An amino acid motif in HLA-DRβ1 distinguishes patients with uveitis in juvenile idiopathic arthritis.

A.J.W. Haasnoot[1,2], M.W. Schilham[3], S.S.M. Kamphuis[4], P.C.E. Hissink Muller[3], A. Heiligenhaus[5], D. Foell[6], R.A. Ophoff[7], K. Minden[8], T.R.D.J. Radstake[2,9], A.I. Den Hollander[10], T.H.C.M. Reinards[3], S. Hiddingh[2], N. Schalij-Delfos[11], E.P.A.H. Hoppenreijs[12], M.A.J. van Rossum[13], C. Wouters[14], R.K. Saurenmann[15], N. Wulffraat[16], ICON-JIA study group, R. ten Cate[3], J.H. de Boer[1]*, S.L. Pulit[17]*, J.J.W. Kuiper[1,2]*

*) JB, SP, JK equally contributed as senior authors

1. Department of Ophthalmology, University Medical Center Utrecht, The Netherlands.
2. Laboratory of Translational Immunology, University Medical Center Utrecht, The Netherlands.
3. Department of Paediatrics/Paediatric Rheumatology, Leiden University Medical Centre, The Netherlands
4. Department of Pediatrics/Pediatric Rheumatology, Erasmus MC Sophia Children's Hospital, Rotterdam, The Netherlands
5. Department of Ophthalmology, Ophtha Lab at St. Franziskus-Hospital, Muenster, Germany; Department of Ophthalmology, University of Duisburg-, Essen, Germany.
6. Department of Pediatric Rheumatology and Immunology, University of Muenster, Germany
7. Department of Psychiatry, Brain Center Rudolf Magnus, University Medical Center Utrecht, The Netherlands. Department of Human Genetics, David Geffen School of Medicine, University of California, USA. Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, University of California, USA.
8. German Rheumatism Research Center and Charité Universitätsmedizin Charité, Berlin, Germany.
9. Department of Rheumatology and Clinical Immunology, University Medical Centre Utrecht, Utrecht, The Netherlands.
10. Department of Ophthalmology, Donders Institute of Brain, Cognition, and Behaviour, Radboud University Medical Center, Nijmegen, the Netherlands. Department of Human Genetics, Donders Institute of Brain, Cognition, and Behaviour, Radboud University Medical Center, Nijmegen, the Netherlands.
11. Department of Ophthalmology, Leiden University Medical Centre, The Netherlands
12. Department of Paediatrics/Paediatric Rheumatology, Radboud University Nijmegen Medical Centre, The Netherlands
13. Department of Paediatrics/Paediatric Rheumatology, Emma Children's Hospital AMC, Amsterdam, The Netherlands. Department of Paediatric Rheumatology, Jan van Breemen Institute, Amsterdam, The Netherlands
14. Department of Pediatric Immunology, University Hospitals Leuven, Belgium
15. University Children's Hospital, Pediatric Rheumatology, Zurich, Switzerland
16. Department of Paediatric Rheumatology, University Medical Center Utrecht, The Netherlands.
17. Department of Neurology, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, The Netherlands.

Correspondence: Jonas Kuiper, MD, PhD. Department of Ophthalmology. University Medical Center Utrecht, Utrecht, The Netherlands. j.j.w.kuiper@umcutrecht.nl

Abstract

Objectives

Uveitis is a visually-debilitating disorder that affects up to 30% of children with the most common forms of juvenile idiopathic arthritis (JIA). The disease mechanisms predisposing only a subgroup of children to uveitis are unknown. To identify genetic susceptibility loci for uveitis in JIA, we conducted a genome-wide association study totalling 522 JIA cases.

Methods

Two cohorts of JIA patients with ophthalmological follow-up were separately genotyped and then imputed using a genome-wide imputation reference panel, and an HLA-specific reference panel used for imputing amino acids and HLA types in the major histocompatibility complex (MHC). After imputation, we performed genome-wide and MHC-specific analyses. We used a reverse immunology approach to model antigen presentation at 13 common HLA-DRB1 allotypes.

Results

We identified the amino acid serine at position 11 (serine-11) in *HLA-DRB1* as associated to increased risk of uveitis (OR = 2.60, $p = 5.43 \times 10^{-10}$). We found the serine-11 signal to be specific to females ($p_{females} = 7.61 \times 10^{-10}$, $p_{males} = 0.18$). Serine-11 resides in the YST-motif in the peptide binding groove of the HLA-DRB1 protein; all three amino acids are in perfect linkage disequilibrium and show identical association to disease. Quantitative prediction of binding affinity revealed that discernable peptide-binding preferences distinguish HLA-DRB1 allotypes with the YST-motif.

Conclusion

Our findings highlight a genetically distinct, sexually-dimorphic feature of JIA-uveitis compared to JIA without uveitis in *HLA-DRB1*. The association indicates the potential involvement for antigen presentation by HLA-DRB1 in the development of uveitis in JIA.

Words: 245

Introduction

Juvenile idiopathic arthritis (JIA) is the most common chronic rheumatic disease in childhood,[1,2] affecting approximately 16-150 per 100,000 individuals.[3] As many as 1 in 3 children that suffer from the most common JIA categories, oligoarticular and polyarticular rheumatoid factor-negative JIA, develop uveitis, a chronic inflammatory eye disease, which is the most frequent extra-articular manifestation of JIA.[4,5] Uveitis threatens the sight of those affected, and can result in complications including cataracts, glaucoma, and finally hypotony.[4] Because it typically afflicts the young (<7 years old), uveitis dramatically impacts the quality of life of patients in both childhood and as an adult.[4,6]

Early detection and adequate ophthalmological management of uveitis in JIA are critical to prevent significant, sight-threatening complications.[7–9] Despite its severity, uveitis is typically insidious in onset and often becomes symptomatic only after irreversible damage has occurred. Consequently, rigorous ophthalmological screening of all JIA patients is required for early detection and treatment of uveitis associated with JIA.[10] Nonetheless, severe ocular complications may already be present at time of diagnosis of uveitis. Several clinical factors have been reported to be associated with increased risk of developing uveitis: JIA category, young age (<7 years old) at JIA onset; the presence of antinuclear antibodies (ANA), indicating abnormal immune response; and elevated erythrocyte sedimentation rate (ESR) at JIA onset, indicating inflammation.[10–12] The role of these risk factors is especially pronounced in the uveitis-prone categories oligoarticular and polyarticular rheumatoid factor (RF)-negative JIA.

Both JIA and JIA-associated uveitis (JIA-uveitis) are multifactorial autoimmune disorders; each has a strong, though not fully penetrant, genetic predisposition.[4,13] Both conditions are complex and driven by biological and environmental factors.[4,13] There have been no reports of observed Mendelian inheritance patterns of uveitis in JIA-affected families,[4] further indicating a complex trait architecture. Genome-wide association studies in JIA (combined with uveitis) have revealed a number of loci associated to diseases,[14–16] including multiple independent loci in the major histocompatibility complex (MHC) which collectively explain >20% of the phenotypic variation in JIA.[16] A number of HLA types have also been identified as increasing risk of uveitis compared to controls.[4,17–19] However, genome-wide genetic markers that distinguish JIA without uveitis from JIA-uveitis remain elusive, and the exact disease mechanisms predisposing only a subgroup of children to developing uveitis are unknown. Due to the severity of uveitis and the complications surrounding its detection, biomarkers to predict the development of uveitis would help in early detection of the presence of uveitis in high-risk children without frequent, extensive, and costly screenings for all JIA patients.

Here, we performed genotyping in 192 uveitis cases and 330 JIA patients without uveitis for at least 4 years, with the aim of identifying those genetic variants that segregate more commonly in JIA patients with uveitis compared to those without uveitis. To identify risk variants in a high-density set of polymorphisms, we performed two sets of imputation: one using a sequencing-based reference panel, to scan for risk-increasing alleles genome-wide;

and a second, with an MHC-specific reference panel, to identify additional SNPs, amino acids, and HLA types that confer risk to disease. We identified serine (S) at position 11 in *HLA-DRB1* as highly associated to increased risk of uveitis (odds ratio (OR) = 2.60 [95% CI: 1.92 - 3.52], $p = 5.46 \times 10^{-10}$). Sex-specific analysis revealed that the association at serine-11 in *HLA-DRB1* was specific to females only ($p_{females} = 7.61 \times 10^{-10}$, $p_{males} = 0.18$). Quantitative prediction of peptide binding revealed that the presence of serine-11, which is located in the peptide-binding groove, affects overall peptide-binding preferences of common HLA-DRB1 allotypes.

Results

We performed our analysis in two phases that were then jointly analyzed to improve power for locus discovery.[20] In Phase 1, we genotyped 137 uveitis cases, 247 JIA without uveitis samples, and 398 population-level controls without JIA or uveitis, all collected from within the Netherlands. Samples were genotyped on the HumanOmniExpress24 array, containing ~700,000 markers. Samples and SNPs underwent quality control (QC) (**Materials and Methods**) to remove those samples and SNPs with low-quality genotyping data. After QC, 126 uveitis cases, 231 non-uveitis JIA samples, and 394 population-level controls remained (**Table 1**). We prephased the genotyping data using SHAPEIT2[21–23] and imputed the phased haplotypes with IMPUTE2,[24–26] using a reference panel constructed from the 2,504 global samples included in the 1000 Genomes Project Phase 3.[27] In parallel to genome-wide imputation, we extracted the MHC from the genotype data and imputed additional SNPs, amino acids and HLA types using the SNP2HLA pipeline[28] (**Materials and Methods**).

In Phase 2, an additional 192 samples: 77 uveitis cases and 115 JIA without uveitis samples, collected from Germany, Belgium and Switzerland were genotyped on the same array. Samples and SNPs underwent QC nearly identical to that of Phase 1 (**Materials and Methods**). After QC, 66 uveitis cases and 99 JIA without uveitis samples remained for analysis. The cleaned genotyping data was prephased and imputed both genome-wide and within the MHC, precisely as in Phase 1.

Genome-wide association testing

For genome-wide association testing in the Phase 1 and Phase 2 data, we kept those SNPs with a minor allele frequency (MAF) > 1% and with an imputation quality score > 0.7; >9.2M variants were available in both the Phase 1 and Phase 2 after imputation and filtering (**Supplementary Table 1**). For MHC association testing, we analyzed only those variants with MAF > 1%. No imputation quality filter was applied to the MHC-specific imputed data; 97.7% and 96.4% of MHC variants in Phase 1 and Phase 2, respectively, had an imputation quality score > 0.9 (**Supplementary Table 1**).

We performed genome-wide association testing across the autosomal genome, using a logistic regression framework assuming an additive model in PLINK 1.9[29,30], correcting for sex and the first two principal components (**Supplementary Figures 1 and 2**). Our data enabled three possible comparison groups: (a) JIA cases and population-level controls, (b) uveitis cases and population-level controls, and (c) uveitis cases and JIA without uveitis. As

the goal of our analysis was to investigate similarities and differences in the genetic architecture of JIA and uveitis, and because far larger GWAS of JIA and population-level controls have already been performed,[31] we performed the first two genome-wide association studies (GWAS) as a means of data quality control only (**Supplementary Figure 3**) and account for them when considering multiple testing burden. We focus here on the results of the last comparison: uveitis cases and JIA without uveitis samples.

We performed separate GWAS in Phase 1 and Phase 2, comparing 330 JIA without uveitis samples to 192 individuals with uveitis, and then combined the summary-level results in an inverse variance-weighted fixed effects meta-analysis in METAL.[32] In the genome-wide scan, the single signal achieving genome wide significance (set at $p < 2 \times 10^{-8}$, after adjusting for three phenotype comparisons) resided in the MHC (**Supplementary Figures 1 and 2**).

Association testing and fine-mapping in the MHC

To identify the amino acids or HLA types driving the genome-wide association signal, we performed a mega-analysis across the imputed MHC data. We first merged imputation dosages from Phase 1 and Phase 2, and then used PLINK 1.9[30] to perform logistic regression, assuming an additive model and correcting for the top 5 principal components, sex, and analysis phase (i.e., Phase 1 or Phase 2). This 'mega-analysis' approach is theoretically and empirically highly similar to inverse variance-weighted meta-analysis,[33,34] We also performed a meta-analysis across Phase 1 and Phase 2 and found that, indeed, the odds ratios derived from mega-analysis and meta-analysis in the MHC were highly concordant (Pearson's r = 0.95, **Supplementary Figure 4**). The mega-analysis allows for the additional advantage of allowing for conditional analysis on any associated variants across the full dataset.

Of the SNPs, amino acids, and classical alleles tested in the MHC, the strongest association was observed for the presence of either serine (S) or aspartic acid (D) at position 11 in *HLA-DRB1* (OR = 2.59 [95% CI: 1.92 - 3.50], $p = 4.80 \times 10^{-10}$; **Figure 1**, **Table 2**). To identify which of the possible residues at position 11 were explaining the signal, we performed an omnibus test of all but one of the 6 alleles present at *HLA-DRB1* position 11. We found that the goodness of fit of the omnibus test far exceed that of a null model including only sex, principal components and genotyping batch (likelihood ratio test $p = 1.5 \times 10^{-9}$). To identify if a single *HLA-DRB1* position 11 allele was driving the top association signal (as opposed to all possible alleles together, as modeled by the omnibus test), we conditioned the top signal on the dosages of the *HLA-DRB1* serine variant and then the *HLA-DRB1* aspartic acid variant as an additional covariate in the logistic regression model (**Figure 1**, **Table 2**, **Supplementary Table 2**). After conditioning on serine at position 11, the association signal dropped substantially (p = 0.069), while conditioning on aspartic acid at position 11 left the association signal essentially unchanged (odds ratio = 2.60 [95% CI: 1.92 - 3.52], $p = 5.43 \times 10^{-10}$), indicating that serine explains the bulk of the signal.

Serine at position 11 in *HLA-DRB1* is positioned in the middle of what is known as the YST-motif[35,36] in *HLA-DRB1*; the motif is comprised of tyrosine (Y) at position 10, serine (S) at position 11, and threonine (T) at position 12. All three amino acids are in perfect linkage

disequilibrium (LD) with one another (i.e., LD = 1 between all residue pairs). Additionally, the residues in the YST-motif are also in perfect LD with a fourth residue, presence of serine (S) or glycine (G) at position 13 in *HLA-DRB1* (OR = 2.47 [95% CI: 1.84 - 3.32, $p = 1.44 \times 10^{-9}$ in the MHC analysis). Thus, all four amino acid configurations show statistically identical association to uveitis (**Table 2**). To further decipher which specific residue(s) were driving the association in uveitis, we performed a series of likelihood ratio tests. In brief, a likelihood ratio test compares the fit of two models to the observed data and indicates which of the models, if either, best fits the data. We found that the S (or Y or T) amino acids best fit the data ($p = 1.48 \times 10^{-10}$, compared to a model containing only PCs and sex as independent variables). Including the S/G residue at position 13 as well as the YST motif only modestly improved the fit of the model (likelihood ratio test p = 0.043, compared to a model containing serine at position 11 only); neither the S nor G residues alone at position 13 improved the model (p = 0.70 for both tests). Thus, we concluded that serine at position 11 (or tyrosine-10 or threonine-12) was the primary driver of the observed association signal.

Sexual dimorphism in JIA-associated uveitis

Since uveitis is epidemiologically known to be more prevalent in females with oligoarticular JIA,[37] we explored if the primary signal in uveitis showed evidence of sexual dimorphism. In testing female samples only, we found that the association with *HLA-DRB1* was specific to females (serine at position 11, OR = 3.30 [95% CI: 2.26 - 5.83], $p = 7.61 \times 10^{-10}$) and was not found in males (OR = 0.73 [95% CI: 0.44 - 1.19], p = 0.21; **Table 2 and Supplementary Figure 5**). Although the most significant association with uveitis in females was mapped to amino acid position 233 in the cytoplasmic domain of *HLA-DRB1* (presence of threonine (T), OR = 0.30 [95% CI: 0.20 - 0.41], $p = 3.50 \times 10^{-10}$; **Table 2**), this position is tightly correlated (r^2=0.98) with tyrosine (Y) at position 10, serine (S) at position 11, threonine (T) at position 12 (the YST-motif), and position 13 in *HLA-DRB1* (**Supplementary Figure 6**), and consequently yield nearly identical associations (**Table 2**).

To further explore the potential for a sex-specific effect at *HLA-DRB1* position 11, we ran the association test at serine-11 across all samples, including an interaction term between sex and the imputed dosage at serine-11. We found a significant effect of the interaction term (OR = 2.25 [95% CI:1.22-4.15] p = 0.0094), indicating that the serine-11 signal in *HLA-DRB1* was indeed a sex-specific effect.

In Silico Peptide Binding to HLA-DRβ1

Polymorphisms in the beta chain specify the peptide binding preference of the HLA-DR molecule. Serine at position 11, as well as the adjacent associated amino acids at position 10 - 13, are located in the bottom of the antigen-binding groove of the HLA-DRB1 protein (**Figure 2**), suggesting that different peptide-binding preferences of HLA-DRB1 may confer risk for developing uveitis. To explore if the presence of serine at position 11 affects peptide-MHCII interactions, we compared the binding affinity for relevant peptides to a panel of representative *DRB1* alleles using the *NetMHCIIpan* server.[38] Briefly, the neural

network-based *NetMHCIIpan* algorithm is capable of reliably detecting differences between peptide-binding repertoires of highly similar MHC class II molecules. We compared the predicted binding affinity for 13 common HLA-DRB1 allotypes (representing 79% of *DRB1* alleles in cases) using a large panel of >80,000 peptides based on human iris proteome data (**Materials and Methods** and **Supplementary Table 3**).

To identify *HLA-DRB1* alleles that encode proteins with similar predicted binding preferences, we performed unsupervised hierarchical clustering. Clustering of the peptide affinity data discerned two major clusters of classical alleles strikingly similar to the distribution of serine at position 11 (**Supplementary Figure 7**). We observed that the average binding affinity of the peptide panel was higher for the *HLA-DRB1* alleles that encode serine at position 11 versus alleles that have other amino acids at this position (Wilcoxon signed-rank test, $p = 3.44 \times 10^{-136}$, **Supplementary Table 4**). To compare these two clusters, we then computed the ratio of the average binding affinity of the 6 HLA-DRB1 allotypes that contain serine-11 and the 7 that have other amino acids at this position (**Supplementary Table 4**). Since MHC class II molecules at the cell surface present repertoires that are skewed in favor of high affinity binders,[39] we selected for peptides with an (IC_{50}) affinities <500 nM (an affinity of <500 nM is routinely used as a threshold for potential immunogenicity) or <50nM (strong binding peptides) for HLA-DRB1 allotypes with serine-11 (**Materials and Methods**). The data indicated that peptides with an intermediate or high affinity for allotypes containing serine-11 have less affinity to allotypes that contain other amino acids at position 11 (**Supplementary Table 4**).

Previously-described associations to JIA and uveitis

A number of HLA types and amino acid positions have been demonstrated to be associated to JIA, uveitis and related autoimmune diseases through genome-wide association and candidate gene studies.[31,40,41] We looked up 10 previously-associated HLA types and amino acids in our GWAS of (1) JIA and unaffected controls and (2) uveitis and unaffected controls. Seven of the 10 associations appeared in our data, and all associations were nominally significant ($p < 0.05$; **Supplementary Table 5** and **Supplementary Figure 8**).

Discussion

Through interrogation of imputed HLA types and amino acids, we have identified the amino acid serine at position 11 in the *HLA-DRB1* gene as strongly associated to increased risk of uveitis in female patients with JIA. This association indicates that, though JIA with uveitis and JIA without uveitis share clinical features and genetic risk factors, there is genetic architecture specific to uveitis[4,42,43] These findings suggest that the two phenotypes are likely to have biological features and progressions that are partially distinct.[44]

Association testing initially revealed either serine or aspartic acid at position 11 in *HLA-DRB1* as the most-associated variant, a signal we determined was primarily driven by serine. However, the perfect linkage disequilibrium between serine, tyrosine (position 10), and threonine (position 12) makes it impossible to disentangle the three amino acids in a statistical framework; a subset of the amino acids, or all three (that is, the YST-motif) may

be key to disease onset. Importantly, though all three residues are located at the bottom of the HLA-DRB1 peptide-binding groove, only serine-11 is positioned towards, and, thus most likely interacting with, binding peptide epitopes (**Figure 2**). Intriguingly, previous work has identified serine-11 as the strongest risk factor in seronegative rheumatoid arthritis (RA). Some JIA patients, including those with uveitis, might be categorized as seronegative RA by the time they reach adulthood, and seronegative RA is considered to genetically mirror the uveitis-prone JIA categories.[31] In contrast, serine-11 is highly protective against seropositive RA, a biologically distinct form of RA in which uveitis is not common (<1% of cases).[45,46]

Sex-specific analyses in the MHC further revealed the *HLA-DRB1* signal was unique to female samples and that the signal was entirely absent when the analysis was restricted to male samples only. Although in-depth data on sex-specific difference in the epidemiology of uveitis associated with JIA are sparse, the typical bilateral chronic anterior uveitis occurs more frequently in females.[37,47] This genotype-by-sex interaction informs an exciting field of future research that aims to elucidate how the uveitis risk mechanisms underlying the *HLA-DRB1* signal may be mediated by sex-specific factors, such as hormone regulation. Further, the sex-specific pattern of genetic association may help to resolve previously reported evidence of sexual dimorphism with respect to severity or disease course in children with JIA-associated uveitis.[48] Lastly, the finding indicates that sex stratification may be beneficial for future clinical trials, as some therapeutic agents may be more efficacious in females or males only.[49,50]

HLA-DRβ1, like other MHC class II molecules, is essential to immunity and orchestrates downstream immune responses by presenting a dynamic cargo of thousands of different peptides for scrutiny by T helper cells.[51] Both JIA and uveitis are considered diseases mediated by T helper cell subsets.[4] The strong association of an amino acid motif (tagged by serine-11) in *HLA-DRB1* suggests that predetermined peptide preferences may affect the regulation of T helper cells. To investigate the immunological impact of the motif in *HLA-DRB1*, we used a "reverse immunology" approach[38] (**Materials and Methods**), performed using data from proteins experimentally observed in human iris.[52] Since JIA-uveitis patients frequently have circulating ANAs directed to iris tissues, we specifically focused on nuclear proteins to provide a disease-relevant dataset for modelling. Although the large panel of putative antigens that we tested was by no means exhaustive (**Supplementary Table 3**), the experiment demonstrated that the_presence of a serine-11 motif in *HLA-DRB1* is accompanied by distinguished changes in binding affinity in the resulting protein (**Supplementary Table 4, Supplementary Figures 7**), suggesting that antigen presentation by the HLA-DRβ1 protein may underpin the uveitis susceptibility. Risk alleles of *HLA-DRB1* that encode serine-11 (e.g. *HLA-DRB1*11*, **Supplementary Table 2**) may communicate distinct 'peptidomes' that influence T helper cells and significantly increase the likelihood of downstream immune responses (e.g., ANA) to the eye. Further, ANA-positivity is modestly correlated with presence of serine-11 in the uveitis cases studied here (Pearson's r= 0.28, $P = 1.0 \times 10^{-4}$). Functional studies using HLA-proteomic approaches will be necessary to experimentally validate and systematically dissect the complex downstream effects of serine-11 on tissue-specific antigen presentation by HLA-DRβ1 in uveitis.

The relatively high frequency of serine-11 in the JIA cases that did not develop uveitis (**Table 2**) indicates the likely involvement of additional (epi)genetic and environmental factors. Indeed, genes outside the MHC have been implicated in JIA-uveitis, including a polymorphism in *VTCN1*[53] and variants near the immune genes *TRAF1* and *C5*.[54] More than half of patients with JIA-uveitis display elevated titers of anti–parvovirus B19 antibody in ocular fluids,[43,55] and human parvovirus B19 infection can cause various clinical features, including arthritis, uveitis and ANA production.[56,57] In addition, immunity to Parvo-B19 infection is strongly associated with *HLA-DRB1* and involves T helper cells. These observations collectively outline a complex interplay of both biologic and environmental factors in JIA-uveitis.[58–61]

Though we were able to identify the serine-11 association in this collection of JIA-uveitis and JIA without uveitis samples, our analyses leave unanswered questions as to the genetic underpinnings of uveitis, and potential sex-based differences in their pathogenesis. Given our sample size, we were well powered to detect common (frequency >10%) and highly-penetrant (odds ratio > 3, **Supplementary Figure 9**) variation that associates with increased uveitis risk compared to JIA individuals without uveitis. However, we are underpowered to identify common, modestly-penetrant variation - the hallmark genetic feature of complex phenotypes[62,63] - that may further distinguish the two phenotypes. In the current study, we included only patients for whom we were able to obtain detailed ophthalmological follow-up for ≥ 4 year (±90% of uveitis develops within 4 years after onset).[64] While imperative to revealing the serine-11 association, this strategy limited case ascertainment. Future genetic studies will likely benefit from broader inclusion of cases, including essential, detailed phenotype information,[7,65] in order to identify additional genetic associations that are shared across or distinct to particular uveitis subtypes of gender.[66] In addition, clinical studies - including clinical trials - may be improved by being well-powered in both sexes to enable sex-specific investigations.

While much progress has been made in understanding the biology of JIA,[16,31] our understanding of the genetic landscape of uveitis is more limited. Only a handful of candidate gene studies[41] and HLA-specific analyses[4,17–19] have investigated uveitis as a phenotype separate from all JIA. A recent GWAS identified *HLA-DRB1*1501* as a risk factor for of non-anterior uveitis,[67] and recent studies in children with anterior uveitis in JIA point toward a role for infiltrated B plasma cells[68] , CD4+ T helper cells[4] , and changes in the ocular fluid microenvironment.[43,69,70] Future, larger genetic studies comparing JIA-uveitis and JIA without uveitis, with and without sex-specific analysis, will likely expand our understanding of uveitis pathogenesis in a number of ways. Broad case inclusion will elucidate whether subtype-specific architectures exist and larger samples will improve statistical power to identify more modest-effect disease-modifying variants both within and outside of the MHC. For example, this may substantiate the association of additional HLA alleles (such as *HLA-DPB1*) with uveitis in JIA (**Figure 1**).

Larger studies may also help in revealing clinically-useful biomarkers. Risk factors such as young age and presence of ANA, assumed to reflect aberrant immune activation,[37,40,64,71] indicate patients potentially at higher risk of uveitis but are in no way predictive of disease outcome. An easily-testable biomarker such as a genotype would be highly impactful in the

management of uveitis, as it would help in overcoming the challenge of detecting the insidious onset of disease[7,48,65]. Intriguingly, after deeper examination of the phenotypic data, we found that 99% of female uveitis cases in our study carry at least one copy of the variant coding for serine-11. The only female uveitis case who did not carry serine-11 in *HLA-DRB1* appeared to suffer from ANA-negative oligoarthritis with mild vitritis and peripheral multifocal choroiditis in the absence of anterior segment inflammation; this ocular finding is atypical for JIA, and thus, according to our inclusion criteria, this sample was likely improperly included at cohort collection. Prospective studies in larger populations, including detailed clinical evaluation of development of uveitis and secondary uveitis phenotypes, will be necessary to dissect the potential of serine-11 or other genotypes as biomarkers for disease risk. Regardless, the current study justifies further genetic analysis as a potentially powerful instrument in moving analysis of the genetic architecture of uveitis towards accurate and efficient clinical decision tools.

The results of our study represent a key step in understanding the pathogenesis of uveitis in JIA, helping to discern biologically shared and distinct features of JIA-uveitis and JIA without uveitis. Future work will allow us to further disentangle the two phenotypes and evaluate shared and distinct disease etiology. Unraveling the pathogenesis of the two diseases will improve understanding of the key pathways that trigger disease. By pinpointing and understanding these mechanisms, we can potentially identify clinically-useful biomarkers that stratify patients for disease risk in order to improve personalized care, and catalyze future lines of research in precision medicine. Collectively, this work will advance our progress towards treating and preventing sight-threatening complications of uveitis in children with JIA.

Material and methods

Patient collection

JIA and uveitis samples

A total of 576 patients with oligoarticular extended, oligoarticular persistent or rheumatoid factor negative polyarticular JIA, of whom 214 (37%) had developed uveitis, were enrolled in this study. These patients were collected from two independent cohorts: (1) a Dutch and cohort, and (2) a cohort comprised of samples collected in Germany, Belgium and Switzerland. For simplicity, we will refer to these as the Phase one and Phase 2 sample sets, respectively. Three-hundred eighty four patients were included in the Phase 1 samples (137 uveitis samples) and 192 patients were included in the Phase 2 samples (77 uveitis samples).

JIA was diagnosed according to the criteria of the International League of Associations for Rheumatology (ILAR), or by former criteria (e.g. European League Against Rheumatism (EULAR)).[2,72] All patients were screened by an ophthalmologist according to the guidelines of the Academy of Pediatrics and patients with no clinical signs of uveitis had an ophthalmologic follow-up of at least 4 years after onset of JIA.[73] Patients with at least trace cells or more in the anterior chamber and treated with at least topical steroids during ophthalmologic examinations, were diagnosed with JIA-associated (anterior) uveitis.

DNA material of JIA patients with and without uveitis from Phase 1 were collected at the University Medical Center Utrecht, University Medical Center Leiden, Erasmus Medical Center Rotterdam, Academic Medical Center Amsterdam and Radboud University Medical Center Nijmegen (all based in the Netherlands). Samples from Phase 2 were collected within the ICON study and provided by the ICON biobank at the Westfälische Wilhelms-Universität Münster (ICON-JIA Study, Germany), the University Hospitals Leuven (Belgium), the University Children's Hospital at Zurich (Switzerland).

Population-level control samples

Genotype data from 394 unrelated and unaffected Dutch samples were used as population controls and had been previously genotyped using the same platform as the JIA and uveitis samples contained in this study.[74]

Data collection

Information such as JIA subtype, sex and uveitis status were collected from all patients; presence of antinuclear antibodies (ANA) and age at JIA onset could be collected from the majority of patients (85%) from samples included in Phase 1.

This study was approved by the local Institutional Review Boards and is in compliance with Helsinki principles. Informed consent was obtained from all participating patients if they

were 18 years or older, from both parents and patients if they were 12-18 years of age, and from parents only if they were younger than 12 years old.

Genotyping and quality control

All samples were genotyped using the Infinium HumanOmniExpress-24v1.1 and v1.2 arrays, which contain ~700,000 SNPs. For Phase 1, all JIA cases (with and without uveitis) were genotyped together; unaffected controls were genotyped on the same array, but as a separate batch. All Phase 2 samples were genotyped together.

Sample-level quality control

Before performing imputation, we ran sample-level and SNP-level data quality control, following standard genome-wide association quality control steps. All steps, thresholds, and number of removed samples and SNPs are summarized in **Supplementary Table 6**. Briefly, we used PLINK 1.9[30] to first check sample-level missingness across all samples and removed all samples with missingness >5%. Next, for the remainder of sample-level quality control, we reduced the dataset to only high-quality SNPs: SNPs with minor allele frequency (MAF) > 10%; SNPs outside the MHC (chromosome 6), lactase locus (chromosome 2), and the inversions on chromosomes 8 and 17; SNPs with missingness < 0.1%; and SNPs linkage disequilibrium (LD) pruned at an LD (r^2) threshold of 0.2.

With this high-quality set of SNPs, we performed principal component analysis (PCA) using EIGENSTRAT.[75] We defined samples to be of European ancestry if their principal component (PC) 1 and PC 2 values were within 6 standard deviations of the European-ancestry populations included in the HapMap 3 dataset (the CEU population, samples of Northern and Western European ancestry living in Utah; and the TSI population, Toscans living in Italy). In addition, we calculated inbreeding coefficients for all samples and removed samples further than 3 standard deviations from the distribution. Related samples (pi-hat > 0.125) were also removed. Finally, using all SNPs available on chromosome X, we performed a sex check to identify samples with mismatching genotypic and phenotypic sex; samples with mismatching genotype and phenotype information were removed from the analysis.

SNP-level quality control

We dropped failing samples from the data and then proceeded to perform SNP QC. All SNPs with missingness >5% were removed, as were SNPs with frequency <1% (too rare to be analyzed in the logistic regression framework, given the sample size). SNPs out of Hardy-Weinberg equilibrium in controls or across the full dataset were removed ($p < 1 \times 10^{-6}$ in Phase 1; $p < 1 \times 10^{-3}$ in Phase 2), as were SNPs with significant differential missingness between cases and controls ($p < 5 \times 10^{-2}$ in Phase 1; $p < 1 \times 10^{-3}$ in Phase 2).

Once sample- and SNP-level QC were complete and the failing samples and SNPs had been removed from the data, we performed a second iteration of PCA to ensure that cases and controls were overlapping in PCA space (i.e., there was no obvious population stratification in the data) (**Supplementary Figure 10**).

Prephasing and imputation

Prephasing

Once sample- and SNP-level were complete, the Phase 1 and Phase 2 data were phased separately, as they were genotyped and QC'd separately, using SHAPEIT2.[22] As the sample size was >100 samples, prephasing was run without a reference panel, per the SHAPEIT2 recommendations.[22]

Genome-wide imputation

Following the prephasing, we imputed all samples using the imputation reference panel constructed through whole-genome sequencing of 2,504 samples in 1000 Genomes Project (Phase 3).[27] Briefly, the 1000 Genomes Project Phase 3 data is comprised of samples collected from the Americas, Africa, East Asia, Europe and South Asia. Samples were whole-genome sequenced at ~80x coverage across the exome and at ~4x coverage in non-exome regions. We imputed the prephased haplotypes using IMPUTE2[23,76]; the data were phased in windows 5 megabases long, with a buffer region of 250kb. We set the effective sample size (Ne) to 20,000, per IMPUTE2 recommendations, and set k_hap to 1,000, as our sample was of European ancestry and the 1000 Genomes Phase 3 imputation reference panel contains ~1,000 haplotypes from European-ancestry individuals.

Imputation of HLA classical alleles and amino acids

To impute amino acids and HLA alleles in the MHC, we used the SNP2HLA pipeline.[28] In short, the SNP2HLA pipeline uses a reference panel assembled through HLA typing of 5,225 European-ancestry individuals collected by the Type 1 Diabetes Genetics Consortium.[77] The panel includes SNPs and amino acids in the MHC, as well as HLA types from the Class I (*HLA-A*, *HLA-B*, and *HLA-C*) and Class II (*HLA-DRB1*, *HLA-DQA1*, *HLA-DQB1*, *HLA-DPA1*, and *HLA-DPB1*) HLA genes. HLA types are imputed to 2- and 4-digit resolution. The SNP2HLA pipeline uses BEAGLE[78] to phase the data and then impute.

Post-imputation, we checked the sum of the dosages across each HLA gene for each individual. Samples with dosages >2.5 at any one of the HLA genes (induced through imprecision in the imputation) were dropped from further analysis.

Association testing

Genome-wide association testing

To perform genome-wide association testing, we coded JIA samples with uveitis as cases and JIA samples without uveitis as controls. First, we performed a genome-wide association study (GWAS) in Phase 1 and Phase 2 individually, to check the overall behavior of the genome-wide test statistics ($lambda_{Phase1}$ = 1.008, $lambda_{Phase2}$ = 1.017). GWAS were performed using PLINK 1.9[30] using an additive logistic regression model, correcting for the top two principal components and sex. Data were then meta-analyzed using METAL.[32] In the Phase 1 data, additional GWAS were performed comparing samples with JIA and uveitis to unaffected controls, as well as JIA samples without uveitis to unaffected controls; results

were concordant with studies previously performed in these phenotypes (**Supplementary Table 5**).

Association testing in the MHC

Similar to the genome-wide association testing, additive logistic regression in the Phase 1 and Phase 2 data was first performed separately to check the overall behavior of the data. Then, the dosages were merged together, and logistic regression was performed on the dosage data using PLINK1.9,[29,30] correcting for the top 5 principal components, sex, and phase. To ensure that this mega-analysis approach was appropriate, we additionally performed an inverse variance-weighted meta-analysis of the two phases, and found that the results were highly concordant (Pearson's r of genome-wide betas = 0.95). To identify potential independent signals within the MHC, we performed a follow-up regression by conditioning on the top-most variant (**Figure 1**).

In silico peptide binding prediction to *HLA-DRβ1*

Mass spectrometric proteome data from human iris tissues (2,959 nonredundant proteins) was used as a representative source of proteins present in iris tissue.[52] JIA-uveitis patients commonly have antinuclear antibodies (ANA) directed to iris tissues, thus, to we focused on nuclear iris proteins to provide a potentially disease-relevant dataset. Protein accession numbers were extracted to filter in UniProt (Universal Protein Resource) for proteins with a subcellular location annotation limited to the nucleus.[79] We selected 147 proteins (**Supplementary Table 3**) that fulfilled these criteria and their full length amino acid sequences were fed into the neural network of the *netMHCIIpan3.1* server. The netMHCIIpan algorithm[38] was selected for its high quality performance in peptide binding prediction for any MHC class II allele for which the sequence is identified, including those for whom only few experimental peptide binding affinities (limited allele-specific training data) are available. We ascertained this by testing the predicted binding affinity of several hundreds of experimentally identified 15-mer peptide ligands eluted from HLA-DRβ1:01,[80,81] which revealed that >95% were correctly picked up by netMHCIIpan as ligands for DRβ1:01 (IC_{50} <500nM, data not shown).

Next, we tested the predicted affinities of all 83,686 overlapping 15-mer peptides from the selected 147 proteins in *netMHCIIpan3.1*. that tentatively could be presented by HLA-DR for binding to representative four-digit alleles of *HLA-DRB1* (**01:01, *03:01, *04:01, *07:01, *08:01, *09:01, *10:01, *11:01, *12:01, *13:01, *14:01, *15:01,* and **16:01*). These *DRB1* alleles account for 79% of *HLA-DRβ1* alleles in the JIA cases. The IC50 (affinity), data was log-transformed to a value between 0 and 1 using: $1-\log(IC_{50}nM)/\log(50{,}000)$.[38] To categorize *HLA-DRB1* alleles that encode allotypes with similar predicted binding preferences, we performed unsupervised hierarchical clustering. Heatmaps were created based on the Euclidean distance measure and the Ward's linkage method using the MetaboAnalyst server.[82] We computed the ratio of the average binding affinity of HLA-DRβ1 molecules that contain Serine-11 in the peptide-binding groove (**03:01, *08:01, *11:01, *12:01, *13:01,* and **14:01*) over the average binding affinity of HLA-DRβ1 proteins that have other amino acids at this position (**01:01, *04:01, *07:01, *09:01, *10:01, *15:01,*

and **16:01*) as a measure for the overall difference in predicted binding affinity for each peptide. A frequency distribution of the ratio of the binding affinities was plotted for the entire set of peptides (including low and nonbinding peptides), potential ligands for DRβ1 with a binding affinity stronger than half maximal inhibitory concentration (IC50) of 500 nM, or peptides with high binding affinity for Serine-11 encoding alleles(IC_{50}<50nM).

Tables and Figures

Table 1 | Samples included in Phase 1 and Phase 2 of the analysis. Sample numbers are shown both pre- and post-quality control (QC). The subset of JIA cases that also have uveitis (JIA-uveitis) are also indicated. The proportion of patients that are positive for antinuclear antibodies (ANA+) as well as age and sex distribution are shown.

	Number of samples	
	JIA patients	JIA-uveitis (%)
Pre-QC		
Phase 1	384	137 (36)
Phase 2	192	77 (40)
Total	576	214 (37)
Post-QC		
Phase 1	357	126 (35)
Phase 2	165	66 (40)
Total	522	192 (37)
ANA+	204/295*	68/85 (80%)*
Female (%)	364	136 (71%)
Male (%)	158	56(29%)

*number of cases with available data for ANA testing.

Table 2 | Association results for amino acids in *HLA-DRB1*. The top results from the mega-analysis in uveitis and JIA without uveitis in *HLA-DRB1*. All reported results correspond to presence of the given amino acid residue(s). Presence of serine (Ser) or aspartic acid (Asp) at position 11 in *HLA-DRB1* (imputation info score = 1.11) was the top hit after initial association testing. Conditioning on either aspartic acid or serine revealed that the amino acids at positions 10-13, all well-imputed (imputation info = 1.08) and in perfect linkage disequilibrium, explain the bulk of the initial signal.

***DRβ1* position**	**Amino acid residue(s)**	**Classical HLA alleles**	**Samples** (all, females, males)	**Mega-analysis**		
				Freq. Case | Cntl	**OR [95% CI]**	**P-value**
Initial association testing						
11	Ser or Asp	*03,*08,*09, *11,*12,*13,*14	All	0.79 | 0.59	2.59 [1.92 - 3.50]	4.80×10^{-10}
11	Ser	*03,*08,*11, *12,*13,*14	All	0.77 | 0.57	2.47 [1.84 - 3.32]	1.44×10^{-9}
Conditioning on aspartic acid (Asp) at position 11						
11	Ser or Asp	*03,*08,*09, *11,*12,*13,*14	All	0.79 | 0.59	2.60 [1.92 - 3.52]	5.43×10^{-10}
11	Ser	*03,*08,*11, *12,*13,*14		0.77 | 0.57	2.60 [1.92 - 3.52]	5.46×10^{-10}
Conditioning on serine (Ser) at position 11						
11	Ser or Asp	*03,*08,*09, *11,*12,*13,*14	All	0.79 | 0.59	2.36 [0.93 - 5.98]	0.069
Sex-specific association testing						
233	Thr	*01,*04,*07,*08, *09,*10,*15, *16	Females	0.18 | 0.44	0.30 [0.20 - 0.43]	3.50×10^{-10}
			Males	0.37 | 0.46	0.73 [0.44 - 1.20]	0.210
11	Ser or Asp	*03,*08,*09, *11,*12,*13,*14	Females	0.84 | 0.59	3.48 [2.35 - 5.16]	4.92×10^{-10}
			Males	0.67 | 0.56	1.52 [0.92 - 2.51]	0.100
11	Ser	*03,*08,*11, *12,*13,*14	Females	0.82 | 0.57	3.30 [2.26 - 4.83]	7.61×10^{-10}
			Males	0.64 | 0.55	1.41 [0.86 - 2.31]	0.177
Conditioning on threonine (Thr) at position 233						
11	Ser or Asp	*03,*08,*09, *11,*12,*13,*14	Females	0.84 | 0.59	1.98 [0.73 - 5.32]	0.178
			Males	0.67 | 0.56	3.14 [0.70 - 14.15]	0.136
11	Ser	*03,*08,*11, *12,*13,*14	Females	0.82 | 0.57	0.84 [0.10 - 6.94]	0.878
			Males	0.64 | 0.55	2.83 [0.16 - 51.47]	0.482
Sex-interaction association testing						
11	Ser*Sex	*03,*08,*11, *12,*13,*14	All	--	2.25 [1.22 - 4.15]	0.0096

Association statistics for tyrosine (position 10), threonine (position 12) and serine or glycine (position 13) are identical to the association statistics reported for serine (position 11) due to linkage disequilibrium.

Figure 1 | Association and conditional testing in *HLA-DRB1*. A. Initial association testing in the MHC revealed a genome-wide significant signal at *HLA-DRB1* position 11 (presence of serine (S) or aspartic acid (D), purple diamond). **B.** Conditioning on the presence of aspartic acid at position 11 left the association signal essentially unchanged (green and white diamonds, grey outline); presence of serine remained the strongest association (pale purple diamond). Conditioning on the presence of serine at position 11 (aquamarine and white diamonds, black outline), dramatically mitigated the association signal, indicating that presence of serine explains the bulk of the association at *DRB1* position 11. Presence of lysine (K) at position 69 in *HLA-DPB1* remained modestly associated (dark purple diamond). **C.** Conditioning on lysine at position 69 in *HLA-DPB1* removes the remainder of the signal (rs6457109, dark purple diamond, p = 0.0024).

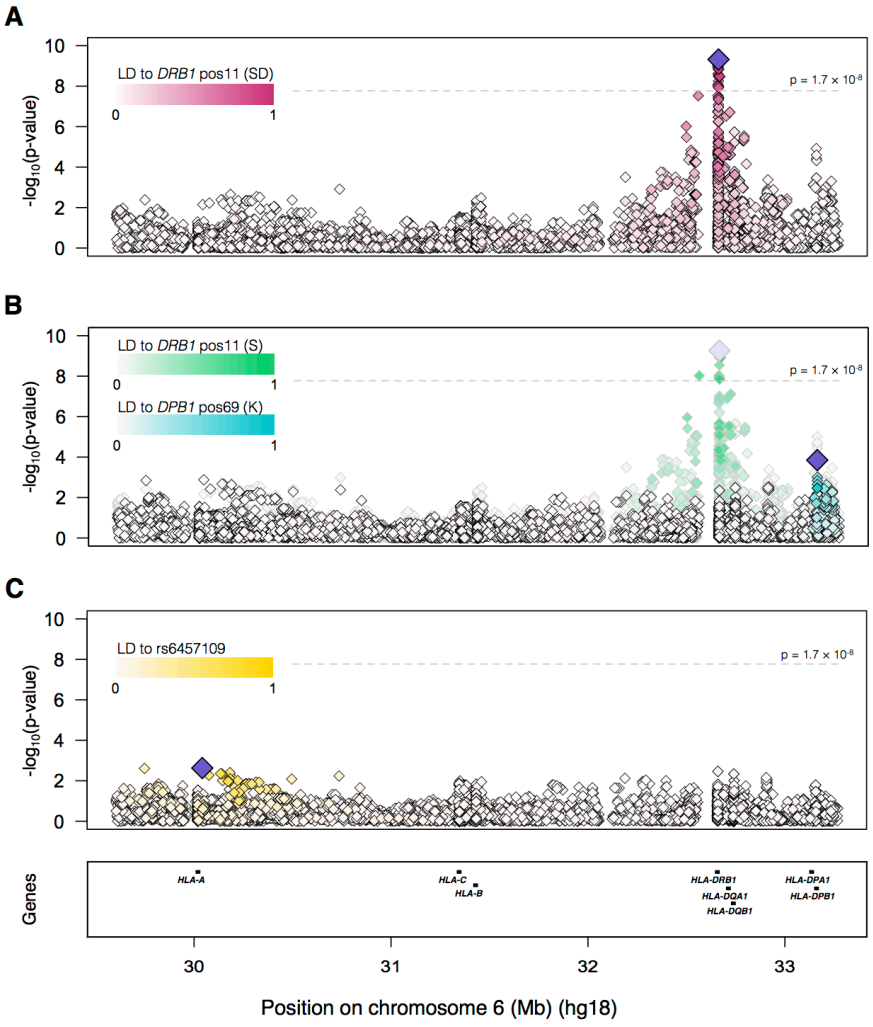
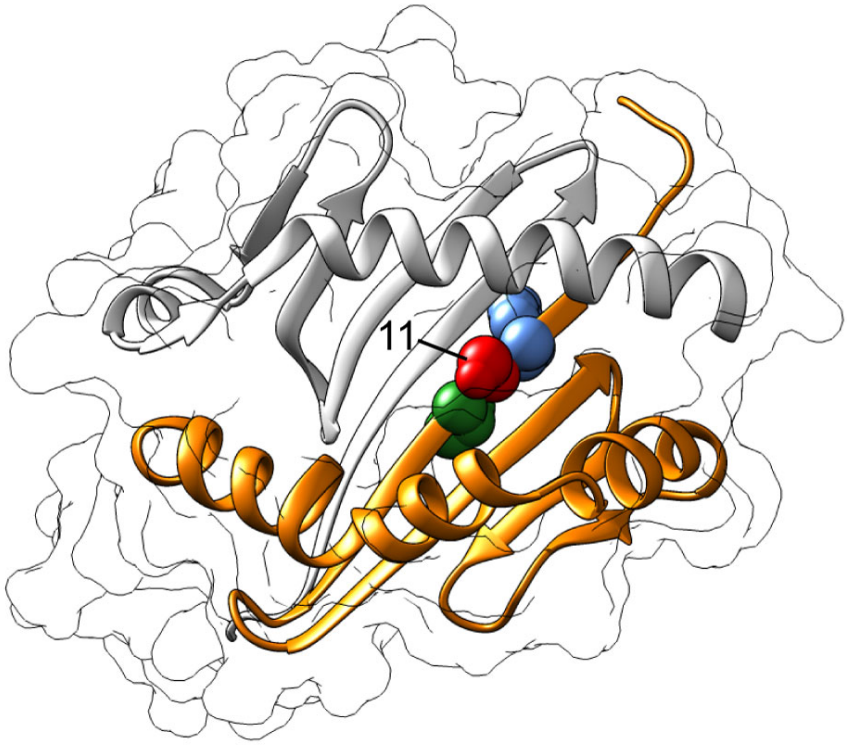


Figure 2 | Three-dimensional ribbon model for HLA-DR (Protein Data Bank entry: 3pdo)**.** The molecule is positioned to provide a view from the top of the peptide-binding groove. The beta-chain (DRß) is highlighted in orange. Amino acid serine at position 11 (red) is located in the bottom center of the peptide-binding groove of HLA-DRß1. Adjacent amino acid tyrosine at position 10 (green) and threonine at position 12 (blue) and serine at position 11 are displayed as spheres. 3D structure was produced using UCSF Chimera.[83]



References

1. Borchers, A. T. *et al.* Juvenile idiopathic arthritis. *Autoimmun. Rev.* **5,** 279–298 (2006).
2. Petty, R. E. *et al.* International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. *J. Rheumatol.* **31,** 390–392 (2004).
3. Ravelli, A. & Martini, A. Juvenile idiopathic arthritis. *Lancet* **369,** 767–778 (2007).
4. Sen, E. S., Dick, A. D. & Ramanan, A. V. Uveitis associated with juvenile idiopathic arthritis. *Nat. Rev. Rheumatol.* **11,** 338–348 (2015).
5. Kesen, M. R., Setlur, V. & Goldstein, D. A. Juvenile idiopathic arthritis-related uveitis. *Int. Ophthalmol. Clin.* **48,** 21–38 (2008).
6. Haasnoot, A.-M. J. W. *et al.* Impact of Juvenile Idiopathic Arthritis Associated Uveitis in Early Adulthood. *PLoS One* **11,** e0164312 (2016).
7. Gregory, A. C., 2nd *et al.* Risk factors for loss of visual acuity among patients with uveitis associated with juvenile idiopathic arthritis: the Systemic Immunosuppressive Therapy for Eye Diseases Study. *Ophthalmology* **120,** 186–192 (2013).
8. de Boer, J., Wulffraat, N. & Rothova, A. Visual loss in uveitis of childhood. *Br. J. Ophthalmol.* **87,** 879–884 (2003).
9. Tugal-Tutkun, I., Havrlikova, K., Power, W. J. & Foster, C. S. Changing patterns in uveitis of childhood. *Ophthalmology* **103,** 375–383 (1996).
10. Cassidy, J. *et al.* Ophthalmologic examinations in children with juvenile rheumatoid arthritis. *Pediatrics* **117,** 1843–1845 (2006).
11. Haasnoot, A. J. W. *et al.* Erythrocyte sedimentation rate as baseline predictor for the development of uveitis in children with juvenile idiopathic arthritis. *Am. J. Ophthalmol.* **159,** 372–7.e1 (2015).
12. Heiligenhaus, A., Heinz, C., Edelsten, C., Kotaniemi, K. & Minden, K. Review for disease of the year: epidemiology of juvenile idiopathic arthritis and its associated uveitis: the probable risk factors. *Ocul. Immunol. Inflamm.* **21,** 180–191 (2013).

13. Kalinina Ayuso, V. *et al.* Pathogenesis of juvenile idiopathic arthritis associated uveitis: the known and unknown. *Surv. Ophthalmol.* **59,** 517–531 (2014).
14. Hinks, A. *et al.* Identification of a novel susceptibility locus for juvenile idiopathic arthritis by genome-wide association analysis. *Arthritis & Rheumatism* **60,** 258–263 (2009).
15. Thompson, S. D. *et al.* Genome-wide association analysis of juvenile idiopathic arthritis identifies a new susceptibility locus at chromosomal region 3q13. *Arthritis & Rheumatism* **64,** 2781–2791 (2012).
16. Hinks, A. *et al.* Dense genotyping of immune-related disease regions identifies 14 new susceptibility loci for juvenile idiopathic arthritis. *Nat. Genet.* **45,** 664–669 (2013).
17. Giannini, E. H. *et al.* Longitudinal analysis of HLA associated risks for iridocyclitis in juvenile rheumatoid arthritis. *J. Rheumatol.* **18,** 1394–1397 (1991).
18. Melin-Aldana, H. *et al.* Human leukocyte antigen-DRB1*1104 in the chronic iridocyclitis of pauciarticular juvenile rheumatoid arthritis. *J. Pediatr.* **121,** 56–60 (1992).
19. Du, L., Kijlstra, A. & Yang, P. Immune response genes in uveitis. *Ocul. Immunol. Inflamm.* **17,** 249–256 (2009).
20. Skol, A. D., Scott, L. J., Abecasis, G. R. & Boehnke, M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat. Genet.* **38,** 209–213 (2006).
21. O'Connell, J. *et al.* A General Approach for Haplotype Phasing across the Full Spectrum of Relatedness. *PLoS Genet.* **10,** (2014).
22. Delaneau, O., Marchini, J. & Zagury, J.-F. A linear complexity phasing method for thousands of genomes. *Nature Methods* **9,** 179–181 (2011).
23. Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. & Abecasis, G. R. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat. Genet.* **44,** 955–959 (2012).
24. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method

for genome-wide association studies by imputation of genotypes. *Nat. Genet.* **39,** 906–913 (2007).

25. Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **5,** (2009).
26. Marchini, J. & Howie, B. Genotype imputation for genome-wide association studies. *Nat. Rev. Genet.* **11,** 499–511 (2010).
27. Auton, A. *et al.* A global reference for human genetic variation. *Nature* **526,** 68–74 (2015).
28. Jia, X. *et al.* Imputing Amino Acid Polymorphisms in Human Leukocyte Antigens. *PLoS One* **8,** (2013).
29. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81,** 559–575 (2007).
30. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4,** 1–16 (2015).
31. Hinks, A. *et al.* Fine-mapping the MHC locus in juvenile idiopathic arthritis (JIA) reveals genetic heterogeneity corresponding to distinct adult inflammatory arthritic diseases. *Ann. Rheum. Dis.* (2016). doi:10.1136/annrheumdis-2016-210025
32. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26,** 2190–2191 (2010).
33. Lin, D. Y. & Zeng, D. Meta-analysis of genome-wide association studies: no efficiency gain in using individual participant data. *Genet. Epidemiol.* **34,** 60–66 (2010).
34. Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium. Genome-wide association study identifies five new schizophrenia loci. *Nat. Genet.* **43,** 969–976 (2011).
35. Kriener, K., O'hUigin, C., Tichy, H. & Klein, J. Convergent evolution of major histocompatibility complex molecules in humans and New World monkeys.

Immunogenetics **51,** 169–178 (2000).

36. Bengtsson, M., Jansson, I. E., Danielsson, F., Henrysson, H. & Källsten, K. Identification of a novel HLA DRB1 exon 2 sequence, DRB1*1345. *Tissue Antigens* **59,** 159–161 (2002).
37. Saurenmann, R. K. *et al.* Prevalence, risk factors, and outcome of uveitis in juvenile idiopathic arthritis: a long-term followup study. *Arthritis Rheum.* **56,** 647–657 (2007).
38. Andreatta, M. *et al.* Accurate pan-specific prediction of peptide-MHC class II binding affinity with improved binding core identification. *Immunogenetics* **67,** 641–650 (2015).
39. Ferrante, A. Thermodynamics of Peptide-MHC Class II Interactions: Not all Complexes are Created Equal. *Front. Immunol.* **4,** 308 (2013).
40. Zulian, F. *et al.* Early predictors of severe course of uveitis in oligoarticular juvenile idiopathic arthritis. *J. Rheumatol.* **29,** 2446–2453 (2002).
41. Zeggini, E. *et al.* Association of HLA-DRB1*13 with susceptibility to uveitis in juvenile idiopathic arthritis in two independent data sets. *Rheumatology* **45,** 972–974 (2006).
42. Swart, J. F., de Roock, S. & Prakken, B. J. Understanding inflammation in juvenile idiopathic arthritis: How immune biomarkers guide clinical strategies in the systemic onset subtype. *Eur. J. Immunol.* **46,** 2068–2077 (2016).
43. Haasnoot, A.-M. J. W. *et al.* Ocular Fluid Analysis in Children Reveals Interleukin-29/Interferon-λ1 as a Biomarker for Juvenile Idiopathic Arthritis-Associated Uveitis. *Arthritis Rheumatol* **68,** 1769–1779 (2016).
44. Grassi, A., Corona, F., Casellato, A., Carnelli, V. & Bardare, M. Prevalence and outcome of juvenile idiopathic arthritis-associated uveitis and relation to articular disease. *J. Rheumatol.* **34,** 1139–1145 (2007).
45. Generali, E., Cantarini, L. & Selmi, C. Ocular Involvement in Systemic Autoimmune Diseases. *Clin. Rev. Allergy Immunol.* **49,** 263–270 (2015).
46. Vignesh, A. P. P. & Srinivasan, R. Ocular manifestations of rheumatoid arthritis and their correlation with anti-cyclic citrullinated peptide antibodies. *Clin. Ophthalmol.* **9,** 393–397

(2015).

47. Moradi, A., Amin, R. M. & Thorne, J. E. The role of gender in juvenile idiopathic arthritis-associated uveitis. *J. Ophthalmol.* **2014,** 461078 (2014).
48. Kalinina Ayuso, V., Ten Cate, H. A. T., van der Does, P., Rothova, A. & de Boer, J. H. Male gender and poor visual outcome in uveitis associated with juvenile idiopathic arthritis. *Am. J. Ophthalmol.* **149,** 987–993 (2010).
49. Whitley, H. & Lindsey, W. Sex-based differences in drug activity. *Am. Fam. Physician* **80,** 1254–1258 (2009).
50. Regitz-Zagrosek, V. Therapeutic implications of the gender-specific aspects of cardiovascular disease. *Nat. Rev. Drug Discov.* **5,** 425–438 (2006).
51. Gutierrez-Arcelus, M., Rich, S. S. & Raychaudhuri, S. Autoimmune diseases - connecting risk alleles with molecular traits of the immune system. *Nat. Rev. Genet.* **17,** 160–174 (2016).
52. Zhang, P. *et al.* Defining the proteome of human iris, ciliary body, retinal pigment epithelium, and choroid. *Proteomics* **16,** 1146–1153 (2016).
53. Alberdi-Saugstrup, M. *et al.* Non-HLA gene polymorphisms in juvenile idiopathic arthritis: associations with disease outcome. *Scand. J. Rheumatol.* 1–8 (2017).
54. Pers, Y.-M. *et al.* Association of TRAF1-C5 with risk of uveitis in juvenile idiopathic arthritis. *Joint Bone Spine* (2016). doi:10.1016/j.jbspin.2016.04.014
55. de Groot-Mijnes, J. D. F. *et al.* Antibody Production against B19 Virus in Ocular Fluid of JIA-Associated Uveitis Patients. *Ophthalmology* **122,** 1270–1272.e1 (2015).
56. Kerr, J. R. Pathogenesis of parvovirus B19 infection: host gene variability, and possible means and effects of virus persistence. *J. Vet. Med. B Infect. Dis. Vet. Public Health* **52,** 335–339 (2005).
57. Maini, R. & Edelsten, C. Uveitis associated with parvovirus infection. *Br. J. Ophthalmol.* **83,** 1403–1404 (1999).
58. Kerr, J. R., Mattey, D. L., Thomson, W., Poulton, K. V. & Ollier, W. E. R. Association of

symptomatic acute human parvovirus B19 infection with human leukocyte antigen class I and II alleles. *J. Infect. Dis.* **186,** 447–452 (2002).

59. Kerr, J. R. *et al.* Evidence for the role of demyelination, HLA-DR alleles, and cytokines in the pathogenesis of parvovirus B19 meningoencephalitis and its sequelae. *J. Neurol. Neurosurg. Psychiatry* **73,** 739–746 (2002).
60. Kumar, A. *et al.* Granzyme B mediated function of Parvovirus B19-specific CD4(+) T cells. *Clin Transl Immunology* **4,** e39 (2015).
61. Franssila, R. *et al.* T helper cell-mediated interferon-gamma expression after human parvovirus B19 infection: persisting VP2-specific and transient VP1u-specific activity. *Clin. Exp. Immunol.* **142,** 53–61 (2005).
62. Visscher, P. M., Brown, M. A., McCarthy, M. I. & Yang, J. Five years of GWAS discovery. *Am. J. Hum. Genet.* **90,** 7–24 (2012).
63. Manolio, T. A. Bringing genome-wide association findings into clinical use. *Nat. Rev. Genet.* **14,** 549–558 (2013).
64. Heiligenhaus, A. *et al.* Prevalence and complications of uveitis in juvenile idiopathic arthritis in a population-based nation-wide study in Germany: suggested modification of the current screening guidelines. *Rheumatology* **46,** 1015–1019 (2007).
65. Ozdal, P. C., Vianna, R. N. G. & Deschênes, J. Visual outcome of juvenile rheumatoid arthritis-associated uveitis in adults. *Ocul. Immunol. Inflamm.* **13,** 33–38 (2005).
66. Han, B. *et al.* Fine mapping seronegative and seropositive rheumatoid arthritis to shared and distinct HLA alleles by adjusting for the effects of heterogeneity. *Am. J. Hum. Genet.* **94,** 522–532 (2014).
67. Márquez, A. *et al.* New insights into the genetic component of non-infectious uveitis through an Immunochip strategy. *J. Med. Genet.* **54,** 38–46 (2017).
68. Kalinina Ayuso, V., van Dijk, M. R. & de Boer, J. H. Infiltration of Plasma Cells in the Iris of Children With ANA-Positive Anterior Uveitis. *Invest. Ophthalmol. Vis. Sci.* **56,** 6770–6778 (2015).

69. Kalinina Ayuso, V. *et al.* Intraocular biomarker identification in uveitis associated with juvenile idiopathic arthritis. *Invest. Ophthalmol. Vis. Sci.* **54,** 3709–3720 (2013).

70. Sijssens, K. M. *et al.* Cytokines, chemokines and soluble adhesion molecules in aqueous humor of children with uveitis. *Exp. Eye Res.* **85,** 443–449 (2007).

71. Kotaniemi, K., Kautiainen, H., Karma, A. & Aho, K. Occurrence of uveitis in recently diagnosed juvenile chronic arthritis: a prospective study. *Ophthalmology* **108,** 2071–2075 (2001).

72. Berntson, L. *et al.* Incidence of juvenile idiopathic arthritis in the Nordic countries. A population based study with special reference to the validity of the ILAR and EULAR criteria. *J. Rheumatol.* **30,** 2275–2282 (2003).

73. American Academy of Pediatrics Section on Rheumatology and Section on Ophthalmology: Guidelines for ophthalmologic examinations in children with juvenile rheumatoid arthritis. *Pediatrics* **92,** 295–296 (1993).

74. Kuiper, J. J. W. *et al.* A genome-wide association study identifies a functional ERAP2 haplotype associated with birdshot chorioretinopathy. *Hum. Mol. Genet.* **23,** 6081–6087 (2014).

75. Price, A. L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38,** 904–909 (2006).

76. Howie, B., Marchini, J., Stephens, M. & Chakravarti, A. Genotype imputation with thousands of genomes. *G3: Genes|Genomes|Genetics* **1,** 457–470 (2011).

77. Rich, S. S. *et al.* The Type 1 Diabetes Genetics Consortium. *Ann. N. Y. Acad. Sci.* **1079,** 1–8 (2006).

78. Browning, B. L. & Browning, S. R. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am. J. Hum. Genet.* **84,** 210–223 (2009).

79. Pundir, S., Martin, M. J. & O'Donovan, C. UniProt Protein Knowledgebase. *Methods Mol. Biol.* **1558,** 41–55 (2017).

80. Heiskanen, O. Ruptured intracranial arterial aneurysms of children and adolescents. Surgical and total management results. *Childs. Nerv. Syst.* **5,** 66–70 (1989).

81. Clement, C. C. *et al.* The Dendritic Cell Major Histocompatibility Complex II (MHC II) Peptidome Derives from a Variety of Processing Pathways and Includes Peptides with a Broad Spectrum of HLA-DM Sensitivity. *J. Biol. Chem.* **291,** 5576–5595 (2016).

82. Xia, J. & Wishart, D. S. Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst. *Nat. Protoc.* **6,** 743–760 (2011).

83. Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25,** 1605–1612 (2004).

Funding
This study was funded by: The Dr. F.P. Fischer Stichting, Amersfoort; The ODAS Stichting, the Landelijke Stichting Voor Blinden en Slechtzienden, Utrecht; the Stichting Nederlands Oogheelkundig Onderzoek (SNOO), Rotterdam, the Netherlands. Funding for the ICON-JIA study group: Research grant of the Federal Ministry of Education and Research [FKZ 01ER0812, FKZ 01ER0813; FKZ 01ER0828].

Acknowledgements
The following ICON-JIA study group collaborators have contributed to the DNA samples: Universitätsmedizin Charité Berlin, Tilmann Kallinich; Medizinische Hochschule Hannover, Angelika Thon; Universität Tübingen, Jasmin Kümmerle-Deschner; Prof.-Hess-Kinderklinik Bremen, Hans-Iko Huppertz; Asklepios Kinderklinik Sankt Augustin, Gerd Horneff; Olgahospital Stuttgart, Anton Hospach; Kinderkrankenhaus der Stadt Köln, Kirsten Mönkemöller; Deutsches Zentrum für Kinder- und Jugendrheumatologie Garmisch-Partenkirchen, Johannes-Peter Haas; St. Joseph-Stift Sendenhorst, Gerd Ganser; Kinderrheumatologische Praxis am AK Eilbek Hamburg, Ivan Foeldvari; Deutsches Rheuma-Forschungszentrum Berlin, Jens Klotsche.

Contributors JJWK, JHB, and SLP led the study. AJWH, SLP and JJWK wrote the paper. SLP, AJWH, and JJWK performed the data and statistical analysis. JHB, RC, MWS, TRDJR, SSMK, PCEHM, AJWH, CW, TS, DF, RAO, TRDJR, AIDH, SH, NSD, EPAHH, MAJR, and NW contributed to the patient ascertainment, sample collection and/or genotyping. All authors reviewed the final manuscript.