Juvenile Idiopathic Arthritis CD4+ T cell Regulation Identified by Integration of Chromatin Access and Transcriptional Activity

Introduction

Juvenile idiopathic arthritis is the most common chronic inflammatory arthropathy in children, encompassing a spectrum of disorders defined by persistent synovitis of unknown etiology beginning before the age of 16. Its clinical heterogeneity is reflected in distinct subtypes, including oligoarticular JIA, which predominantly affects the knees and ankles and typically presents with involvement of four or fewer joints; polyarticular JIA, which often targets the small joints of the hands and wrists and involves five or more joints within the first six months; and systemic JIA, characterized by quotidian fevers, evanescent rash, and variable progression to chronic polyarthritis. Although these subtypes differ in their clinical manifestations and typical joint distributions, they share underlying immune dysregulation that promotes chronic synovial inflammation and ultimately joint damage.

CD4+ T cells are central coordinators of adaptive immunity and play critical roles in the initiation and maintenance of juvenile idiopathic arthritis. These cells exhibit substantial phenotypic diversity, encompassing classical Th1 cells that produce interferon-gamma and promote macrophage activation, Th17 cells that secrete interleukin-17 and drive neutrophil recruitment and osteoclastogenesis, and regulatory T cells (Tregs) that maintain peripheral tolerance by suppressing autoreactive responses. The differentiation of naïve CD4+ T cells into these effector or regulatory lineages is orchestrated by a complex interplay of antigenic stimulation, co-stimulatory signals, cytokine milieus such as IL-6, IL-23, and TGF-β, and transcription factor networks including T-bet, RORγt, and FOXP3.

In juvenile idiopathic arthritis, accumulating evidence indicates that even under basal, unstimulated conditions, CD4+ T cells from patients exhibit transcriptional and epigenetic abnormalities that may prime them for exaggerated inflammatory responses upon subsequent activation. Peripheral blood CD4+ T cells isolated from children with active JIA have been shown to display altered chromatin accessibility patterns and transcriptional profiles, suggesting that disease-associated immune programming extends beyond local joint microenvironments and may reflect systemic immune imprinting. Such basal alterations in chromatin landscapes and gene expression could lower the threshold for pathogenic activation, sustain memory-like pathogenic states, or facilitate rapid effector differentiation upon encountering inflammatory cues. These features are likely influenced by both intrinsic genetic risk factors and extrinsic inflammatory history, though the integrated regulatory networks governing these poised states are not fully understood.

While much of the current research in JIA focuses on the heightened responses of CD4+ T cells under inflammatory conditions or upon cytokine stimulation, it remains critical to characterize the basal molecular states of these cells to understand how they are epigenetically and transcriptionally primed for disease. Identifying the transcription factors and chromatin elements that maintain aberrant resting-state programs could illuminate fundamental mechanisms by which genetic susceptibility and prior immune challenges converge to establish a pathogenic immune set point. This, in turn, may help explain why certain JIA subtypes favor distinct patterns of joint involvement and exhibit variable trajectories of persistence or remission.

To address these questions, we conducted an integrated multi-omic analysis of RNA-seq and ATAC-seq data from peripheral blood CD4+ T cells of children with active juvenile idiopathic arthritis under basal, unstimulated conditions. Employing the Taiji framework, we reconstructed regulatory networks to identify key transcription factors and gene circuits that shape the resting-state phenotypes of CD4+ T cells in JIA. Our analyses revealed basal transcriptional programs associated with proinflammatory cytokine readiness, increased proliferative and migratory signatures, and enrichment of regulatory modules intersecting with known JIA genetic susceptibility loci. These computational predictions were supported by functional assays that demonstrated altered cytokine expression, enhanced proliferation, and migration even in the absence of overt external stimulation.

Together, these findings highlight the importance of basal immune programming in the pathogenesis of juvenile idiopathic arthritis. They underscore how intrinsic transcriptional and epigenetic landscapes in CD4+ T cells may set the stage for the persistent synovitis characteristic of JIA subtypes. By illuminating the regulatory networks that establish these poised pathogenic states, this work provides a foundation for developing therapeutic strategies aimed at resetting maladaptive immune baselines, offering potential avenues to modulate disease activity before overt inflammatory escalation occurs.

Methods

Data acquisition and preprocessing

Multi-omic datasets were obtained from the NCBI Gene Expression Omnibus under accession number GSE164213, comprising peripheral blood CD4+ T cells from children with active juvenile idiopathic arthritis. This dataset included ATAC-seq for chromatin accessibility, RNA-seq for gene expression profiling, and HiChIP data to infer three-dimensional chromatin architecture. Processed ATAC-seq bedGraph files were downloaded from GEO and converted to narrowPeak format using MACS3 (v3.0.0a6) with default parameters optimized for ATAC-seq, which models local Poisson noise to identify significant open chromatin regions. For RNA-seq, raw count matrices were retrieved, and Ensembl transcript IDs were mapped to official gene symbols using the biomaRt R package, supplemented by custom readr and plyr scripts for streamlined batch conversion and to handle unmatched entries, yielding standardized gene-level expression tables. Furthermore, HiChIP-derived chromatin contacts were processed with EpiTensor, which integrates epigenomic signals to improve detection of regulatory interactions. We selected the top 10% most confident long-range contacts to prioritize biologically meaningful enhancer-promoter and regulatory loops for downstream network construction.

Regulatory network construction with Taiji

We used the Taiji framework to integrate chromatin accessibility, gene expression, and 3D contact data into a unified transcriptional regulatory network. Taiji connects transcription factors (TFs) to target genes through accessible regulatory regions constrained by three-dimensional chromatin contacts, and computes influence scores to rank TFs by their predicted regulatory impact.

Inputs included ATAC narrowPeak files, normalized RNA expression matrices, and high-confidence HiChIP contact files, specified in a custom YAML configuration. Taiji was run on a high-performance computer cluster, completing in approximately five days.

Pathway enrichment and network analysis

We performed pathway and gene ontology enrichment using ReactomePA and ClusterProfiler on genes identified as targets within the reconstructed networks, revealing enrichment of pathways such as cytokine signaling, cell cycle, and migration. Influence scores from Taiji were visualized using pheatmap, generating ranked heatmaps to highlight the most central transcription factors and regulatory modules. Additionally, we explored network topologies with igraph to depict hierarchical TF-target relationships and regulatory clusters.

Pairwise Wilcoxon tests were conducted to compare TF influence scores across different anatomical sites (hand vs. knee, hand vs. hip, knee vs. hip) within healthy controls, assessing whether certain TFs showed context-dependent prominence potentially related to tissue-specific disease patterns.

Computational environment and reproducibility

All analyses were performed in a dedicated conda-managed R environment (r_bio_env) to ensure reproducibility and consistent package management. This environment was created with:

```
conda create -n r_bio_env -c bioconda -c conda-forge r-base=4.2.2 bioconductor-reactomepa
bioconductor-clusterprofiler -y
conda activate r_bio_env
```

Additional R packages such as readr, plyr, pheatmap, and igraph were installed directly within R using `install.packages` or `BiocManager::install`. Data import, manipulation, and visualization were performed using readr for fast file reading, plyr for data restructuring, and standard R plotting tools. Code and environment details were maintained in version-controlled scripts to ensure full reproducibility.

Summary of objectives

This integrative approach aimed to identify transcription factors and regulatory circuits that establish basal pathogenic states in JIA CD4+ T cells. By combining chromatin accessibility, gene expression, and 3D genome architecture, we systematically characterized networks that may prime these cells for inflammatory activation, offering insights into mechanisms sustaining chronic synovitis and informing strategies to reset maladaptive immune baselines.