Published in final edited form as:

Curr Opin Rheumatol. 2014 September ; 26(5): 579–584. doi:10.1097/BOR.0000000000000094.

Genetics of JIA: New tools bring new approaches

Halima Moncrieffe¹, Sampath Prahalad², and Susan D. Thompson¹

¹Division of Rheumatology and the Center for Autoimmune Disease Genomics and Etiology, Cincinnati Children's Hospital Medical Center, 3333 Burnet Avenue, Cincinnati, OH 45229

²Division of Pediatric Rheumatology, Department of Pediatrics and Genetics, Emory University School of Medicine, 1760 Haygood Dr, Atlanta, GA 30322

Abstract

Purpose of the review—In Juvenile Idiopathic Arthritis (JIA), there are now more than 25 regions represented by single nucleotide polymorphisms (SNPs) that show strong genetic associations. The causal variants and corresponding functions have not yet been defined for the majority of these regions. Here, we review current JIA association findings and the recent progress in the annotation of non-coding portion of the human genome as well as the new technologies necessary to apply this knowledge to JIA association findings.

Recent findings—An international collaboration was able to amass sufficient numbers of JIA and control samples to identify significantly robust genetic associations for JIA. The ENCODE Project and the NIH Roadmap Epigenetics Program have now annotated more than 80% of the noncoding genome, important in understanding the impact of risk loci, the majority of which fall outside of protein coding regions. Recent technological advances in high throughput sequencing, chromatin structure determination, transcription factor and enhancer binding site mapping and genome editing will likely provide a basis for understanding JIA genetic risk.

Summary—Understanding the role of genetic variation in the etiology of JIA will provide insight for disease mechanism and may explain disease heterogeneity between JIA subtypes and between autoimmune diseases in general.

Keywords

juvenile idiopathic arthritis; functional variation; causal risk variants; complex disease

Introduction

The cause of Juvenile Idiopathic Arthritis (JIA) remains unknown, but likely includes complex interactions between genes and environmental exposures resulting in dysregulation of the immune system. Genomics, a relatively new scientific discipline, driven by rapidly advancing technologies, has enabled new approaches to understand complex disease. Over

Corresponding author: Susan D. Thompson, Address: Division of Rheumatology and The Center for Autoimmune Disease Genomics and Etiology, ML4010, Cincinnati Children's Hospital Medical Center, 3333 Burnet Avenue, Cincinnati, OH 45229, Phone number: 513-636-3899, Susan.Thompson@cchmc.org.

No conflicts of interest

the last few years, genome-wide association studies (GWAS) in multiple autoimmune diseases, including JIA, have identified hundreds of genomic loci harboring risk variants that typically are represented by single-nucleotide polymorphisms (SNPs) and may often represent markers in linkage disequilibrium (LD) with the causal variant(s). While risk loci can map to protein coding regions (2% of genome) and affect protein function by altering protein sequence, translation rate, or alternative splicing of exons, it is more likely that functional variants lie in noncoding regions of the genome. Noncoding variants can affect regulatory elements, chromatin architecture or noncoding RNA (ncRNA) and modulate gene expression in a tissue-specific, disease state manner. Here we review current JIA genetics findings and discuss the application of post-GWAS technologies to understand the role of genetic variation in JIA pathobiology.

Genome-wide association studies

GWAS have been greatly successful at identifying common variants associated with susceptibility to many complex traits. In a recent meta-analysis of genome-wide studies of adult rheumatoid arthritis (RA), Okada et al. analyzed data representing 29,880 RA cases and 73,758 controls after imputation which allowed the evaluation of 10 million SNPs [1]. Using bioinformatics methods based on functional annotation, expression quantitative loci and pathway analysis and including overlapping loci reported for human primary immunodeficiency, somatic mutations in hematological malignancies and knockout phenotypes in mice, 98 biological candidate genes at 101 risk loci were identified. Of interest, this study demonstrated that these candidate genes are targets of approved RA therapeutic agents. These results are important for several reasons. First, they shed light on fundamental genes and pathways contributing to RA pathogenesis. Second, they provide empirical evidence that understanding the genetics of a disease can provide important information for drug discovery. Finally, identification of genes that overlap between JIA and phenotypes such as RA or SLE may facilitate the application of new pharmacological agents.

To date, there have been two JIA GWAS. The first included all JIA subtypes with 279 cases in the discovery cohort and 321 cases for validation and was limited to 100,000 SNP markers [2]. The HLA region showed the strongest association, and the second strongest association was a SNP in the VTCN1 gene, which encodes the co-stimulatory molecule B7-H4. VTCN1 variants were also associated with RA in a Dutch cohort [3]. Furthermore, a recent study of 272 children with JIA, found that VTCN1 SNP rs10923223 and JIA subtype were the strongest independent predictors of disease course, suggesting a role for VTCN1 in both JIA and RA pathogenesis [4].

Another GWAS of 814 oligoarticular and polyarticular JIA cases and 3058 controls reported 11 SNPs that were also genotyped in a large replication cohort of over 1700 JIA cases [5]. The discovery cohort revealed evidence of association of JIA with chromosomal region 3q13 which includes CD80, KTELC1 and c3orf1 genes with 1 SNP in region significantly associated in the replication cohort. Another association (rs6479891) was reported near JMJD1C, a gene that encodes a histone demethylase. Of interest, gene expression data on 68 JIA cases and 23 controls showed cis expression quantitative trait locus associations for

SNPs in C3orf1 and JMJD1C. By correlating gene expression findings with genetic association results, this study suggests novel genes as plausible JIA candidates.

A consortium to investigate shared loci identified in GWAS across autoimmune disorders developed a custom genotyping array called the Immunochip [6]. The Immunochip has almost 200,000 SNPs, including dense coverage of the MHC region, and ~180 loci with strong statistical evidence of association with one or more of 12 autoimmune diseases [6] and can be done at a fraction of the cost of a genome-wide SNP array. Studies utilizing the Immunochip have been successful at identifying loci associated with many autoimmune disorders including Celiac disease [7], inflammatory bowel disease [8], RA [9], ankylosing spondylitis [10] and JIA [11]. In total, 250,000 samples were genotyped and providing a future opportunity for a large meta-analysis to potentially discern genetic contributions to sub-phenotypes present in more than one disease type- as well as features related to autoimmunity in general. To facilitate such comparisons, a curated and integrated set of datasets and tools was developed (www.immunobase.org). Noteworthy, a combined analysis of Immunochip data with published GWAS data allowed the identification of 163 loci associated with IBD. [8].

For JIA, an international consortium was formed to maximize sample size which included 2816 cases with oligoarticular or RF-negative polyarticular JIA and 13056 controls [11]. In addition to confirming the 3 loci that had previously been associated at genome-wide levels of significance (HLA, PTPN22, PTPN2), 14 new loci were identified, including IL-2 pathway members IL2RA, IL2RB, IL2-IL21 and related loci, SH2B3 and STAT4, highlighting the importance of IL-2 pathway in JIA pathogenesis. In addition 11 loci were found at suggestive levels of significance. Of the 17 regions that reached genome wide significance, linkage disequilibrium patterns and functional annotation provided strong evidence to localize the signal to a single gene in 8 loci. Many of the loci are shared with other autoimmune disease. Thus, the JIA Immunochip consortium has demonstrated the benefits of international collaboration by successfully identifying a large number of variants predisposing to the most common forms of JIA [11].

HLA associations

It has been known for decades that certain HLA-DRB1 alleles encoding the shared epitope (consensus amino acid sequence in positions 70 to 74 in the β 1 subunit of the HLA DR molecule) were strongly associated with the risk of RA[12]. Recently, Raychaudhuri et al. used SNP genotype data from 5018 cases with CCP-positive RA and 14984 controls of European descent to refine the association between HLA and RA [13]. Using a large reference panel of individuals of European descent, they imputed classical HLA allele genotypes as well as their corresponding amino acids. Amino acid position 11 of HLA-DR β 1 showed the strongest association ($p < 1 \times 10^{-581}$), with valine or leucine conferring a high risk of disease. HLA-DR β 1 amino acids at positions 71 and 74 were also associated with RA. After conditioning on positions 11, 71 and 74, no residual association at other HLA-DR β 1 amino acids was observed. Amino acids at position 9 in HLA-B and HLA-DP β 1 also demonstrated association after conditioning on these amino acids and those at positions 11, 71 and 74 in HLA-DR β 1, there were no residual signals of association across the MHC,

suggesting that the association between MHC and risk of RA can be almost completely explained by these five positions. This discovery highlights that the application of novel analytical techniques can improve our understanding of disease. A similar approach is being taken using Immunochip data to help unravel JIA disease-specific associated amino acids. Interestingly, DRB1 susceptibility alleles are not shared between JIA and RA and in fact, DRB1*04:01 which contains a valine at amino acid 11 is protective in JIA [14]. Fine mapping genetic risk across the HLA region at the amino acid level may provide important clues to understanding antigens that potentially trigger disease in susceptible individuals and lead to a better understanding of JIA etiology.

Rare variants in complex diseases

The loci identified by GWAS, are based on genotyping of common SNPs (minor allele frequency > 0.05), typically only explain a small fraction of the disease heritability. This is especially true for JIA, where 28 associated loci explain about 18% of disease risk [11]. Likewise in RA, >100 associated loci explain only about 35% of disease risk [1, 15]. Many explanations for this missing heritability have been suggested, including much larger numbers of variants of smaller effect yet to be found; rarer variants (possibly with larger effects) that are not present on genotyping arrays or structural variants poorly captured by existing arrays [16, 17]. Supporting the rare variant hypothesis, resequencing studies targeting established RA risk loci identified numerous rare non-synonymous variants for a number of loci, and exclusive to RA patient samples [18].

For JIA and other complex diseases, next generation sequencing (NGS) offers the possibly to directly detect infrequent variants in the many different genes contributing to disease. NGS can be applied to the whole genome, the exome (all exons) or targeted to specific loci relevant to disease. NGS data comprise numerous short read sequences 50–150 bases long which overlap with each other and can be aligned to a reference sequence [19]. NGS is also known as deep or massively parallel sequencing. Errors in the short read sequences occur using this high-throughput approach, but are corrected by overlapping sequences reads such that each base is typically represented by 15 to 50 overlapping reads, representing the “sequencing depth”. Typical experiments targeting the whole exome require 30 million reads for each sample, to achieve this desired depth for all genes.

While conceptually possible, there are barriers for employing NGS for genome-wide, rare variant association studies including sequencing costs and the requirement for vast sample sizes ($\sim 25,000$ cases) and the development of new analytic pipelines [20]. Furthermore, existing single variant tests cannot distinguish between causal and LD-induced association and may be underpowered for identifying rare or low-frequency associated variants, yet using variant grouping methods will not identify individual causal variants. A method that combines LD with known genome-wide association signals and functional conservation has been proposed as a solution to accurately detect causal variants [20].

Expression quantitative expression loci (eQTL)

Gene transcript and protein abundance may be directly modified by polymorphisms in regulatory elements, which are referred to as expression quantitative trait loci. eQTLs may

act in cis (locally) or trans (at a distance) to a gene. Gene expression is tightly regulated and dependent on the context of relevant cell or tissue types in a particular biological state and highlighted in recent publications [21, 22]. This context-dependent specificity means that previous studies using lymphoblastoid cell lines and other tissues, although providing important insights, may not capture the *in vivo* activity of particular variants. Analyzing SNP associations in the context of gene transcription increases the capacity to detect disease-causing genomic loci.

Non-coding genetic risk variants

The ENCODE project (<http://www.genome.gov/Encode/>) has provided a wealth of information to define the functional DNA elements in the human genome in a cell-type specific manner [23]. Work by the ENCODE consortium has demonstrated fundamental new characteristics of the human genome and assigned biochemical functions for ~80% of the genome [23-25]. ENCODE data indicates about three-quarters of the human genome is capable of being transcribed, thus challenging the concept of a gene [26].

The term ‘epigenome’ meaning ‘on or above genomes’ was coined to describe genome-wide chromatin modifications including DNA methylation, histone methylation, chromatin accessibility, small and long non-coding RNA and 3D chromatin interactions [27]. JIA risk loci are most commonly found in the non-coding regions of the genome and may alter the epigenome. The epigenetic features related to disease are even more complex than genomic features because hundreds of epigenomes relate to the single genome of an individual. These epigenomes, correspond to various types of epigenomic marks that differ by developmental state and tissue type. Genome-wide measurements of chromatin conformation [28] and characterization of new long-range interactions between genes using 5C (chromosome conformation capture carbon copy) technology [24] are just two of the new valuable resources available from ENCODE as a comparator when investigating the mechanisms by which variation in the JIA genome contributes to disease. The International Human Epigenome Consortium is mapping 1000 reference epigenomes (<http://www.ihec-epigenomes.org/>). Datasets for human blood CD4+ T lymphocytes and CD14+ monocytes in healthy individuals are complete, while analysis of CD4+ T lymphocytes from lupus and scleroderma patients is in progress. Similar work is planned for RA (<http://ihec-epigenomes.org/outcomes/datasets/>) but not yet for JIA. Nonetheless, information from normal human blood/tissues and adult autoimmune diseases will serve as a useful resource in understanding disease mechanisms.

Studies relating epigenomes to disease can be termed “eGWAS”: epigenome-wide association studies [29]. In JIA, a genome-scale analysis interrogated the methylation state of promoter regions in purified peripheral blood CD4+ T cells from methotrexate naïve patients and healthy age- and sex- matched controls [30]. Altered methylation can impact on gene activation and silencing. Reduced methylation in the JIA IL-32 regulatory region was found, which may impact on isoform switching (IL-32 γ to IL-32 β switching was associated with reduced chronic inflammation in RA patients) [31]. Angiogenic functions for IL-32 have also been described [32] consistent with mechanisms reported in JIA pathogenesis [33], adding further significance to eGWAS findings. Genome-wide methylation studies

have provided interesting findings in lupus related to IFN pathways [34, 35] and within the HLA locus in RA [36].

Most human RNA transcripts are not translated to protein, but instead represent non-coding RNAs (ncRNAs) with important biological roles in modulating epigenetic regulation [37]. There are many types of ncRNA including lncRNA (>200 nucleotides long) and microRNAs (miRNA ~22 nucleotides long). lncRNA interact with chromatin regulators to guide them to chromatin [38] while miRNAs function as post-transcription repressors [39]. Both ncRNAs mediate gene expression in a sequence-dependent manner. Data from a cohort of 150 enthesitis-related arthritis (ERA) patients demonstrated that disease susceptibility was associated with miRNA-146a and a polymorphism in its target gene IRAK1 [40]. Given the reported hypoxic environment in the JIA joint, hypoxia-induced miRNAs could be important players in regulating JIA gene expression. For example, the hypoxia-regulated microRNA miR-210 negatively regulates the Th17 pathway [41] that has pathogenic roles in the JIA joint [42]. Furthermore for RA, decreased expression of miR-146a and miR-155 were found to contribute to an abnormal Treg phenotype [43]. The role of miRNA in the balance of inflammation and T cell regulation in JIA remain important areas of investigation.

Transcription factors bind to thousands of regulatory elements including promoters, enhancers, silencers and insulators. By combining chromatin immunoprecipitation (ChIP) and massively parallel sequencing, ChIP-Seq can be used to accurately survey interactions between protein, DNA, and RNA and specifically identify these regulatory elements. This is important since the casual variant may impact transcription factor (TF) binding by creating or destroying canonical TF binding sites leading to altered gene expression. A systematic comparison of the models that predict TF binding sites completed by Weirauch et al. [44] resulted in improved algorithms to narrow potential candidates to be tested using wet lab techniques which measure protein-DNA interactions such as electrophoretic mobility shift assay (EMSA) and high-throughput systems such as protein binding microarrays [45].

Genome editing

Demonstrating the variant's function in disease pathology is an important step to understanding disease mechanism and identifying new therapeutic targets. Introducing precise changes in the DNA sequence is termed genome editing, and is used to artificially recreate the variants in models systems relevant to disease. Multiple artificial nuclease systems can be used for genome editing in both cells and model organisms. For example, zinc finger nucleases (ZFN) are site-specific enzymes that introduce DNA double-strand breaks enabling gene disruption or gene insertion and are widely used [46-48] Transcription activator-like effector nucleases (TALENs) can also be programed to target specific DNA sequences and do this by shuffling amino acid recognition motifs [49-51]. An approach using a heteroduplex assay in HEK293 T cells was recently developed to alleviate challenging steps in screening for nuclease modifications [52].

Genome editing with the CRISPR-Cas (clustered regularly interspaced short palindromic repeats/CRISPR-associated) is considered a simpler alternative to the systems described

above and relies on RNA rather than protein sequence to target specific DNA sequences [53]. CRISPR RNAs guide the Cas9 endonuclease to complementary DNA sequences for subsequent cleavage. Sequence specificity for Cas9-mediated DNA cleavage requires the genomic locus of interest to be followed by a NGG protospacer adjacent motif (PAM). Ideally, these RNA guides are designed to not have off-target effects. To facilitate this, a genome-wide binding map for the CRISPR endonuclease has been developed [54]. CRISPR technologies have been successfully used to create a gene-knockout library to screen for novel genes involved in bacterial toxin-resistance [55], generate mice containing reporter and conditional alleles [56] and modify human cells [57, 58]. These genome editing tools will facilitate screening of JIA risk variants to assess roles in disease causality and function and perhaps find a replacement for the term “idiopathic” in JIA.

Conclusion

GWAS studies fail to capture all genetic variation related to disease, and do not address the causal nature of genetic variation. NGS continues to fuel the discovery of rare variants and eventually may replace SNP based association studies. The recent annotation of the human genome provides a foundation for functional validation of associated risk variants. Bioinformatic integration of genetic association findings with the developing set of functional annotations in the context of linkage disequilibrium can be used to predict function and guide experimentation to address long standing questions related to disease mechanisms, causation, and disease heterogeneity and ultimately translated into new tools for the clinician and patient.

Acknowledgments

NIH Funding: P01AR048929, P30AR047363

Bibliography

1. Okada Y, Wu D, Trynka G, Raj T, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature*. 2014; 506(7488):376–81. This is a comprehensive GWAS meta-analysis of RA investigating over 100,000 subjects. This study identifies 98 candidate genes contributing to risk of RA. As proof of principle, the genes are found in pathways that are targets of current therapy of RA. [PubMed: 24390342]
2. Hinks A, Barton A, Shephard N, Eyre S, et al. Identification of a novel susceptibility locus for juvenile idiopathic arthritis by genome-wide association analysis. *Arthritis Rheum*. 2009; 60(1): 258–63. [PubMed: 19116933]
3. Daha NA, Lie BA, Trouw LA, Stoeken G, et al. Novel genetic association of the VTCN1 region with rheumatoid arthritis. *Ann Rheum Dis*. 2012; 71(4):567–71. [PubMed: 22323440]
4. Albers HM, Reinards TH, Brinkman DM, Kamphuis SS, et al. Genetic variation in VTCN1 (B7-H4) is associated with course of disease in juvenile idiopathic arthritis. *Ann Rheum Dis*. 2013
5. Thompson SD, Marion MC, Sudman M, Ryan M, et al. Genome-wide association analysis of juvenile idiopathic arthritis identifies a new susceptibility locus at chromosomal region 3q13. *Arthritis Rheum*. 2012; 64(8):2781–91. [PubMed: 22354554]
6. Cortes A, Brown MA. Promise and pitfalls of the Immunochip. *Arthritis Res Ther*. 2011; 13(1):101. [PubMed: 21345260]
7. Trynka G, Hunt KA, Bockett NA, Romanos J, et al. Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. *Nat Genet*. 2011; 43(12): 1193–201. [PubMed: 22057235]

- 8■. Jostins L, Ripke S, Weersma RK, Duerr RH, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature*. 2012; 491(7422):119–24. This is a GWAS meta-analysis of IBD. 163 candidate genes associated with IBD are identified demonstrating overlap with related phenotypes such as AS and Psoriasis and the level of complexity in understanding genetic determinants of autoimmunity. [PubMed: 23128233]
9. Eyre S, Bowes J, Diogo D, Lee A, et al. High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. *Nat Genet*. 2012; 44(12):1336–40. [PubMed: 23143596]
10. Cortes A, Hadler J, Pointon JP, Robinson PC, et al. Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. *Nat Genet*. 2013; 45(7):730–8. [PubMed: 23749187]
- 11■. Hinks A, Cobb J, Marion MC, Prahalad S, et al. Dense genotyping of immune-related disease regions identifies 14 new susceptibility loci for juvenile idiopathic arthritis. *Nat Genet*. 2013; 45(6):664–9. The JIA Immunochip Consortium published the findings of Immunochip analysis of over 2800 subjects with polyarticular and oligoarticular JIA. Seventeen loci were associated at genome wide levels of significance, and 11 additional loci were identified at suggestive levels of association. Highlights important pathways involved in JIA pathogenesis. [PubMed: 23603761]
12. Gregersen PK, Silver J, Winchester RJ. The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. *Arthritis Rheum*. 1987; 30(11):1205–13. [PubMed: 2446635]
- 13■. Raychaudhuri S, Sandor C, Stahl EA, Freudenberg J, et al. Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. *Nat Genet*. 2012; 44(3):291–6. This report exemplifies application of modern analytical tools to old problems. This report narrows the MHC association with antibody positive RA to 5 amino acid positions in the HLA DRβ1, HLA B and HLA DPβ1 chains. [PubMed: 22286218]
14. Hollenbach J, Thompson SD, Bugawan T, Ryan M, et al. Juvenile idiopathic arthritis and HLA class I and class II interaction and age-at-onset effects. *Arthritis Rheum*. 2010; 42(6):1781–91. [PubMed: 20191588]
15. Gusev A, Bhatia G, Zaitlen N, Vilhjalmsdottir BJ, et al. Quantifying missing heritability at known GWAS loci. *PLoS genetics*. 2013; 9(12):e1003993. [PubMed: 24385918]
16. Manolio TA, Collins FS, Cox NJ, Goldstein DB, et al. Finding the missing heritability of complex diseases. *Nature*. 2009; 461(7265):747–53. [PubMed: 19812666]
17. Chang D, Keinan A. Predicting signatures of “synthetic associations” and “natural associations” from empirical patterns of human genetic variation. *PLoS computational biology*. 2012; 8(7):e1002600. [PubMed: 22792059]
18. Diogo D, Kurreeman F, Stahl EA, Liao KP, et al. Rare, low-frequency, and common variants in the protein-coding sequence of biological candidate genes from GWASs contribute to risk of rheumatoid arthritis. *Am J Hum Genet*. 2013; 92(1):15–27. [PubMed: 23261300]
19. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012; 491(7422):56–65. [PubMed: 23128226]
- 20■. Zuk O, Schaffner SF, Samocha K, Do R, et al. Searching for missing heritability: designing rare variant association studies. *Proc Natl Acad Sci U S A*. 2014; 111(4):E455–64. Genetic association studies based on exome or genome level sequence are necessary to identify rare variants contributing to disease. A useful framework is reported for the design and analysis of these challenging datasets. [PubMed: 24443550]
21. Lee MN, Ye C, Villani AC, Raj T, et al. Common genetic variants modulate pathogen-sensing responses in human dendritic cells. *Science*. 343(6175):1246980. [PubMed: 24604203]
22. Fairfax BP, Humburg P, Makino S, Naranbhai V, et al. Innate Immune Activity Conditions the Effect of Regulatory Variants upon Monocyte Gene Expression. *Science*. 2014; 343(6175)
- 23■. Bernstein BE, Birney E, Dunham I, Green ED, et al. Encode Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012; 489(7414):57–74. Summarizes the findings of the Encyclopedia of DNA Elements (ENCODE) project and the utility of these functional annotations in understanding human disease. More indepth reports of specific ENCODE work is found in this same issue of *Nature*. [PubMed: 22955616]

24. Sanyal A, Lajoie BR, Jain G, Dekker J. The long-range interaction landscape of gene promoters. *Nature*. 2012; 489(7414):109–13. [PubMed: 22955621]
25. Thurman RE, Rynes E, Humbert R, Vierstra J, et al. The accessible chromatin landscape of the human genome. *Nature*. 2012; 489(7414):75–82. [PubMed: 22955617]
26. Djebali S, Davis CA, Merkel A, Dobin A, et al. Landscape of transcription in human cells. *Nature*. 2012; 489(7414):101–8. [PubMed: 22955620]
27. Sarda S, Hannenhalli S. Next-Generation Sequencing and Epigenomics Research: A Hammer in Search of Nails. *Genomics & informatics*. 2014; 12(1):2–11. [PubMed: 24748856]
28. Maurano MT, Humbert R, Rynes E, Thurman RE, et al. Systematic Localization of Common Disease-Associated Variation in Regulatory DNA. *Science*. 2012; 337(6099):1190–5. [PubMed: 22955828]
29. Maunakea AK, Chepelev I, Zhao K. Epigenome mapping in normal and disease States. *Circ Res*. 2010; 107(3):327–39. [PubMed: 20689072]
30. Ellis JA, Munro JE, Chavez RA, Gordon L, et al. Genome-scale case-control analysis of CD4+ T-cell DNA methylation in juvenile idiopathic arthritis reveals potential targets involved in disease. *Clinical epigenetics*. 2012; 4(1):20. This is the first genome-wide epigenetic study for JIA and reports methylation differences in the regulatory region of the functionally relative IL-32 gene in JIA patient CD4+ T cells compared to healthy controls. [PubMed: 23148518]
31. Heinhuis B, Koenders MI, van de Loo FA, Netea MG, et al. Inflammation-dependent secretion and splicing of IL-32{gamma} in rheumatoid arthritis. *Proc Natl Acad Sci U S A*. 2011; 108(12):4962–7. [PubMed: 21383200]
32. Nold-Petry CA, Rudloff I, Baumer Y, Ruvo M, et al. IL-32 promotes angiogenesis. *J Immunol*. 2014; 192(2):589–602. [PubMed: 24337385]
33. Barnes MG, Aronow BJ, Luyrink LK, Moroldo MB, et al. Gene expression in juvenile arthritis and spondyloarthropathy: pro-angiogenic ELR+ chemokine genes relate to course of arthritis. *Rheumatology (Oxford)*. 2004; 43(8):973–9. [PubMed: 15150433]
34. Absher DM, Li X, Waite LL, Gibson A, et al. Genome-wide DNA methylation analysis of systemic lupus erythematosus reveals persistent hypomethylation of interferon genes and compositional changes to CD4+ T-cell populations. *PLoS genetics*. 2013; 9(8):e1003678. [PubMed: 23950730]
35. Coit P, Jeffries M, Altorok N, Dozmorov MG, et al. Genome-wide DNA methylation study suggests epigenetic accessibility and transcriptional poising of interferon-regulated genes in naive CD4+ T cells from lupus patients. *J Autoimmun*. 2013; 43:78–84. [PubMed: 23623029]
36. Liu Y, Aryee MJ, Padyukov L, Fallin MD, et al. Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. *Nat Biotech*. 2013; 31(2):142–7.
37. Peschansky VJ, Wahlestedt C. Non-coding RNAs as direct and indirect modulators of epigenetic regulation. *Epigenetics : official journal of the DNA Methylation Society*. 2014; 9(1):3–12. [PubMed: 24739571]
38. Rinn JL, Kertesz M, Wang JK, Squazzo SL, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell*. 2007; 129(7):1311–23. [PubMed: 17604720]
39. Bartel DP. MicroRNAs: target recognition and regulatory functions. *Cell*. 2009; 136(2):215–33. [PubMed: 19167326]
40. Singh S, Rai G, Aggarwal A. Association of microRNA-146a and its target gene IRAK1 polymorphism with enthesitis related arthritis category of juvenile idiopathic arthritis. *Rheumatol Int*. 2014 epub 2014/04/11.
41. Wang H, Flach H, Onizawa M, Wei L, et al. Negative regulation of Hif1a expression and TH17 differentiation by the hypoxia-regulated microRNA miR-210. *Nat Immunol*. 2014; 15(4):393–401. [PubMed: 24608041]
42. Nistala K, Moncrieffe H, Newton KR, Varsani H, et al. Interleukin-17-producing T cells are enriched in the joints of children with arthritis, but have a reciprocal relationship to regulatory T cell numbers. *Arthritis Rheum*. 2008; 58(3):875–87. [PubMed: 18311821]

43. Zhou Q, Haupt S, Kreuzer JT, Hammitzsch A, et al. Decreased expression of miR-146a and miR-155 contributes to an abnormal Treg phenotype in patients with rheumatoid arthritis. *Ann Rheum Dis*. 2014 epub 2014/02/25. This article demonstrates a mechanism of microRNA regulation in rheumatoid arthritis Treg function. Treg in patients with active RA had increased STAT1 activity which is a direct target of miR-146a.
44. Weirauch MT, Cote A, Norel R, Annala M, et al. Evaluation of methods for modeling transcription factor sequence specificity. *Nat Biotech*. 2013; 31(2):126–34.
45. Bulyk ML. Protein binding microarrays for the characterization of DNA-protein interactions. *Advances in biochemical engineering/biotechnology*. 2007; 104:65–85. [PubMed: 17290819]
46. Bhakta MS, Henry IM, Ousterout DG, Das KT, et al. Highly active zinc-finger nucleases by extended modular assembly. *Genome research*. 2013; 23(3):530–8. [PubMed: 23222846]
47. Kim S, Lee MJ, Kim H, Kang M, et al. Preassembled zinc-finger arrays for rapid construction of ZFNs. *Nat Methods*. 2011; 8(1):7. [PubMed: 21191366]
48. Wood AJ, Lo TW, Zeitler B, Pickle CS, et al. Targeted Genome Editing Across Species Using ZFNs and TALENs. *Science*. 2011; 333(6040):307. [PubMed: 21700836]
49. Joung JK, Sander JD. TALENs: a widely applicable technology for targeted genome editing. *Nat Rev Mol Cell Biol*. 2013; 14(1):49–55. [PubMed: 23169466]
50. Deng D, Yan C, Pan X, Mahfouz M, et al. Structural Basis for Sequence-Specific Recognition of DNA by TAL Effectors. *Science*. 2012; 335(6069):720–3. [PubMed: 22223738]
51. Kim Y, Kweon J, Kim A, Chon JK, et al. A library of TAL effector nucleases spanning the human genome. *Nature biotechnology*. 2013; 31(3):251–8.
52. Nerys-Junior A, Costa LC, Braga-Dias LP, Oliveira M, et al. Use of the heteroduplex mobility assay and cell sorting to select genome sequences of the CCR5 gene in HEK 293T cells edited by transcription activator-like effector nucleases. *Genetics and molecular biology*. 2014; 37(1):120–6. [PubMed: 24688299]
53. Mali P, Yang L, Esvelt KM, Aach J, et al. RNA-guided human genome engineering via Cas9. *Science*. 2013; 339(6121):823–6. [PubMed: 23287722]
54. Wu X, Scott DA, Kriz AJ, Chiu AC, et al. Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat Biotech*. 2014 advance online publication. Genetic editing requires precise targeting of DNA. This work catalogs CRISPR binding sites across the genome and is an important resource in experimental design to avoid off target edits.
55. Zhou Y, Zhu S, Cai C, Yuan P, et al. High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells. *Nature*. 2014 advance online publication. This article demonstrates a novel use for CRISPR/Cas9 genome editing in functional genomics by generating a library of gene knockouts to screen for bacterial toxins in the human HeLa cell line.
56. Yang H, Wang H, Shivalila CS, Cheng AW, et al. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell*. 2013; 154(6):1370–9. [PubMed: 23992847]
57. Cho SW, Kim S, Kim JM, Kim JS. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nature biotechnology*. 2013; 31(3):230–2.
58. Jinek M, East A, Cheng A, Lin S, et al. RNA-programmed genome editing in human cells. *eLife*. 2013; 2:e00471. [PubMed: 23386978]

Key Points

- The discovery of JIA- associated risk variants has accelerated in the past few years with genome-wide and ImmunoChip assays revealing novel loci in both coding and noncoding regions of the genome.
- The ENCODE project and other large scale initiatives to annotate the non-coding portion of the genome have created a “user's guide” for the human genome.
- JIA associated risk loci can be analyzed bioinformatically, in the context of this functional annotation to predict biological impact.
- Functional validation is required for designating variants as “disease causing”, and facilitated by the availability of genome editing tools such as CRISPR technology to artificially recreate the variant in a model system relevant for disease.