



Integrative Analysis of Transcriptome-Wide Association Study and Gene-Based Association Analysis Identifies In Silico Candidate Genes Associated with Juvenile Idiopathic Arthritis

Shuai Liu^{1,2,†}, Weiming Gong^{1,2,†}, Lu Liu^{1,2}, Ran Yan^{1,2}, Shukang Wang^{1,2,*} and Zhongshang Yuan^{1,2,*}

- ¹ Department of Biostatistics, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan 250012, China
- ² Institute for Medical Dataology, Cheeloo College of Medicine, Shandong University, Jinan 250012, China

Correspondence: wsk2001@sdu.edu.cn (S.W.); yuanzhongshang@sdu.edu.cn (Z.Y.)

+ These authors contributed equally to this work.

Abstract: Genome-wide association study (GWAS) of Juvenile idiopathic arthritis (JIA) suffers from low power due to limited sample size and the interpretation challenge due to most signals located in non-coding regions. Gene-level analysis could alleviate these issues. Using GWAS summary statistics, we performed two typical gene-level analysis of JIA, transcriptome-wide association studies (TWAS) using FUnctional Summary-based ImputatiON (FUSION) and gene-based analysis using eQTL Multi-marker Analysis of GenoMic Annotation (eMAGMA), followed by comprehensive enrichment analysis. Among 33 overlapped significant genes from these two methods, 11 were previously reported, including TYK2 ($P_{\text{FUSION}} = 5.12 \times 10^{-6}$, $P_{\text{eMAGMA}} = 1.94 \times 10^{-7}$ for whole blood), IL-6R $(P_{\text{FUSION}} = 8.63 \times 10^{-7}, P_{\text{eMAGMA}} = 2.74 \times 10^{-6}$ for cells EBV-transformed lymphocytes), and Fas $(P_{\text{FUSION}} = 5.21 \times 10^{-5}, P_{\text{eMAGMA}} = 1.08 \times 10^{-6}$ for muscle skeletal). Some newly plausible JIAassociated genes are also reported, including IL-27 ($P_{FUSION} = 2.10 \times 10^{-7}$, $P_{eMAGMA} = 3.93 \times 10^{-8}$ for Liver), LAT ($P_{\text{FUSION}} = 1.53 \times 10^{-4}$, $P_{\text{eMAGMA}} = 4.62 \times 10^{-7}$ for Artery Aorta), and MAGI3 $(P_{\text{FUSION}} = 1.30 \times 10^{-5}, P_{\text{eMAGMA}} = 1.73 \times 10^{-7}$ for Muscle Skeletal). Enrichment analysis further highlighted 4 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways and 10 Gene Ontology (GO) terms. Our findings can benefit the understanding of genetic determinants and potential therapeutic targets for JIA.

Keywords: juvenile idiopathic arthritis; transcriptome-wide association study; gene-based association analysis; enrichment analysis

1. Introduction

Juvenile idiopathic arthritis (JIA) is one of the most common rheumatic diseases characterized by arthritis in childhood [1]. It can cause damage to multiple organs such as arthrosis, heart, lung, skin, and eyes, and is an important cause of disability in children under the age of 16 [2]. Previous studies have shown that the concordance rate of JIA in monozygotic twins and the relative risk of the disease in the siblings of JIA patients are higher [3,4], highlighting the important role of genetic factors in the development of JIA [5,6]. Therefore, it is of great significance to probe the complex genetic association of JIA to better understand the genetic mechanisms and investigate potential intervention targets.

Genome-wide association studies (GWAS) have successfully identified 22 risk loci associated with JIA [7–9], however, these GWASs suffer from either small sample size or low case proportions, which may be presumably due to the difficulty in accurate JIA diagnosis and the lack of pathognomonic features [10]. So far, the JIA GWAS with relatively larger sample size and the largest case proportions only includes 3305 JIA cases and 9196 controls [9]. In addition, most genetic variants identified from GWAS of JIA are located in non-coding regions [11], leading to the difficulty in explaining the association signals.



Citation: Liu, S.; Gong, W.; Liu, L.; Yan, R.; Wang, S.; Yuan, Z. Integrative Analysis of Transcriptome-Wide Association Study and Gene-Based Association Analysis Identifies In Silico Candidate Genes Associated with Juvenile Idiopathic Arthritis. *Int. J. Mol. Sci.* **2022**, 23, 13555. https:// doi.org/10.3390/ijms232113555

Academic Editors: Giuseppe Novelli and Wajid Zaman

Received: 21 August 2022 Accepted: 2 November 2022 Published: 4 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). On the other hand, statistically, GWAS often provide JIA-associated single nucleotide polymorphisms (SNPs), the effects of which are too weak to be detected. Gene-level analysis can not only aggregate many SNPs with small effects to improve the power, but provide good biological interpretations, which is more straightforward to be translated into clinical practice.

With the increase of publicly available GWAS summary data and the well-developed efficient tools, it is feasible to conduct the gene-level analysis for JIA [12]. There are two typical gene-level association analysis methods with different model assumptions, one is transcriptome-wide association studies (TWAS), which has shown great promise in interpreting the GWAS signals and is powerful in detecting the association between the gene expression level and the complex disease [13,14]. Recently, one TWAS analysis of JIA has been conducted, however, it only involves two tissues with lower JIA case proportions [15], which may lead to the power loss and the insufficiency in capturing the tissue information related with JIA. The other is multi-marker analysis, which can assign SNPs to genes based on physical proximity and further conduct gene-based association analysis. Both methods, though have different statistical principles, can produce the gene-level *p* values. We would like to emphasize that using different gene-level methods with different model statistical principles to obtain the common genes can avoid the risk of false discoveries from using single method. Actually, the results from different analysis could complement to each other [16,17].

In the present study, using the GWAS summary data of JIA with relatively larger sample size and case proportions (3305 cases and 9196 controls), we performed TWAS analysis and the gene-based association analysis to identify the tissue-specific JIA-associated genes. We used the false discovery rate (FDR) correction on each tissue to declare the significant genes. Finally, we overlapped the genes significantly detected from these two gene-level methods, and performed the enrichment analysis for these overlapped genes on the Metascape website to identify the significant Gene Ontology (GO) terms as well as the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways.

2. Methods

2.1. Study Design, Data Source, and Quality Control

The analysis flowchart of this study is shown in Figure 1. We used the GWAS of JIA from a large-scale meta-analysis [9]. We obtained the largest GWAS summary statistics of JIA from the NHGRI-EBI GWAS Catalog (Study Accession Code GCST90010715), where the JIA cases were diagnosed according to The International League Against Rheumatism (ILAR) criteria. The data were restricted to European ancestry with stringent quality control as described previously [9]. Briefly, the GWAS of JIA initially recruited 4520 UK JIA patients and 9965 healthy individuals. Individuals with call rate less than 0.98 and discrepancy between genetically predicted sex and database record were removed. In addition, SNPs that were non-autosomal, had a call rate <0.98 or a minor allele frequency (MAF) <0.01 were further excluded. About 12,501 individuals (3305 cases and 9196 healthy controls) were finally remained. For the summary data, we further excluded the major histocompatibility complex (MHC) region (chromosome 6: 25–35 Mb) due to its complex structure, restricted to biallelic SNPs and removed SNPs with duplicated or missing rs ID for subsequent analyses. Totally 7,415,262 SNPs are finally included. Bearing in mind that using different methods with different model assumptions to obtain the overlapped genes can avoid the risk of false discoveries from using single method, we here applied two gene-level approaches with distinct principles, TWAS analysis and gene-based association analysis, as parallel analyses to obtain the common JIA-associated genes for enrichment analysis.



Figure 1. The flowchart of integrative analysis of FUSION and eMAGMA. ABBR: JIA, Juvenile idiopathic arthritis; GWAS, genome-wide association study; MHC, major histocompatibility complex; GTEx, Genotype-Tissue Expression Project; LD, linkage disequilibrium; EUR, European; FUSION, functional summary-based imputation; eQTL, expression quantitative trait loci; eMAGMA, eQTL Multi-marker Analysis of GenoMic Annotation; FDR, false discovery rate; KEGG, Kyoto Encyclopedia of Genes and Genomes; GO, Gene Ontology.

2.2. TWAS Analysis

TWAS aim to integrate GWAS and expression quantitative trait loci (eQTL) studies to identify tissue-specific gene-trait associations [18], which has shown great promise both in interpreting GWAS findings and in elucidating the underlying disease mechanisms. We here adopted FUnctional Summary-based ImputatiON (FUSION) method (http://gusevlab.org/projects/fusion/, accessed on 8 April 2022) to conducted TWAS analysis of JIA. FUSION is the most commonly used TWAS method and has shown great promise in large-scale integrative omics data analysis [14]. Once inputting the GWAS summary data and expression weight, FUSION will impute the gene expressions in GWAS, and then perform an association analysis between the predicted gene expression and JIA. We selected Genotype-Tissue Expression Project (GTEx) v8 with pre-computed gene expression weights from totally 49 tissues. Using all the tissues may introduce the nuisance information and increase the computation burden; previous studies recommend using an expression panel from only plausible disease-related tissues in TWAS [13]. Here, we determine the analyzed tissue based on not only previous studies [2,19], but also the clinical symptoms and involved organs of JIA, such as hepatomegaly, splenomegaly, anemia, disseminated intravascular coagulation, arthritis, rash, mesenteric lymphadenopathy, pericarditis, and pneumonia. We finally selected 18 tissues for analysis including artery aorta, artery coronary, artery tibial, cells EBV-transformed lymphocytes, colon sigmoid, colon transverse, heart atrial appendage, heart left ventricle, liver, lung, muscle skeletal, nerve tibial, skin not sun exposed suprapubic, skin sun exposed lower leg, small intestine terminal ileum, spleen, stomach, whole blood. We used the 1000 Genomes European panel as a linkage disequilibrium (LD) reference data, and obtained tissue-specific *p* value for each gene across different tissues. We finally performed FDR correction on each tissue, and genes with FDR less than 0.05 are declared to be significant.

2.3. Gene-Based Association Analysis

We used eQTL Multi-marker Analysis of GenoMic Annotation (eMAGMA) method (https://github.com/eskederks/eMAGMA-tutorial, accessed on 28 March 2022) to conduct gene-based association analysis for JIA. eMAGMA follows the same statistical framework as MAGMA, which is based on a multiple linear principal component regression model and can provide better statistical performance in gene-based association analysis [16]. Besides this, eMAGMA can further utilize tissue-specific cis-eQTL information to assign SNPs to putative genes, providing more biologically meaningful and interpretable results [17]. Gene-based analysis using eMAGMA typically involves two stages, annotation and gene analysis [16]. We, in the annotation stage, used the same tissues as that in above TWAS analysis and directly used the GTEx v8-based annotation files provided on eMAGMA's website. In the gene analysis stage, we used the 1000 Genomes European panel as the reference panel and tested the association between the annotated genes and JIA. We further performed FDR correction on each tissue, and selected significant genes with FDR less than 0.05.

2.4. Gene Set Enrichment Analysis

We conducted gene set enrichment analysis for the overlapped genes that were significantly identified by both TWAS and eMAGMA analysis. Specifically, these overlapped significant genes were subjected to GO term and KEGG pathway enrichment analysis on the Metascape website (https://metascape.org/gp/index.html#/main/step1, accessed on 26 April 2022) to better understand the biological mechanisms. Metascape essentially utilizes the hypergeometric test and Benjamini-Hochberg *p* value correction algorithm to identify all ontology terms. A large number of terms would make the results redundant and complicate the interpretation, Kappa consistency test was thus performed and terms with Kappa > 0.3 were grouped into a cluster, and the most statistically significant term in the cluster was selected to represent the cluster [20]. The parameters of Min Overlap, *p* Value Cutoff, and Min Enrichment are set to be the default values, respectively. In addition, we also made a protein–protein interaction (PPI) network for the overlapped significant genes on the STRING website (https://cn.string-db.org/, accessed on 15 September 2022).

3. Results

3.1. TWAS Analysis

We analyzed all genes involved in the 18 tissues, among which 275 tissue-specific genes were significantly detected by FUSION with FDR less than 0.05. Note that LAT is significantly detected at the border line ($p = 1.53 \times 10^{-4}$ and FDR = 5.40×10^{-2} for Artery Aorta). These TWAS significant genes included some established JIA-associated genes that have been reported previously, such as CCDC101 ($p = 5.82 \times 10^{-8}$ and FDR = 1.64×10^{-4} for Muscle Skeletal), CLN3 ($p = 5.82 \times 10^{-8}$ and FDR = 2.53×10^{-4} for Whole Blood), ERAP2 ($p = 5.49 \times 10^{-6}$ and FDR = 2.16×10^{-3} for Cells EBV-transformed lymphocytes), LNPEP ($p = 3.53 \times 10^{-6}$ and FDR = 2.36×10^{-3} for Artery Tibial). The consistent results with previous GWAS partly indicates the correctness of the FUSION analysis. All significant genes in FUSION analysis were displayed in Supplementary Table S1.

3.2. Gene-Based Association Analysis Results

Similarly, we analyzed all genes involved in the 18 tissues, among which 380 tissuespecific genes were significantly detected by eMAGMA with FDR less than 0.05. These included some well-known JIA-associated genes, such as SGF29 ($p = 2.22 \times 10^{-8}$ and FDR = 3.38×10^{-5} for Whole Blood), ANKRD55 ($p = 3.09 \times 10^{-8}$ and FDR = 1.08×10^{-4} for Spleen), ATP8B2 ($p = 7.53 \times 10^{-7}$ and FDR = 6.59×10^{-4} for Stomach), PTPN2 ($p = 3.96 \times 10^{-6}$ and FDR = 2.75×10^{-3} for Spleen). Again, all these genes are included in the 22 risk loci identified by previous GWAS illustrate the eMAGMA results more reliable. All significant genes from eMAGMA analysis were summarized in Supplementary Table S2.

3.3. Gene Set Enrichment Analysis

We intersected the significant genes from both TWAS analysis and eMAGMA genebased association analysis according to different tissues, and found a total of 132 tissuespecific genes. A total of 33 unique genes were further identified after removing the duplicated ones, among which 11 genes have been reported in previous studies [7–9], such as TYK2 ($P_{FUSION} = 5.12 \times 10^{-6}$, $P_{eMAGMA} = 1.94 \times 10^{-7}$ for Whole Blood), IL (Interleukin)-6R ($P_{FUSION} = 8.63 \times 10^{-7}$, $P_{eMAGMA} = 2.74 \times 10^{-6}$ for Cells EBV-transformed lymphocytes), and Fas ($P_{FUSION} = 5.21 \times 10^{-5}$, $P_{eMAGMA} = 1.08 \times 10^{-6}$ for Muscle Skeletal). The remaining newly discovered 22 genes are APOBR, ATXN2L, GSDMB, IKZF3, IL27, KIAA1109, LAT, LMAN2L, MAGI3, NFATC2IP, NUPR1, ORMDL3, PHTF1, PSMB7, RGS14, SBK1, SH2B1, STEAP1B, TNFSF15, TUFM, UXS1, ZNF197. Among them, IL-27 ($P_{FUSION} = 2.10 \times 10^{-7}$, $P_{eMAGMA} = 3.93 \times 10^{-8}$ for Liver), LAT ($P_{FUSION} = 1.53 \times 10^{-4}$, $P_{eMAGMA} = 4.62 \times 10^{-7}$ for Artery Aorta) and MAGI3 ($P_{FUSION} = 1.30 \times 10^{-5}$, $P_{eMAGMA} = 1.73 \times 10^{-7}$ for Muscle Skeletal) are the novel genes that are more likely to be associated with JIA. Detailed information for 33 genes were summarized in Table 1.

Gene Symbol	Chromosome	Gene Start (bp)	Gene End (bp)	P _{FUSION}	FDR _{FUSION}	P _{eMAGMA}	FDR _{eMAGMA}
IL6R	1	154,377,669	154,441,926	8.63×10^{-7}	$7.80 imes 10^{-4}$	$2.74 imes 10^{-6}$	1.56×10^{-3}
MAGI3	1	113,933,371	114,228,545	$1.30 imes 10^{-5}$	$7.35 imes 10^{-3}$	$1.73 imes10^{-7}$	$2.62 imes 10^{-4}$
PHTF1	1	114,239,453	114,302,111	$6.55 imes10^{-7}$	$1.21 imes 10^{-3}$	$2.93 imes10^{-7}$	$5.60 imes10^{-4}$
LMAN2L	2	97,371,666	97,405,801	$4.16 imes10^{-5}$	$1.69 imes10^{-2}$	$1.56 imes10^{-5}$	$6.48 imes 10^{-3}$
UXS1	2	106,709,759	106,810,795	$1.39 imes10^{-4}$	$4.78 imes10^{-2}$	$1.39 imes10^{-4}$	$3.52 imes 10^{-2}$
ZNF197	3	44,626,380	44,689,963	$7.92 imes10^{-5}$	$3.28 imes 10^{-2}$	$2.04 imes10^{-5}$	$1.23 imes10^{-2}$
KIAA1109	4	123,073,488	123,283,913	$1.51 imes 10^{-5}$	$9.73 imes10^{-3}$	$2.06 imes 10^{-5}$	$1.44 imes 10^{-2}$
ERAP2	5	96,211,643	96,255,420	$5.49 imes10^{-6}$	$2.16 imes10^{-3}$	$1.03 imes10^{-6}$	$8.68 imes10^{-4}$
LNPEP	5	96,271,098	96,373,219	$3.53 imes10^{-6}$	$2.36 imes10^{-3}$	$1.50 imes10^{-6}$	$9.68 imes10^{-4}$
RGS14	5	176,784,838	176,799,602	$1.02 imes 10^{-6}$	$1.97 imes10^{-3}$	$1.07 imes10^{-5}$	$9.71 imes10^{-3}$
AHI1	6	135,604,670	135,818,914	$1.98 imes 10^{-6}$	$2.61 imes 10^{-3}$	$3.31 imes10^{-5}$	$1.10 imes 10^{-2}$
STEAP1B	7	22,459,063	22,672,544	$1.23 imes10^{-5}$	$5.29 imes10^{-3}$	$2.86 imes10^{-5}$	$2.07 imes10^{-2}$
PSMB7	9	127,115,745	127,177,723	$4.84 imes10^{-5}$	$2.28 imes 10^{-2}$	$1.10 imes10^{-4}$	$3.84 imes10^{-2}$
TNFSF15	9	117,546,915	117,568,406	$4.03 imes10^{-5}$	$1.60 imes 10^{-2}$	$4.09 imes10^{-5}$	$1.61 imes 10^{-2}$
ACTA2	10	90,694,831	90,751,147	$5.64 imes10^{-6}$	$4.05 imes10^{-3}$	$1.03 imes10^{-6}$	$7.17 imes10^{-4}$
FAS	10	90,750,414	90,775,542	$5.21 imes 10^{-5}$	$2.32 imes 10^{-2}$	$1.08 imes10^{-6}$	$8.93 imes10^{-4}$
APOBR	16	28,505,970	28,510,291	$6.52 imes10^{-7}$	$8.61 imes10^{-4}$	$1.25 imes 10^{-6}$	$9.48 imes10^{-4}$
ATP2A1	16	28,889,726	28,915,830	$5.31 imes10^{-7}$	$1.18 imes 10^{-3}$	$3.66 imes10^{-7}$	$5.87 imes10^{-4}$
ATXN2L	16	28,834,356	28,848,558	$8.15 imes10^{-6}$	$5.37 imes10^{-3}$	$1.03 imes10^{-6}$	$7.17 imes10^{-4}$
CLN3	16	28,477,983	28,506,896	$5.82 imes10^{-8}$	$2.53 imes10^{-4}$	$2.43 imes10^{-8}$	$3.38 imes10^{-5}$
IL27	16	28,510,683	28,523,372	$2.10 imes10^{-7}$	$4.10 imes10^{-4}$	$3.93 imes10^{-8}$	$1.30 imes10^{-4}$
LAT	16	28,996,147	29,002,104	$1.53 imes10^{-4}$	$5.40 imes10^{-2}$	$4.62 imes10^{-7}$	$6.13 imes10^{-4}$
NFATC2IP	16	28,962,128	28,978,418	$1.93 imes10^{-8}$	$9.20 imes10^{-5}$	$8.07 imes10^{-8}$	$1.84 imes10^{-4}$
NUPR1	16	28,548,606	28,550,495	$6.39 imes10^{-8}$	$3.04 imes10^{-4}$	$1.94 imes10^{-8}$	$4.69 imes10^{-5}$
SBK1	16	28,303,840	28,335,170	$2.55 imes 10^{-7}$	$4.39 imes10^{-4}$	$7.70 imes10^{-7}$	$7.45 imes 10^{-4}$

Gene Symbol	Chromosome	Gene Start (bp)	Gene End (bp)	P _{FUSION}	FDR _{FUSION}	P _{eMAGMA}	FDR _{eMAGMA}
SH2B1	16	28,857,921	28,885,526	$5.91 imes 10^{-9}$	$5.13 imes 10^{-5}$	$7.76 imes 10^{-7}$	$1.10 imes 10^{-3}$
SULT1A1	16	28,616,903	28,634,946	$1.54 imes10^{-9}$	$8.17 imes10^{-6}$	$8.27 imes10^{-8}$	$1.25 imes10^{-4}$
SULT1A2	16	28,603,264	28,608,430	$7.19 imes10^{-11}$	$7.63 imes10^{-7}$	$3.63 imes10^{-8}$	$7.79 imes10^{-5}$
TUFM	16	28,853,732	28,857,729	$2.12 imes10^{-6}$	$2.05 imes 10^{-3}$	$4.10 imes10^{-7}$	$7.41 imes10^{-4}$
GSDMB	17	38,060,848	38,076,107	$6.69 imes10^{-5}$	$2.64 imes10^{-2}$	$2.22 imes 10^{-5}$	$1.16 imes10^{-2}$
IKZF3	17	37,921,198	38,020,441	$4.01 imes10^{-5}$	$1.69 imes10^{-2}$	$2.70 imes10^{-5}$	$9.35 imes 10^{-3}$
ORMDL3	17	38,077,294	38,083,854	$4.99 imes10^{-5}$	$1.38 imes 10^{-2}$	$1.64 imes10^{-5}$	$6.66 imes 10^{-3}$
TYK2	19	10,461,209	10,491,352	$5.12 imes10^{-6}$	$4.05 imes10^{-3}$	$1.94 imes10^{-7}$	$2.01 imes10^{-4}$

Table 1. Cont.

22 novel gene were shown in bold. Gene position (start and end) are based on GRCh37/hg19 by Ensembl.

We further performed KEGG and GO enrichment analysis on the overlapped 33 significant genes on the Metascape website, respectively. For KEGG enrichment analysis, we totally found four significant KEGG pathways (Figure 2), including Th17 cell differentiation $(p = 5.83 \times 10^{-6})$, cytokine-cytokine receptor interaction $(p = 2.92 \times 10^{-4})$, spinocerebellar ataxia ($p = 5.11 \times 10^{-4}$), and Rap1 signaling pathway ($p = 1.55 \times 10^{-3}$). For GO enrichment analysis, we totally found ten significant GO terms (Figure 2), including signaling receptor complex adaptor activity ($p = 2.25 \times 10^{-5}$), apoptotic signaling pathway ($p = 2.09 \times 10^{-4}$), cytokine receptor binding ($p = 2.14 \times 10^{-4}$), regulation of blood pressure ($p = 1.16 \times 10^{-3}$), positive regulation of protein phosphorylation ($p = 1.18 \times 10^{-3}$), inflammatory response $(p = 2.44 \times 10^{-3})$, steroid metabolic process $(p = 2.66 \times 10^{-3})$, regulation of autophagy $(p = 5.96 \times 10^{-3})$, protein homodimerization activity $(p = 6.40 \times 10^{-3})$, and regulation of MAPK cascade ($p = 6.47 \times 10^{-3}$). In addition, chord graphs (Figure 3) were also depicted to show the most significant enrichment pathways of KEGG and GO and visualize the targeting relationship between significant genes and significant pathways, so as to visualize which pathways every gene is enriched to, and which genes are enriched in each pathway. PPI network (Figure 4) shows the interaction between proteins.







Figure 3. Chord graphs for four significant KEGG pathways and ten significant GO terms. (a) Chord graph of four significant KEGG pathways. (b) Chord graph of ten significant GO terms. For each panel, the right semicircle represented significant pathways or terms, and the left semicircle represented the genes enriched in these pathways or terms.



Figure 4. Protein–protein interaction (PPI) network for 33 overlapped genes by STRING. Each circle represents a protein, a line between proteins indicates PPI, line thickness indicates the strength of data support.

4. Discussion

In the present study, we performed a comprehensive large-scale gene-level analysis using co-complementary methods and successfully detected 33 common genes, including 11 previously reported genes such as TYK2, IL-6R, and Fas [7–9], and 22 novel potential genes such as IL-27, LAT, and MAGI3. Enrichment analysis suggested important role of pathways involving Th17 cell differentiation and Rap1 signaling pathway, followed by PPI network illustrating the protein–protein interactions. All these findings provide novel insights into the potential molecular mechanisms underlying the development of JIA.

TYK2, IL-6R, and Fas appear more frequently in the significant KEGG pathways and Go terms, and the expression products of these three genes have been shown to play important roles in the inflammatory and immune responses of JIA [21–23]. Tyrosine kinase 2 encoded by gene TYK2 is a part of janus kinase (JAK), which mediates the activation of signal transducers and activators of transcription (STAT) proteins. That is, TYK2 may play a role in autoimmunity and inflammation through abnormal expression in JAK-STAT pathway, thus leading to JIA [21,24,25]. IL-6R encodes part of the interleukin 6 receptor, as a pro-inflammatory cytokine, IL-6 is significantly elevated in the serum of JIA patients. Inhibition of IL-6R expression reduces IL-6 and IL-6 receptor binding, thereby reducing inflammation and immune responses in JIA patients [22,26]. Fas can induce T-cell apoptosis by binding to the Fas Ligand (FasL), so decreased gene Fas expression may lead to the accumulation of activated T-cells and cause autoimmune diseases [27].

IL-27 is involved in encoding the synthesis of IL-27, a cytokine that plays a role in innate immunity and whose primary function is to promote pro-inflammatory Th1 differentiation and inhibit anti-inflammatory Th2 responses [28–32]. IL-27 promotes Th1 cell differentiation, which in turn produces a large amount of the proinflammatory cytokine interferon- γ (IFN- γ) to play a pathogenic role in JIA [32,33]. The promotion of Th1 differentiation and the inhibition of Th2 differentiation by IL-27 is also dependent on the action of STAT1 [28,31–33], that is, the role of IL-27 is also involved in the JAK-STAT signaling pathway. Therefore, if the expression of gene IL-27 is inhibited, the inflammatory response of JIA patients can be correspondingly reduced.

LAT encodes a protein called T-cell activation adaptor [34]. T-cell receptor (TCR) signaling is an important process in T-cell development and its activation in the periphery [35]. LAT is a key signaling hub connecting TCRs to trigger downstream T-cell responses. If LAT gene expression is reduced, peripheral T-cell development and numbers are inhibited. Decreased numbers of T cells are prone to lead to immunodeficiency and autoimmune diseases [36], and patients with JIA are likely to have decreased autoimmune function due to lack of LAT.

MAGI3 encodes Membrane-associated guanylate kinase, WW and PDZ domaincontaining protein 3. Abnormal expression of MAGI3 may affect Notch signaling and thus affect bone and joint development in children [37]. In addition, MAGI3 is also a risk gene for rheumatoid arthritis (RA), Graves' disease, and other autoimmune diseases, indicating it can also cause JIA by affecting the human immune system [38,39].

Enrichment analysis suggested important role of pathways involving Th17 cell differentiation and Rap1 signaling pathway. The Th17 cell differentiation pathway is involved in inflammation and bone destruction, IL-27, IL-6R, LAT, TYK2 are all on this pathway. Th17 cells are actively differentiated and mainly secrete IL-17, which not only promotes the production of inflammatory cytokines in the JAK-STAT signaling pathway, but also catalyzes the maturation of osteoclasts, leading to osteopenia and joint damage [40,41]. The activation of inflammatory cytokines and osteopenia together lead to arthritis, and so inhibiting the differentiation of Th17 cells may inhibit and treat JIA to a certain extent [42–44].

The Rap1 signaling pathway plays a central role in the functional outcome of T-cell stimulation [45]. Rap1 is a T-cell receptor proximal signaling protein, and its abnormal expression may lead to abnormal T cells [46]. T cells activate macrophages and synovial stromal cells pleiotropically through cell-to-cell contact and interleukin production, leading to synovitis and joint destruction in RA [47,48]. The pathogenic behavior of the above T

cells is caused by Rap1 inactivation, and sustained Rap1 signaling in T cells can effectively reduce the incidence and severity of arthritis [49,50]. Therefore, activation and enhancement of Rap1 signaling also contribute to the prevention and treatment of JIA.

Our study is not without limitations. First, we only focused on European ancestry due to the current large-scale GWASs of JIA was only available for European population. The findings cannot be directly generalized to other ethnic population. Second, the results from data analysis are often less reliable than that from serious and high-cost experiments, which is likely to be the gold standard in biomedicine studies. However, the data analysis is still valuable. For example, it is often hard to pre-specify the experimental target under a hypothesis-free approach, data analysis can help to narrow down the candidate experimental target list and provide the evidence of target priority. In addition, the current analysis pipeline can be easily extended, an alternative way is to search for restriction endonuclease (RE) sites in the non-coding regions and gain insights through RE digestion patterns [51,52]. Third, our findings are obtained from a joint analysis of all subtypes of JIA, and there is no guarantee that the conclusions would be valid for any subtype. Finally, the results must necessarily be confirmed by experiments in the laboratory, given that all the analysis are essentially in silico.

5. Conclusions

In summary, we performed gene-level analysis as well as enrichment analysis on the largest GWAS summary data of JIA. We identified novel JIA-associated genes including IL-27, LAT, and MAGI3, and highlight the important role of Th17 cell differentiation, Rap1 signaling pathway in the development of JIA. Our results can provide new insights into the pathogenic mechanisms as well as potential therapeutic targets of JIA; however, further studies are still required to validate these findings.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms232113555/s1.

Author Contributions: Conceptualization, Z.Y. and S.W.; methodology, W.G. and S.L.; software, W.G. and S.L.; formal analysis, W.G. and S.L.; writing—original draft preparation, S.L.; writing—review and editing, Z.Y., S.W., W.G., S.L., L.L., and R.Y.; supervision, Z.Y. and S.W.; funding acquisition, Z.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (No. 81872712 and No. 82173624), the Natural Science Foundation of Shandong Province (No. ZR2019ZD02), the Cheeloo Young Talent Program of Shandong University, and the Taishan Scholar Project of Shandong Province, all awarded to Z.Y.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The GWAS summary data analyzed during the current study are available in the NHGRI-EBI GWAS Catalog (https://www.ebi.ac.uk/gwas/downloads/summary-statistics, accessed on 9 November 2021) (Study Accession Code GCST90010715). The FUSION software and GTEx v8 gene expression datasets used during the current study are available in the FUSION website (http://gusevlab.org/projects/fusion/, accessed on 8 April 2022). The eMAGMA software and GTEx v8 annotion datasets used during the current study are available in the eMAGMA website (https://github.com/eskederks/eMAGMA-tutorial, accessed on 28 March 2022).

Acknowledgments: We would like to thank GWAS Catalog for providing us with the JIA GWAS summary data.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Cimaz, R. Systemic-onset juvenile idiopathic arthritis. *Autoimmun. Rev.* **2016**, *15*, 931–934. [CrossRef] [PubMed]
- 2. Ravelli, A.; Martini, A. Juvenile idiopathic arthritis. *Lancet* 2007, 369, 767–778. [CrossRef]

- 3. Giancane, G.; Alongi, A.; Ravelli, A. Update on the pathogenesis and treatment of juvenile idiopathic arthritis. *Curr. Opin. Rheumatol.* **2017**, *29*, 523–529. [CrossRef] [PubMed]
- Cobb, J.E.; Hinks, A.; Thomson, W. The genetics of juvenile idiopathic arthritis: Current understanding and future prospects. *Rheumatology* 2014, 53, 592–599. [CrossRef] [PubMed]
- Hersh, A.O.; Prahalad, S. Immunogenetics of juvenile idiopathic arthritis: A comprehensive review. J. Autoimmun. 2015, 64, 113–124. [CrossRef]
- 6. Moncrieffe, H.; Prahalad, S.; Thompson, S.D. Genetics of juvenile idiopathic arthritis: New tools bring new approaches. *Curr. Opin. Rheumatol.* **2014**, *26*, 579–584. [CrossRef]
- Thompson, S.D.; Sudman, M.; Ramos, P.S.; Marion, M.C.; Ryan, M.; Tsoras, M.; Weiler, T.; Wagner, M.; Keddache, M.; Haas, J.P.; 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. [CrossRef]
- Hinks, A.; Cobb, J.; Marion, M.C.; Prahalad, S.; Sudman, M.; Bowes, J.; Martin, P.; Comeau, M.E.; Sajuthi, S.; Andrews, R.; et al. Dense genotyping of immune-related disease regions identifies 14 new susceptibility loci for juvenile idiopathic arthritis. *Nat. Genet.* 2013, 45, 664–669. [CrossRef]
- Lopez-Isac, E.; Smith, S.L.; Marion, M.C.; Wood, A.; Sudman, M.; Yarwood, A.; Shi, C.; Gaddi, V.P.; Martin, P.; Prahalad, S.; et al. Combined genetic analysis of juvenile idiopathic arthritis clinical subtypes identifies novel risk loci, target genes and key regulatory mechanisms. *Ann. Rheum. Dis.* 2021, *80*, 321–328. [CrossRef]
- 10. Giancane, G.; Consolaro, A.; Lanni, S.; Davi, S.; Schiappapietra, B.; Ravelli, A. Juvenile Idiopathic Arthritis: Diagnosis and Treatment. *Rheumatol. Ther.* **2016**, *3*, 187–207. [CrossRef]
- 11. Tak, Y.G.; Farnham, P.J. Making sense of GWAS: Using epigenomics and genome engineering to understand the functional relevance of SNPs in non-coding regions of the human genome. *Epigenetics Chromatin* **2015**, *8*, 57. [CrossRef]
- Li, B.; Ritchie, M.D. From GWAS to Gene: Transcriptome-Wide Association Studies and Other Methods to Functionally Understand GWAS Discoveries. *Front. Genet.* 2021, 12, 713230. [CrossRef]
- Wainberg, M.; Sinnott-Armstrong, N.; Mancuso, N.; Barbeira, A.N.; Knowles, D.A.; Golan, D.; Ermel, R.; Ruusalepp, A.; Quertermous, T.; Hao, K.; et al. Opportunities and challenges for transcriptome-wide association studies. *Nat. Genet.* 2019, 51, 592–599. [CrossRef]
- 14. Gusev, A.; Ko, A.; Shi, H.; Bhatia, G.; Chung, W.; Penninx, B.W.; Jansen, R.; de Geus, E.J.; Boomsma, D.I.; Wright, F.A.; et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* **2016**, *48*, 245–252. [CrossRef]
- Xu, J.; Ma, J.; Zeng, Y.; Si, H.; Wu, Y.; Zhang, S.; Shen, B. A Cross-Tissue Transcriptome-Wide Association Study Identifies Novel Susceptibility Genes for Juvenile Idiopathic Arthritis in Asia and Europe. *Front. Immunol.* 2022, 13, 941398. [CrossRef]
- De Leeuw, C.A.; Mooij, J.M.; Heskes, T.; Posthuma, D. MAGMA: Generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* 2015, 11, e1004219. [CrossRef]
- 17. Gerring, Z.F.; Mina-Vargas, A.; Gamazon, E.R.; Derks, E.M. E-MAGMA: An eQTL-informed method to identify risk genes using genome-wide association study summary statistics. *Bioinformatics* **2021**, *37*, 2245–2249. [CrossRef]
- 18. Zhu, H.; Zhou, X. Transcriptome-wide association studies: A view from Mendelian randomization. Quant. Biol. 2020. [CrossRef]
- Hahn, Y.S.; Kim, J.G. Pathogenesis and clinical manifestations of juvenile rheumatoid arthritis. *Korean J. Pediatr.* 2010, 53, 921–930. [CrossRef]
- Zhou, Y.; Zhou, B.; Pache, L.; Chang, M.; Khodabakhshi, A.H.; Tanaseichuk, O.; Benner, C.; Chanda, S.K. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* 2019, 10, 1523. [CrossRef]
- Tao, J.H.; Zou, Y.F.; Feng, X.L.; Li, J.; Wang, F.; Pan, F.M.; Ye, D.Q. Meta-analysis of TYK2 gene polymorphisms association with susceptibility to autoimmune and inflammatory diseases. *Mol. Biol. Rep.* 2011, *38*, 4663–4672. [CrossRef] [PubMed]
- 22. Yokota, S. Interleukin 6 as a therapeutic target in systemic-onset juvenile idiopathic arthritis. *Curr. Opin. Rheumatol.* 2003, 15, 581–586. [CrossRef]
- 23. Van Parijs, L.; Abbas, A.K. Role of Fas-mediated cell death in the regulation of immune responses. *Curr. Opin. Immunol.* **1996**, *8*, 355–361. [CrossRef]
- Kyogoku, C.; Tsuchiya, N. A compass that points to lupus: Genetic studies on type I interferon pathway. *Genes Immun.* 2007, 8, 445–455. [CrossRef]
- 25. Richter, M.F.; Dumenil, G.; Uze, G.; Fellous, M.; Pellegrini, S. Specific contribution of Tyk2 JH regions to the binding and the expression of the interferon alpha/beta receptor component IFNAR1. *J. Biol. Chem.* **1998**, 273, 24723–24729. [CrossRef]
- 26. Akioka, S. Interleukin-6 in juvenile idiopathic arthritis. Mod. Rheumatol. 2019, 29, 275–286. [CrossRef]
- Tadaki, H.; Saitsu, H.; Kanegane, H.; Miyake, N.; Imagawa, T.; Kikuchi, M.; Hara, R.; Kaneko, U.; Kishi, T.; Miyamae, T.; et al. Exonic deletion of CASP10 in a patient presenting with systemic juvenile idiopathic arthritis, but not with autoimmune lymphoproliferative syndrome type IIa. *Int. J. Immunogenet.* 2011, *38*, 287–293. [CrossRef]
- 28. Yoshida, H.; Hunter, C.A. The immunobiology of interleukin-27. Annu. Rev. Immunol. 2015, 33, 417–443. [CrossRef]
- 29. Meka, R.R.; Venkatesha, S.H.; Dudics, S.; Acharya, B.; Moudgil, K.D. IL-27-induced modulation of autoimmunity and its therapeutic potential. *Autoimmun. Rev.* 2015, *14*, 1131–1141. [CrossRef]
- 30. Iwasaki, Y.; Fujio, K.; Okamura, T.; Yamamoto, K. Interleukin-27 in T cell immunity. Int. J. Mol. Sci. 2015, 16, 2851–2863. [CrossRef]
- Aparicio-Siegmund, S.; Garbers, C. The biology of interleukin-27 reveals unique pro- and anti-inflammatory functions in immunity. *Cytokine Growth Factor Rev.* 2015, 26, 579–586. [CrossRef]

- 32. Raphael, I.; Nalawade, S.; Eagar, T.N.; Forsthuber, T.G. T cell subsets and their signature cytokines in autoimmune and inflammatory diseases. *Cytokine* **2015**, *74*, 5–17. [CrossRef]
- Takeda, A.; Hamano, S.; Yamanaka, A.; Hanada, T.; Ishibashi, T.; Mak, T.W.; Yoshimura, A.; Yoshida, H. Cutting edge: Role of IL-27/WSX-1 signaling for induction of T-bet through activation of STAT1 during initial Th1 commitment. *J. Immunol.* 2003, 170, 4886–4890. [CrossRef]
- Zhang, W.; Sloan-Lancaster, J.; Kitchen, J.; Trible, R.P.; Samelson, L.E. LAT: The ZAP-70 tyrosine kinase substrate that links T cell receptor to cellular activation. *Cell* 1998, 92, 83–92. [CrossRef]
- Bacchelli, C.; Moretti, F.A.; Carmo, M.; Adams, S.; Stanescu, H.C.; Pearce, K.; Madkaikar, M.; Gilmour, K.C.; Nicholas, A.K.; Woods, C.G.; et al. Mutations in linker for activation of T cells (LAT) lead to a novel form of severe combined immunodeficiency. J. Allergy Clin. Immunol. 2017, 139, 634–642.e635. [CrossRef]
- Keller, B.; Zaidman, I.; Yousefi, O.S.; Hershkovitz, D.; Stein, J.; Unger, S.; Schachtrup, K.; Sigvardsson, M.; Kuperman, A.A.; Shaag, A.; et al. Early onset combined immunodeficiency and autoimmunity in patients with loss-of-function mutation in LAT. *J. Exp. Med.* 2016, *213*, 1185–1199. [CrossRef]
- Zuo, C.; Huang, Y.; Bajis, R.; Sahih, M.; Li, Y.P.; Dai, K.; Zhang, X. Osteoblastogenesis regulation signals in bone remodeling. Osteoporos. Int. 2012, 23, 1653–1663. [CrossRef]
- Liu, Y.Q.; Liu, Y.; Zhang, Q.; Xiao, T.; Deng, H.W. Identification of Novel Pleiotropic SNPs Associated with Osteoporosis and Rheumatoid Arthritis. *Calcif. Tissue Int.* 2021, 109, 17–31. [CrossRef]
- Kus, A.; Szymanski, K.; Peeters, R.P.; Miskiewicz, P.; Porcu, E.; Pistis, G.; Sanna, S.; Naitza, S.; Płoski, R.; Medici, M.; et al. The association of thyroid peroxidase antibody risk loci with susceptibility to and phenotype of Graves' disease. *Clin. Endocrinol.* 2015, *83*, 556–562. [CrossRef]
- Kotake, S.; Udagawa, N.; Takahashi, N.; Matsuzaki, K.; Itoh, K.; Ishiyama, S.; Saito, S.; Inoue, K.; Kamatani, N.; Gillespie, M.T.; et al. IL-17 in synovial fluids from patients with rheumatoid arthritis is a potent stimulator of osteoclastogenesis. *J. Clin. Investig.* 1999, 103, 1345–1352. [CrossRef]
- Hueber, A.J.; Asquith, D.L.; Miller, A.M.; Reilly, J.; Kerr, S.; Leipe, J.; Melendez, A.J.; McInnes, I.B. Mast cells express IL-17A in rheumatoid arthritis synovium. *J. Immunol.* 2010, 184, 3336–3340. [CrossRef]
- 42. Okamoto, K.; Takayanagi, H. Osteoclasts in arthritis and Th17 cell development. *Int. Immunopharmacol.* **2011**, *11*, 543–548. [CrossRef]
- 43. Schett, G.; Teitelbaum, S.L. Osteoclasts and Arthritis. J. Bone Miner. Res. 2009, 24, 1142–1146. [CrossRef]
- 44. Sato, K.; Takayanagi, H. Osteoclasts, rheumatoid arthritis, and osteoimmunology. *Curr. Opin. Rheumatol.* **2006**, *18*, 419–426. [CrossRef]
- 45. Remans, P.H.; Wijbrandts, C.A.; Sanders, M.E.; Toes, R.E.; Breedveld, F.C.; Tak, P.P.; van Laar, J.M.; Reedquist, K.A. CTLA-4IG suppresses reactive oxygen species by preventing synovial adherent cell-induced inactivation of Rap1, a Ras family GTPASE mediator of oxidative stress in rheumatoid arthritis T cells. *Arthritis Rheum.* 2006, 54, 3135–3143. [CrossRef]
- 46. Thomas, R.; Turner, M.; Cope, A.P. High avidity autoreactive T cells with a low signalling capacity through the T-cell receptor: Central to rheumatoid arthritis pathogenesis? *Arthritis Res. Ther.* **2008**, *10*, 210. [CrossRef]
- 47. Sakaguchi, S.; Benham, H.; Cope, A.P.; Thomas, R. T-cell receptor signaling and the pathogenesis of autoimmune arthritis: Insights from mouse and man. *Immunol. Cell Biol.* **2012**, *90*, 277–287. [CrossRef]
- Zhernakova, A.; Stahl, E.A.; Trynka, G.; Raychaudhuri, S.; Festen, E.A.; Franke, L.; Westra, H.J.; Fehrmann, R.S.; Kurreeman, F.A.; Thomson, B.; et al. Meta-analysis of genome-wide association studies in celiac disease and rheumatoid arthritis identifies fourteen non-HLA shared loci. *PLoS Genet.* 2011, 7, e1002004. [CrossRef]
- Abreu, J.R.; Krausz, S.; Dontje, W.; Grabiec, A.M.; de Launay, D.; Nolte, M.A.; Tak, P.P.; Reedquist, K.A. Sustained T cell Rap1 signaling is protective in the collagen-induced arthritis model of rheumatoid arthritis. *Arthritis Rheum.* 2010, 62, 3289–3299. [CrossRef]
- Liu, G.; Jiang, Y.; Chen, X.; Zhang, R.; Ma, G.; Feng, R.; Zhang, L.; Liao, M.; Miao, Y.; Chen, Z.; et al. Measles contributes to rheumatoid arthritis: Evidence from pathway and network analyses of genome-wide association studies. *PLoS ONE* 2013, *8*, e75951. [CrossRef]
- Kalia, V.C.; Kumar, R.; Kumar, P.; Koul, S. A Genome-Wide Profiling Strategy as an Aid for Searching Unique Identification Biomarkers for Streptococcus. *Indian J. Microbiol.* 2016, 56, 46–58. [CrossRef]
- 52. Kalia, V.C.; Raju, S.C.; Purohit, H.J. Genomic analysis reveals versatile organisms for quorum quenching enzymes: Acylhomoserine lactone-acylase and -lactonase. *Open Microbiol. J.* **2011**, *5*, 1. [CrossRef]