

**ASSOCIATION WITH HLA-DR β 1 POSITION 37 DISTINGUISHES JUVENILE
DERMATOMYOSITIS FROM ADULT-ONSET MYOSITIS**

Authors: Claire T. Deakin,^{*1,2,3} John Bowes,^{4,5} Lisa G. Rider,⁶ Frederick W. Miller,⁶ Lauren M. Pachman,⁷ Helga Sanner,^{8,9} Kelly Rouster-Stevens,¹⁰ Gulnara Mamyrova,¹¹ Rodolfo Curiel,¹¹ Brian M. Feldman,¹² Adam M. Huber,¹³ Ann M. Reed,¹⁴ Heinrike Schmeling,¹⁵ Charlotte G. Cook,¹ Lucy R. Marshall,^{1,2,3} Meredyth G. Ll. Wilkinson,^{1,2,3} Stephen Eyre,^{4,5} Soumya Raychaudhuri,^{#4,5,16,17} and Lucy R. Wedderburn,^{#1,2,3} on behalf of the Juvenile Dermatomyositis Cohort and Biomarker Study, the Childhood Myositis Heterogeneity Study Group, and the Myositis Genetics Consortium (MYOGEN)

*Corresponding author: Dr Claire T. Deakin

Address: Infection, Immunity and Inflammation Research and Teaching Department

UCL Great Ormond Street Institute of Child Health

30 Guilford Street

London WC1N 1EH

United Kingdom

Telephone: +44 207 905 2671

Email: c.deakin@ucl.ac.uk

1. Infection, Immunity and Inflammation Research and Teaching Department, UCL Great Ormond Street Institute of Child Health, London, UK
2. Centre for Adolescent Rheumatology Versus Arthritis at UCL, UCL Hospital and Great Ormond Street Hospital, London, UK
3. NIHR Biomedical Research Centre at Great Ormond Street Hospital, London, UK
4. Centre for Genetics and Genomics Versus Arthritis, Centre for Musculoskeletal Research, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK
5. National Institute of Health Research Manchester Biomedical Research Centre, Manchester Academic Health Science Centre, Manchester University NHS Foundation Trust, Manchester, UK
6. Environmental Autoimmunity Group, Clinical Research Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Bethesda, Maryland, USA
7. Ann & Robert H. Lurie Children's Hospital of Chicago, Northwestern University, Feinberg School of Medicine, Chicago, Illinois, USA
8. Department of Rheumatology, University of Oslo, Oslo, Norway
9. Oslo New University College, Oslo, Norway
10. Emory University School of Medicine, Atlanta, Georgia, USA
11. Division of Rheumatology, George Washington University School of Medicine and Health Sciences, Washington, DC, USA
12. Division of Rheumatology, Department of Pediatrics, The Hospital for Sick Children, Toronto, Ontario, Canada
13. IWK Health Centre and Dalhousie University, Halifax, Nova Scotia, Canada
14. Pediatrics, Duke University, Durham, North Carolina, USA
15. Alberta Children's Hospital and Department of Pediatrics, University of Calgary, Calgary, Alberta, Canada

16. Department of Medicine, Brigham and Women's Hospital and Harvard Medical School,
Boston, Massachusetts, USA

17. Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge,
Massachusetts, USA

The authors wish it to be known that, in their opinion, the final 2 authors should be regarded as joint

Senior Authors

UNCORRECTED MANUSCRIPT

ABSTRACT

Objectives:

Juvenile dermatomyositis (JDM) is a rare, severe autoimmune disease and the most common idiopathic inflammatory myopathy (IIM) of children. JDM and adult-onset dermatomyositis (DM) have similar clinical, biological and serological features, although these features differ in prevalence between childhood-onset and adult-onset disease, suggesting age of disease onset may influence pathogenesis. Therefore, a JDM-focused genetic analysis was performed using the largest collection of JDM samples to date.

Methods:

Caucasian JDM samples (n=952) obtained via international collaboration were genotyped using the Illumina HumanCoreExome chip. Additional non-assayed HLA loci and genome-wide SNPs were imputed.

Results:

*HLA-DRB1*03:01* was confirmed as the classical HLA allele most strongly associated with JDM (OR 1.66; 95% CI 1.46, 1.89; $p=1.4\times 10^{-14}$), with an independent association at *HLA-C*02:02* (OR=1.74; 95% CI 1.42, 2.13, $p=7.13\times 10^{-8}$). Analyses of amino acid positions within *HLA-DRB1* indicated the strongest association was at position 37 (omnibus $p=3.3\times 10^{-19}$), with suggestive evidence this association was independent of position 74 (omnibus $p=5.1\times 10^{-5}$), the position most strongly associated with adult-onset DM. Conditional analyses also suggested the association at position 37 of *HLA-DRB1* was independent of some alleles of the Caucasian HLA 8.1 ancestral haplotype (AH8.1) such as *HLA-DQB1*02:01* (OR=1.62; 95% CI 1.36, 1.93; $p=8.70\times 10^{-8}$), but not *HLA-DRB1*03:01* (OR=1.49; 95% CR 1.24, 1.80; $p=2.24\times 10^{-5}$). No associations outside the HLA region were identified.

Conclusions:

Our findings confirm previous associations with AH8.1 and *HLA-DRB1*03:01*, *HLA-C*02:02* and identify a novel association with amino acid position 37 within *HLA-DRB1* which may distinguish JDM from adult DM.

UNCORRECTED MANUSCRIPT

INTRODUCTION

Juvenile dermatomyositis (JDM) is a rare, severe autoimmune disease and the most prevalent idiopathic inflammatory myopathy (IIM) with proximal muscle weakness and skin rash as typical features. Clinical features are heterogeneous, and can include serious complications such as calcinosis, ulceration, treatment-resistant rash and involvement of major organs, including gut, lungs and brain. Although some patients achieve remission following standard disease management, which consists of long-term immunosuppression with glucocorticoids, methotrexate and other medications, others respond poorly.

While JDM and adult-onset dermatomyositis (DM) share similar clinical and biological features, there are differences in prevalence of clinical features(1). The incidence of JDM is approximately one-tenth of the incidence of DM(2). DM can be associated with cancer, but this has not been reported in JDM. Conversely, calcinosis is a major cause of morbidity in JDM, but has a lower prevalence in DM. The prevalence of myositis-specific autoantibodies (MSA), which are linked to different clinical features of disease, also differs between the adult and juvenile forms of the disease. Anti-NXP2 is one of the more abundant MSA in JDM (reported in 20-25% of patients (3–5)), but has a lower prevalence in DM (reported in 1.6%-17% of different patient populations (6–8)). The most prevalent MSA in DM, anti-Jo-1, is rare in JDM. These differences in the distribution of MSA and clinical features suggest an influential role for age of disease onset in the pathogenesis of disease. However, little is known at the mechanistic level about the influence of age on JDM phenotypes and pathogenesis. Knowledge about how disease mechanisms differ between patient subgroups and interact with patient age to result in different complications may enable targeting of novel molecular pathways, more accurate modelling of life-long risk and more stratified therapeutic approaches to address this risk.

Candidate gene and genome-wide studies of myositis have established the strongest genetic association within the Caucasian 8.1 ancestral haplotype (AH8.1; HLA A1-B8-DR3-DQ2) of the major histocompatibility complex (MHC), also associated with many other immune-mediated diseases (9–13). Distinct human leukocyte antigen (HLA) alleles have been identified as associated with serological

subphenotypes of myositis in different ethnic populations. Most notably associations between the development of anti-Jo-1 autoantibodies and *HLA-DRB1*03:01*, *HLA-DQB1*02:01* and *HLA-B*08*, consistent with AH8.1, have been identified in Caucasian and African-American patients (14–16). In adult myositis, gene-environment interactions have been found between *HLA-DRB1*03*, smoking and the presence of anti-Jo-1 autoantibodies, and between *HLA-DRB1*11:01* and anti-HMGCR-positive statin-induced immune-mediated necrotising myopathy (17,18). To date, the rarity of JDM has meant that candidate gene studies in JDM have been small and sub-group analyses of JDM in genome-wide studies have had limited statistical power relative to other myositis phenotypes. The aim of this research was to identify novel genetic loci associated with JDM using a larger cohort of patients.

RESULTS

Samples and genotyping quality control

Samples of Caucasian ancestry (n=952) were obtained via international collaboration including samples from the UK Juvenile Dermatomyositis Cohort & Biomarker Study (JDCBS), the Childhood Myositis Heterogeneity Study Group, and the Myositis Genetics Consortium (MYOGEN; Table 1). Many of these cases have contributed to previous analyses (10,19). Demographic features are described in Supplementary Table S1. After genotyping and stringent QC, n=178,164 SNPs remained (Supplementary Table S2) on n=851 JDM samples (Supplementary Tables S3 & S4). The proportion of phenotypic variance explained by these markers was estimated as 0.18 (0.02).

*HLA-DRB1*03:01* confirmed as allele most strongly associated with JDM

Case-control analysis of assayed SNPs confirmed the strongest association with JDM was within the HLA region (Figure 1; Supplementary Figure S1). Analysis of imputed markers within the HLA region indicated the strongest association was with SNP rs3117103 (OR=1.87; 95% CI 1.64, 2.13; $p=1.79 \times 10^{-20}$) and the classical allele *HLA-DRB1*03:01*, consistent with previous reports (9,10). The OR for *HLA-DRB1*03:01* itself was 1.66 (95% CI 1.46, 1.89; $p=1.4 \times 10^{-14}$; Supplementary Table S5).

Conditioning on this allele, an independent association with *HLA-C*02:02* was identified (OR=1.74; 95% CI 1.42, 2.13, $p=7.13\times 10^{-8}$) as reported previously (10). The next most associated allele was *HLA-B*44:02*, but the p-value for this association was above the threshold for statistical significance (OR=1.31, 95% CI 1.15, 1.50; $p=6.04\times 10^{-5}$). No alleles within *HLA-A* were associated with JDM (all p-values above 0.08).

Analysis of amino acid positions within *HLA-DRB1* identifies position 37 as the most significantly associated with JDM

Since the strongest association with JDM was identified within *HLA-DRB1*, further analysis sought to resolve this association at the level of amino acid positions. Position 37 was the most strongly associated with disease (omnibus $p=3.3\times 10^{-19}$; Figure 2 & Table 2). Residues at position 37 are located within the P9 pocket of the antigen binding groove (Supplementary Figure S2) (20,21). Relative to Tyr-37, Ser-37 was the most significantly associated residue at this position (OR=0.65; 95% CI 0.57, 0.75; $p=7.34\times 10^{-10}$), followed by Phe-37 (OR=0.63; 95% CI 0.54, 0.75; $p=1.14\times 10^{-7}$) (Table 3). Asn-37, which is found on *HLA-DRB1*03:01*, was not associated relative to Tyr-37 (OR=1.13; 95% CI 1.06, 1.19; $p=0.05$). Conditioning on position 37, there was suggestive evidence of possible independent associations at positions 74 ($p=5.1\times 10^{-5}$) and 26 ($p=5.9\times 10^{-5}$), although the p-values for these associations were not significant (Table 2).

To evaluate whether position 37 explained the association within *HLA-DRB1* more convincingly than the classical *HLA-DRB1*03:01* allele and AH8.1, conditional analyses were performed. Conditioning on *HLA-DRB1*03:01*, the association of position 37 was above the threshold for significance (omnibus $p=4.42\times 10^{-5}$), although there may be weak evidence of an independent association. After conditioning on all residues at position 37, there was evidence of independent effects for *HLA-DQB1*02:01* within AH8.1 (OR=1.62; 95% CI 1.36, 1.93; $p=8.70\times 10^{-8}$) and *HLA-C*02:02* (OR=1.72; 95% CI 1.40, 2.10; $p=1.58\times 10^{-7}$). However, the effect for *HLA-DRB1*03:01* did not meet the threshold for significance (OR=1.49; 95% CR 1.24, 1.80; $p=2.24\times 10^{-5}$). Taken together, these results indicate an effect of position

37 that is independent of some of the previously reported effects within AH8.1 and at *HLA-C*02:02*, but that is not independent of *HLA-DRB1*03:01*.

Genome-wide imputation identifies possible loci associated at a suggestive level of significance

Following genome-wide imputation, there were no additional loci identified at the genome-wide level of statistical significance ($p=5\times 10^{-8}$). Two loci with $p<1\times 10^{-6}$ and minor allele frequency (MAF) over 0.05 were identified (Table 4; Figure 3). rs6501160 is an intronic variant within *TMEM114*, a transmembrane protein. rs6892006 is an intronic variant within *MEF2C-AS1*, an anti-sense RNA gene (Supplementary Figure S3). Associations with clinical phenotypes have not been reported for either variant. Although additional loci were identified at $p<5\times 10^{-8}$ (Supplementary Table S6, Supplementary Figure S4), these had MAF between 0.01 and 0.05 and may represent imputation artefacts.

DISCUSSION

This analysis represents the largest international genetic study of JDM to date. We found that the strongest association was in the HLA region within *HLA-DRB1*03:01*, consistent with previous studies (9,10,19). We confirmed an independent association with *HLA-C*02:02*, and found some evidence of a possible additional independent association at *HLA-B*44:02*, which would need to be confirmed in a future study. Interestingly, we did not find evidence of an association with *HLA-A*, even though associations within this gene are well known for multiple other autoimmune diseases including rheumatoid arthritis (RA), juvenile idiopathic arthritis (JIA), psoriatic arthritis and type I diabetes (22–25).

In this analysis focused on juvenile-onset disease, we found amino acid position 37 within *HLA-DRB1* had the strongest association with JDM, with Ser-37 and Phe-37 having protective effects relative to Tyr-37. We also found evidence that the association with position 37 was independent of AH8.1, a well-established association with myositis, but there was only weak evidence of independence from *HLA-DRB1*03:01*, and so this needs to be followed up in future studies with greater numbers of both

juvenile-onset and adult-onset patients using the same genotyping chip. Amino acid position 37 is within the P9 pocket of the antigen-binding groove (20,21). Substitution of Ser-37 to Tyr-37 has been shown to be sufficient to alter alloantigenicity and stimulate a T cell response (26). Tyr-37 promotes a stronger response to streptococcal protein (27).

The previous ImmunoChip analysis of adult and juvenile myositis combined identified the strongest association within *HLA-DRB1* at amino acid position 74 (10). Subsequent analyses of that dataset identified position 74 as having an association with anti-Jo1, anti-PM/Scl and anti-cN1A autoantibodies (28), which are less prevalent in juvenile-onset disease. We found weak evidence of independent association at position 74 after conditioning on position 37, above the threshold for significance and so this finding needs to be followed up with greater numbers of patients. Nonetheless, it raises an intriguing possibility that children and adults may have different dominant autoantigenic peptides being presented in the antigen-binding groove, although we recognise multiple antigens with different allelic associations are likely to be involved in disease. While the ImmunoChip analyses benefitted from a relatively large cohort (n=2,544), the number of JDM patients was approximately half that of our analysis (n=493). These patients are represented in our analysis, along with additional cases. It may be that the effects of juvenile-onset disease were obscured by larger number of adult-onset patients and other phenotypes present in the combined cohort. Future work and greater numbers are required to better dissect the genetic influence on age effects in juvenile and adult-onset disease, as well as on clinical and serological heterogeneity.

Associations between other positions within *HLA-DRB1* and autoimmune disease have been reported. For example, position 11 has the strongest effect in RA, with independent effects at positions 71 and 74 (29). In JIA, position 13 has the strongest effect, although in systemic JIA the strongest effect is at position 58 (23,30). Positions 71/74 have also been implicated in anti-fibrillarin-positive systemic sclerosis, Crohn's disease, multiple sclerosis, type I diabetes and Grave's disease (31–35). In systemic lupus erythematosus (SLE), Ser-1, Phe47 and Ala71 are associated with disease (36). Position 37 has been found to be associated with primary sclerosing cholangitis, ulcerative colitis in Asians, ACPA⁺ RA in Han Chinese, SLE in Asians, and psoriasis vulgaris in Taiwanese (20,21,37–39). Interestingly, a

recent report of Japanese patients with RA identified an association between position 37 and younger age of disease onset (defined as 16-30 years of age), but not with older age of onset (defined as over 60 years of age) (40). At present it is not well understood how these differing positional effects relate to the spectrum of phenotypes represented in autoimmune disease, and how these effects interact with ethnicity and age of disease onset.

The major limitation of our study is limited statistical power for identifying associations outside the HLA region, which is a practical challenge for rare disease research. Although our combined cohort represents the largest-ever assembled international cohort of Caucasian patients with JDM, it is a relatively small sample size for a genetic study. As such, the findings of suggestive associations will need to be confirmed in future studies with greater numbers of patients. Nonetheless, we were able to confirm the findings of previous studies in a larger cohort dedicated to JDM. Future studies with greater numbers may better define the relationship between the associations at *HLA-DRB1*03:01* and amino acid positions within *HLA-DRB1* such as position 37. Loci outside the HLA region identified at genome-wide and suggestive levels of significance in previous analyses (including *PTPN22*, *UBE2L3*, *CD28*, *TRAF6*, *STAT4*) were not replicated in this study, although signals at these loci were not specifically tested for. This may reflect the smaller sample size of this study and the dominance of adult patients in those analyses whose characteristics may differ from paediatric disease (10). JDM is a heterogeneous phenotype, however, the small sample size restricted our ability to analyse more clinically homogeneous subgroups, such as autoantibody subgroups. It may be that future developments in methodologies for genetic analysis will enable further insights to be derived from this dataset. Many of the patients in this study overlapped with a previous analysis, which identified associations with major MSA subtypes, including anti-Jo-1, anti-PM/Scl, anti-cN1A, anti-Mi-2 and anti-TIF1 γ (28)(28)(28)(28)(28). As more patients become available for inclusion in genetic studies, analyses of further MSA subtypes may become possible. It will also be critical to study genetic associations with MSA subtypes in different ethnic populations.

Analysis following genome-wide imputation identified two possible suggestive loci, an intronic variant within *TMEM114* and an intronic variant within *MEF2C-AS1*. *TMEM114* is a glycosylated

transmembrane protein, and knowledge of its cellular function is limited, although missense mutations in this gene and a chromosomal translocation in its promoter are associated with congenital and juvenile cataract disorders, respectively (41). *MEF2C-AS1* is a non-coding anti-sense RNA gene with no known function. It is unclear how these loci relate to JDM, but it may be these alleles function as epigenetic marks.

In summary, we have confirmed the association between JDM and *HLA-DRB1*03:01* and shown that within *HLA-DRB1*, position 37 is most strongly associated with disease in a population of patients with juvenile-onset myositis.

MATERIALS AND METHODS

Genotyping and genotype calling

Genotyping of Caucasian JDM samples (n=952) was performed using the Illumina HumanCoreExome chip across 3 batches at a single centre (University College London). This cohort included the majority of JDM cases included in previous analyses, as well as additional cases recruited subsequent to those analyses including cases from more centres (10,19). GenomeStudio 2.0 (version 2.0.4 of the Genotyping Module and the GenTrain 3.0 Cluster Algorithm) was used for genotype clustering and calling for each separate batch. Samples with less than 90% call rate were excluded, and genotype clustering and calling were repeated, before data were exported in PLINK format for quality control.

Data on healthy individuals (n=12,474) of European ancestry who had also been genotyped using the HumanCoreExome chip as part of the International Age-Related Macular Degeneration Genetic Consortium (IAMDGC) were obtained from dbGaP (dbGAP Study Accession phs001039.v1.p1).

Quality control

Quality control (QC) of markers and samples was performed as described previously (24). The following steps were done separately for each of the 3 batches of JDM data and the control data using

PLINK 1.07 (42). Mitochondrial and Y chromosome SNPs were excluded. SNPs were also excluded if they had elevated missing rates (over 2%), had a MAF below 1% and deviated significantly from Hardy-Weinberg equilibrium (HWE; threshold of $p < 0.0001$). Samples with elevated missing genotypes (over 5%) and outlying heterozygosity rates (above or below 5 standard deviations from the mean rate). Data were aligned to the Haplotype Reference Consortium (HRC) before merging for further quality control using the HRC checking tool (43).

Duplicated or related individuals in the merged dataset were identified using identity-by-descent, performed on a linkage disequilibrium (LD)-pruned dataset of 65,862 SNPs with minor allele frequency over 5%. A PI_HAT threshold of 0.2 was used, with the individual with the most missing data excluded. Principal component analysis (PCA) was performed using PLINK 1.9 (44), to evaluate population stratification in the LD-pruned data merged with the International HapMap 3 data (45). Cases and controls were retained if they were within 10 standard deviations of the mean value for the first 2 principal components (PCs) for the HapMap CEU population (Figure S5).

Following quality control, there were 178,164 SNPs for $n=851$ JDM cases and $n=12,232$ controls (Supplementary Tables S1-S4).

Analysis of assayed markers

Autosomal markers were analysed for association with JDM using logistic regression in PLINK with adjustment for the first 10 PCs to control for population stratification. Manhattan plots were generated using R version 6.3 and the “ggplot2” package (version 3.3.2). The proportion of phenotypic variance explained by the SNPs, and a standard error for that estimated proportion, were estimated using the genome-based restricted maximum likelihood (GREML) method using GTAC version 1.93.3, assuming an approximate prevalence of 0.00004 for JDM (2,46,47).

Imputation of HLA loci

Classical HLA alleles, amino acids and SNPs within the HLA region were imputed using SNP2HLA (version 1.0.3) and the Type 1 Diabetes Genetics Consortium reference panel ($n=5,225$) (48). QC of

imputed markers used the following criteria: imputation information score (R^2) over 0.9, MAF over 0.01 and significant departures from HWE in controls ($p < 0.001$). Imputed markers were coded as present or absent.

Analysis of HLA and amino acid positions

Case-control analysis of all imputed HLA was performed assuming an additive model using logistic regression in PLINK, also with adjustment for the first 10 PCs, to identify the most strongly associated locus. Clinical covariates were not adjusted for. For HLA analyses, p-values below 6.8×10^{-6} were considered statistically significant, using a Bonferroni correction for the number of imputed markers as reported previously (49). The 2- and 4-digit classical alleles identified by this locus were subsequently conditioned on in further logistic regression analyses to identify any independent associations.

Amino acid positions were interrogated using likelihood ratio tests (LRTs) in R as follows. At each position, logistic regression models were fitted with all residues in the model, except the residue that was most prevalent in the controls and served as the reference. The first 10 PCs were adjusted for in each of these models. To evaluate whether amino acid positions were associated with disease, likelihood ratio tests were performed to compare the model that was fitted for the residues at each position against a null model, which comprised the first 10 PCs only. The effect of each amino acid position evaluated using a LRT is represented by the LRT p-value, but there is no estimated effect size for the amino acid position generated by this test. Allele frequencies, odds ratios and p-values for amino acid residues at key positions are also reported. Possible independent effects were identified by conditional analysis as above.

Genome-wide imputation and analysis

Genome-wide imputation of SNPs was performed using the Michigan Imputation Server (version 1.2.4) and the HRC reference panel (version r1.1; Supplementary Figure S6) (43,50). Imputed SNPs with MAF below 0.01 or imputation information score below 0.5 were filtered out.

Imputed SNPs were analysed using SNPTEST version 2.5.4-beta3 (51), with adjustment for the first 10 PCs and weighting for the imputation information score to account for imputation uncertainty. Since imputation artefacts are enriched in rare variants, a stringent MAF threshold of 0.05 was used during analysis, although loci with MAF 0.01-0.05 are also reported in Supplementary material.

UNCORRECTED MANUSCRIPT

ACKNOWLEDGEMENTS

This work was funded by a TMA Fellowship and a Cure JM grant to CTD. CTD would like to thank the study administrators and research assistants at each international centre who assisted with extraction of clinical data and preparation and shipment of DNA samples, including: Dr Terry O'Hanlon and Payam Noroozi Farhadi (NIH), Siri Tennebø Flåm (University of Oslo); Gabrielle Morgan (Northwestern University); Dwaraka Veeramreddy (Alberta Children's Hospital); Hayyah Clairman (Hospital for Sick Children, Toronto).

CTD would also like to thank Dr Socrates Varakliotis (Centre for Adolescent Rheumatology Versus Arthritis at UCL, UCLH and GOSH) for setting up remote access to a Linux desktop computer which enabled analyses to continue during the COVID-19 pandemic. CTD thanks Drs. Michael Ombrello and Elaine Remmers for useful comments on the manuscript.

Funding for the UK JDM Cohort and Biomarker study has been provided by generous grants from the Wellcome Trust UK [085860]; Action Medical Research UK [SP4252]; the Myositis Support Group UK, Arthritis Research UK now Versus Arthritis [14518, 20164, 21593]; the Henry Smith Charity and Great Ormond Street Children's Charity [V1268]; Tiny Hearts Society, the Myositis Association, Cure JM, the Medical Research Council [MR/N003322/1], and the National Institute for Health Research (NIHR) via the Translational Research Collaboration (TRC) Rare Diseases and the NIHR-Biomedical Research Centre at GOSH. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

The Juvenile Dermatomyositis Research Group would like to thank all of the patients and their families who contributed to the Juvenile Dermatomyositis Cohort & Biomarker Study & Repository. We thank all local research coordinators and principal investigators who have made this research possible. The JDRG members were as follows:

Dr Kate Armon, Mr Joe Ellis-Gage, Ms Holly Roper, Ms Vanja Briggs and Ms Joanna Watts (Norfolk and Norwich University Hospitals), Dr Liza McCann, Mr Ian Roberts, Dr Eileen Baildam, Ms Louise

Hanna, Ms Olivia Lloyd, Susan Wadeson and Michelle Andrews (The Royal Liverpool Children's Hospital, Alder Hey, Liverpool), Dr Phil Riley, Ms Ann McGovern, Verna Cuthbert (Royal Manchester Children's Hospital, Manchester), Dr Clive Ryder, Mrs. Janis Scott, Mrs. Beverley Thomas, Professor Taunton Southwood, Dr Eslam Al-Abadi and Ruth Howman (Birmingham Children's Hospital, Birmingham), Dr Sue Wyatt, Mrs Gillian Jackson, Dr Mark Wood, Dr Tania Amin, Dr Vanessa VanRooyen and Ms Deborah Burton, Louise Turner & Sarah Hanson (Leeds General Infirmary, Leeds), Dr Joyce Davidson, Dr Janet Gardner-Medwin, Dr Neil Martin, Ms Sue Ferguson, Ms Liz Waxman and Mr Michael Browne, Ms Roisin Boyle, Ms Emily Blyth (The Royal Hospital for Sick Children, Yorkhill, Glasgow), Dr Mark Friswell, Professor Helen Foster, Mrs Alison Swift, Dr Sharmila Jandial, Ms Vicky Stevenson, Ms Debbie Wade, Dr Ethan Sen, Dr Eve Smith, Ms Lisa Qiao, Mr Stuart Watson and Ms Claire Duong (Great North Children's Hospital, Newcastle), Dr Helen Venning, Dr Rangaraj Satyapal, Mrs Elizabeth Stretton, Ms Mary Jordan, Dr Ellen Mosley, Ms Anna Frost, Ms Lindsay Crate, Dr Kishore Warriar, Stefanie Stafford, Kelly Sandhu & Tracey Dandy (Queens Medical Centre, Nottingham), Professor Lucy Wedderburn, Dr Clarissa Pilkington, Dr Nathan Hasson, Dr Muthana Al-Obadi, Dr Giulia Varnier, Dr Sandrine Lacassagne, Mrs Sue Maillard, Mrs Lauren Stone, Ms Elizabeth Halkon, Ms Virginia Brown, Ms Audrey Juggins, Dr Sally Smith, Mrs Sian Lunt, Ms Elli Enayat, Mrs Hemlata Varsani, Miss Laura Kassoumeri, Miss Laura Beard, Miss Katie Arnold, Mrs Yvonne Glackin, Ms Stephanie Simou, Dr Beyerley Almeida, Dr Kiran Nistala, Dr Raquel Marques, Dr Claire Deakin, Dr Parichat Khaosut, Ms Stefanie Dowle, Dr Charalampia Papadopoulou, Dr Shireena Yasin, Dr Christina Boros, Dr Meredyth Wilkinson, Dr Chris Piper, Mrs Cerise Johnson-Moore, Ms Lucy Marshall, Ms Kathryn O'Brien, Ms Emily Robinson (Great Ormond Street Hospital, London), Dr Kevin Murray (Princess Margaret Hospital, Perth, Western Australia) Dr Coziana Ciurtin, Dr John Ioannou, Mrs Caitlin Clifford and Ms Linda Suffield (University College London Hospital, London) Ms Helen Lee, Ms Sam Leach, Ms Helen Smith, Dr Anne-Marie McMahon, Ms Heather Chisem, Jeanette Hall, Ruth Kingshott and Maxine Mutten (Sheffield's Children's Hospital, Sheffield); Dr Nick Wilkinson, Ms Emma Inness, Ms Eunice Kendall, Mr David Mayers, Ruth Etherton, Danielle Miller and Dr Kathryn Bailey (Oxford University Hospitals, Oxford); Dr Jacqui Clinch, Ms Natalie Fineman, Ms Helen Pluess-Hall and Suzanne Sketchley (Bristol Royal Hospital for Children, Bristol); Dr Joyce

Davidson, Margaret Connon and Ms Lindsay Vallance (Royal Aberdeen Children's Hospital); Dr Kirsty Haslam, Charlene Bass-Woodcock, Trudy Booth and Ms Louise Akeroyd (Bradford Teaching Hospitals); Dr Alice Leahy, Amy Collier, Rebecca Cutts, Emma Macleod, Dr Hans De Graaf, Dr Brian Davidson, Sarah Hartfree, Danny Pratt, Elizabeth Fofana and Lorena Caruana (University Hospital Southampton) and all the Children, Young people and their families who have contributed to this research.

The Childhood Myositis Heterogeneity Study Group is supported by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences. The Childhood Myositis Heterogeneity Study Group thanks the following contributors to this study:

Daniel Albert (Dartmouth-Hitchcock Medical Center), Kathy Amoroso (Carilion Roanoke Memorial Hospital), Bitu Arabshahi (Fairfax Inova Medical Center), Edsel Arce-Hernandez (Valley Children's Healthcare), Alan Baer (Johns Hopkins University School of Medicine), Susan Ballinger (Riley Children's Hospital, Indiana University), Lilliana Barillas (Albany Medical Center), Mara Becker (Children's Mercy, Kansas City), Catherine April Bingham (Penn State Children's Hospital), John Bohnsack (University of Utah), Jon M Burnham (Children's Hospital of Philadelphia), Ruy Carrasco (Child Neurology Consultants of Austin), Victoria Cartwright (Legacy Emanuel Medical Center), Gail Cawkwell (Johns Hopkins All Children's Hospital), Wendy De La Pena (Loma Linda University), Fatma Dedeoglu (Boston Children's Hospital), Barbara Anne Eberhard (Northwell Health), Kaleo Ede (Phoenix Children's Hospital), Barbara Edelheit (University of Connecticut), Andrew Eichenfield (New York-Presbyterian Hospital), Terri Finkel (University of Colorado), Robert Fuhlbrigge (Boston Children's Hospital), Abraham Gedalia (Children's Hospital New Orleans), Stephen George (University of Maryland), Harry Gewanter (Virginia Commonwealth University), Ellen Goldmuntz (Children's National Medical Center), Donald Goldsmith (St. Christopher's Hospital for Children, Drexel University), Beth Gottlieb (Northwell Health), Brent Graham (Vanderbilt University), Thomas Griffin (Carolinas Medical Center), Hilary Haftel (University of Michigan), Melissa Hawkins-Holt (University of Maryland), Theresa Hennon (John R. Oishei Children's Hospital, University of Buffalo), Michael Henrickson (Cincinnati Children's Hospital), Aimee Hersh (University of Utah), Gloria Higgins

(Nationwide Children's Hospital), Lisa Imundo (Columbia University), Christi Inman (University of Utah), Anna Jansen (NIEHS, NIH), Rita Jerath (Medical College of Georgia), Olcay Y Jones (Walter Reed National Military Medical Center), Lawrence Jung (Children's National Medical Center), Ildy Katona (Uniformed University of the Health Sciences), Philip Kahn (New York University), Hanna Kim (NIAMS, NIH), Susan Kim (University of California, San Francisco), Yukiko Kimura (Hackensack Meridian School of Medicine), Daniel Kingsbury (Legacy Emanuel Medical Center), W Patrick Knibbe (St Luke's Clinic), Bianca Lang (Dalhousie University), Melissa Lerman (Children's Hospital of Philadelphia), Carol Lindsley (University of Kansas), Daniel Lovell (Cincinnati Children's Hospital), Azadeh Majlessi (Loma Linda University), Gulnara Mamyrova (George Washington University), Paul McCarthy (Yale University), Diana Milojevic (Johns Hopkins All Children's Hospital), Stephen R. Mitchell (Georgetown University), Renee Modica (Arnold Palmer Hospital for Children, University of Florida), Chihiro Morishima (University of Washington), Simona Nativ (Atlantic Health System), Edward Oberle (Nationwide Children's Hospital), Kiem Oen (University of Manitoba), Judyann Olson (Children's Wisconsin, Medical College of Wisconsin), Elif Oral (University of Michigan), Barbara Ostrov (Penn State Children's Hospital), Lauren M Pachman (Lurie Children's Hospital, Northwestern University), Murray Passo (The Medical University of South Carolina), Maria Perez (Cook Children's Medical Center), Marilyn Punaro (University of Texas at Dallas), C. Eglia Rabinovich (Duke University), Linda Ray (University of Mississippi), Robert Rennebohm (Cleveland Clinic), Rafael Rivas-Chacon (Nicklaus Children's Hospital, University of Miami), Tova Ronis (Children's National Hospital), Margalit Rosenkranz (University of Pittsburgh Medical Center), Deborah Rothman (Albany Medical College), Sara Sabbagh (NIAMS, NIH), Adam Schiffenbauer (NIEHS), Bracha Shaham (Children's Hospital of Los Angeles, University of Southern California), Robert Sheets (Rady Children's Hospital, University of California at San Diego), Susan Sheno (Seattle Children's, University of Washington), David Sherry (Seattle Children's, University of Washington), Abigail Smukler (The Toledo Clinic), Jennifer Soep (University of Colorado), Sangeeta Sule (Children's National Medical Center), Robert Sundel (Boston Children's Hospital), Stacey Tarvin (Riley Hospital for Children, University of Indiana), Mohamad Tolaymat (University of Florida), Cagri Toruner (Texas Children's Hospital), Rita Volochayev (NIEHS, NIH), Jennifer Wargula (Seattle

Children's), Peter Weiser (University of Alabama), Andrew White (Washington University), Patience White (George Washington University), Andrew Zeft (Cleveland Clinic), Lawrence Zemel (Connecticut Children's Medical Center)

Conflict of interest statement: The authors have declared that no conflict of interest exists

UNCORRECTED MANUSCRIPT

REFERENCES

1. Tansley, S. and Wedderburn, L.R. (2015) Comparing and contrasting clinical and serological features of juvenile and adult-onset myositis. *Current Opinion in Rheumatology*, **27**, 601–607.
2. Meyer, A., Meyer, N., Schaeffer, M., Gottenberg, J.-E., Geny, B. and Sibilia, J. (2015) Incidence and prevalence of inflammatory myopathies: a systematic review. *Rheumatology*, **54**, 50–63.
3. Tansley, S.L., Betteridge, Z.E., Shaddick, G., Gunawardena, H., Arnold, K., Wedderburn, L.R. and McHugh, N.J. (2014) Calcinosis in juvenile dermatomyositis is influenced by both anti-NXP2 autoantibody status and age at disease onset. *Rheumatology (Oxford, England)*, **53**, 2204–2208.
4. Rider, L.G., Shah, M., Mamyrova, G., Huber, A.M., Rice, M.M., Targoff, I.N. and Miller, F.W. (2013) The myositis autoantibody phenotypes of the juvenile idiopathic inflammatory myopathies. *Medicine*, **92**, 223–243.
5. Espada, G., Maldonado Cocco, J.A., Fertig, N. and Oddis, C. V. (2009) Clinical and serologic characterization of an Argentine pediatric myositis cohort: Identification of a novel autoantibody (anti-mj) to a 142-kda protein. *Journal of Rheumatology*, **36**, 2547–2551.
6. Ichimura, Y., Matsushita, T., Hamaguchi, Y., Kaji, K., Hasegawa, M., Tanino, Y., Inokoshi, Y., Kawai, K., Kanekura, T., Habuchi, M., *et al.* (2012) Anti-NXP2 autoantibodies in adult patients with idiopathic inflammatory myopathies: possible association with malignancy. *Annals of the rheumatic diseases*, **71**, 710–713.
7. Fiorentino, D.F., Chung, L.S., Christopher-Stine, L., Zaba, L., Li, S., Mammen, A.L., Rosen, A. and Casciola-Rosen, L. (2013) Most Patients With Cancer-Associated Dermatomyositis Have Antibodies to Nuclear Matrix Protein NXP-2 or Transcription Intermediary Factor 1 γ . *Arthritis & Rheumatism*, **65**, 2954–2962.

8. Ceribelli, A., Fredi, M., Taraborelli, M., Cavazzana, I., Franceschini, F., Quinzanini, M., Tincani, A., Ross, S.J., Chan, J.Y.F., Pauley, B.A., *et al.* (2012) Anti-MJ/NXP-2 autoantibody specificity in a cohort of adult Italian patients with polymyositis/dermatomyositis. *Arthritis Research and Therapy*, **14**, R97.
9. Miller, F.W., Chen, W., O'Hanlon, T.P., Cooper, R.G., Vencovsky, J., Rider, L.G., Danko, K., Wedderburn, L.R., Lundberg, I.E., Pachman, L.M., *et al.* (2015) Genome-wide association study identifies HLA 8.1 ancestral haplotype alleles as major genetic risk factors for myositis phenotypes. *Genes and immunity*, **16**, 470–480.
10. Rothwell, S., Cooper, R.G., Lundberg, I.E., Miller, F.W., Gregersen, P.K., Bowes, J., Vencovsky, J., Danko, K., Limaye, V., Selva-O'Callaghan, A., *et al.* (2016) Dense genotyping of immune-related loci in idiopathic inflammatory myopathies confirms HLA alleles as the strongest genetic risk factor and suggests different genetic background for major clinical subgroups. *Annals of the rheumatic diseases*, **75**, 1558–1566.
11. Arnett, F.C., Targoff, I.N., Mimori, T., Goldstein, R., Warner, N.B. and Reveille, J.D. (1996) Interrelationship of major histocompatibility complex class II alleles and autoantibodies in four ethnic groups with various forms of myositis. *Arthritis and Rheumatism*, **39**, 1507–1518.
12. O'Hanlon, T.P., Carrick, D.M., Targoff, I.N., Arnett, F.C., Reveille, J.D., Carrington, M., Gao, X., Oddis, C. v., Morel, P.A., Malley, J.D., *et al.* (2006) 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. *Medicine*, **85**, 111–127.
13. Chinoy, H., Salway, F., Fertig, N., Shephard, N., Tait, B.D., Thomson, W., Isenberg, D.A., Oddis, C. v., Silman, A.J., Ollier, W.E.R.R., *et al.* (2006) In adult onset myositis, the presence of interstitial lung disease and myositis specific/associated antibodies are governed by HLA class II haplotype, rather than by myositis subtype. *Arthritis research & therapy*, **8**, R13.

14. Arnett, F.C., Targoff, I.N., Mimori, T., Goldstein, R., Warner, N.B. and Reveille, J.D. (1996) Interrelationship of major histocompatibility complex class II alleles and autoantibodies in four ethnic groups with various forms of myositis. *Arthritis and Rheumatism*, **39**, 1507–1518.
15. O’Hanlon, T.P., Carrick, D.M., Targoff, I.N., Arnett, F.C., Reveille, J.D., Carrington, M., Gao, X., Oddis, C. V., Morel, P.A., Malley, J.D., *et al.* (2006) 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. *Medicine*, **85**, 111–127.
16. O’Hanlon, T.P., Rider, L.G., Mamyrova, G., Targoff, I.N., Arnett, F.C., Reveille, J.D., Carrington, M., Gao, X., Oddis, C. V., Morel, P.A., *et al.* (2006) HLA polymorphisms in African Americans with idiopathic inflammatory myopathy: Allelic profiles distinguish patients with different clinical phenotypes and myositis autoantibodies. *Arthritis & Rheumatism*, **54**, 3670–3681.
17. Chinoy, H., Adimulam, S., Marriage, F., New, P., Vincze, M., Zilahi, E., Kapitány, A., Gyetvai, A., Ekholm, L., Novota, P., *et al.* (2012) Interaction of HLA-DRB1*03 and smoking for the development of anti-Jo-1 antibodies in adult idiopathic inflammatory myopathies: A European-wide case study. *Annals of the Rheumatic Diseases*, **71**, 961–965.
18. Mammen, A.L., Gaudet, D., Brisson, D., Christopher-Stine, L., Lloyd, T.E., Leffell, M.S. and Zachary, A.A. (2012) Increased frequency of DRB1*11:01 in anti-HMG-CoA reductase-associated autoimmune myopathy. *Arthritis Care & Research*, **64**, 1233–1237.
19. Miller, F.W., Cooper, R.G., Vencovský, J., Rider, L.G., Danko, K., Wedderburn, L.R., Lundberg, I.E., Pachman, L.M., Reed, A.M., Ytterberg, S.R., *et al.* (2013) Genome-wide association study of dermatomyositis reveals genetic overlap with other autoimmune disorders. *Arthritis and Rheumatism*, **65**, 3239–3247.
20. Guo, J., Zhang, T., Cao, H., Li, X., Liang, H., Liu, M., Zou, Y., Zhang, Y., Wang, Y., Sun, X., *et al.* (2019) Sequencing of the MHC region defines HLA-DQA1 as the major genetic risk for

- seropositive rheumatoid arthritis in Han Chinese population. *Annals of the Rheumatic Diseases*, **78**, 773–780.
21. Hov, J.R., Kosmoliaptsis, V., Traherne, J.A., Olsson, M., Boberg, K.M., Bergquist, A., Schrupf, E., Bradley, J.A., Taylor, C.J., Lie, B.A., *et al.* (2011) Electrostatic modifications of the human leukocyte antigen-DR P9 peptide-binding pocket and susceptibility to primary sclerosing cholangitis. *Hepatology*, **53**, 1967–1976.
 22. Han, B., Diogo, D., Eyre, S., Kallberg, H., Zhernakova, A., Bowes, J., Padyukov, L., Okada, Y., González-Gay, M.A., Rantapää-Dahlqvist, S., *et al.* (2014) Fine mapping seronegative and seropositive rheumatoid arthritis to shared and distinct HLA alleles by adjusting for the effects of heterogeneity. *American journal of human genetics*, **94**, 522–532.
 23. Hinks, A., Bowes, J., Cobb, J., Ainsworth, H.C., Marion, M.C., Comeau, M.E., Sudman, M., Han, B., Becker, M.L., Bohnsack, J.F., *et al.* (2017) Fine-mapping the MHC locus in juvenile idiopathic arthritis (JIA) reveals genetic heterogeneity corresponding to distinct adult inflammatory arthritic diseases. *Annals of the Rheumatic Diseases*, **76**, 765–772.
 24. Bowes, J., Budu-Aggrey, A., Huffmeier, U., Uebe, S., Steel, K., Hebert, H.L., Wallace, C., Massey, J., Bruce, I.N., Bluett, J., *et al.* (2015) Dense genotyping of immune-related susceptibility loci reveals new insights into the genetics of psoriatic arthritis. *Nature Communications*, **6**, 1–11.
 25. Nejentsev, S., Howson, J.M.M., Walker, N.M., Szszko, J., Field, S.F., Stevens, H.E., Reynolds, P., Hardy, M., King, E., Masters, J., *et al.* (2007) Localization of type 1 diabetes susceptibility to the MHC class I genes HLA-B and HLA-A. *Nature*, **450**, 887–892.
 26. Kaneko, T. and Obata, F. (1996) Allogeneic recognition of HLA-DRB1(*)0406 by T cells with HLA-DRB1(*)0403: Role of amino acid residue 37 on the β sheet in T cell recognition. *Immunobiology*, **195**, 261–270.

27. Chen, Y.Z., Matsushita, S. and Nishimura, Y. (1997) A single residue polymorphism at DR β 37 affects recognition of peptides by T cells. *Human Immunology*, **54**, 30–39.
28. Rothwell, S., Chinoy, H., Lamb, J.A., Miller, F.W., Rider, L.G., Wedderburn, L.R., McHugh, N.J., Mammen, A.L., Betteridge, Z.E., Tansley, S.L., *et al.* (2019) Focused HLA analysis in Caucasians with myositis identifies significant associations with autoantibody subgroups. *Annals of the Rheumatic Diseases*, **78**, 996–1002.
29. Raychaudhuri, S., Sandor, C., Stahl, E.A., Freudenberg, J., Lee, H.-S., Jia, X., Alfredsson, L., Padyukov, L., Klareskog, L., Worthington, J., *et al.* (2012) Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. *Nature genetics*, **44**, 291–296.
30. Ombrello, M.J., Remmers, E.F., Tachmazidou, I., Grom, A., Foell, D., Haas, J.P., Martini, A., Gattorno, M., Özen, S., Prahalad, S., *et al.* (2015) HLA-DRB1*11 and variants of the MHC class II locus are strong risk factors for systemic juvenile idiopathic arthritis. *Proceedings of the National Academy of Sciences of the United States of America*, **112**, 15970–15975.
31. Goyette, P., Boucher, G., Mallon, D., Ellinghaus, E., Jostins, L., Huang, H., Ripke, S., Gusareva, E.S., Annese, V., Hauser, S.L., *et al.* (2015) High-density mapping of the MHC identifies a shared role for HLA-DRB1*01:03 in inflammatory bowel diseases and heterozygous advantage in ulcerative colitis. *Nature Genetics*, **47**, 172–179.
32. Patsopoulos, N.A., Barcellos, L.F., Hintzen, R.Q., Schaefer, C., van Duijn, C.M., Noble, J.A., Raj, T., Gourraud, P.-A., Stranger, B.E., Oksenberg, J., *et al.* (2013) Fine-Mapping the Genetic Association of the Major Histocompatibility Complex in Multiple Sclerosis: HLA and Non-HLA Effects. *PLoS Genetics*, **9**, e1003926.
33. Menconi, F., Osman, R., Monti, M.C., Greenberg, D.A., Concepcion, E.S. and Tomer, Y. (2010) Shared molecular amino acid signature in the HLA-DR peptide binding pocket predisposes to both autoimmune diabetes and thyroiditis. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 16899–16903.

34. Ban, Y., Davies, T.F., Greenberg, D.A., Concepcion, E.S., Osman, R., Oashi, T. and Tomer, Y. (2004) Arginine at position 74 of the HLA-DR β 1 chain is associated with Graves' disease. *Genes and Immunity*, **5**, 203–208.
35. Gourh, P., Safran, S.A., Alexander, T., Boyden, S.E., Morgan, N.D., Shah, A.A., Mayes, M.D., Doumatey, A., Bentley, A.R., Shriner, D., *et al.* (2020) HLA and autoantibodies define scleroderma subtypes and risk in African and European Americans and suggest a role for molecular mimicry. *Proceedings of the National Academy of Sciences of the United States of America*, **117**, 552–562.
36. Langefeld, C.D., Ainsworth, H.C., Graham, D.S.C., Kelly, J.A., Comeau, M.E., Marion, M.C., Howard, T.D., Ramos, P.S., Croker, J.A., Morris, D.L., *et al.* (2017) Transancestral mapping and genetic load in systemic lupus erythematosus. *Nature Communications*, **8**, 1–18.
37. Han, B., Akiyama, M., Kim, K.K., Oh, H., Choi, H., Lee, C.H., Jung, S., Lee, H.S., Kim, E.E., Cook, S., *et al.* (2018) Amino acid position 37 of HLA-DR β 1 affects susceptibility to Crohn's disease in Asians. *Human Molecular Genetics*, **27**, 3901–3910.
38. Molineros, J.E., Looger, L.L., Kim, K., Okada, Y., Terao, C., Sun, C., Zhou, X., Raj, P., Kochi, Y., Suzuki, A., *et al.* (2019) Amino acid signatures of HLA Class-I and II molecules are strongly associated with SLE susceptibility and autoantibody production in Eastern Asians. *PLOS Genetics*, **15**, e1008092.
39. Jee, S.-H., Tsai, T.-F., Tsai, W.-L., Liaw, S.-H., Chang, C.-H. and HU, C.-Y. (1998) HLA-DRB1*0701 and DRB1*1401 are associated with genetic susceptibility to psoriasis vulgaris in a Taiwanese population. *British Journal of Dermatology*, **139**, 978–983.
40. Oka, S., Furukawa, H., Shimada, K., Hashimoto, A., Komiya, A., Tsunoda, S., Saisho, K., Tsuchiya, N., Katayama, M., Shinohara, S., *et al.* (2019) Association of HLA-DRB1 genotype with younger age onset and elder age onset rheumatoid arthritis in Japanese populations. *Medicine*, **98**, e18218.

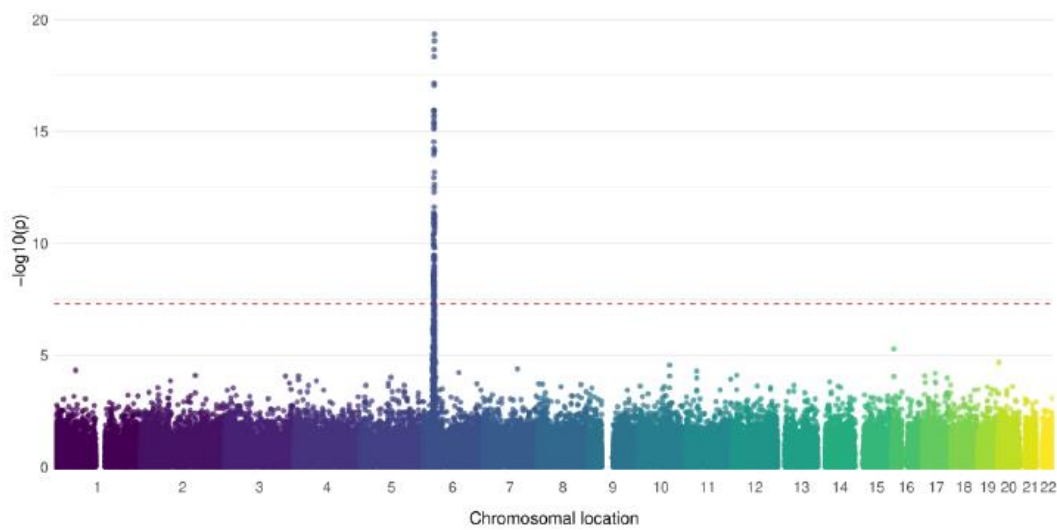
41. Jamieson, R. V., Farrar, N., Stewart, K., Perveen, R., Mihelec, M., Carette, M., Grigg, J.R., McAvoy, J.W., Lovicu, F.J., Tam, P.P.L., *et al.* (2007) Characterization of a familial t(16;22) balanced translocation associated with congenital cataract leads to identification of a novel gene, TMEM114, expressed in the lens and disrupted by the translocation. *Human Mutation*, **28**, 968–977.
42. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., De Bakker, P.I.W., Daly, M.J., *et al.* (2007) PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, **81**, 559–575.
43. McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A.R., Teumer, A., Kang, H.M., Fuchsberger, C., Danecek, P., Sharp, K., *et al.* (2016) A reference panel of 64,976 haplotypes for genotype imputation. *Nature Genetics*, **48**, 1279–1283.
44. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M. and Lee, J.J. (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, **4**, 7.
45. Altshuler, D.M., Gibbs, R.A., Peltonen, L., Schaffner, S.F., Yu, F., Dermitzakis, E., Bonnen, P.E., De Bakker, P.I.W., Deloukas, P., Gabriel, S.B., *et al.* (2010) Integrating common and rare genetic variation in diverse human populations. *Nature*, **467**, 52–58.
46. Yang, J., Benyamin, B., McEvoy, B.P., Gordon, S., Henders, A.K., Nyholt, D.R., Madden, P.A., Heath, A.C., Martin, N.G., Montgomery, G.W., *et al.* (2010) Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, **42**, 565–569.
47. Lee, S.H., Wray, N.R., Goddard, M.E. and Visscher, P.M. (2011) Estimating missing heritability for disease from genome-wide association studies. *American Journal of Human Genetics*, **88**, 294–305.

48. Jia, X., Han, B., Onengut-Gumuscu, S., Chen, W.-M., Concannon, P.J., Rich, S.S., Raychaudhuri, S. and de Bakker, P.I.W. (2013) Imputing Amino Acid Polymorphisms in Human Leukocyte Antigens. *PLoS ONE*, **8**, e64683.
49. Okada, Y., Han, B., Tsoi, L.C., Stuart, P.E., Ellinghaus, E., Tejasvi, T., Chandran, V., Pellett, F., Pollock, R., Bowcock, A.M., *et al.* (2014) Fine mapping major histocompatibility complex associations in psoriasis and its clinical subtypes. *American Journal of Human Genetics*, **95**, 162–172.
50. Das, S., Forer, L., Schönherr, S., Sidore, C., Locke, A.E., Kwong, A., Vrieze, S.I., Chew, E.Y., Levy, S., McGue, M., *et al.* (2016) Next-generation genotype imputation service and methods. *Nature Genetics*, **48**, 1284–1287.
51. Marchini, J., Howie, B., Myers, S., McVean, G. and Donnelly, P. (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. *Nature Genetics*, **39**, 906–913.

UNCORRECTED MANUSCRIPT

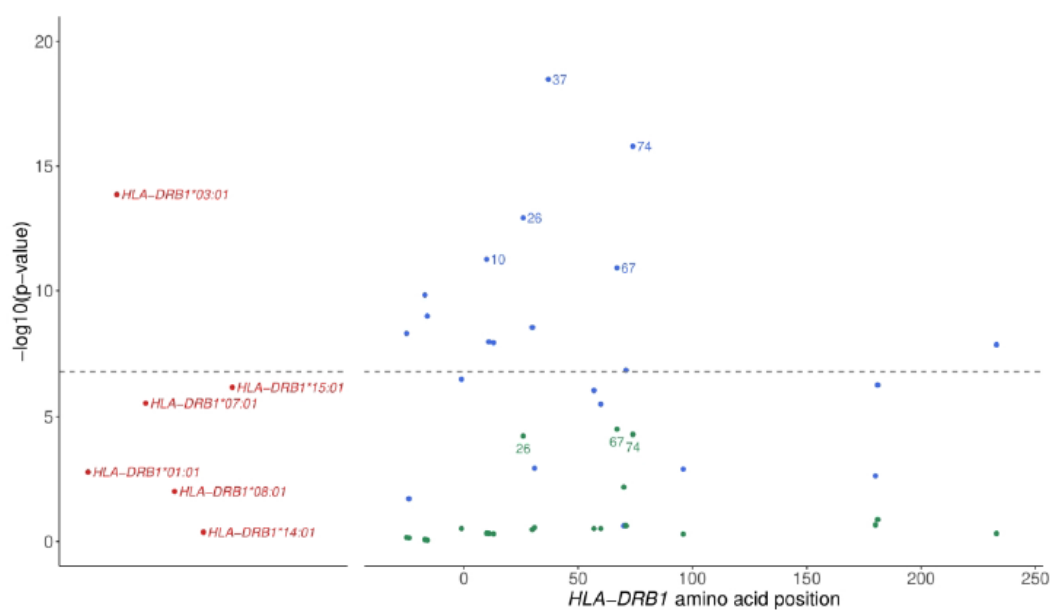
FIGURE LEGENDS

Figure 1. Manhattan plot of the association of assayed SNPs with JDM. SNPs (n=178,164) were available for n=851 JDM samples and n=12,232 controls of Caucasian origin. Association was tested using logistic regression, with the first 10 principal components included as covariates to account for population stratification. The red dotted line indicates genome-wide level of significance (5×10^{-8}).



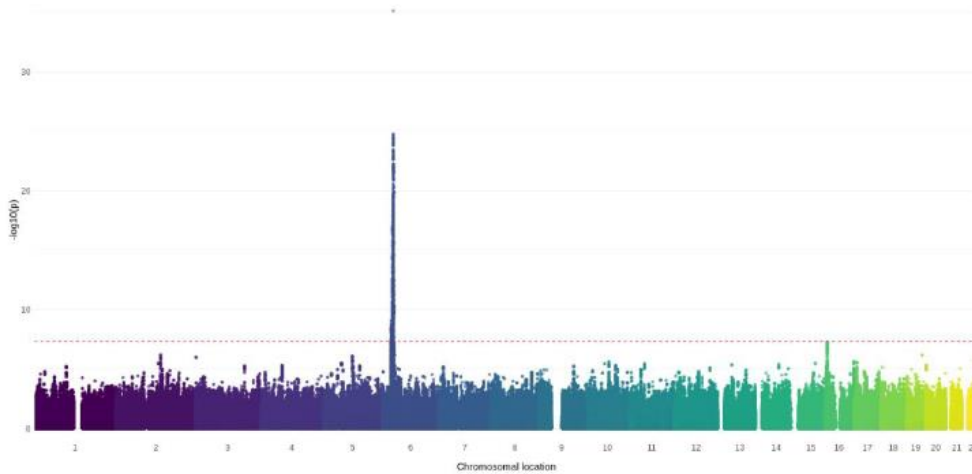
UNCORRECTED

Figure 2. Representation of the p-values for omnibus tests of individual amino acid positions within *HLA-DRB1* for association with JDM. Blue circles represent p-values for amino acid positions tested alone. Green circles represent p-values for amino acid positions after conditioning on position 37. Red circles represent p-values for 4-digit classical alleles. The grey dashed line represents the threshold for statistical significance ($p=6.8 \times 10^{-6}$).



UNCORRECTED

Figure 3. Manhattan plot of the association of imputed SNPs with JDM. Loci were filtered according to minor allele frequency of 0.05. Data were imputed for n=851 JDM samples and n=12,232 controls of Caucasian origin. The first 10 principal components were adjusted for during analysis to account for population stratification. SNPs were weighted by the information score to account for imputation uncertainty. The red dotted line indicates genome-wide level of significance (5×10^{-8}). The degree of transparency of each data-point represents the R^2 value for imputation accuracy, with more solid colours representing higher certainty.



TABLES

Table 1. Sources of samples from Caucasian patients with JDM

Country	Source	Genotyped samples (n=952)	Genotyped samples after quality control (n=851)
United Kingdom	UK Juvenile Dermatomyositis Cohort & Biomarker Study	365	326
United States	National Institute of Environmental Health Sciences	262	224
	Northwestern University	140	116
	Emory University	41	31
	George Washington University	32	30
	Mayo Clinic	13	12
Canada	The Hospital for Sick Children, Toronto	24	24
	IWK Health Centre	16	16
	Alberta University	10	9
Norway	Oslo University Hospital	49	46

Table 2. Analysis of amino acid positions within *HLA-DRB1* associated with JDM^a

Position	χ^{2b}	P-value	Effects independent of position 37	
			χ^2	P-value
37	89.2	3.3×10^{-19}	-	-
74	76.7	1.6×10^{-16}	22.5	5.1×10^{-05}
26	59.5	1.2×10^{-13}	19.5	5.9×10^{-05}
10	47.5	5.4×10^{-12}		
67	46.0	1.2×10^{-11}		
-17	41.0	1.5×10^{-10}		
-16	37.5	9.1×10^{-10}		
30	39.5	2.6×10^{-9}		
-25	34.4	4.5×10^{-9}		
11	40.2	9.9×10^{-9}		
13	40.0	1.1×10^{-08}		
233	32.3	1.3×10^{-8}		
71	27.8	1.4×10^{-7}		
-1	26.1	3.2×10^{-7}		
181	28.9	5.3×10^{-7}		
57	24.2	8.8×10^{-7}		
60	25.4	3.1×10^{-6}		

^aAlpha threshold for statistical significance in the MHC region taken as 6.8×10^{-6} (49)(49)(49)(49)(49)

^bTest statistics compared using a chi-squared distribution

Table 3. Association of residues within position 37 with JDM, with effects of classical HLA alleles linked to these residues^{a,b}

Residue	Classical HLA allele linked to residue	Allele frequency in JDM	Allele frequency in Controls	Odds Ratio	95% Confidence Interval	p-value
Tyr-37	-	0.37	0.31	-	-	-
Ser-37	-	0.21	0.26	0.65	0.57, 0.75	7.3×10 ⁻¹⁰
-	<i>HLA-DRB1*01:01</i>	0.10	0.12	0.76	0.65, 1.46	1.7×10 ⁻³
-	<i>HLA-DRB1*15:01</i>	0.11	0.14	0.66	0.57, 0.78	6.5×10 ⁻⁷
Phe-37	-	0.11	0.15	0.63	0.54, 0.75	1.1×10 ⁻⁷
-	<i>HLA-DRB1*07:01</i>	0.09	0.13	0.67	0.57, 0.79	2.8×10 ⁻⁶
-	<i>HLA-DRB1*14:01</i>	0.02	0.02	0.86	0.60, 1.24	0.42
Asn-37	-	0.31	0.24	1.13	1.06, 1.19	0.05
-	<i>HLA-DRB1*03:01</i>	0.19	0.12	1.66	1.46, 1.89	1.4×10 ⁻¹⁴

^aSummary statistics for individual residues derived from the omnibus model for position 37, with Tyr-37 used as the reference residue. Summary statistics for classical 4-digit HLA alleles derived from case-control analysis of imputed HLA data. Alleles linked to these residues at position 37 but not represented were filtered out during QC.

^bAlpha threshold for statistical significance in the MHC region taken as 6.8×10^{-6} (49)(49)(49)(49)(49).

Table 4. Genome-wide imputed loci displaying potential association with JDM, with p-value cut-offs of 1×10^{-6} and minor allele frequency cut-off of 0.01^a

Rsid ^b	Chromosome	Position	Allele A	Allele B	MAF	OR	95% CI	p-value	Imputation R ² ^c	Information measure ^d	Association information measure ^e
rs650160	16	8607139	G	A	0.27	0.72	0.65, 0.81	5.8×10^{-8}	0.91	0.91	0.91
rs6892006	5	88450541	T	G	0.13	0.68	0.58, 0.79	8.7×10^{-7}	0.87	0.87	0.89

^aLoci with p-values below a suggestive level of significance (1×10^{-6}) are displayed for minor allele frequency (MAF) cut-offs of 0.05 and 0.01; only additional loci captured by the lower MAF cut-off are shown for that cut-off

^bRsid, Reference SNP cluster ID. There were no assayed markers within 100,000 kb of either of these loci

^cSquared Correlation of imputation of genotypes with true unmeasured genotypes, as estimated by Minimac3 during imputation via the Michigan Imputation Server:

https://genome.sph.umich.edu/wiki/Minimac3_Info_File#Rsq

^dImpute INFO measure calculated by SNPTTEST:

https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html#info_measures

^eRelative information measure about parameters of the model fitted during association testing:

https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.v2.p

ABBREVIATIONS

Abbreviation	Definition
ACPA	Anti-citrullinated protein antibodies
AH8.1	Caucasian 8.1 ancestral haplotype (HLA A1-B8-DR3-DQ2)
Anti-cN1A	Anti-cytosolic 5'-nucleotidase 1A autoantibodies
Anti-HMGCR	Anti-3-hydroxy-3-methylglutaryl-CoA reductase autoantibodies
Anti-Jo-1	Anti-histidyl tRNA synthetase autoantibodies
Anti-Mi-2	Autoantibodies against the Mi-2 antigen of the nucleosome remodeling-deacetylase complex
Anti-NXP2	Anti-nuclear matrix protein-2 autoantibodies; also known as anti-MJ
Anti-PM/Scl	Autoantibodies against the 75-kDa and/or 100-kDa subunits of the PM/Scl human exosome complex
Anti-TIF1 γ	Anti-transcription intermediary factor 1 γ autoantibodies
CEU	Caucasian
HLA	Human leukocyte antigen
HRC	Haplotype reference consortium
HWE	Hardy-Weinberg equilibrium
IAMDGC	International Age-related Macular Degeneration Genetics Consortium
IIM	Idiopathic inflammatory myopathy
JDCBS	Juvenile Dermatomyositis Cohort and Biomarker Study
JDM	Juvenile dermatomyositis
JIA	Juvenile idiopathic arthritis
LD	Linkage disequilibrium
LRT	Likelihood ratio test
MAF	Minor allele frequency
<i>MEF2C-AS1</i>	<i>Myocyte enhancer factor 2C antisense RNA 1</i>
MHC	Major histocompatibility complex
MSA	Myositis-specific autoantibody
MYOGEN	Myositis Genetics Consortium
OR	Odds ratio
PC	Principal component
PCA	Principal component analysis
QC	Quality control
RA	Rheumatoid arthritis
SLE	Systemic lupus erythematosus
SNP	Single nucleotide polymorphism
<i>TMEM114</i>	<i>Transmembrane protein 114</i>