Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

HHS Public Access
Author manuscript
*Nat Genet*. Author manuscript; available in PMC 2013 December 01.

Published in final edited form as:
*Nat Genet*. 2013 June ; 45(6): 664–669. doi:10.1038/ng.2614.

# Dense genotyping of immune-related disease regions identifies 14 new susceptibility loci for juvenile idiopathic arthritis

**Anne Hinks**[1,2,22,*], **Joanna Cobb**[1,2,22], **Miranda C. Marion**[3,4,22], **Sampath Prahalad**[5,6], **Marc Sudman**[7], **John Bowes**[1,2], **Paul Martin**[1,2], **Mary E. Comeau**[3,4], **Satria Sajuthi**[3,4], **Robert Andrews**[8], **Milton Brown**[5], **Wei-Min Chen**[9], **Patrick Concannon**[9], **Panos Deloukas**[8], **Sarah Edkins**[8], **Stephen Eyre**[1,2], **Patrick M. Gaffney**[10], **Stephen L. Guthery**[11,12], **Joel M. Guthridge**[10], **Sarah E. Hunt**[8], **Judith A. James**[10], **Mehdi Keddache**[13], **Kathy L. Moser**[10], **Peter A. Nigrovic**[14,15], **Suna Onengut-Gumuscu**[9], **Mitchell L. Onslow**[7], **Carlos D. Rosé**[14,15], **Stephen S. Rich**[9], **Kathryn J.A. Steel**[1,2], **Edward K. Wakeland**[16], **Carol A. Wallace**[17], **Lucy R. Wedderburn**[18], **Patricia Woo**[19], **Boston Children's JIA Registry**[20], **British Society of Paediatric and Adolescent Rheumatology (BSPAR) study group**[20], **Childhood arthritis prospective study (CAPS)**[20], **Childhood Arthritis Response to Medication Study (CHARMS)**[20], **German society for Pediatric Rheumatology (GKJR)**[20], **JIA gene expression study**[20], **NIAMS JIA genetic registry**[20], **TREAT study**[20], **United Kingdom Juvenile Idiopathic Arthritis Genetics Consortium (UKJIAGC)**[20], **John F. Bohnsack**[11,12], **Johannes Peter Haas**[21], **David N. Glass**[7], **Carl D. Langefeld**[3,4], **Wendy Thomson**[1,2,23], and **Susan D. Thompson**[7,23]

[1]Arthritis Research UK Epidemiology Unit, Stopford Building, Manchester Academic Health Science Centre, The University of Manchester, Manchester, M13 9PT, UK.

[2]National Institute for Health Research, Manchester Musculoskeletal Biomedical Research Unit, Central Manchester University Hospitals National Health Service Foundation Trust, Manchester Academic Health Science Centre, Manchester, M13 9PT, UK.

[3]Department of Biostatistical Sciences, Wake Forest School of Medicine, Winston-Salem, NC 27157, USA.

[4]Center for Public Health Genomics, Wake Forest School of Medicine, Winston-Salem, NC 27157, USA.

[5]Department of Pediatrics, Emory University School of Medicine, Atlanta, GA 30322, USA.

---

Users may view, print, copy, download and text and data- mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: http://www.nature.com/authors/editorial_policies/license.html#terms

[*]Correspondence should be addressed to Anne Hinks (Anne.Hinks@manchester.ac.uk)..
[20]A list of members and affiliations is provided in the **Supplementary Note**.
[22]These authors contributed equally to this work
[23]These authors jointly directed this work

Author Contributions
S.D.T, W.T, C.D.L, S.P, M.C.M, J.C and A.H led the study. A.H, J.C, M.C.M, C.D.L, S.P, W.T and S.D.T wrote the paper. A.H, J.C, C.D.L, M.C.M, M.S, S.P, J.B, M.E.C and S.S performed the data and statistical analysis. A.H and P.M performed the bioinformatic analysis. D.N.G, J.P.H, J.F.B, R.A, M.B, W-M.C, P.C, P.D, S.Edkins, S.Eyre, P.M.G, S.L.G, J.M.G, S.E.H, J.A.J, M.K, K.L.M, P.A.N, S.O-G, M.L.O, C.D.R, S.S.R, K.J.A.S, E.K.W, C.A.W, L.R.W and P.W contributed primarily to the patient ascertainment, sample collection and/or genotyping. All authors reviewed the final manuscript.

Competing financial interest
The authors declare no competing financial interests.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

[6]Department of Human Genetics, Emory University School of Medicine, Atlanta, GA 30322, USA.

[7]Division of Rheumatology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229, USA.

[8]The Wellcome Trust Sanger Institute, Cambridge, CB10 1SA, UK.

[9]Center for Public Health Genomics, University of Virginia, Charlottesville, Virginia, VA22908, USA.

[10]Arthritis & Clinical Immunology Program, Oklahoma Medical Research Foundation, Oklahoma City, OK 73104, USA.

[11]Department of Pediatrics, University of Utah, Salt Lake City, UT 84132, USA.

[12]School of Medicine, University of Utah, Salt Lake City, UT 84132, USA.

[13]Division of Human Genetics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229, USA.

[14]Division of Rheumatology, Immunology, and Allergy, Brigham and Women's Hospital, Boston, MA 02115, USA.

[15]Division of Immunology, Boston Children's Hospital, Boston, MA 02115, USA.

[16]Department of Immunology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA.

[17]Department of Pediatrics, Seattle Children's Hospital and Research Institute, Seattle, Washington, WA 98105, USA.

[18]Rheumatology Unit, UCL Institute of Child Health, 30 Guilford Street, London, WC1N 1EH, UK.

[19]Division of Infection and Immunity, 3rd Floor Rayne Institute, University College London, 5 University St, London, WC1E 6JF, UK.

[21]German Centre for Rheumatology in Children and Young People, D-82467 Garmisch-Partenkirchen, Germany.

## Abstract

Analysis of the ImmunoChip single nucleotide polymorphism (SNP) array in 2816 individuals, comprising the most common subtypes (oligoarticular and RF negative polyarticular) of juvenile idiopathic arthritis (JIA) and 13056 controls strengthens the evidence for association to three known JIA-risk loci (*HLA, PTPN22* and *PTPN2*) and has identified fourteen risk loci reaching genome-wide significance ($p < 5 \times 10^{-8}$) for the first time. Eleven additional novel regions showed suggestive evidence for association with JIA ($p < 1 \times 10^{-6}$). Dense-mapping of loci along with bioinformatic analysis has refined the association to one gene for eight regions, highlighting crucial pathways, including the IL-2 pathway, in JIA disease pathogenesis. The entire ImmunoChip loci, HLA region and the top 27 loci ($p < 1 \times 10^{-6}$) explain an estimated 18%, 13% and 6% risk of JIA, respectively. Analysis of the ImmunoChip dataset, the largest cohort of JIA cases investigated to date, provides new insight in understanding the genetic basis for this childhood autoimmune disease.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Juvenile idiopathic arthritis (JIA) is the most common chronic rheumatic disease of childhood and describes a group of clinically heterogeneous arthritides which begin before the age of 16 years, persist for at least 6 weeks and have an unknown cause [1]. It is established that there is a strong genetic contribution to the risk of JIA, with a sibling risk ratio of $\approx$ 11.6 [2] and an increased risk for other autoimmune diseases for families of JIA patients [3]. Using the International League of Associations for Rheumatology (ILAR) criteria, JIA can be divided into subtypes based on clinical features [4]. A recent genome wide association study (GWAS) identified a number of JIA susceptibility regions [5, 6]. Additional loci have also been identified through candidate gene association studies and confirmed in multiple, independent studies [7-14]. However, to date only three loci reach genome-wide significance thresholds (*HLA, PTPN22* and *PTPN2*) [5].

Many confirmed and nominally associated JIA susceptibility loci show association with other autoimmune diseases [5]. This striking overlap of autoimmune disease susceptibility loci may occur where the same variants contribute to multiple diseases or it may be that different variants in the same gene lead to different autoimmune disease. Thus, dense mapping of as many of the susceptibility loci as possible will be important in order to understand how individual variants contribute to the risk of disease. To facilitate these efforts a custom Illumina Infinium genotyping array has been designed by the ImmunoChip Consortium based on confirmed risk loci from 12 autoimmune diseases [15], not including JIA. The chip includes dense coverage of the extended HLA region and 186 non-HLA loci [15]. In this study, we report on analysis of the ImmunoChip in 2816 individuals with oligoarticular or rheumatoid factor (RF) negative polyarticular JIA and 13056 controls post quality control (**Supplementary Table 1 and 2**). There is overlap in the samples used in this study and in previous genetic studies of JIA [5-12, 14]; further detail can be found in the online methods. Restriction to these two subtypes (~70% of JIA cases) reduces phenotypic heterogeneity. Given that JIA is a complex genetic disorder that shares risk loci with other autoimmune diseases, the ImmunoChip provides a unique opportunity to discover novel JIA-risk loci. In addition, the dense coverage for many regions allows for fine-mapping analysis to identify possible causal variants and help inform future studies into the functional role of JIA-risk loci.

After stringent data quality control (**Supplementary Table 3**) 123,003 SNPs with MAF 1% were available for analysis. The inflation factor ($\lambda_{GC}$) (calculated using a set of SNPs included on the ImmunoChip for a study investigating the genetic basis for reading and writing ability) for this study was $\lambda_{GC}$=1.265, $\lambda_{GC1000}$=1.057. Seventeen of the 187 autoimmune regions investigated were significantly associated with oligoarticular and RF negative polyarticular JIA ($p < 5 \times 10^{-8}$) (**Figure 1, Table 1 and Supplementary Fig. 1**). These data strengthen the associations for three established JIA susceptibility loci (*HLA, PTPN22* and *PTPN2*) and provide evidence for an additional 14 regions which reach genome-wide significance for the first time. Among the three established associations, the most significant associations were observed within the *MHC* region (**Figure 2**). Specifically, rs7775055 ($MAF_{controls}$=2%) provided the strongest evidence of association with JIA (OR = 6.01, $p = 3.14 \times 10^{-174}$). In addition, stepwise logistic regression identified 14 SNPs that showed separate effects in the region (**Supplementary Table 4**). The most significant SNP,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

rs7775055 tags the *DRB1*0801-DQA1*0401-DQB1*0402* haplotype, which has been consistently implicated as conferring risk to JIA [16, 17], however other haplotypes have also been associated with JIA. The *HLA* SNP rs7775055 showed a highly significant difference in SNP allele frequencies between the two subtypes (**Supplementary Table 5**). The association was stronger in the oligoarticular subtype compared to the RF negative polyarthritis subtype which is consistent with previous studies showing differences in HLA associations between the two subtypes [16, 17]. Further analysis at the amino acid level is necessary to fully understand this complex region in JIA and its subtypes. The most significant association outside the *MHC* region is with rs6679677 (OR = 1.59, $p = 3.19 \times 10^{-25}$) on chromosome 1p13.2, which contains the *PTPN22* gene; rs6679677 is in linkage disequilibrium (LD) ($r^2$=1) with rs2476601, the SNP previously associated with JIA [5, 7] and implicated as the *PTPN22* causal variant [18]. We also confirmed association to *PTPN2*, with rs2847293 (OR = 1.31, $p = 1.44 \times 10^{-12}$) which lies in the intergenic region 3′ of the *PTPN2* gene and is in LD ($r^2$=0.94) with rs1893217, a SNP previously associated with oligoarticular and RF negative polyarticular JIA [5]. Stepwise logistic regression including the most significant SNP in the *PTPN2* region as a covariate suggests that there is a uncommon variant, rs149850873, ($MAF_{controls}$= 2%) that confers an independent secondary effect in the region (**Supplementary Table 6 and Supplementary Fig. 2**).

Of the 14 loci confirmed as novel JIA susceptibility loci in this study at the genome wide significance level ($p < 5 \times 10^{-8}$) (**Figure 1, Table 1 and Supplementary Fig. 1**), five (*STAT4, ANKRD55, IL2/IL21, IL2RA* and *SH2B3/ATXN2*) have supportive evidence with JIA susceptibility from previous studies. The most significant SNP in the *STAT4* region (rs10174238) is in high LD with a SNP (rs7574865) previously reported in JIA [5, 8, 10] and other autoimmune diseases [19]. However, stepwise logistic regression analysis suggests two additional independent effects (rs45539732 and rs13029532), which are located within the adjacent *STAT1* gene (**Supplementary Table 6 and Supplementary Fig. 3**). Notably rs45539732 is an uncommon SNP ($MAF_{controls}$ = 3%).

There were 11 additional regions showing suggestive ($p <1 \times 10^{-6}$ and $p > 5 \times 10^{-8}$) evidence for association with oligoarticular and RF negative polyarticular JIA (**Table 2**), of which four have supportive evidence from previous studies (*COG6, CCR1/CCR3, C3orf1/CD80, AFF3/LONRF2*).

We imputed across the non-HLA JIA risk loci identified in this study using the 1000 Genomes Project (online methods) (**Table 1, Table 2** and **Supplementary Fig. 1**). We found only modest differences between the p-values of the top genotyped SNP compared to the top imputed SNP. We note two regions that are minor exceptions, the *PRM1/C16orf75* and the *C5orf56/IRF1* region **(Supplementary Fig. 1)**. For the latter region the top imputed SNP lies within the *C5orf56* gene. The lack of a substantial gain of information from imputation of the regions is consistent with other reports on the performance of ImmunoChip imputation20, 21. This likely is due to the dense fine mapping of most of the regions on the ImmunoChip.

Of the top 17 regions, that reach genome-wide significance, 13 regions are densely mapped on ImmunoChip. LD patterns and functional annotation provide strong evidence that the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

signal localizes to a single gene in eight cases (*PTPN2, IL2RA, STAT4, IL2RB* and *ZFP36L1* based on LD patterns and *PTPN22, SH2B3/ATXN2* and *TYK2* based on the most significant SNP being a non-synonymous coding variant) (Table 3, Supplementary Table 7 and Supplementary Fig. 1), however further functional analysis is required for confirmation.

All but one of the variants which reached genome-wide significance were common (>5% MAF). One variant, a non-synonymous coding variant within the *TYK2* gene had a low allele frequency ($MAF_{controls}$ = 5%). In addition a couple of the secondary effects in *PTPN2* and *STAT4* were uncommon.

For three regions (*TYK2, SH2B3/ATXN2* and *LTBR*) the most significant SNP (or a SNP in $r^2$>0.9) lies within a coding region and are therefore strong candidates for the causal variant. For *SH2B3/ATXN2*, the same variant has also been associated with celiac disease (CeD) [22], vitiligo [23], RA [24], type 1 diabetes (T1D) [25] and multiple sclerosis (MS) [26]. The *TYK2* SNP (rs34536443) is also the lead SNP in the region in RA [27], primary biliary cirrhosis [20] and psoriasis [28]. Other regions (*IL6R, ZFP36L, IL2/IL21, UBE2L3, LTBR* and *C3orf1/CD80*) contain SNPs that show evidence for high mammalian conservation (17-way vertebrate conservation) [29] or have a high regulatory potential score (**Table 3**) calculated using alignments of seven mammalian genomes [30]. There is eQTL evidence for the associated SNPs in *LTBR, UBE2L3* and *LNPEP* (**Table 3**). The SNP in *LNPEP*, rs27290, is also in LD ($r^2$=0.78) with rs2248374, a SNP which lies within a splice site for *ERAP2* [31]. The rs2248374-G allele results in a spliced *ERAP2* mRNA which encodes a truncated protein. For JIA the rs2248374-G minor allele showed protective association (OR = 0.76, $p = 1.8 \times 10^{-7}$).

*IL2RA*, the *IL2/IL21* region and *IL2RB* are now all considered confirmed susceptibility loci for JIA and implicate an important role for the IL-2 pathway in JIA disease pathogenesis. This pathway plays a vital role in T cell activation and development as well as a key role in maintenance of immune tolerance through the dependence of regulatory T cells on IL-2. Other confirmed JIA loci identified here are related to this pathway, *SH2B3* is an adaptor protein involved in T cell activation and *STAT4* is a transcription factor important in T cell differentiation.

We next considered the top non-HLA SNP associations separately for each JIA subtype (oligoarticular and RF negative polyarticular JIA). Only one region showed evidence for differential association, the *C5orf56/IRF1* region, where the association was limited to the oligoarticular subtype of JIA. All other regions showed associations with similar effect sizes and direction of effect (**Supplementary Table 5**).

As expected, many of the JIA-associated regions shown in **Table 1 and Table 2** are also associated with other autoimmune diseases (**Supplementary Table 8**) with the same SNP, or a highly correlated SNP associated in the same direction (assessed by comparing with information from the Catalogue of published GWAS and recent publications investigating the ImmunoChip in other autoimmune diseases [20-22, 27, 28, 32, 33]). We find a strong overlap with RA loci, which is not surprising due to the clinical similarities with JIA, and is consistent with previous studies [8, 10, 34]. In addition, there is notable overlap with T1D and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

CeD. Some regions (IL2/IL21, C5orf56/IRF1, IL2RB, ATP8B2/IL6R, Chr13q14, CCR1/CCR3, RUNX3 and *C3orf1/CD80*) show association with other autoimmune diseases but their top SNP is not highly correlated with our top JIA SNP. Some regions have not been previously associated by GWAS or ImmunoChip. In depth analysis of the results across all ImmunoChip studies will be of great value to understanding the contributions of the individual loci to the various diseases.

This study of 2816 JIA cases is the largest collaborative cohort study of JIA to date, and includes samples from across the United States, United Kingdom and Germany. The power derived from this cohort plus the large control sample size, combined with the comprehensive coverage for SNPs in regions implicated in autoimmune disease on the ImmunoChip has substantially increased our power to detect association. In setting the statistical threshold at stringent genome-wide significance levels ($p < 5 \times 10^{-8}$) we report 14 new loci. In addition, a second tier of 11 regions with suggestive evidence for association ($p < 1 \times 10^{-6}$) has been identified that are plausible candidates as risk factors but require validation. While this study dramatically increases the number of susceptibility loci identified for JIA, additional genetic risk factors likely remain to be discovered, which is supported by the QQ plot (**Supplementary Fig. 4**) that suggests there are residual associations after removing the above implicated regions. In addition we calculated that the entire ImmunoChip loci, the HLA region and the top 27 loci explain an estimated 18%, 13% and 6% of risk of JIA, respectively. This also suggests there must be other regions of the genome that harbor additional JIA-risk loci. In summary, this analysis of ImmunoChip has substantially enhanced our understanding of the genetic component of JIA, increasing the number of confirmed JIA loci from 3 to 17. The dense mapping of confirmed regions has narrowed down the regions to take forward into future functional studies. Importantly, these studies allow us to begin to understand where JIA fits in the spectrum of autoimmune diseases and identified a number of novel genes and pathways as potential targets for future therapeutic intervention.

## Online Methods

### Subjects

All cohorts comprised individuals from populations of European descent from the US, UK and Germany.

The post QC US cohorts comprised 1596 US oligoarthritis and RF negative polyarthritis JIA patients and 4048 US controls. Less than one half of these cases have already been included in a genome-wide association study and previously described [5-6]. Notably, 95 of these patients were from multiplex pedigrees such that for each pedigree one RF negative polyarthritis or oligoarticular JIA case was randomly selected for genotyping. Clinics enrolling the JIA patients for Cincinnati-based studies (listed in order of number contributed) were located in Cincinnati, OH; Atlanta, GA; Columbus, OH; Little Rock, AR; Long Island, NY; Chicago, IL; Dover, DE; Salt Lake City, UT; Cleveland, OH; Philadelphia, PA; Toledo, OH; Nashville, TN; Milwaukee, WI; and Charleston, SC. Additional DNA from JIA cases collected independently by investigators in Salt Lake City, UT (314 cases, where about 75% overlap with replication cohort in previous GWAS

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

studies [5-6]) and Boston, MA (13 cases) or enrolled as part of the Trial of Early Aggressive Therapy in Juvenile Idiopathic Arthritis (TREAT) study (clinical trials identifier NCT00443430) (22 cases) were made available for genotyping in Cincinnati.

The US controls were derived from four sources: 793 healthy children without known major health conditions recruited from the geographical area served by Cincinnati Children's Hospital Medical Center (CCHMC) and 119 healthy adults collected at CCHMC. Previous JIA GWAS studies have include about 75% of only the pediatric controls, 484 healthy adult controls from Utah screened for autoimmune diseases and all were included in the replication cohort of previous GWAS studies [5-6]. 848 healthy adult controls collected at the Oklahoma Medical Research Foundation; and 1804 healthy US adult controls from the Genotype and Phenotype registry (www.gapregistry.org) and the NIDDK IBD Genetics Consortium. Healthy controls from the Oklahoma Medical Research Foundation (OMRF) were provided by the Lupus Family Registry and Repository (LFRR)[39] and the Oklahoma Immune Cohort (OIC). Each individual completed the Connective Tissue Disease Screening Questionnaire (CSQ)[40] and individuals with a "probable" systemic rheumatic disease were excluded. Each individual was enrolled into these studies after appropriate written consent and IRB approval by the OMRF and the University of Oklahoma Health Sciences Center. Healthy controls were also provided from the University of Minnesota SLE sibship collection[41] and these subjects were enrolled after appropriate written consent and IRB approval by the University of Minnesota.

The US collections and their use in genetic studies have been approved by the Institutional Review Board of CCHMC and each collaborating center.

The post QC UK cohort comprised 772 UK oligoarthritis and RF negative polyarthritis JIA patients from five sources: The British Society for Paediatric and Adolescent Rheumatology (BSPAR) National Repository of JIA; a group of UK patients with long-standing JIA, described previously [42]; a cohort collected as part of the Childhood Arthritis Prospective Study (CAPS), a prospective inception cohort study of JIA cases from 5 centers across UK[43]; a cohort of children recruited for the SPARKS-CHARM (Childhood Arthritis Response to Medication) study, who fulfill ILAR criteria for JIA and are about to start new disease-modifying medication for active arthritis[44] and an ongoing collection of UK cases, the UK JIA Genetics Consortium (UKJIAGC). There is overlap in the JIA cases used in this study and in previous UK candidate gene studies of JIA [7, 9-12]. JIA cases were classified according to ILAR criteria [4]. All UK JIA cases were recruited with ethical approval and provided informed consent [North-West Multi-Centre Research Ethics Committee (MREC 99/8/84), the University of Manchester Committee on the Ethics of Research on Human Beings and National Research Ethics Service (NRES 02/8/104)]. The 8530 UK controls comprised the shared UK 1958 Birth cohort and UK Blood Services Common Controls. The collection was established as part of the WTCCC[45].

The post QC German Cohort comprised 448 German oligoarthritis and RF negative polyarthritis JIA patients and 478 controls. These cases have already been included as a replication cohort in a genome-wide association study and previously described [5-6]. These patients were recruited from the German Center for Rheumatology in Children and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Adolescents, Garmisch-Partenkirchen; the Department of Pediatrics, University of Tübingen; Children's Rheumatology Unit Sendenhorst, Germany; and the Department of Pediatrics, University of Prague, Czech Republic. JIA was determined retrospectively by chart review. German population-based control samples were prepared from cord blood obtained from healthy newborns in the Survey of Neonates in Pomerania (SNiP) consortium [46]. The respective Institutional Review Boards approved the collection of these samples and participation in this study. Demographic breakdown of the cohorts is shown in **Supplementary Table 1**.

### Genotyping and quality controls

Samples were genotyped using ImmunoChip, a custom-made Illumina Infinium array, described previously[22]. Genotyping was performed according to Illumina's protocols at labs in Hinxton, UK, Manchester, UK, Cincinnati, US, Utah, US, Charlottesville, US and New York, US. The Illumina GenomeStudio GenTrain2.0 algorithm was used to recluster all 15872 samples together.

SNPs were excluded if they had a call rate <98% and a cluster separation score of <0.4. Samples were then excluded for call rate <98% across 178203 markers or if there were inconsistencies between recorded and genotype inferred gender. Duplicates and first- or second-degree relatives were also removed. Principal component (PC) analysis was computed, using Eigensoft v4.2 (http://www.hsph.harvard.edu/faculty/alkes-price/software/) [47, 48], on the samples, merged with HapMap phase 2 individuals (CEU, YRI and CHB) as reference populations, to identify genetic outliers. PC analysis was performed on a subset of SNPs, removing SNPs in known regions of high linkage disequilibrium (LD), with MAF < 0.05 and pruned for LD between markers. To maximize genetic homogeneity within the samples the initial PC analysis was followed by five subsequent PC analyses where at each iteration individuals 5 standard deviations from the mean were removed. The PCs from the 5$^{th}$ iteration were used as covariates in the logistic regression analysis. A SNP was removed from the primary analysis if it exhibited significant differential missingness between cases and controls (p<0.05), had significant departure from Hardy-Weinberg equilibrium (p<0.001 in controls) or had a MAF<0.01.

### Statistical analysis

To test for an association between a SNP and case/control status, a logistic regression analysis was computed using the 5 PCs as covariates. The primary inference was based on the additive genetic model, unless there was significant lack-of-fit to the additive model (P<0.05). If there was evidence of a departure from an additive model, then inference was based on the most significant of the dominant, additive and recessive genetic models. The additive and recessive models were computed only if there were at least 10 and 20 individuals homozygous for the minor allele, respectively. For analysis of the X chromosome, the data analysis was first stratified by gender followed by a meta-analysis. The genomic control inflation factor ($\lambda_{GC}$) was calculated using a set of SNPs included on the ImmunoChip for a study investigating the genetic basis for reading and writing ability (submitted by J.C.Barrett). We visually inspected the cluster plots for the most associated SNPs in the regions to confirm the genotyping quality. Additionally, concordance of

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

genotyping data was compared with data previously generated on other platforms. A subset of cases had high resolution HLA genotyping. These data were used to investigate if the SNPs with the strongest statistical associations with JIA were in high linkage disequilibrium with classical HLA alleles/haplotypes. To investigate subtype effects the two main subtypes (Oligoarticular JIA and RF negative polyarticular JIA) were compared separately against the same controls. Disease association heterogeneity was tested by testing for significant differences in SNP allele frequencies between the two subtypes. To determine how many independent associations were within a genomic region, a manual stepwise procedure (i.e., forward selection with backward elimination, entry and exit criteria of $P<0.0001$) was computed [49]. Specifically, for each region which reached genome-wide significance, the top SNP was included as a covariate and the association statistics re-calculated. SNPs were allowed to enter and exit models in this stepwise fashion until no additional SNPs met a significance threshold of $P<0.0001$. The stepwise procedure was modified slightly in the greater MHC region to have an entry and exit criteria of $P<0.00001$. These statistical analyses were performed using PLINK v1.07[50] and SNPGWA version 4.0 (www.phs.wfubmc.edu).

The cumulative variance explained by common SNP variation was estimated using a variance component model and restricted maximum likelihood estimation as implemented in the program GCTA [51] and adjusting for the PCs as covariates and using Yang's correction factor (c=0 from formula 9) for imperfect LD with causal variants. Estimates are based on SNPs that had <1% missing genotypes and stringent relatedness threshold of 0.025.

We computed SNP genotype imputation across the regions of the ImmunoChip. We used the program SHAPEIT (www.shapeit.fr/) to pre-phase our ImmunoChip data and IMPUTE2 (https://mathgen.stats.ox.ac.uk/impute/impute_v2.html) with the 1000 Genomes Phase 1 integrated reference panel to impute the SNP genotypes. To account for phase uncertainty, we tested for association using SNPTEST (https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html). Only genotyped SNPs of high quality were used to inform imputation. Imputed SNP quality was assessed using the information score (> 0.5) and the confidence score (> 0.9).

Regional plots of association and adjusting for the strongest SNP association was computed using Locuszoom (http://csg.sph.umich.edu/locuszoom/) [52].

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

We thank Paul Gilbert for preparing UK JIA case samples for genotyping and Mary Ryan for preparing US JIA and the Cincinnati local control samples. Genotyping of the US JIA, German JIA and respective control collections was supported by RC1-AR-058587 and U01-AI-067150S1. In addition, patient recruitment and DNA preparation in the US was largely funded by N01-AR-42272, P01-AR-048929 and P30-AR-473639, with contributions from the Arthritis Foundation, The Val A. Browning Charitable Foundation in Salt Lake City and the Marcus Foundation Inc. in Atlanta, GA as well as NIH grants K23-AR-50177 and R01-AR-060893. The Federal Ministry of Education and Research, Germany (BMBF grants 01GM0907 and 01 ZZ 0403 supported patient recruitment and sample preparation in Germany. Genotyping of the UK JIA cases samples was supported by the Arthritis Research UK

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

grant reference number 17552. Sparks CHARMS was funded by Sparks UK, reference 08ICH09, and the Big Lottery Fund UK, reference RG/1/010135231. The study is on the UK Medicines for Children Research Network (MCRN) portfolio. We acknowledge support from the Wake Forest School of Medicine Center for Public Health Genomics and from the NIH for computing resources and data analysis (R01-AR-057106).

Sample recruitment was supported in part by Grants Number N01AR62277, GM103510, AI082714, AR053483 from NIAMS/NIGMS/NIAID/NIH. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of these institutes or NIH.

We thank Jeffrey Barrett and Chris Wallace for the SNP selection. We would like to thank the WTSI Genotyping Facility and in particular Emma Gray, Sue Bumpstead, Doug Simpkin and Hannah Blackburn for typing the UK samples.

We acknowledge use of DNA from the UK Blood Services collection of common controls (UKBS-CC collection), which is funded by the Wellcome Trust grant 076113/C/04/Z and by US National Institute for Health Research program grant to the National Health Service Blood and Transplant (RP-PG-0310-1002). We acknowledge the use of DNA from the British 1958 Birth Cohort Collection, which is funded by the UK Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/01. Genotyping of control samples was supported, in part, by grants from the Juvenile Diabetes Research Foundation International (JDRF) and the NIH (U01 DK062418).

We thank Peter K. Gregersen at the Feinstein Institute for providing U.S. control genotyping from the Genotype and Phenotype registry (www.gapregistry.org) supported by National Institutes of Health grant RC2AR059092. We thank the NIDDK IBD Genetics Consortium for providing North American control genotyping supported by the National Institutes of Health grants DK062431, DK062422, DK062420, DK062432, DK062423, DK062413, and DK062429.

We gratefully acknowledge contributions from physicians at CCHMC and collaborating clinics. We also acknowledge the assistance of Sandy Kramer, Bronte Clifford and Lori Ponder for patient recruitment and coordination of clinical information at Cincinnati Children's Hospital Medical Center, the University of Utah and at Emory University, respectively. The Cincinnati normal control DNA collection was supported and made available by Cincinnati Children's Hospital Medical Center.

## Reference List

1. Ravelli A, Martini A. Juvenile idiopathic arthritis. Lancet. 2007; 369:767–778. [PubMed: 17336654]
2. Prahalad S, et al. Quantification of the familial contribution to juvenile idiopathic arthritis. Arthritis Rheum. 2010; 62:2525–2529. [PubMed: 20506132]
3. Prahalad S, Shear ES, Thompson SD, Giannini EH, Glass DN. Increased prevalence of familial autoimmunity in simplex and multiplex families with juvenile rheumatoid arthritis. Arthritis Rheum. 2002; 46:1851–1856. [PubMed: 12124869]
4. Petty RE, et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J. Rheumatol. 2004; 31:390–392. [PubMed: 14760812]
5. Thompson SD, et al. The susceptibility loci juvenile idiopathic arthritis shares with other autoimmune diseases extend to PTPN2, COG6, and ANGPT1. Arthritis Rheum. 2010; 62:3265–3276. [PubMed: 20722033]
6. Thompson SD, et al. Genome-wide association analysis of juvenile idiopathic arthritis identifies a new susceptibility locus at chromosomal region 3q13. Arthritis Rheum. 2012; 64:2781–2791. [PubMed: 22354554]
7. Hinks A, et al. Association between the PTPN22 gene and rheumatoid arthritis and juvenile idiopathic arthritis in a UK population: Further support that PTPN22 is an autoimmunity gene. Arthritis Rheum. 2005; 52:1694–1699. [PubMed: 15934099]
8. Prahalad S, et al. Variants in TNFAIP3, STAT4, and C12orf30 loci associated with multiple autoimmune diseases are also associated with juvenile idiopathic arthritis. Arthritis Rheum. 2009; 60:2124–2130. [PubMed: 19565500]
9. Hinks A, et al. Association of the IL2RA/CD25 gene with juvenile idiopathic arthritis. Arthritis Rheum. 2009; 60:251–257. [PubMed: 19116909]

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript




10. Hinks A, et al. Overlap of disease susceptibility loci for rheumatoid arthritis and juvenile idiopathic arthritis. Ann. Rheum. Dis. 2010; 69:1049–1053. [PubMed: 19674979]
11. Hinks A, et al. Association of the CCR5 gene with juvenile idiopathic arthritis. Genes Immun. 2010; 11:584–589. [PubMed: 20463745]
12. Hinks A, et al. Association of the AFF3 gene and IL2/IL21 gene region with juvenile idiopathic arthritis. Genes Immun. 2010; 11:194–198. [PubMed: 20072139]
13. Albers HM, et al. Association of the autoimmunity locus 4q27 with juvenile idiopathic arthritis. Arthritis Rheum. 2009; 60:901–904. [PubMed: 19248117]
14. Prahalad S, et al. Association of two functional polymorphisms in the CCR5 gene with juvenile rheumatoid arthritis. Genes Immun. 2006; 7:468–475. [PubMed: 16775617]
15. Cortes A, Brown MA. Promise and pitfalls of the Immunochip. Arthritis Res. Ther. 2011; 13:101. [PubMed: 21345260]
16. Hollenbach JA, et al. Juvenile idiopathic arthritis and HLA class I and class II interactions and age-at-onset effects. Arthritis Rheum. 2010; 62:1781–1791. [PubMed: 20191588]
17. Thomson W, et al. Juvenile idiopathic arthritis classified by the ILAR criteria: HLA associations in UK patients. Rheumatology. (Oxford). 2002; 41:1183–1189. [PubMed: 12364641]
18. Vang T, et al. Autoimmune-associated lymphoid tyrosine phosphatase is a gain-of-function variant. Nat. Genet. 2005; 37:1317–1319. [PubMed: 16273109]
19. Remmers EF, et al. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N. Engl. J. Med. 2007; 357:977–986. [PubMed: 17804842]
20. Liu JZ, et al. Dense fine-mapping study identifies new susceptibility loci for primary biliary cirrhosis. Nat. Genet. 2012; 44:1137–1141. [PubMed: 22961000]
21. Juran BD, et al. Immunochip analyses identify a novel risk locus for primary biliary cirrhosis at 13q14, multiple independent associations at four established risk loci and epistasis between 1p31 and 7q32 risk variants. Hum. Mol. Genet. 2012; 21:5209–5221. [PubMed: 22936693]
22. Trynka G, et al. Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. Nat. Genet. 2011; 43:1193–1201. [PubMed: 22057235]
23. Jin Y, et al. Genome-wide association analyses identify 13 new susceptibility loci for generalized vitiligo. Nat. Genet. 2012; 44:676–680. [PubMed: 22561518]
24. Coenen MJ, et al. Common and different genetic background for rheumatoid arthritis and coeliac disease. Hum. Mol. Genet. 2009; 18:4195–4203. [PubMed: 19648290]
25. Barrett JC, et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat. Genet. 2009; 41:703–707. [PubMed: 19430480]
26. Alcina A, et al. The autoimmune disease-associated KIF5A, CD226 and SH2B3 gene variants confer susceptibility for multiple sclerosis. Genes Immun. 2010; 11:439–445. [PubMed: 20508602]
27. Eyre S, et al. High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. Nat. Genet. 2012; 44:1336–1340. [PubMed: 23143596]
28. Tsoi LC, et al. Identification of 15 new psoriasis susceptibility loci highlights the role of innate immunity. Nat. Genet. 2012; 44:1341–1348. [PubMed: 23143594]
29. Siepel A, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. 2005; 15:1034–1050. [PubMed: 16024819]
30. King DC, et al. Evaluation of regulatory potential and conservation scores for detecting cis-regulatory modules in aligned mammalian genome sequences. Genome Res. 2005; 15:1051–1060. [PubMed: 16024817]
31. Andres AM, et al. Balancing selection maintains a form of ERAP2 that undergoes nonsense-mediated decay and affects antigen presentation. PLoS. Genet. 2010; 6:e1001157. [PubMed: 20976248]
32. Cooper JD, et al. Seven newly identified loci for autoimmune thyroid disease. Hum. Mol. Genet. 2012; 21:5202–5208. [PubMed: 22922229]
33. Jostins L, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature. 2012; 491:119–124. [PubMed: 23128233]

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript




34. Hinks A, et al. Investigation of rheumatoid arthritis susceptibility loci in juvenile idiopathic arthritis confirms high degree of overlap. Ann. Rheum. Dis. 2012; 71:1117–1121. [PubMed: 22294642]
35. Stranger BE, et al. Patterns of cis regulatory variation in diverse human populations. PLoS. Genet. 2012; 8:e1002639. [PubMed: 22532805]
36. Dimas AS, et al. Common regulatory variation impacts gene expression in a cell type-dependent manner. Science. 2009; 325:1246–1250. [PubMed: 19644074]
37. Grundberg E, et al. Mapping cis- and trans-regulatory effects across multiple tissues in twins. Nat. Genet. 2012
38. Ramensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. Nucleic Acids Res. 2002; 30:3894–3900. [PubMed: 12202775]
39. Rasmussen A, et al. The lupus family registry and repository. Rheumatology. (Oxford). 2011; 50:47–59. [PubMed: 20864496]
40. Karlson EW, et al. A connective tissue disease screening questionnaire for population studies. Ann. Epidemiol. 1995; 5:297–302. [PubMed: 8520712]
41. Gaffney PM, et al. Genome screening in human systemic lupus erythematosus: Results from a second Minnesota Cohort and combined analyses of 187 sib-pair families. American journal of Human Genetics. 2000; 66:547–556. [PubMed: 10677315]
42. Packham JC, Hall MA. Long-term follow-up of 246 adults with juvenile idiopathic arthritis: functional outcome. Rheumatology. (Oxford). 2002; 41:1428–1435. [PubMed: 12468825]
43. Adib N, et al. Association between duration of symptoms and severity of disease at first presentation to paediatric rheumatology: results from the Childhood Arthritis Prospective Study. Rheumatology. (Oxford). 2008; 47:991–995. [PubMed: 18417527]
44. Moncrieffe H, et al. Generation of novel pharmacogenomic candidates in response to methotrexate in juvenile idiopathic arthritis: correlation between gene expression and genotype. Pharmacogenet. Genomics. 2010; 20:665–676. [PubMed: 20827233]
45. The Wellcome Trust Case Control consortium Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007; 447:661–678. [PubMed: 17554300]
46. Beyersdorff A, et al. Survey of Neonates in Pomerania (SniP): a population based analysis of the mothers' quality of life after delivery with special relations to their social integration. Int. J. Public Health. 2008; 53:87–95. [PubMed: 18681337]
47. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. PLoS. Genet. 2006; 2:e190. [PubMed: 17194218]
48. Price AL, et al. Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. 2006; 38:904–909. [PubMed: 16862161]
49. Hosmer, DW.; Lemeshow, S. Applied Logistic Regression. New York: 1989.
50. Purcell S, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 2007; 81:559–575. [PubMed: 17701901]
51. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am. J. Hum. Genet. 2011; 88:76–82. [PubMed: 21167468]
52. Pruim RJ, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics. 2010; 26:2336–2337. [PubMed: 20634204]

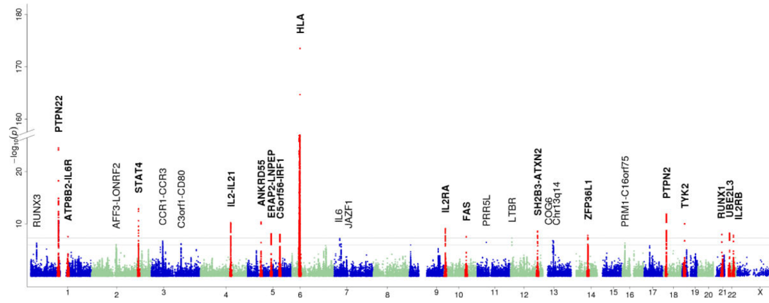


**Figure 1. Manhattan plot of association statistics for oligoarticular and RF negative polyarticular juvenile idiopathic arthritis risk loci**

The upper dashed black line indicates the threshold for genome wide significance ($p < 5 \times 10^{-8}$, loci reaching this threshold are highlighted in bold font and individual SNPs mapping to these loci are shown in red. The lower dashed grey line indicates the threshold for suggestive association ($p < 1 \times 10^{-6}$ and $p > 5 \times 10^{-8}$), loci reaching this threshold are labeled in non-bold font.



Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

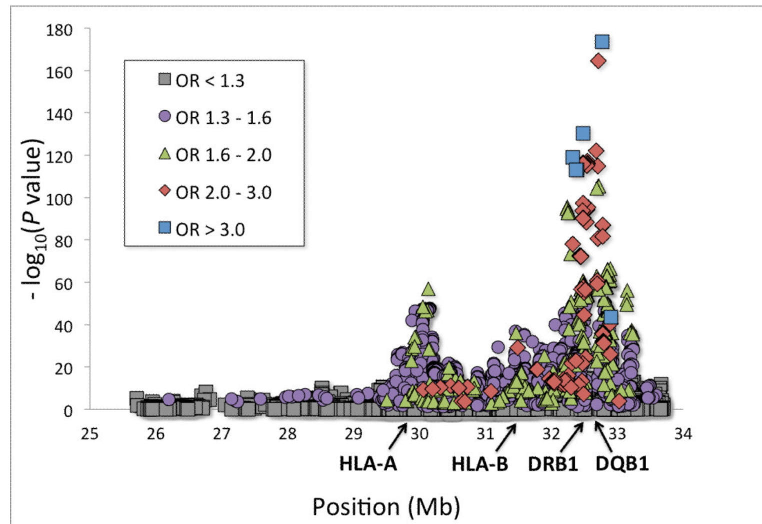


**Figure 2. Association results for the HLA region (chromosome 6, 25-34 Mb)**

SNPs are color-coded by odds ratio (OR) strata.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

## Table 1

Regions reaching genome-wide significant association with oligoarticular and RF negative polyarticular juvenile idiopathic arthritis.

| Gene Region | Chr | Position* | Most significant SNP | Minor allele | MAF controls (n=13056) | MAF cases (n=2816) | Best p-value | Model | Odds ratio | 95% confidence intervals | SNP position |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *HLA-DQB1/HLA-DQA2* | 6 | 32657916 | rs7775055 | G | 0.02 | 0.12 | $3.14 \times 10^{-174}$ | Dominant | 6.01 | 5.3-6.81 | Intergenic |
| *PTPN22* | 1 | 114303808 | rs6679677 | A | 0.1 | 0.14 | $3.19 \times 10^{-25}$ | Additive | 1.59 | 1.45-1.73 | Intergenic |
| *STAT4* | 2 | 191973034 | rs10174238 | G | 0.23 | 0.28 | $1.28 \times 10^{-13}$ | Additive | 1.29 | 1.20-1.37 | Intron |
| *PTPN2* | 18 | 12782448 | rs2847293 | A | 0.17 | 0.2 | $1.44 \times 10^{-12}$ | Additive | 1.31 | 1.22-1.41 | Intergenic |
| *ANKRD55* | 5 | 55440730 | rs71624119 | A | 0.25 | 0.2 | $4.40 \times 10^{-11}$ | Additive | 0.78 | 0.73-0.84 | Intron |
| | | 55442249 | rs10213692# | C | 0.25 | 0.2 | $2.73 \times 10^{-11}$ | Additive | 0.79 | 0.74-0.8 | Intron |
| *IL2/IL21* | 4 | 123387600 | rs1479924 | G | 0.29 | 0.24 | $6.24 \times 10^{-11}$ | Additive | 0.79 | 0.74-0.85 | Intergenic |
| *TYK2* | 19 | 10463118 | rs34536443 | G | 0.05 | 0.03 | $1 \times 10^{-10}$ | Additive | 0.56 | 0.47-0.67 | Coding (NS) |
| *IL2RA* | 10 | 6089841 | rs7909519 | C | 0.11 | 0.08 | $8 \times 10^{-10}$ | Additive | 0.72 | 0.64-0.8 | Intron |
| *SH2B3/ATXN2* | 12 | 111884608 | rs3184504 | A | 0.49 | 0.54 | $2.60 \times 10^{-09}$ | Additive | 1.2 | 1.13-1.27 | Coding (NS) |
| | | 111932800 | rs7137828# | C | 0.49 | 0.54 | $1.61 \times 10^{-09}$ | Additive | 1.20 | 1.13-1.28 | Intron |
| *ERAP2/LNPEP* | 5 | 96350088 | rs27290 | G | 0.44 | 0.47 | $7.5 \times 10^{-09}$ | Dominant | 1.32 | 1.20-1.45 | Intron |
| | | 96357178 | rs27293# | A | 0.44 | 0.47 | $7.37 \times 10^{-09}$ | Dominant | 1.31 | 1.19-1.43 | Intron |
| *UBE2L3* | 22 | 21922904 | rs2266959 | A | 0.19 | 0.22 | $6.2 \times 10^{-09}$ | Dominant | 1.24 | 1.15-1.33 | Intron |
| *C5orf56/IRF1* | 5 | 131813219 | rs4705862 | T | 0.44 | 0.39 | $1.02 \times 10^{-08}$ | Additive | 0.84 | 0.79-0.89 | Intergenic |
| | | 131797547 | rs6894249# | G | 0.39 | 0.35 | $9.73 \times 10^{-10}$ | Dominant | 0.76 | 0.70-0.83 | Intron |
| *RUNX1* | 21 | 36715761 | rs9979383 | G | 0.37 | 0.33 | $1.06 \times 10^{-08}$ | Dominant | 0.78 | 0.72-0.85 | Intergenic |
| | | 36712588 | rs8129030# | T | 0.37 | 0.33 | $5.44 \times 10^{-09}$ | Dominant | 0.78 | 0.71-0.84 | Intergenic |
| *IL2RB* | 22 | 37534034 | rs2284033 | A | 0.44 | 0.39 | $1.55 \times 10^{-08}$ | Additive | 0.84 | 0.79-0.89 | Intron |
| *ATP8B2/IL6R* | 1 | 154364140 | rs11265608 | A | 0.1 | 0.12 | $2.75 \times 10^{-08}$ | Dominant | 1.33 | 1.2-1.47 | Intergenic |
| | | 154379369 | rs72698115# | C | 0.1 | 0.12 | $1.26 \times 10^{-08}$ | Dominant | 1.36 | 1.22-1.52 | Intron |
| *FAS* | 10 | 90762376 | rs7069750 | C | 0.44 | 0.48 | $2.93 \times 10^{-08}$ | Additive | 1.18 | 1.11-1.25 | Intron |
| *ZFP36L1* | 14 | 69253364 | rs12434551 | A | 0.47 | 0.43 | $1.59 \times 10^{-08}$ | Dominant | 0.77 | 0.71-0.85 | Intergenic |

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

| Gene Region | Chr | Position* | Most significant SNP | Minor allele | MAF controls (n=13056) | MAF cases (n=2816) | Best p-value | Model | Odds ratio | 95% confidence intervals | SNP position |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 69260588 | rs3825568# | T | 0.46 | 0.42 | $1.24 \times 10^{-08}$ | Dominant | 0.77 | 0.70-0.84 | 5′UTR |

Chr=chromosome. MAF=minor allele frequency, NS=non-synonymous.

*Coordinates are based on the NCBI37 assembly.

#Imputed SNP results are included when they show a better p-value than the most significant directly genotyped SNP in the region.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

## Table 2

Regions with suggestive significant association with oligoarticular and RF negative polyarticular juvenile idiopathic arthritis ($p < 1 \times 10^{-6}$ and $p > 5 \times 10^{-8}$)

| Gene region | Chr | Position* | Most significant SNP | Minor allele | MAF controls (n=13056) | MAF cases (n=2816) | Best p-value | Model | Odds ratio | 95% confidence intervals | SNP position |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *LTBR* | 12 | 6495275 | rs2364480 | C | 0.25 | 0.28 | $5.10 \times 10^{-08}$ | Additive | 1.2 | 1.12-1.28 | Coding (NS) |
| | | 6493351 | rs10849448# | A | 0.24 | 0.27 | $4.54 \times 10^{-09}$ | Additive | 1.24 | 1.15-1.33 | 5′UTR |
| *IL6* | 7 | 22798080 | rs7808122 | A | 0.44 | 0.48 | $5.80 \times 10^{-08}$ | Additive | 1.19 | 1.11-1.25 | Intergenic |
| | | 22809490 | rs6946509# | T | 0.45 | 0.48 | $3.36 \times 10^{-08}$ | Additive | 1.19 | 1.12-1.26 | Intergenic |
| *COG6* | 13 | 40350912 | rs7993214 | A | 0.35 | 0.31 | $1.61 \times 10^{-07}$ | Additive | 0.84 | 0.79-0.9 | Intergenic |
| | | 40355913 | rs9532434# | T | 0.36 | 0.32 | $4.52 \times 10^{-08}$ | Additive | 0.84 | 0.79-0.89 | Intron |
| Chr13q14 | 13 | 43056036 | rs34132030 | A | 0.32 | 0.29 | $1.77 \times 10^{-07}$ | Additive | 1.18 | 1.11-1.26 | Intergenic |
| *CCR1/CCR3* | 3 | 46253650 | rs79893749 | A | 0.15 | 0.12 | $1.88 \times 10^{-07}$ | Additive | 0.78 | 0.72-0.86 | Intergenic |
| *PRR5L* | 11 | 36363575 | rs4755450 | A | 0.35 | 0.31 | $3.35 \times 10^{-07}$ | Dominant | 0.8 | 0.74-0.87 | Intergenic |
| | | 36343693 | rs7127214# | G | 0.35 | 0.31 | $1.90 \times 10^{-08}$ | Dominant | 0.78 | 0.71-0.85 | Intron |
| *PRM1/C16orf75* | 16 | 11428643 | rs66718203 | C | 0.18 | 0.14 | $4.46 \times 10^{-07}$ | Additive | 0.81 | 0.74-0.88 | Intergenic |
| | | 11471414 | rs11074967# | G | 0.42 | 0.38 | $2.4 \times 10^{-07}$ | Additive | 0.85 | 0.80-0.91 | Intergenic |
| *RUNX3* | 1 | 25197155 | rs4648881 | G | 0.49 | 0.53 | $4.66 \times 10^{-07}$ | Additive | 1.16 | 1.1-1.23 | Intergenic |
| *C3orf1/CD80* | 3 | 119229486 | rs4688013 | A | 0.19 | 0.22 | $6.30 \times 10^{-07}$ | Additive | 1.2 | 1.12-1.29 | Intron |
| | | 119221064 | rs11714843# | A | 0.18 | 0.21 | $3.64 \times 10^{-07}$ | Additive | 1.22 | 1.13-1.31 | Intron |
| *JAZF1* | 7 | 28182306 | rs10280937 | G | 0.11 | 0.13 | $6.60 \times 10^{-07}$ | Additive | 1.25 | 1.15-1.37 | Intron |
| | | 28187344 | rs73300638# | C | 0.11 | 0.14 | $1.12 \times 10^{-07}$ | Additive | 1.28 | 1.17-1.41 | Intron |
| *AFF3/LONRF2* | 2 | 100813499 | rs6740838 | A | 0.39 | 0.43 | $8.83 \times 10^{-07}$ | Dominant | 1.25 | 1.14-1.37 | Intergenic |
| | | 100834217 | rs10194635# | T | 0.39 | 0.43 | $8.10 \times 10^{-07}$ | Dominant | 1.24 | 1.14-1.36 | Intergenic |

Chr=chromosome. MAF=minor allele frequency.

*Coordinates are based on the NCBI37 assembly.

#Imputed SNP results are included when they show a better p-value than the most significant directly genotyped SNP in the region.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

## Table 3

Potential causal SNPs within the JIA risk regions

| Lead SNP | SNP in strong LD ($r^2$>0.9) with the lead SNP | Chr | Position* | $r^2$ with lead SNP | Location | Regulatory potential | Conservation | Functional prediction* | eQTL# |
|---|---|---|---|---|---|---|---|---|---|
| **Genome wide significant SNPs** | | | | | | | | | |
| rs6679677 | rs2476601 | 1 | 114377568 | 1 | Exon of *PTPN22* | 0.14 | 0.999 | benign; tolerated | |
| rs11265608 | rsl205591 | 1 | 154298374 | 1 | Intron of *ATP8B2* | 0.89 | 0 | | |
| rs1479924 | rs13144509 | 4 | 123473487 | 0.94 | Intergenic between *IL2* and *IL21* | 0.17 | 1 | | |
| rs27290 | rs27290 | 5 | 96350088 | - | Intron of *LNPEP* | 0.21 | 0 | | Yes[35-37] |
| rs3184504 | rs3184504 | 12 | 111884608 | - | Exon of *SH2B3* | 0.29 | 0.005 | benign; tolerated | |
| rs12434551 | rs3825568 | 14 | 69260588 | 0.98 | 5′ UTR of *ZFP36L1* | 0.55 | 0.002 | | |
| rs34536443 | rs34536443 | 19 | 10463118 | - | Exon of *TYK2* | 0.40 | 0.19 | probably damaging; deleterious | |
| rs34536443 | rs74956615 | 19 | 10427721 | 1 | Intron of *RAVER1* | 0 | 0.998 | | |
| rs2266959 | rs2266959 | 22 | 21922904 | - | Intron of *UBE2L3* | 0.47 | 0.003 | | |
| rs2266959 | rs2298428 | 22 | 21982892 | 1 | Exon of *YDJC* | 0.37 | 1 | benign; tolerated | |
| rs2266959 | rs4820091 | 22 | 21940189 | 1 | Intron of *UBE2L3* | 0 | 0 | | Yes[35, 37] |
| **Suggestive SNPs** | | | | | | | | | |
| rs4688013 | rs17203104 | 3 | 119139575 | 0.92 | Intergenic between *CDGAP* and *TMEM39A* | 0 | 0.998 | | |
| rs2364480 | rs2364481 | 12 | 6497260 | 1 | Intron of *LTBR* | 0.36 | 0.002 | | |
| rs2364480 | rs2364480 | 12 | 6495275 | - | Exon of *LTBR* | 0.34 | 0.005 | | Yes[35-37] |

SNPs in strong LD ($r^2$>0.9) with the lead SNP on ImmunoChip with evidence for either strong regulatory potential (>0.35)[30] or conservation (>0.998)[29]

*Coordinates are based on the NCBI37 assembly.

*Functional prediction based on PolyPhen[38]

#Data from three studies was considered: lymphoblastoid cell line (LCL) from HapMap3 (Stranger et al, 2012[35]), fibroblast (F), LCL and T-cell(T) from umbilical cords of 75 Geneva Gencord individuals (Dimas et al, 2009[36]) and adipose (A), LCL and skin (S) from 856 healthy female twins of the MuTHER resource (Grundberg et al 2012[37]). Yes if evidence for eQTL ($p < 1 \times 10^{-3}$).