Arthritis & Rheumatology
Vol. 74, No. 8, August 2022, pp 1420–1429
DOI 10.1002/art.42129

American College of Rheumatology
Empowering Rheumatology Professionals

Identification of Novel Loci Shared by Juvenile Idiopathic Arthritis Subtypes Through Integrative Genetic Analysis

Jin Li,[1] Yun R. Li,[2] Joseph T. Glessner,[3] Jie Yang,[1] Michael E. March,[3] Charlly Kao,[3] Courtney N. Vaccaro,[3] Jonathan P. Bradfield,[3] Junyi Li,[1] Frank D. Mentch,[3] Hui-Qi Qu,[3] Xiaohui Qi,[1] Xiao Chang,[3] Cuiping Hou,[3] Debra J. Abrams,[3] Haijun Qiu,[3] Zhi Wei,[4] John J. Connolly,[3] Fengxiang Wang,[3] James Snyder,[3] Berit Flatø,[5] Susan D. Thompson,[6] Carl D. Langefeld,[7] Benedicte A. Lie,[5] Jane E. Munro,[8] Carol Wise,[9] Patrick M. A. Sleiman,[10] and Hakon Hakonarson[10]

Objective. Juvenile idiopathic arthritis (JIA) is the most common chronic immune-mediated joint disease among children and encompasses a heterogeneous group of immune-mediated joint disorders classified into 7 subtypes according to clinical presentation. However, phenotype overlap and biologic evidence suggest a shared mechanistic basis between subtypes. This study was undertaken to systematically investigate shared genetic underpinnings of JIA subtypes.

Methods. We performed a heterogeneity-sensitive genome-wide association study encompassing a total of 1,245 JIA cases (classified into 7 subtypes) and 9,250 controls, followed by fine-mapping of candidate causal variants at each genome-wide significant locus, functional annotation, and pathway and network analysis. We further identified candidate drug targets and drug repurposing opportunities by in silico analyses.

Results. In addition to the major histocompatibility complex locus, we identified 15 genome-wide significant loci shared between at least 2 JIA subtypes, including 10 novel loci. Functional annotation indicated that candidate genes at these loci were expressed in diverse immune cell types.

Conclusion. This study identified novel genetic loci shared by JIA subtypes. Our findings identified candidate mechanisms underlying JIA subtypes and candidate targets with drug repurposing opportunities for JIA treatment.

INTRODUCTION

Juvenile idiopathic arthritis (JIA), the most common chronic immune-mediated joint disease among children, represents a heterogeneous group of immune-mediated diseases that are difficult to diagnose (1). JIA causes severe joint pain, and delays in therapy can result in joint deformities, prompting the need for early genetic or molecular diagnosis.

More than 30 common variant JIA loci have been identified in genome-wide association studies (GWAS); however, the number of JIA loci is far less than that of other autoimmune diseases. To date, GWAS on JIA have been formally performed in seronegative JIA (2) and systemic JIA (3), while other studies show that the loci implicated in seropositive adult rheumatoid arthritis (RA) are applicable to seropositive polyarticular JIA (4). Due to limited sample sizes, it is difficult to investigate the 7 JIA disease subtypes

Ms. Yang's work was supported by the Ministry of Science and Technology of China (grant 2018YFC1313002). Dr. Yun R. Li's work was supported by the NIH (fellowship awards F30 and F32), and the Children's Hospital of Philadelphia Institute For Translational Medicine and Therapeutics Junior Investigator Pilot Grant Program Pilot Research award. Dr. Hakonarson's work was supported by institutional development funds from the Children's Hospital of Philadelphia, the Children's Hospital of Philadelphia Endowed Chair in Genomic Research, and the National Human Genome Research Institute–sponsored Electronic Medical Records and Genomics Network (grant U01-HG006830).

[1]Jin Li, PhD, Jie Yang, MS, Junyi Li, MB, Xiaohui Qi, MB: Department of Cell Biology, the Province and Ministry Co-sponsored Collaborative Innovation Center for Medical Epigenetics, School of Basic Medical Sciences, Tianjin Medical University, Tianjin, China; [2]Yun R. Li, MD, PhD: The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, City of Hope Comprehensive Cancer Center, Duarte, California, and Translational Genomics Research Institute, Phoenix, Arizona; [3]Joseph T. Glessner, PhD, Michael E. March, PhD, Charlly Kao, PhD, Courtney N. Vaccaro, MS, Jonathan P. Bradfield, PhD, Frank D. Mentch, PhD, Hui-Qi Qu, MD, PhD, Xiao Chang, PhD, Cuiping Hou, MS, Debra J. Abrams, MS, Haijun Qiu, PhD, John J. Connolly, PhD, Fengxiang Wang, MD, PhD, James Snyder, BS: The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania; [4]Zhi Wei, PhD: New Jersey Institute of Technology, Newark; [5]Berit Flatø, MD, PhD, Benedicte A. Lie, PhD: Oslo University Hospital, Rikshospitalet, Oslo, Norway; [6]Susan D. Thompson, PhD: University of Cincinnati College of Medicine and Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio; [7]Carl D. Langefeld, PhD: Wake Forest University School of Medicine, Winston-Salem, North Carolina; [8]Jane E. Munro, MD: Royal Children's Hospital, Parkville, Victoria, Australia; [9]Carol Wise, PhD: Scottish Rite for Children, Dallas, Texas; [10]Patrick M. A. Sleiman, PhD, Hakon Hakonarson, MD, PhD: The Children's Hospital of Philadelphia and University of Pennsylvania, Philadelphia.

Drs. Jin Li and Yun R. Li contributed equally to this work.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Fart.42129&file=art42129-sup-0001-Disclosureform.pdf.

Address correspondence to Hakon Hakonarson, MD, PhD, 1515 Civic Center Boulevard, Philadelphia, PA 19104-4399. Email: hakonarson@email.chop.edu.

Submitted for publication March 21, 2021; accepted in revised form March 25, 2022.

defined using the International League of Associations for Rheumatology (ILAR) criteria: enthesitis-related arthritis (ERA), rheumatoid factor (RF)–negative polyarthritis (PA), RF-positive PA, oligoarthritis, psoriatic arthritis (PsA), systemic arthritis, and undifferentiated arthritis (UA) (5).

Biologic and molecular evidence of serum autoantibodies and other molecular biomarkers of immunologic defects suggest that the existing 7 JIA subtypes are heterogeneous but also have overlapping molecular features (6). The RF-positive PA JIA subtype shares serologic features, such as RF, anti–cyclic citrullinated peptide antibodies, anti–mutated citrullinated vimentin antibodies, and genetic loci, with other autoimmune diseases like RA (4,7), resembling predominantly seropositive autoimmune diseases, while ERA and systemic arthritis findings had more characteristics of autoinflammatory diseases (8,9). A transcriptomic study also identified distinct differentially expressed genes and differentially expressed shared genes between the polyarthritis, oligoarthritis, and systemic arthritis subtypes (10). Therefore, the 7 JIA subtypes are not so distinct from a mechanistic or therapeutic standpoint (1,11). Given that a large number of immune pathway modulators are already approved for use in other rheumatologic and immune-mediated disorders, including RA, understanding the genetic basis of JIA subtypes may allow for the early and rapid introduction of already approved drugs to treat disease according to molecular subtype definitions. Studies focusing on JIA subtypes have started to identify molecular differences and similarities (3,4). However, these studies have not systematically examined all 7 subtypes of JIA in terms of genome-wide variants.

In this subset-sensitive GWAS, we identified 15 genome-wide significant loci shared between certain JIA subtypes. We further identified candidate drug targets with drug repurposing opportunities based on genetic associations. The integrative genetic analysis results presented provide new insight into the biologic differences between JIA subtypes and suggest therapeutic targets.

PATIENTS AND METHODS

Dr. Hakonarson will provide access to data upon reasonable request.

Study population. Subjects were recruited for the JIA cohort in the US, Australia, and Norway and the cohort comprised a total of 1,485 patients with arthritis onset at age <16 years (Supplementary Table 1, available on the *Arthritis & Rheumatology* website at http://onlinelibrary.wiley.com/doi/10.1002/art.42129). JIA diagnoses and subtypes were determined according to the revised ILAR criteria (5) and confirmed using the JIA calculator software (12), an algorithm-based tool adapted from the ILAR JIA criteria. Prior to standard quality control procedures and exclusion of patients of non-European ancestry, the JIA cohort was composed of 464 case subjects of self-reported European ancestry from the Texas Scottish Rite Hospital for Children (Dallas, Texas) and Children's Mercy Kansas City Hospital (Kansas City, Missouri), 296 case subjects from the Children's Hospital of Philadelphia (Philadelphia, Pennsylvania), 221 case subjects from the Murdoch Children's Research Institute at the Royal Children's Hospital (Melbourne, Victoria, Australia), and 504 case subjects from Oslo University Hospital (Oslo, Norway). Age- and sex-matched control subjects were identified from the Children's Hospital of Philadelphia Center for Applied Genomics Biobank, ascertained by the exclusion of any patient with any International Classification of Diseases, Ninth Revision, codes for autoimmune disorders or immunodeficiency disorders. A subset of the current study subjects was described in a previous study (see Supplementary Table 1, http://onlinelibrary.wiley.com/doi/10.1002/art.42129) (13). This study did not contain personal medical information about an identifiable living individual.

Ethics statement. Ethics approval for this study was obtained from the Children's Hospital of Philadelphia research ethics Institutional Review Board (approval no. 16-013278) and the ethics boards at other collaborating centers. This study was carried out in accordance with nationally approved guidelines. Written informed assent or consent was obtained from all subjects and/or their legal guardians.

Genotyping. Genomic DNA was extracted from peripheral blood, and we performed sample quality control filtering before and after genotyping using standard methods. In our cohort, all samples were genotyped at the Center for Applied Genomics on HumanHap550 and HumanHap610 BeadChip arrays (Illumina). The single-nucleotide polymorphism (SNP) genotype was defined using BeadStudio (Illumina) with default parameters. To minimize population stratification, only individuals of self-reported European ancestry, further confirmed using principal components analysis, were included in the present study. Details of the principal components analysis are provided below.

Sample and SNP quality control. SNPs with a low genotyping rate (<95%), those with a low minor allele frequency (<0.01), or those with significant departure from the expected Hardy-Weinberg equilibrium ($P < 1 \times 10^{-6}$) were excluded. Samples with a low overall genotyping call rate (<95%) or those determined to be from patients of European ancestry who were considered to be outliers according to principal components analysis (detailed below) were removed. In addition, one of each pair of related individuals, as determined using identity-by-state analysis (PI_HAT > 0.1875), was excluded, with cases preferentially retained when possible. We conducted case–case comparison by performing association testing between the case groups in each of the 4 cohorts. Any SNP with an association indicated by $P < 1 \times 10^{-5}$, which suggests significant differences in allele frequency between 2 cohorts, was excluded from further analyses.

(Supplementary Tables 2 and 3, http://onlinelibrary.wiley.com/doi/10.1002/art.42129).

Principal components analysis. To assess ethnicity, we combined our SNP data together with the HapMap data set to conduct a principal components analysis. We took the set of SNPs common to both data sets and narrowed them down using Plink command "--indep-pairwise 50 10 0.2." We conducted principal components analysis on the data set with the narrowed-down SNPs via Plink (14). K-means clustering was used to group subjects into distinct populations of ethnic origin, and subjects of European ancestry were identified. A principal components analysis was similarly conducted among subjects of European ancestry in our data set again to determine within-population structure (Supplementary Figure 1, http://onlinelibrary.wiley.com/doi/10.1002/art.42129).

Genome-wide SNP imputation. We used Shapelt (15) for whole-chromosome pre-phasing and IMPUTE2 for imputation of the 1000 Genomes Project reference panel (URL: https://mathgen.stats.ox.ac.uk/impute/impute_v2.html [June 2014 haplotype release]). For both, we used parameters suggested by the software developers and described elsewhere (15,16). Imputation was conducted for each 5-Mb regional chunk across the genome, and data were subsequently merged for association testing. Prior to imputation, all SNPs were filtered using the criteria described above. We filtered out SNPs with an Info score <0.8, Hardy-Weinberg equilibrium test with a significance $P < 1 \times 10^{-6}$, and overall minor allele frequency <0.01.

Association analysis. We performed whole-genome association testing using post-imputation genotype probabilities with the score test implemented in SNPTEST software version 2.5. In all analyses, we adjusted for both sex and ancestry with conditioning for sex and the first 9 principal components derived from the Plink principal components analysis, which yielded λ_{gc} values within acceptable limits for all disease subtype cohorts. The extent of population stratification was assessed using a quantile–quantile plot of the test statistics and by calculating inflation factor λ_s.

Heterogeneity-sensitive meta-analysis. To identify genetic loci that were associated with multiple JIA subtypes and determine the subtype combination that each locus was most strongly associated with, "h.types" and "h.traits" in the R statistical software package ASSET (17) were applied to an exhaustive disease subtype model search, which has been described in detail in our previous study and in other studies (18,19). Briefly, the different combinations of JIA subtypes were exhaustively enumerated and tested for associations with each locus. The combination that yielded the most significant association statistics was selected as the best disease subtype model. A score test implemented in R package ASSET was used in the "h.types" approach, with adjustment for covariates in the analysis. In our analyses, the "h.types" and "h.traits" methods yielded similar results. We used the discrete local maximum method of correction for multiple testing across all subtype combinations. The contribution of each non-null study to the shared association was measured using the absolute value of the weighted Z statistics $\left|\sqrt{\pi_k(S)}Z_k\right|$, in which $\pi_k(S) = n_k / \Sigma_{k\varepsilon S} n_k$ represented the sample size of k^{th} subtype relative to the total sample size of the subtypes of the most significantly associated subtype combination S.

Fine-mapping. Fine-mapping was performed using FINEMAP version 1.3.1 (20). We used GWAS summary statistics and SNP Pearson's correlation matrixes calculated from genotyped data from the same individuals as input in FINEMAP. We used the default parameter setting with the maximum number of allowed causal SNPs as 5. Candidate causal SNPs with posterior probability >0.2 and heterogeneity-sensitive GWAS with $P < 10^{-4}$ were selected.

Pathway and protein–protein interaction network analysis. The overrepresentation pathway analysis of the 16 candidate genes at 15 genome-wide significant loci was conducted using a web portal WEB-based gene set analysis toolkit (URL: http://www.webgestalt.org/). The protein–protein interaction network visualization analysis and competitive pathway enrichment analysis based on genome-wide, summary-level data were performed using GSA-SNP2 (URL: https://sites.google.com/view/gsasnp2). Highly correlated adjoining genes were combined based on linkage disequilibrium in the 1000 Genomes European population. The default setting of GSA-SNP2 was used to define each gene region and gene transcript region 20 kb upstream and downstream. The collection of gene set databases included BioCarta, KEGG, the Reactome database, and Molecular Signatures Database Pathway Interaction Database. We used the STRING database for network construction and visualization. The significance threshold was defined as $q < 0.05$ after correction for multiple testing. Significance, defined as a gene score <0.005 and $q < 0.05$, was chosen for constructing a global visual network.

HLA imputation. SNPs within the HLA region, spanning 29–34 Mb on chromosome 6 of the human (hg19) reference genome, were extracted after SNP array data had been quality control filtered. Data from all JIA subjects and controls were imputed together using SNP2HLA software (URL: http://www.broadinstitute.org/mpg/snp2hla/) with the Type 1 Diabetes Genetics Consortium reference panel. We also conducted a case–case comparison of the HLA alleles by performing association testing between the case groups in each of the 4 cohorts. Any HLA alleles with an association indicated by $P < 1 \times 10^{-5}$, which suggested significant differences in allele frequency

between 2 cohorts, were excluded from further analyses. The HLA allele frequencies at a 2-digit level were compared between cases and controls for each JIA subtype, with the odds ratio and *P* value for the association derived from a chi-square test of the 2 × 2 table.

RESULTS

Identification of novel pleiotropic JIA loci through a heterogeneity-sensitive GWAS. Our JIA case–control cohort included 1,485 JIA cases (including all 7 JIA subtypes) and 10,352 controls with no history of any existing autoimmune or immune-mediated disease (Figure 1 and Supplementary Table 1, http://onlinelibrary.wiley.com/doi/10.1002/art.42129). A total of 506,520 genotyped SNPs from 1,245 JIA cases and 9,250 controls passed quality control filtering (Supplementary Figure 1, http://onlinelibrary.wiley.com/doi/10.1002/art.42129).

To optimize study power, after imputation, we performed a heterogeneity-sensitive GWAS (17) across JIA cases in 7 JIA subtypes and the pool of shared control samples. The heterogeneity-sensitive GWAS approach was used to first test every SNP that passed quality control to identify the most strongly associated disease combinations at each SNP, with the discrete local maximum method applied for adjustment for multiple testing. This approach has been successfully used to assess the complex relationships between pediatric autoimmune diseases (18) and neuropsychiatric disorders (19). We included the first 9 principal components as covariates, and the resulting genomic inflation factor for the final heterogeneity-sensitive GWAS results was 1.01 (Supplementary Figure 2 and Supplementary Table 4, http://onlinelibrary.wiley.com/doi/10.1002/art.42129), suggesting that population stratification was well controlled. Our results showed that the majority of the association loci reported in previous studies was replicated in our study. Among the 120 association signals reported in the GWAS Catalog, including those with marginal genome-wide significance ($5 \times 10^{-8} < P < 1 \times 10^{-6}$), 81.0% were replicated in our study at least at a nominal significance level (Supplementary Data 1, http://onlinelibrary.wiley.com/doi/10.1002/art.42129).

We observed 15 loci surpassing the genome-wide significance threshold (Figure 2), in addition to strong association signals at the major histocompatibility complex (MHC) region (the 29–34-Mb region on chromosome 6). Five of the 15 genome-wide significant loci overlapped with previously reported autoimmune disease loci (Supplementary Table 5, http://onlinelibrary.wiley.com/doi/10.1002/art.42129), and the remaining 10 were novel JIA loci (Table 1) (for regional association plots, see Figure 3 and Supplementary Figure 3, http://onlinelibrary.wiley.com/doi/10.1002/art.42129).

The genome-wide significant association of all 15 loci was attributed to associations at a nominal significance level in 2 or more JIA subtypes (Supplementary Figure 4 and Supplementary Table 6 and 7, http://onlinelibrary.wiley.com/doi/10.1002/art.42129), including those shared between systemic arthritis and other JIA subtypes. In addition, we observed 25 SNPs with marginal genome-wide significance in our heterogeneity-sensitive GWAS, with 24 shared between 2 or more JIA subtypes at a nominal significance level (Supplementary Table 8, http://onlinelibrary.wiley.com/doi/10.1002/art.42129). Thus, our findings support a common genetic basis among JIA subtypes in addition to distinct clinical and molecular features.

Replication of the novel JIA loci. After examining UK Biobank data sets, we found a significant, genome-wide association between SNP rs7731626 and RA. Interestingly, SNP rs12203592 is strongly associated with an RF level >16 IU/ml, while in our study rs12203592 was more strongly associated with

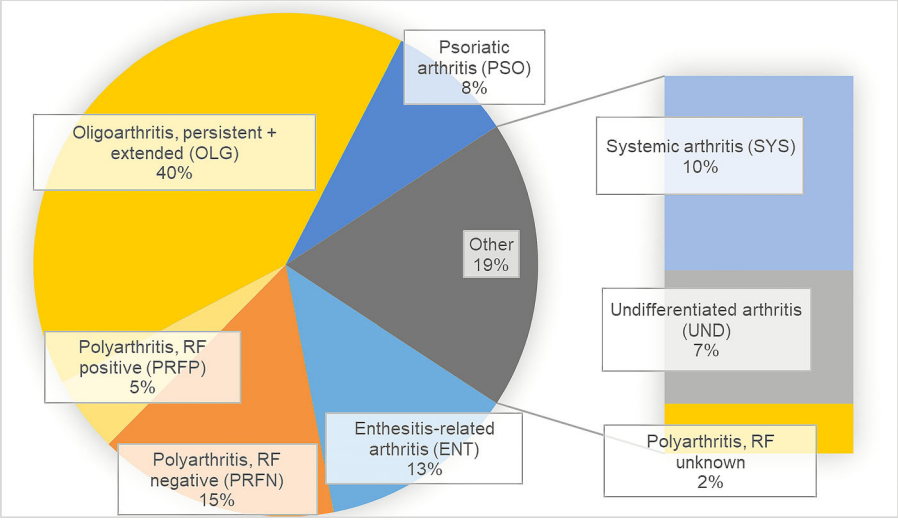

Figure 1. Distribution of juvenile idiopathic arthritis (JIA) subtypes represented in our JIA case–control cohort. RF = rheumatoid factor. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.42129/abstract.

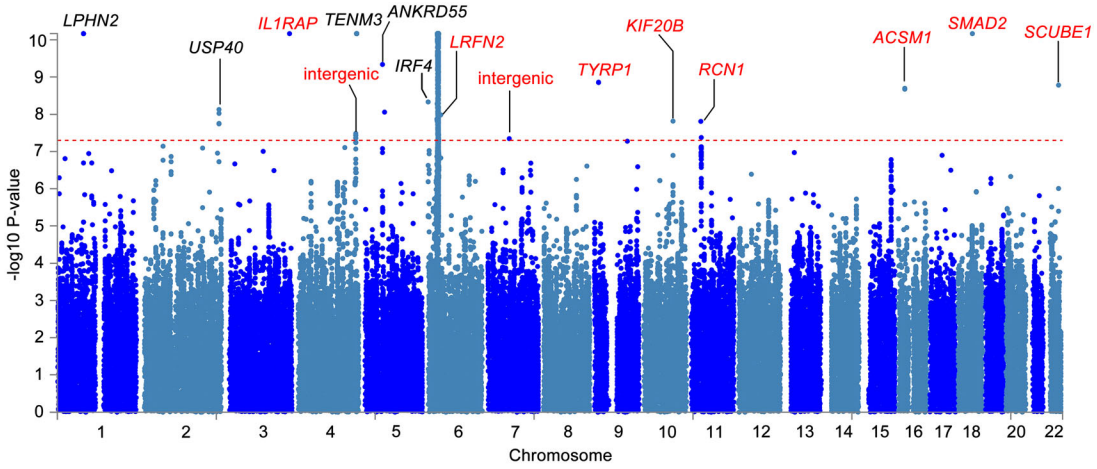

Figure 2. Manhattan plot showing association statistics for the heterogeneity-sensitive genome-wide association study of juvenile idiopathic arthritis subtypes, with adjustment for multiple testing. Candidate gene symbols for genome-wide significant loci are shown, with novel loci indicated in red. Symbols represent individual genes. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.42129/abstract.

RF-negative PA and oligoarthritis than the other JIA subtypes. In addition, a significant, genome-wide association between SNP rs7660520 and psoriasis was shown. In our data set, this SNP was associated with multiple JIA subtypes, including PsA. In the UK Biobank, rs114664970, a novel, genome-wide significant SNP identified in our study, was associated with chronic sinusitis, severe cases of which could be symptomatic of autoimmune diseases (Supplementary Table 9, http://onlinelibrary.wiley.com/doi/10.1002/art.42129). The index SNPs at other loci were also associated with established autoimmune or immune-mediated musculoskeletal system conditions. We examined ImmunoChip (IC) data that had previously been reported by Hinks et al (2). At only 2 regions, there were SNPs within 250 kb upstream/downstream of our heterogeneity-sensitive GWAS SNPs rs12203592 and rs7731626 with $r^2 > 0.5$. These SNPs were associated with JIA according to the IC data (Supplementary Table 10, http://onlinelibrary.wiley.com/doi/10.1002/art.42129).

Fine-mapping and functional annotation of novel loci. We conducted fine-mapping for each novel loci. Ten SNPs were identified as candidate causal variants (Supplementary Table 11, http://onlinelibrary.wiley.com/doi/10.1002/art.42129). We conducted functional annotation of all index SNPs, candidate causal variants, and leading SNPs at marginally genome-wide significant loci in the Encyclopedia of DNA Elements and Roadmap Epigenomics databases and found overlap between these SNPs/loci and chromatin marks or DNase I–hypersensitive sites, likely playing a role in regulating target gene expression (Supplementary Figure 5, http://onlinelibrary.wiley.com/doi/10.1002/art.42129).

We mapped index SNPs to candidate genes according to expression quantitative trait locus (eQTL) data and high-throughput chromosome conformation capture (3C) data (Supplementary Figures 6–8, http://onlinelibrary.wiley.com/doi/10.1002/art.42129), and the most likely candidate genes at genome-wide significant loci are indicated in Figure 2. Highly significant eQTL relationships were observed between heterogeneity-sensitive GWAS SNP gene pairs (rs7731626 and *ANKRD55*, rs7731626 and *IL6ST*, rs12203592 and *IRF4*) in different immune tissue and immune cell types (Supplementary Figures 6 and 7, http://onlinelibrary.wiley.com/doi/10.1002/art.42129). A strong eQTL relationship between rs12795402 and *RCN1* was also reported in the Biobank-based Integrative Omics Studies QTL database (21). The candidate gene or genes for other heterogeneity-sensitive GWAS SNPs were determined according to nominally significant eQTL and/or high-throughput 3C interactions (Supplementary Figure 8, http://onlinelibrary.wiley.com/doi/10.1002/art.42129). According to data from the Database of Immune Cell Expression, Expression quantitative trait loci and Epigenomics database, 14 candidate genes were expressed in diverse immune cell types at medium-to-high levels (Supplementary Figure 9, http://onlinelibrary.wiley.com/doi/10.1002/art.42129), suggesting that both innate immunity and adaptive immunity are involved in JIA pathogenesis.

Association of HLA alleles and JIA subtypes. According to the SNP genotype at the MHC region, we further imputed classic HLA alleles and examined their association with each JIA subtype after quality control filtering (Supplementary Table 12, http://onlinelibrary.wiley.com/doi/10.1002/art.42129). As expected, we found a highly significant association between HLA–B*27 and ERA. Multiple HLA alleles, such as HLA–B*40, HLA–DRB1*04, and HLA–DPB1*02, were significantly associated with >1 JIA subtype (Supplementary Table 13, http://onlinelibrary.wiley.com/doi/10.1002/art.42129). Additional HLA alleles were associated with JIA subtypes that surpassed the multiple testing–adjusted significance threshold. We conducted stepwise

Table 1. Summary statistics of the independent, genome-wide significant loci*

SNP	Chromosome	Position†	Candidate gene(s)	A1	MAF	P_{adj}	Associated subtypes
rs2066363	1	82237577	*LPHN2*	T	0.310	6.82×10^{-11}	ERA, RF-negative PA, RF-positive PA, oligoarthritis, PsA, systemic arthritis, UA
rs144844686	2	234576970	*USP40*	T	0.029	7.40×10^{-9}	RF-negative PA, oligoarthritis, PsA
rs7636581‡	3	189781195	*IL1RAP*, *CLDN1*	A	0.120	6.82×10^{-11}	RF-negative PA, oligoarthritis, PSA
rs13119493‡	4	180911259	Intergenic	G	0.048	3.26×10^{-8}	RF-negative PA, oligoarthritis
rs7660520	4	183745321	*DCTD*, *TENM3*	A	0.270	6.82×10^{-11}	ERA, RF-negative PA, RF-positive PA, oligoarthritis, PsA, systemic arthritis, UA
rs7731626	5	55444683	*ANKRD55*, *IL6ST*	A	0.380	4.62×10^{-10}	RF-negative PA, oligoarthritis, PsA
rs12203592	6	396321	*IRF4*	T	0.190	4.62×10^{-9}	ERA, RF-negative PA, oligoarthritis, systemic arthritis, UA
rs114664970‡	6	40127169	*LRFN2*	C	0.012	1.03×10^{-8}	PsA, systemic arthritis
rs727845‡	7	67607209	Intergenic	G	0.190	4.49×10^{-8}	ERA, RF-negative PA, RF-positive PA, oligoarthritis, UA
rs7042370‡	9	12785073	*TYRP1*	C	0.440	1.39×10^{-9}	ERA, RF-positive PA, oligoarthritis, PsA, systemic arthritis
rs117572873‡	10	91997663	*KIF20B*	G	0.011	1.52×10^{-8}	RF-positive PA, oligoarthritis
rs12795402‡	11	32255936	*RCN1*	C	0.340	4.19×10^{-8}	RF-negative PA, RF-positive PA, oligoarthritis
rs147585949‡	16	20726695	*ACSM1*	A	0.014	1.98×10^{-9}	ERA, RF-negative PA, PsA
rs11663074‡	18	45023793	*SMAD2*	C	0.170	6.82×10^{-11}	ERA, RF-negative PA, RF-positive PA, oligoarthritis, PsA, systemic arthritis, UA
rs138816451‡	22	43649657	*SCUBE1*	A	0.016	1.63×10^{-9}	RF-negative PA, oligoarthritis, PsA, UA

* SNP = single-nucleotide polymorphism; A1 = alternative allele; MAF = minor allele frequency; P_{adj} = adjusted *P* value; ERA = enthesitis-related arthritis; RF = rheumatoid factor; PA = polyarthritis; PsA = psoriatic arthritis; UA = undifferentiated arthritis.
† Position is measured in bp.
‡ Novel locus.

conditional analyses for each JIA subtype. Similar to results reported by Hink et al (22), we observed >2 independent effects across the MHC region in RF-negative PA and oligoarthritis. In addition, multiple independent association signals were detected within the HLA–DRB1 gene (Supplementary Table 14, http://onlinelibrary.wiley.com/doi/10.1002/art.42129).

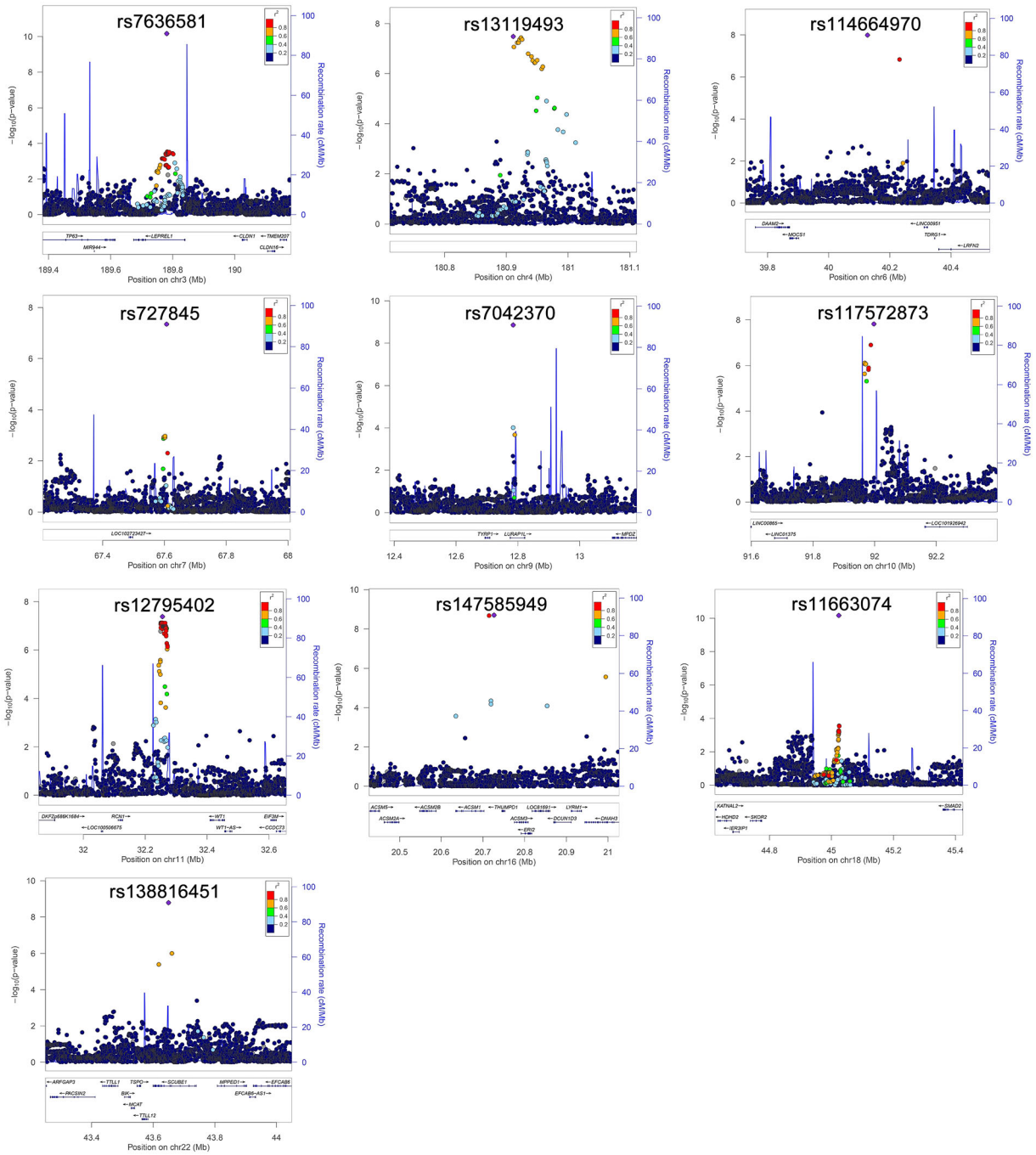

Figure 3. Regional association plots showing novel juvenile idiopathic arthritis genome-wide significant loci. Color-coded symbols represent individual genes showing an association at different thresholds of significance. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.42129/abstract.

Pathway enrichment and network analysis of genome-wide significant loci. To better understand how these loci may contribute to JIA etiology, we performed pathway enrichment analyses and protein–protein interaction network analysis. We first investigated the most likely candidate genes of the 15 genome-wide significant loci using an overrepresentation analysis. The KEGG Th17 cell differentiation pathway was significantly overrepresented, with 4 candidate genes *SMAD2*, *IRF4*, *IL1RAP*, and *IL6ST* in this pathway (Supplementary Table 15, http://onlinelibrary.wiley.com/doi/10.1002/art.42129). Subsequently, in investigating all the heterogeneity-sensitive GWAS results, we found enrichment of 81 KEGG, BioCarta, and Reactome pathways (Supplementary Data 2, http://onlinelibrary.wiley.com/doi/10.1002/art.42129), with pathways related to autoimmune diseases ranking at the top. These top-ranked pathways were mostly driven by the genes at the HLA locus on

chromosome 6. Protein–protein interaction network analysis revealed extensive interaction between immune genes centered on *TNF*, *NOS1*, and HLA genes (Supplementary Figure 10, http://onlinelibrary.wiley.com/doi/10.1002/art.42129).

Identification of candidate drug targets with drug repurposing opportunities among JIA-associated loci. We searched drug target gene databases DrugBank (URL: https://www.drugbank.ca/), DrugCentral (URL: http://drugcentral.org/), and PharmGKB (URL: https://www.pharmgkb.org/) and found that candidate genes at multiple genome-wide significant loci are known targets of existing drugs, including several used for the treatment of arthritis, such as diflunisal, methotrexate, cyclosporine, and diclofenac (Supplementary Table 16, http://onlinelibrary.wiley.com/doi/10.1002/art.42129). These target genes of arthritis drugs support the biologic relevance of our study.

In addition, the high-throughput 3C data revealed evidence of chromatin interaction between rs7636581 and *IL1RAP*, suggesting that *IL1RAP* may be a candidate target gene of this locus. With regard to existing therapies, interleukin-1 (IL-1) antagonists (and IL-6 blockade) have been used for treament of the systemic arthritis subtype and have been transformative in treating this JIA subtype (23). In our analysis, this locus was associated with oligoarthritis and PsA subtypes, suggesting potential application of IL-1 blockade for the treatment of these JIA subtypes in addition to systemic arthritis. Another interesting association was between rs7731626 and *IL6ST*, encoding glycoprotein 130, a coreceptor for many other cytokine receptor complexes besides IL-6 (24). This locus was associated with multiple JIA subtypes, including oligoarthritis, RF-negative PA, and PsA, suggesting that IL-6 blockade may be broadly effective for JIA subtypes. Indeed, tocilizumab is approved by the US Food and Drug Administration for treatment of both systemic arthritis and polyarticular JIA. Patients of both subtypes have shown significant improvements following tocilizumab treatment with confirmed efficacy and safety (25–27).

The association between rs138816451 and *SCUBE1* on chromosome 22 in the JIA subtypes UA, PsA, oligoarthritis, and RF-negative PA was intriguing, as signal peptide, CUB domain, and ECF-like domain containing protein 1 (SCUBE-1) has been implicated in playing a role in angiogenesis and is expressed and bound to the surface of endothelial cells (28). Findings have been reported suggesting that SCUBE-1, SCUBE-3, and vascular endothelial growth factor (VEGF) levels in serum may be a biomarker for angiogenesis (29), which is one of the pathogenic processes involved in psoriasis and arthritis, suggesting that drugs that block angiogenesis may be effective for treating arthritis and psoriasis/PsA (30,31).

DISCUSSION

JIA is a clinically important chronic autoimmune disease among children that causes significant morbidity. However, it has not been as well studied as many other autoimmune diseases, mostly due to sample size limitations and the clinical heterogeneity of JIA. To address these limitations, we conducted a heterogeneity-sensitive GWAS accounting for phenotypic heterogeneity and identified novel pleiotropic loci shared among multiple JIA subtypes. We genetically illustrated how these loci may have joint or disparate effects on JIA disease subtype susceptibility, which sheds light on JIA pathogenesis and the development of targeted therapeutic approaches.

The significant enrichment of Th17 cell differentiation found in pathway analyses highlight the potential importance of this pathway in JIA etiology. Four candidate genes at genome-wide significant loci (*SMAD2*, *IL1RAP*, *IL6ST*, and *IRF4*) are involved in this pathway. SMAD family member 2 plays critical roles in Th1 cell development and in the generation of Th17 cells that drive the development of autoimmune diseases (32,33). The SNP rs80142631 at the *SMAD2* locus has been reported to be associated with eosinophil counts in the European ancestry patient population (34). IL-1 receptor accessory protein (IL1RAP) belongs to the IL-1 receptor complex. NF-κB signaling pathway genes and other genes downstream of IL-1 play critical roles in Th17 cell differentiation. *IL6ST* and *IRF4* have both been associated with autoimmune and autoinflammatory diseases in previous studies, including Crohn's disease and RA. Th17 cells are a lineage of CD4+ T cells that secrete cytokines IL-17A and IL-17F, which are involved in the pathogenesis of both autoimmune diseases and inflammatory diseases (35).

Current JIA therapies mainly include disease-modifying antirheumatic drugs and pain therapies. Targeting the underlying causes of JIA may further enhance the effectiveness of therapeutic strategies, particularly for rare JIA subtypes. IL-1 signaling plays an important role in the regulation of proinflammatory reactions that are involved in various autoinflammatory diseases (36). The association between IL1RAP and oligoarthritis and PsA in our analyses suggests that this locus is a potential therapeutic target that might extend to JIA subtypes besides systemic arthritis. In a pilot clinical trial, IL-1 blockade was shown to improve symptoms among adult patients with PsA (37). Similarly, the association between *IL6ST* and multiple JIA subtypes suggests potential repurposing opportunities for anti–IL-6 for several JIA subtypes. In addition to tocilizumab, sarilumab, another antibody to the IL-6 receptor, has undergone testing for polyarticular JIA and systemic arthritis (38). The association between *SCUBE1* and several JIA subtypes (RF-negative PA, oligoarthritis, PsA, and UA) implicates a potential role of angiogenesis in JIA pathogenesis (30,39). Some antiinflammatory drugs, such as anti–tumor necrosis factor and anti–IL-6, have dual roles in blocking both inflammation and angiogenesis (40,41). Drugs that target angiogenesis/vascularization (e.g., anti-VEGF, anti-TIE2, and anti-angiopoietins) may also have a role in these JIA subtypes, but additional experiments and clinical trials should be conducted to clarify this role (42). The identification of associated variants,

candidate genes, and pathways shared between JIA subtypes may lead to the selection of drug repurposing candidates.

There are several limitations of our study. First, the sample size was limited compared to GWAS on complex human diseases and previous GWAS on JIA (2,11). Second, phenotypic heterogeneity may affect study power. Despite these limitations, in this study, we identified several novel loci, likely owing to the strong genetic contribution to pediatric diseases and improved methodology. Compared to GWAS in adult patients, pediatric disease–expressing phenotypes in early life are typically associated with much stronger gene signals than diseases presenting later in life that are often critically impacted by gene–environment interactions. In addition, our study utilized improved methodology (e.g., heterogeneity-sensitive meta-analysis) and integrative genetic analysis. For example, it is plausible that certain SNPs may either be associated with only some disease subtypes or have opposite effects across JIA subtypes. Thus, using a heterogeneity-sensitive GWAS study method considers JIA subtype heterogeneity, boosts study power, and enables the identification of novel loci associated with JIA.

In summary, we identified novel genetic loci with pleotropic effects across multiple JIA subtypes. Functional annotation indicates that candidate genes at these loci are expressed in diverse immune cell types, which is consistent with their potential role in JIA pathogenesis. In silico analyses suggest that there may be drug repurposing opportunities for rare JIA subtypes, and JIA subtypes may benefit from shared therapeutic approaches according to potential underlying genetic mechanisms.

ACKNOWLEDGMENT

We would like to thank all the patients for their participation in our study.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Hakonarson had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Jin Li, Y. Li, Hakonarson.

Acquisition of data. Abrams, Connolly, Flatø, Thompson, Langefeld, Lie, Munro, Wise, Hakonarson.

Analysis and interpretation of data. Jin Li, Y. Li, Glessner, Yang, March, Kao, Vaccaro, Bradfield, Junyi Li, Mentch, Qu, Qi, Chang, Hou, Qiu, Wei, Wang, Snyder, Sleiman, Hakonarson.

REFERENCES

1. Prakken B, Albani S, Martini A. Juvenile idiopathic arthritis. Lancet 2011;377:2138–49.
2. Hinks A, Cobb J, Marion MC, Prahalad S, Sudman M, Bowes J, et al. Dense genotyping of immune-related disease regions identifies 14 new susceptibility loci for juvenile idiopathic arthritis. Nat Genet 2013;45:664–9.
3. Ombrello MJ, Arthur VL, Remmers EF, Hinks A, Tachmazidou I, Grom AA, et al. Genetic architecture distinguishes systemic juvenile idiopathic arthritis from other forms of juvenile idiopathic arthritis: clinical and therapeutic implications. Ann Rheum Dis 2017;76:906–13.
4. Hinks A, Marion MC, Cobb J, Comeau ME, Sudman M, Ainsworth HC, et al. The genetic profile of rheumatoid factor-positive polyarticular juvenile idiopathic arthritis resembles that of adult rheumatoid arthritis. Arthritis Rheumatol 2018;70:957–62.
5. Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J, et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J Rheumatol 2004;31:390–2.
6. Martini A, Ravelli A, Avcin T, Beresford MW, Burgos-Vargas R, Cuttica R, et al. Toward new classification criteria for juvenile idiopathic arthritis: first steps, Pediatric Rheumatology International Trials Organization International Consensus. J Rheumatol 2019;46:190–7.
7. Zaripova LN, Midgley A, Christmas SE, Beresford MW, Baildam EM, Oldershaw RA. Juvenile idiopathic arthritis: from aetiopathogenesis to therapeutic approaches. Pediatr Rheumatol Online J 2021;19:135.
8. Mistry RR, Patro P, Agarwal V, Misra DP. Enthesitis-related arthritis: current perspectives. Open Access Rheumatol 2019;11:19–31.
9. Cimaz R. Systemic-onset juvenile idiopathic arthritis. Autoimmun Rev 2016;15:931–4.
10. Mo A, Marigorta UM, Arafat D, Chan LH, Ponder L, Jang SR, et al. Disease-specific regulation of gene expression in a comparative analysis of juvenile idiopathic arthritis and inflammatory bowel disease. Genome Med 2018;10:48.
11. López-Isac E, Smith SL, Marion MC, Wood A, Sudman M, Yarwood A, et al. Combined genetic analysis of juvenile idiopathic arthritis clinical subtypes identifies novel risk loci, target genes and key regulatory mechanisms. Ann Rheum Dis 2020;80:321–8.
12. Behrens EM, Beukelman T, Cron RQ. Juvenile idiopathic arthritis classification criteria: loopholes and diagnosis software. J Rheumatol 2007;34:234.
13. Finkel TH, Li J, Wei Z, Wang W, Zhang H, Behrens EM, et al. Variants in CXCR4 associate with juvenile idiopathic arthritis susceptibility. BMC Med Genet 2016;17:24.
14. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–75.
15. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet 2009;5:e1000529.
16. Delaneau O, Coulonges C, Zagury JF. Shape-IT: new rapid and accurate algorithm for haplotype inference. BMC Bioinformatics 2008; 9:540.
17. Bhattacharjee S, Rajaraman P, Jacobs KB, Wheeler WA, Melin BS, Hartge P, et al. A subset-based approach improves power and interpretation for the combined analysis of genetic association studies of heterogeneous traits. Am J Hum Genet 2012;90:821–35.
18. Li YR, Li J, Zhao SD, Bradfield JP, Mentch FD, Maggadottir SM, et al. Meta-analysis of shared genetic architecture across ten pediatric autoimmune diseases. Nat Med 2015;21:1018–27.
19. Cross-Disorder Group of the Psychiatric Genomics Consortium. Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. Cell 2019;179:1469–82.
20. Benner C, Spencer CC, Havulinna AS, Salomaa V, Ripatti S, Pirinen M. FINEMAP: efficient variable selection using summary data from genome-wide association studies. Bioinformatics 2016;32: 1493–501.
21. Zhernakova DV, Deelen P, Vermaat M, van Iterson M, van Galen M, Arindrarto W, et al. Identification of context-dependent expression quantitative trait loci in whole blood. Nat Genet 2017;49:139–45.

22. Hinks A, Bowes J, Cobb J, Ainsworth HC, Marion MC, Comeau ME, 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 2017;76:765–72.
23. Dinarello CA, Simon A, van der Meer JW. Treating inflammation by blocking interleukin-1 in a broad spectrum of diseases. Nat Rev Drug Discov 2012;11:633–52.
24. Silver JS, Hunter CA. Gp130 at the nexus of inflammation, autoimmunity, and cancer [review]. J Leukoc Biol 2010;88:1145–56.
25. Yokota S, Itoh Y, Morio T, Origasa H, Sumitomo N, Tomobe M, et al. Tocilizumab in systemic juvenile idiopathic arthritis in a real-world clinical setting: results from 1 year of postmarketing surveillance follow-up of 417 patients in Japan. Ann Rheum Dis 2016;75:1654–60.
26. Yokota S, Imagawa T, Mori M, Miyamae T, Aihara Y, Takei S, et al. Efficacy and safety of tocilizumab in patients with systemic-onset juvenile idiopathic arthritis: a randomised, double-blind, placebo-controlled, withdrawal phase III trial. Lancet 2008;371:998–1006.
27. Brunner HI, Ruperto N, Zuber Z, Keane C, Harari O, Kenwright A, et al, for the Paediatric Rheumatology International Trials Organisation (PRINTO) and the Pediatric Rheumatology Collaborative Study Group (PRCSG). Efficacy and safety of tocilizumab in patients with polyarticular-course juvenile idiopathic arthritis: results from a phase 3, randomised, double-blind withdrawal trial. Ann Rheum Dis 2015;74:1110–7.
28. Yang RB, Ng CK, Wasserman SM, Colman SD, Shenoy S, Mehraban F, et al. Identification of a novel family of cell-surface proteins expressed in human vascular endothelium. J Biol Chem 2002; 277:46364–73.
29. Capkin AA, Demir S, Mentese A, Bulut C, Ayar A. Can signal peptide-CUB-EGF domain-containing protein (SCUBE) levels be a marker of angiogenesis in patients with psoriasis? Arch Dermatol Res 2017; 309:203–7.
30. Elshabrawy HA, Chen Z, Volin MV, Ravella S, Virupannavar S, Shahrara S. The pathogenic role of angiogenesis in rheumatoid arthritis. Angiogenesis 2015;18:433–48.
31. Cantatore FP, Maruotti N, Corrado A, Ribatti D. Angiogenesis dysregulation in psoriatic arthritis: molecular mechanisms. Biomed Res Int 2017;2017:5312813.
32. Martinez GJ, Zhang Z, Reynolds JM, Tanaka S, Chung Y, Liu T, et al. Smad2 positively regulates the generation of Th17 cells. J Biol Chem 2010;285:29039–43.
33. Malhotra N, Kang J. SMAD regulatory networks construct a balanced immune system. Immunology 2013;139:1–10.
34. Kichaev G, Bhatia G, Loh PR, Gazal S, Burch K, Freund MK, et al. Leveraging polygenic functional enrichment to improve GWAS power. Am J Hum Genet 2019;104:65–75.
35. Maddur MS, Miossec P, Kaveri SV, Bayry J. Th17 cells: biology, pathogenesis of autoimmune and inflammatory diseases, and therapeutic strategies. Am J Pathol 2012;181:8–18.
36. Gabay C, Lamacchia C, Palmer G. IL-1 pathways in inflammation and human diseases. Nat Rev Rheumatol 2010;6:232–41.
37. Jung N, Hellmann M, Hoheisel R, Lehmann C, Haase I, Perniok A, et al. An open-label pilot study of the efficacy and safety of anakinra in patients with psoriatic arthritis refractory to or intolerant of methotrexate (MTX). Clin Rheumatol 2010;29:1169–73.
38. Pelechas E, Voulgari PV, Drosos AA. Clinical evaluation of the safety, efficacy and tolerability of sarilumab in the treatment of moderate to severe rheumatoid arthritis. Ther Clin Risk Manag 2019;15:1073–9.
39. Swidrowska-Jaros J, Smolewska E. A fresh look at angiogenesis in juvenile idiopathic arthritis. Cent Eur J Immunol 2018;43:325–30.
40. Kim Y, Yi H, Jung H, Rim YA, Park N, Kim J, et al. A dual target-directed agent against interleukin-6 receptor and tumor necrosis factor alpha ameliorates experimental arthritis. Sci Rep 2016;6:20150.
41. Cantatore FP, Maruotti N, Corrado A, Ribatti D. Anti-angiogenic effects of biotechnological therapies in rheumatic diseases. Biologics 2017;11:123–8.
42. Schonthaler HB, Huggenberger R, Wculek SK, Detmar M, Wagner EF. Systemic anti-VEGF treatment strongly reduces skin inflammation in a mouse model of psoriasis. Proc Natl Acad Sci U S A 2009;106: 21264–9.