweden).41 42 The HLA-DRB1 genotypes for samples from the UK were deduced from single-nucleotide polymorphisms data using SNP2HLA software.20 43 44

High concordance of imputed data from DNA sequencing and conventional HLA typing techniques was obtained.20

Statistical analyses
This was a cross-sectional, case–control study. Statistical analyses were performed using JMP16 software from SAS. Continuous data between patients and controls were compared by t-tests. The distributions of C4T, C4A, C4B, C4L, and C4S GCN groups in patients with IIM or in each IIM subgroup and controls were analysed by χ² analyses. The GCN groups for each type of C4 genes were segregated dichotomously into low GCN and medium to high GCN groups, and their frequencies compared between case and controls with χ² analyses to compute ORs and 95% CIs. The low GCN groups were defined as follows: C4T=2+3, C4A=0+1, C4B=0+1, C4L=0+1+2 and C4S=0.

Figure 1 The complement system with emphasis on the genotypic and phenotypic diversities of C4A and C4B. (A) Activation and regulation of the human complement system. Activation of zymogens and progression of pathways are shown in red; regulations of activated products in green. A positive feedback of amplification is common for all three activation pathways. (B). Genetic locations for constituents of the C3 convertases for classical and alternative pathways. (C) Segmental duplications with one to five modules of the RP-C4-CYP21-TNX (RCX) in haplotypes at the class III region of the human leucocyte antigen (HLA). (D) Dichotomy of human C4 gene size with the long gene containing endogenous retrovirus HERV-K(C4) in the ninth intron and the short gene without the endogenous retrovirus. (E) Specific polymorphisms leading the isotypic changes for C4A and C4B proteins. (F) Immunofixation experiments showing the quantitative and qualitative diversities of C4A and C4B protein allotypes including deficiencies. (G) The range of polymorphic variants for C4B (left panel) and for C4A (right panel). CNVs, copy number variations.
A Bonferroni’s correction for a C4 genotype with p<0.01 was considered significant to account for five structural variants being investigated for IIM genetic risk individually. For intra-group comparisons of a specific phenotype with a genotype, a value <0.05 was viewed as significant.

**RESULTS**

Comparisons of GCN variations of complement C4 between IIM and controls

**Total C4**

The mean GCN and SD of C4T among patients with IIM was 3.50±0.78, compared with 3.83±0.76 in healthy controls (δ=-0.37, p=6.0×10−46, t-test) (table 2). The most prevalent GCN group for C4T in IIM was 3 copies with a frequency of 40.1%, and those with 5, 6, 7 and 8 copies comprised a total of 9.2% of all IIM. Categorically, the distributions of C4T GCNs in IIM were substantially different from those in healthy controls, with a p-value of 1.4×10−53 (χ² analysis). The OR and 95% CI for IIM subjects with two copies of C4T was 2.26 (1.74 to 2.95), p=2.0×10−9, and those with two or three copies (C4T=2+3) had an OR=2.62 (2.32 to 2.95), p=2.8×10−35 (figure 3A; see also online supplemental figure S2, supplementary results).

Thus, low GCNs of C4T, that is, C4T=2 and C4T=2+3, had similar magnitude of effects on the genetic risk of IIM.

**C4A in IIM**

GCN of C4A varied from 0 to 6 among patients with IIM with a mean of 1.74±0.88, compared with 2.10±0.84 in healthy controls (δ=-0.37, p=6.0×10−46). There were remarkable increases in the frequencies of C4A low GCN groups and decreases in medium and high GCN groups in IIM (p=6.5×10−36). While 40.1% of patients with IIM had two copies of C4A genes, those with 0 and 1 copy constituted 4.2% and 38.6% of patients, respectively (figure 2). Patients with 3–6
copies (high GCN) of C4A together had a combined frequency of 17.1%. The OR was 2.49 (1.76–3.54, \( p=3.6 \times 10^{-7} \)) for C4A=0 and 2.82 (2.48–3.21, \( p=2.9 \times 10^{-57} \)) for C4A=0+1 (figure 3A).

The magnitudes of the effects of low C4A GCNs on IIM were similar to that observed in C4T=2 and C4T=2+3.

C4B in IIM
Unlike C4T and C4A, C4B copy number group distribution in IIM was almost identical to that observed in healthy controls, which ranged between 0 and 5. Close to two-thirds of the patients with IIM (65.8%) had two copies of C4B, while 21.1% and 27.1% had 0 and 1 copy, respectively. Patients with 3, 4 and 5 copies of C4B constituted a total frequency of 4.9%.

Long genes (C4L) in IIM
The copy number of C4L varied from 0 to 8 in patients with IIM. The mean C4L GCN in IIM was 2.41\( \pm \)1.13, which was significantly lower than that in healthy controls (2.94\( \pm \)1.08, \( p=1.7 \times 10^{-54} \)). The distribution of GCN groups for C4L was different from that of controls (\( p=2.6 \times 10^{-52} \)). The combined frequency for low GCN of long genes (C4L=0+1+2) in IIM was 55.6%, compared with 33.4% in healthy controls (OR=2.50 (2.21–2.82), \( p=2.7 \times 10^{-49} \) (figure 3A). Decreasing GCNs of C4L elevated the ORs for IIM: 2.55 (2.15–3.03, \( p=9.4 \times 10^{-27} \)) for C4L=0+1 and 2.72 (1.85–4.02, \( p=5.1 \times 10^{-7} \)) for C4L=0. The frequency of long genes among total C4 decreased from 74.6% in controls to 63.2% in IIM (C4L/C4T, \( p=2.1 \times 10^{-15} \)).
the largest impact on PM with OR=3.16 (2.65–3.75). Low GCN had the greatest impact on IBM with OR=3.70 (2.72–5.04). Low GCN had the lowest impact on DMD, with OR=2.1 (2.0–2.2). Low GCN had the greatest impact on JDM with OR=3.79 (2.56–5.65). Low GCN had the largest effects on PM and IBM, with ORs of 2.80 and 2.88, respectively.

Figure 3 Forest plots of ORs for low copy number groups for C4T, C4A, long genes (C4L) as risk factors (A), and for C4B and short genes (C4S) as protective factors (B) in IIM and subgroups. A single exception was that low copy number C4B was also a risk factor of IBM. Notice the partial dominance of low GCNs of total C4 (C4T=2 and C4T=2+3) and C4A deficiencies (C4A=0 and C4A=0+1) on conferring risk of IIM and its subgroups DM, PM, IBM and JDM. The ORs in panel A are shown in log-scale. C4A, acidic isotype of complement C4; C4B, basic isotype of complement C4; C4L, long form of C4 gene with human endogenous retrovirus HERV-K(C4); C4S, short form of C4 gene without integration of the retrovirus HERV-K(C4); C4T, total copy number of C4 genes; DM, dermatomyositis; GCNs, gene copy numbers; IBM, inclusion body myositis; IIM, idiopathic inflammatory myopathies; JDM, juvenile dermatomyositis; PM, polymyositis.

Short genes (C4S) in IIM
The copy number of C4S in IIM varied from 0 to 5. The mean copy number was 1.06±0.72, which was higher than that in healthy controls (0.90±0.77, p=3.0×10−12). More than half of the patients with IIM had a single copy of C4S (53.5%). The frequency of subjects lacking C4S (C4S=0) was significantly reduced from 32.3% in controls to 21.1% in IIM (OR=0.56 (0.49–0.65), p=8.8×10−17).

C4 GCN variations among subgroups of IIM
Compared with controls, the four IIM subgroups had lower mean GCNs of C4A in the range of 1.70 to 1.82 but they were not distinguishable among themselves (table 2 and figure 2). Patients with IBM were unusual for having lower GCNs of C4B (1.59±0.65) than other IIM subgroups. In the other three subgroups, lower C4T GCN was primarily attributable to the decreased GCN of C4A.

As shown in figure 3, the effect sizes of C4T=2+3, C4A=0+1 and C4L=0+1+2 on IIM subgroups were similar, with ORs ranging between 2.1 and 3.7. Low C4T GCN had the greatest impact on IBM with OR=3.70 (2.72–5.04). Low C4A GCN had the largest impact on PM with OR=3.16 (2.65–3.75). Low C4L GCN had largest effects on PM and IBM, with ORs of 2.80 and 2.88, respectively.

C4 GCN variations among patients with IIM with and without MSA or MAA
We compared the mean age at diagnosis, sex and C4 GCN variations between patients with IIM with and without myositis-related autoantibodies (table 3). Patients with anti-Jo1, anti-PM/Scl and MAA in general had younger age of disease diagnosis between 43 and 49 years old. Patients with IIM who tested positive for MSA or MAA were more likely to be women (70%–75%). Patients with anti-Jo1 and anti-PM/Scl consistently had the lowest mean GCNs of C4T, C4A and C4L.

Except for anti-Jo1, patients with MSA presented with similar C4 or C3 plasma protein concentrations than those without. In contrast, patients with MAA had significantly lower levels of C4 and C3 than those without MAA (C4: 275.1±100.0 vs 330.9±105.4 mg/L, p=2.1×10−12; C3: 1188.0±309.8 vs 1335.2±283.9 mg/L, p=1.2×10−8). With regards to specific autoantibodies, patients with anti-PM/Scl and anti-Ro each had significantly lower C4 and C3 protein levels than those without these autoantibodies (figure 4A,B). Patients with MAA (83.6±30.6 vs 98.0±31.6 mg/L, p=1.1×10−7), anti-PM/Scl (86.1±26.7 vs 95.2±31.1 mg/L, p=0.03) and anti-Ro (85.0±30.1 vs 95.4±30.8 mg/L, p=0.014) had significantly lower C4 protein yield per gene (C4P/G) than those without these autoantibodies. No significant differences were observed on plasma protein levels of C4 and C4P/G between women and men (figure 4D,E).

Logistic regression analyses of HLA-DRB1*03 and C4A deficiency in genetic risk of IIM and IIM-related autoantibodies
Among healthy control subjects, 26.1% were HLA-DRB1*03 positive, compared with 56.1% in patients with IIM, which translated into an OR of 3.68 (2.94–4.60, p=2.6×10−12) in IIM (table 4). The distribution of HLA-DRB1*03 was uneven among subgroups of IIM, which varied from 75.4% in patients with...
Here we investigated complement C4 genetic diversity in patients with IIM of European descent and matched healthy controls. Our data consistently showed that low copy numbers of C4T and C4L, and C4A deficiency are highly significant risk factors for IIM and its major subgroups, with medium to large effect sizes or ORs between 1.7 and 3.7. Compared with healthy controls, patients with IIM had 0.28 to 0.58 fewer mean gene copies of C4T, C4A or C4L. The C4T=2 group yielded similar risks as the C4T=2+3 group, and the C4A=0 group had similar risk as the C4A=0+1 group. The similar magnitudes of ORs suggested that there were ‘dominant’ effects for low GCN of total C4 (ie, C4T=2 and C4T=2+3) and C4A deficiency (C4A=0 and C4A=0+1) on the risk of IIM, which is analogous to IBM with an OR of 8.71 (5.48–13.8) to 59.5% in patients with PM with an OR of 4.16 (3.15–5.48), 47.6% in patients with DM with an OR of 2.57 (1.90–3.49) and 45.5% in patients with JDM with an OR of 2.36 (1.56–3.79).

We performed logistic regression to investigate the relative roles of C4A deficiency and HLA-DRB1*03 as independent risk factors for IIM and subgroups. The results are shown in table 4. It was found that (1) C4A deficiency and C4 gene size variation were independent risk predictors of JDM and DM and (2) HLA-DRB1*03 and C4A deficiency and GCN of C4T were independent risk factors for PM and IBM. Moreover, HLA-DRB1*03 and C4A deficiency interacted to increase the risk of PM. We also performed intragroup logistic regression analyses to identify independent predictors of IIM-related autoantibodies. Complement C4 or C3 protein or C4P/G, HLA-DRB1*03 and/or HLA-DRB1*15, C4A deficiency or C4A GCN range of variations were risk factors for various myositis-associated autoantibodies except for MSA in general. For patients with MSA, genetic factors such as HLA-DRB1*03, GCNs of C4B and C4L were independent predictors.

**DISCUSSION**

Continued
to when homozygous and heterozygous mutants exhibit the same phenotype in Mendelian genetics. Such phenomena are in stark contrast to those observed in the genetics of human systemic lupus erythematosus (SLE), in which low GCNs of C4T or homozygous C4A deficiency (C4T=2, OR=6.51; C4A=0, OR=5.27) exerted substantially greater risks than those with C4T=3 (OR=1.32) or heterozygous C4A deficiency (C4A=1, OR=1.61).23 24 46 Parallel analyses of C4 structural variants between cases and controls recruited from each geographic location yielded similar results as presented for the entire IIM cohort, which are analogous to replication studies (online supplemental table S1).

Complement-mediated destruction leading to vasculopathy in dermatomyositis has been well-established,1 6 47 48 and we and others have demonstrated C4A genetic deficiency or low GCN of C4T or homozygous C4A deficiency (C4T=2, OR=6.51; C4A=0, OR=5.27) exerted substantially greater risks than those with C4T=3 (OR=1.32) or heterozygous C4A deficiency (C4A=1, OR=1.61).23 24 46 Parallel analyses of C4 structural variants between cases and controls recruited from each geographic location yielded similar results as presented for the entire IIM cohort, which are analogous to replication studies (online supplemental table S1).

Complement-mediated destruction leading to vasculopathy in dermatomyositis has been well-established,1 6 47 48 and we and others have demonstrated C4A genetic deficiency or low GCN of C4T or homozygous C4A deficiency (C4T=2, OR=6.51; C4A=0, OR=5.27) exerted substantially greater risks than those with C4T=3 (OR=1.32) or heterozygous C4A deficiency (C4A=1, OR=1.61).23 24 46 Parallel analyses of C4 structural variants between cases and controls recruited from each geographic location yielded similar results as presented for the entire IIM cohort, which are analogous to replication studies (online supplemental table S1).

The prevalence of low GCNs of C4T and C4L, C4A deficiency, and the presence of myositis-related autoantibodies in these diseases suggests that additional humoral immune effectors play a role in the pathophysiology of PM and IBM. IBM is unique as it has low GCNs in C4T, C4A and C4B. In a study of anti-Ro/anti-La patients with autoimmune diseases including myositis, Lundtoft and colleagues observed low GCNs of C4A in Scandinavian patients.49

It is worthy pointing out that the effects of GCN variation for C4S and C4B were opposite to those of C4L or C4A in JDM, DM and PM, which suggests different functions of C4S and C4B compared with C4L and C4A. Indeed, short C4 genes associate with higher C4 protein production50 51 and activated C4B protein generates faster activation of complement pathways52; long C4 genes associate with attenuated C4 protein production but possibly engage in antisense defence against viral infections.26 27 50 Moreover, activated C4A has greater efficiency to bind to immune complexes for clearance and protection against autoimmunity.26 27 50–53 We postulate that activated C4A and C4B proteins interfere and balance each other’s effects physiologically to achieve optimum defence against infections and autoimmunity, and mitigate collateral damages due to complement-mediated injuries of self-tissue.26 52 54

While IIM typically does not feature dramatic longitudinal fluctuations of plasma C3 and C4 protein levels with disease activity as is the case in SLE,54 intragroup analyses revealed that patients with IIM with anti-Jo-1, MAA in
general, anti-PM/Scl or anti-Ro had significantly lower mean complement protein levels than those without these autoantibodies. Immune complexes formed by autoantibodies and self-antigens in IIM could activate and consume complement, leading to ‘depressed’ C4 and C3 plasma protein levels that were seen here and by others. Lower GCNs of C4T, C4L or C4A in IIM would be among the causes for lower C4 protein levels. Data on C4P/G yield, elevated levels of activation products in the plasma such as C4a, C3a and C5a, or cell-bound complement inactivation products such as erythrocyte-C4d and erythrocyte-C3d would help distinguish whether lower protein levels are due to genetic insufficiency or protein turnover. 33-36 It is of interest to note that except for Jo-1, most MSA were not associated with lower complement levels in circulation, although MAA did. Moreover, patients with JDM have MSA such as anti-TIF1γ and anti-NXP2 37 38 and their relationship with complement activation is yet to be investigated.

In a study of complement in schizophrenia, SLE and Sjögren’s syndrome, it was suggested that C4 exhibited a sex-biased expression differences including in cerebrospinal fluid.39 40 41 We did not detect differential expression of C4 protein in EDTA-plasma between men and women among patients with IIM in this work or in previous studies.42 43 44 We did not detect differences in C4 GCN variations between female and male patients for DM, PM and JDM (online supplemental table S2). However, IBM is a male-dominant disease and we observed slightly higher frequencies of low GCNs for C4L, lower proportions of C4A or C4L among C4T in men compared with women.

The relative roles of HLA class II variants including DRB1*03, DQA1*05, DQB1*02 and C4A deficiency on genetic predisposition to autoimmune diseases such as IIM are an unsolved enigma. 17 47 59 Multivariate logistic regression analyses revealed that C4A deficiency was an independent risk factor for DM and JDM and that HLA-DRB1*03 was a prominent risk factor for IBM, while C4A deficiency and HLA-DRB1*03 contributed independently and interactively to an increased risk of PM. Further analyses of DRB1, DQA1, DQB1 variants and GCNs of C4 revealed the presence of both risk and protective factors in each gene on the predisposition of IIM subgroups and autoantibodies (online supplemental figure S3 and online supplemental table S4).

In summary, our results demonstrated that low GCNs for C4T, C4A and C4L played significant roles in increasing the risk of IIM. The relationship between C4A deficiency and HLA-DRB1*03, which are closely linked, is complex and intriguing. It will be important going forward to carefully interrogate the mechanisms by which HLA-DRB1*03 and C4A deficiency contribute to autoimmunity and IIM. Finally, intragroup analyses showed that patients with IIM with certain autoantibodies presented with lower protein levels of complement C3 and C4. This effect was more notable for MAA than for MSA, which is worthy of investigations. Our findings have broad implications in the assessment and treatment of IIM and other autoimmune diseases.

### Table 4 Logistic regression models for genetic predictors in IIM, subgroups and autoantibodies

<table>
<thead>
<tr>
<th>Predictors</th>
<th>$\chi^2$</th>
<th>OR (95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Genetic predictors for IIM and subgroups</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IIM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HLA-DQB1*03</td>
<td>7.32</td>
<td>1.04 (0.77 to 1.41)</td>
<td>0.65</td>
</tr>
<tr>
<td>C4A=0+1</td>
<td>2.46</td>
<td>1.06 (0.91 to 1.24)</td>
<td>0.29</td>
</tr>
<tr>
<td>C4A=0+2</td>
<td>5.67</td>
<td>1.65 (1.29 to 2.12)</td>
<td>0.0002</td>
</tr>
<tr>
<td>DM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HLA-DQB1*03</td>
<td>3.92</td>
<td>1.45 (1.13 to 1.85)</td>
<td>0.0012</td>
</tr>
<tr>
<td>C4A=0+1</td>
<td>4.38</td>
<td>1.64 (1.13 to 2.39)</td>
<td>0.006</td>
</tr>
<tr>
<td>C4A=0+2</td>
<td>9.49</td>
<td>3.08 (2.17 to 4.40)</td>
<td>1.94E-3</td>
</tr>
<tr>
<td>PM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HLA-DQB1*03</td>
<td>10.7</td>
<td>3.22 (2.39 to 4.34)</td>
<td>2.87E-3</td>
</tr>
<tr>
<td>C4A=0+1</td>
<td>11.8</td>
<td>3.39 (2.32 to 5.00)</td>
<td>2.5E-3</td>
</tr>
<tr>
<td>C4A=0+2</td>
<td>18.9</td>
<td>6.19 (4.15 to 9.34)</td>
<td>1.02E-4</td>
</tr>
<tr>
<td><strong>B. Predictors for myositis-related autoantibodies in patients with IIM</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predictors</td>
<td>$\chi^2$</td>
<td>OR (95% CI)</td>
<td>P value</td>
</tr>
<tr>
<td>------------</td>
<td>---------</td>
<td>-------------</td>
<td>---------</td>
</tr>
<tr>
<td>MSA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HLA-DQB1*03</td>
<td>3.92</td>
<td>1.45 (1.13 to 1.85)</td>
<td>0.0012</td>
</tr>
<tr>
<td>C4A=0+1</td>
<td>5.50</td>
<td>2.35 (1.49 to 3.73)</td>
<td>5.7E-3</td>
</tr>
<tr>
<td>C4A=0+2</td>
<td>11.5</td>
<td>3.11 (1.81 to 5.32)</td>
<td>1.3E-3</td>
</tr>
<tr>
<td>JDM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HLA-DQB1*03</td>
<td>3.92</td>
<td>1.45 (1.13 to 1.85)</td>
<td>0.0012</td>
</tr>
<tr>
<td>C4A=0+1</td>
<td>5.50</td>
<td>2.35 (1.49 to 3.73)</td>
<td>5.7E-3</td>
</tr>
<tr>
<td>C4A=0+2</td>
<td>11.5</td>
<td>3.11 (1.81 to 5.32)</td>
<td>1.3E-3</td>
</tr>
</tbody>
</table>

**Double asterisks (**) between two predictors indicated interactions.**

AUC, area under the curve; C4, acidic isoform of complement C4; C4B, basic isoform of complement C4; C4L, long form of C4 gene with human endogenous retrovirus HERV-K(C4L); C4P, G protein per cell copy; C4T, short form of C4 gene without integration of the retrovirus HERV-K(C4T); C4, total copy number of C4 gene; DM, dermatomyositis; GCN, gene copy number; HLA, human leukocyte antigen; IIM, inclusion body myositis; IBM, idiopathic inflammatory myositis; JDM, juvenile dermatomyositis; MAA, myasthenia-associated autoantibodies; MSA, myasthenia-specific autoantibodies; PM, polymyositis; PM/Scl, polymyositis/scleroderma; Ro, Scl-70; SSc, systemic sclerosis.
Myositis

1 Division of Informatics, Imaging and Data Sciences, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK
2 Faculty of Science and Engineering, Manchester Metropolitan University, Manchester, UK
3 Environmental Autoimmunity Group, Clinical Research Branch, National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health, Bethesda, MD, USA
4 Veteran’s Affairs Medical Center, University of Oklahoma Health Sciences Center, and Oklahoma Medical Research Foundation, Oklahoma City, OK, USA
5 Department of Transfusion Medicine, NIH Clinical Center, National Institutes of Health, Bethesda, MD, USA
6 University of Mississippi Medical Center, Jackson, Mississippi, USA
7 Department of Microbiology and Immunology, Loyola University Chicago, Maywood, IL, USA
8 Division of Biostatistics, The Ohio State University, Columbus, Ohio, USA
9 Division of Population Health, Health Services Research and Primary Care, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK

Twitter Hector Chinoy @drhetchinoy

Acknowledgements We are indebted to volunteers and patients with myositis who contributed blood samples for this study, and to Bi Zhou and Alex Sepeda for technical assistance. We thank Drs Daniella Schwartz and Pravitt Gough for valuable comments on the manuscript.

Contributors Project design: C-YY, DZ, JAL, SR, LP, IEL, HNN, YLYW, SA, CHS, JV, LGR, and FWM. Experimental studies: DZ, EHK, KEM, AD, GRY, SC, YLYW and C-YY performed experiments and collected research data on C4 GGN variations and protein levels in patients with IM and controls. SR, YLYW, KEM, EL, AP, WED and NAF performed experiments or bioinformatic analysis to determine HLA-DRB1 genetic polymorphisms. Subject/patient diagnosis, recruitment and clinical studies: RAA, YY, CH, GMM, DZ, EHK, SC, LGR, FWM, IEL, LP, IV, JAL and C-YY contributed to subject/patient identifications, recruitment and assembly of study cohorts for collaborative studies. C-YY, DZ, EHK, SC, LGR, FWM and JAL drafted the initial manuscript. All authors participated in revisions and approved the current version of this manuscript. Drs Daniele Zhou and Chack-Yung Yu take the responsibility for the overall content and data integrity.

Funding This study is supported in parts by: National Institutes of Health (NIH) grants R21 AR070509, R01 AR073311 from the National Institute of Arthritis, Musculoskeletal and Skin Diseases/Unicef Kennedy Shriver National Institute of Child Health Development/National Institute of Allergy and Infectious Diseases (C-YY), and a research grant from the CureJM Foundation (C-YY); the Intramural Research Program of the National Institute of Environmental Health Sciences of the NIH (FWM and LGR); funding from the Medical Research Council (IRM/N003322/1), Versus Arthritis Programme Grant 18474, and Myositis UK (HC and JAL). JV acknowledges funding from the European Reference Network for Rare Neuromuscular Diseases EURO-NMDC. This report includes independent research supported by the NIHR Biomedical Research Centre Funding Scheme.

Disclaimer The views expressed in this publication are those of the authors and are not necessarily those of the NIH, the National Institute for Health Research or the Department of Health in the UK.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, conduct, reporting or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by the Institutional Review Board (IRB) at the Nationwide Children’s Hospital, and the Committee of the Myositis Genetics Consortium (MYOGEN). Our IRB approval number is IRB14-00544. MHC and CNVs in Disease Associations. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: creativecommons.org/licenses/by/4.0/.

ORCID iDs
Danlei Zhou http://orcid.org/0000-0001-9061-1224
Hector Chinoy http://orcid.org/0000-0001-6492-1288
Frederick W Miller http://orcid.org/0000-0003-2831-9593
Ingrid E Lundberg http://orcid.org/0000-0002-6068-9212
Leonid Padyukov http://orcid.org/0000-0003-2950-5670
Janine A Lamb http://orcid.org/0000-0002-7248-0591
Chack-Yung Yu http://orcid.org/0000-0002-5218-7503

REFERENCES


36.

35.

34.

33.

32.

31.

30.

29.

28.

27.

26.

25.

24.

23.

22.

21.

20.

19.

18.

17.

16.

15.

14.

13.

12.

11.

10.

9.

8.

7.

6.

5.

4.

3.

2.

1.

...