The genetics of juvenile idiopathic arthritis: current understanding and future prospects

Joanna E. Cobb¹,², Anne Hinks¹ and Wendy Thomson¹,²

Abstract

Understanding the genetic risk of JIA, a relatively rare chronic disease, is a challenging task, but recent research in this field has shown great advances. This review summarizes the current understanding of the genetic architecture of JIA susceptibility and proposes where this work is heading in the coming years. Insights into how we might progress this relatively understudied field will be provided, highlighting how the field will move towards the ultimate goals of predicting long-term disease outcomes at onset, predicting drug response, and move towards more targeted treatment options for children with JIA.

Key words: genetics, juvenile idiopathic arthritis, autoimmune disease, predicting drug response.

Juvenile idiopathic arthritis as a complex genetic disease

Although a relatively rare disease, JIA is the most common of the childhood rheumatic diseases [1]. JIA has autoimmune and inflammatory features and appears complex in nature, with both environmental and multiple genetic risk components [2]. Evidence for the heritability of JIA historically comes from family studies, which have shown that JIA has a sibling relative risk (r_s) ranging from 15 to 30, similar to that of type 1 diabetes (T1D) [3-5]. More recently, estimates using a large cohort of 862 JIA patients linked to the Utah Population Database place the prevalence of JIA in Caucasian populations at 1.2 per 1000, with a likely more accurate r_s of 11.6 [6]. Autoimmune diseases (AIDs) such as JIA also appear to have multiple and overlapping environmental risk factors, possibly including infection and vitamin D deficiency [2]. However, the role of these environmental factors in JIA risk is not well characterized and are not the focus of this review.

It is thought that for most complex diseases, such as JIA, there will be many genomic regions contributing relatively small amounts to overall disease risk [7]. A key goal in finding these genetic associations is to improve our understanding of the molecular mechanisms and pathophysiological pathways behind these diseases. Ultimately it is hoped this will enable the production of tailor-made treatment strategies based on individual genotypes, which in turn should lead to a marked improvement in clinical management.

Over the last 30 years, genetic discovery has progressed through various phases primarily based on the availability and advancement of genotyping technologies, with association studies being the preferred approach today. This method allows for the detection of variants with very small disease influence due to the increased statistical power obtained using unrelated individuals and large numbers of genetic markers [8]. In the past, association studies were primarily conducted using a candidate gene approach, where a specific genomic region is selected for investigation based on evidence from similar (often monogenic) diseases, and the physiological plausibility of that gene in the condition of interest, e.g. investigation of the IL2RA gene in JIA [9].

As technology has improved and genotyping costs have been reduced, the complex disease field has shifted towards genome-wide association studies (GWASs). Currently GWASs involve genotyping cases and controls for ~1 million markers across the genome and are well powered to detect common variants of small effect sizes [10]. Standardized GWAS protocols require researchers to replicate findings and impose a stringent significance threshold (P ≤ 5 x 10⁻⁸). Reassuringly, many of the markers identified in GWASs are in biologically meaningful genes and known pathways. This approach has identified ~1500 signals at P ≤ 5 x 10⁻⁸ attributable to 237 traits (http://www.genome.gov/GWAStudies). Abundant among
these signals are the >200 loci associated with AIDs [11]. Indeed, an interesting outcome of GWASs in AIDs is the observation that these AID association signals are often occurring in overlapping genomic regions [12]. For example, the AFF3 gene is a transcription factor preferentially expressed in lymphoid tissue that was first associated in a GWAS of T1D and has since shown involvement in JIA and RA [13–15]. Although this overlap was identified in limited examples prior to the GWAS era, the introduction of GWASs has facilitated the identification of more widespread sharing of genes associated with multiple AIDs, resulting in emerging networks of common pathogenic pathways involved in the causation of autoimmunity.

**Strategies to investigate JIA susceptibility**

Given the challenges associated with JIA genetics resulting from the relative rarity and compounded by the clinical heterogeneity of the disease (there are currently seven ILAR subtypes [16]), researchers used varied strategies in an attempt to uncover the genetic basis of JIA susceptibility [17]. These include selecting genes based on expression profiling results [18], those previously associated with other AIDs [19, 20] and GWASs [21, 22]. Using these approaches, until recently only three genomic loci showed a confirmed association with JIA susceptibility at genome-wide significance (HLA, PTPN22 and PTPN2) [21, 23].

The large contribution of the HLA region, on chromosome 6 within the MHC, to AID risk and JIA susceptibility has been known for some time, with both class I (HLA-A2 and HLA-B27) and class II (HLA-DRB1 and HLA-DP) HLA alleles showing association with various subtypes of JIA [24, 25], as reviewed previously [26, 27]. Refining the association signals in this region and understanding their role in JIA is still a work in progress [28, 29]. The HLA region is estimated to explain ~8–13% of the total variation in JIA susceptibility, confirming that there are still many non-HLA loci to be identified [22, 28].

Focussing on these non-HLA genes, the approach of testing loci identified in other AIDs proved successful for JIA. Earlier studies were often only powered to detect genomic regions with larger effect sizes and only PTPN22 and PTPN2 had been confirmed to the genome-wide significance threshold [23]. Multiple other genes (e.g. STAT4 and IL2RA) have been implicated in earlier candidate gene studies [17, 23]. More recently, the findings have included a 32 bp insertion/deletion in the CCR5 gene, which is important in the recruitment of T helper cells to the synovium [30] and is associated with RA, T1D and coeliac disease [31–33]; associations in the CD247 gene, which plays a role in T cell activation and signalling [20] and associations in the AFF3 gene [14].

JIA is lagging behind other AIDs in taking advantage of the rapidly advancing GWAS era, with only two small GWASs published from 2008 to 2010 and one recent larger GWAS [21–23]. As a result, few uniquely JIA or non-AID-associated loci have been identified to date; however, it is expected that this will change as more GWASs, with larger sample sizes, emerge. Two novel regions (CD80-KTELC1 and JMJD1C) were identified in a recent GWAS of 814 oligoarticular and RF-negative polyarticular cases and 3058 controls, a finding that was replicated in multiple independent populations (1744 cases, 7010 controls) [22]. These two associations represent a common pathogenic pathway that play a key role in regulating gene expression [22]. A number of GWASs investigating JIA susceptibility are currently under way and future meta-analyses combining these with already published GWASs will increase the study sample sizes and improve the power to detect the relatively small effect sizes commonly seen in complex diseases.

It is anticipated that this, along with the establishment of various international research networks, such as the International Childhood Arthritis Genetics (INCHARGE) Consortium, which is facilitating larger studies moving forward, will result in a significant increase in the number of known JIA regions and our understanding of the genetic susceptibility of the disease.

**Utilizing AID overlap proves hugely successful for JIA**

Evidence for overlapping AID signals prompted a single fine-mapping experiment for AIDs to maximize genotyping efficiency and cost. The Immunochip genetic association project aims to fine-map validated AID loci and perform deep replication on loci with evidence from many AIDs, including T1D, multiple sclerosis, Crohn’s disease, coeliac disease and RA. Loci for fine-mapping were included on the chip if association evidence met genome-wide thresholds, giving rise to ~185 validated loci [34]. Single nucleotide polymorphisms (SNPs) were selected from the February 2010 1000 Genome Project release, if not rare, with additional SNPs included based on private sequencing data [34]. This approach resulted in >196 000 SNPs being included on the Illumina Immunochip design. The first articles to report findings from the Immunochip project have recently been published, including those investigating RA, autoimmune thyroid disease, coeliac disease, primary biliary cirrhosis and psoriasis [35–40]. These studies confirm the large genetic overlap in AIDs, show this approach successfully identifies regions not previously associated for some AIDs and refines the associations of previously known disease loci.

Given the previous evidence for the considerable overlap between the genetic associations for JIA and other AIDs, we are in a unique position to take advantage of the Immunochip. A consortium of researchers from the UK and USA have genotyped ~2800 oligoarticular and RF-negative polyarticular JIA cases on the Immunochip [28]. This study has greatly improved our understanding of the genetic susceptibility to JIA.

In the HLA region on chromosome 6, an uncommon SNP (rs7775055, minor allele frequencycontrols = 2%,
odds ratio (OR) = 6.01, \( P = 3.14 \times 10^{-175} \) confirming susceptibility to oligoarticular and RF-negative polyarticular JIA has been identified tagging the DRB1*0801-DQA1*0401-DQB1*0402 haplotype [28]. Although this haplotype has previously been implicated in these JIA subtypes [24, 25], when the oligoarticular and RF-negative polyarticular cases from the Immunochip study were analysed separately a significantly different OR for rs7775055 was found, highlighting the likely subtype-specific differences in the HLA region. The HLA region remains costly and technically difficult to accurately genotype due to its strong linkage disequilibrium (LD) patterns, dense polymorphisms and structural properties. To facilitate work in this uniquely challenging region, researchers have developed imputation tools targeting SNPs and amino acids, with some exciting results [41–43]. By implementing one of these approaches in RA, Raychaudhuri et al. [43] found three amino acid positions in HLA-DRB1 and single amino acid polymorphisms in HLA-B and HLA-DPB1 almost completely explain the MHC association to RA risk. This work has interesting implications for JIA, and using this SNP imputation approach suggests a comprehensive study of the HLA region is now feasible for each of the JIA subtypes.

Additionally, the Immunochip study has confirmed other previously implicated JIA risk loci such as STAT4 and IL2RA and provides evidence for new regions such as ANKRD55 and TYK2. Some exciting insights into JIA risk pathways have been made, e.g. the IL-2 pathway is now implicated in JIA, with multiple gene regions (IL2RA, IL2/IL21 and IL2RB) confirmed susceptibility loci for JIA [28]. This pathway is important in T cell activation, and the dependence of regulatory T cells on IL-2 suggests its importance in the maintenance of immune tolerance.

The Immunochip study has dramatically increased our knowledge of JIA susceptibility loci from only a handful of regions to 16 non-HLA regions that have shown an association with the oligoarticular and RF-negative polyarticular subtypes of JIA at the genome-wide significance threshold (PTPN22, STAT4, ANKRD55, IL2-IL21, TYK2, IL2RA, SH2B3-ATXN2, ERAP2-LNPEP, UBE2L3, CSOrf6-IRF1, RUNX1, IL2RB, ATTP8B2-IL6R, FAS, ZFP36L1) [28]. A further 11 regions were identified in the JIA Immunochip study at \( P < 1 \times 10^{-6} \) and as such represent valid candidates for further investigation (LTBR, IL6, COG6, Chr13q14, CCR1-CCR3, PRR5L, PRM1-RM12, RUNX3, TIMMCD1-CDB80, JAZF1, AFF3-LONRF2) [28]. An additional five regions have been previously associated in more than one cohort of JIA (ANGPT1, VTCN1, TNFAIP3, TAF1/C5, JMD1/C0) [17, 22, 23, 44, 45] and also require validation.

Recent articles suggest the phenomenon of overlapping autoimmune loci is not surprising, with \(~40\%\) of genes associated with immune-mediated phenotypes being pleiotropic [12, 46]. This is further supported by the finding that disease-associated variants systematically alter transcription factor recognition sequences, which in related categories of disease such as AIDs form networks of linked genes, suggesting that shared genetic liability may underlie related categories of disease [47]. In light of the Immunochip findings, performing similar analyses utilizing this higher-density SNP data may enhance our ability to place JIA genetics in the context of all AID genetics and help identify common disease pathways, leading to common treatment options for AIDs.

Despite the great advances in our understanding of complex diseases since the introduction of GWAS technology, for JIA, as for many other diseases, we can currently only explain very little of the genetic risk of disease. This has traditionally been calculated using family data, which are unavailable in the larger case–control studies typically conducted in recent times. The development of statistical methods to calculate heritability at a population level by considering all SNPs simultaneously from GWAS data has lead to the ability to explain 20–45\% of phenotypic variability of traits such as height and BMI [48, 49]. This approach has recently been used to estimate that approximately one-third of JIA risk is attributable to common genetic variation [22, 28], highlighting the importance of searching for additional genomic loci outside AID regions and suggesting a role for other factors such as rare genetic variants and gene–gene interactions.

**JIA subtype associations**

Childhood arthritis is in fact distinct disease subtypes classified under the umbrella term of JIA. It is expected that the known heterogeneity in clinical presentation and long-term outcomes of JIA will be due in some part to underlying genetic differences between the ILAR subtypes. Limited genetic research has been performed on the individual subtypes due to the inevitable reductions in sample size and the corresponding decrease in power to detect weaker genetic effects, and issues regarding multiple testing. With this in mind, it is important to continue collecting larger cohorts of JIA for subtype-focused studies.

There have been small candidate gene studies focussing on particular JIA subtypes. For example, a SNP in the IL23R gene known to be involved in adult psoriasis and PsA has recently been seen in only the juvenile PsA subtype of JIA [50]. Similarly, a SNP in the ERAP1 gene, known to be involved in the adult condition AS, has been associated with only the similar enthesis-related arthritis (ERA) subtype [50]. Another study in the ERA subtype investigated polymorphisms within the IL1T gene and did not find them to be associated with susceptibility to this subtype in their Indian population, although this was a relatively small sample size of 94 cases and only tested two SNPs and one variable number tandem repeat at this locus [51].

It is thought that systemic JIA is particularly different from the other subtypes due to its lack of a strong MHC association, presence of cytokine dysregulation and various innate immune system abnormalities [52]. It has already been seen that systemic JIA is associated with SNPs within genes such as IL10 [53], IL6 [17, 54, 55] and SLC26A2 [56]. While our understanding of the genetic susceptibility to oligoarticular and RF-negative polyarticular JIA is rapidly improving due to recent focus and large,
well-powered studies [28], systemic JIA remains a relatively poorly understood subtype. However, this is being addressed in a large multi-national GWAS of systemic JIA that is currently under way. While there is clearly a need for increased focus on the genetics of the rarer JIA subtypes, our current understanding already begins to paint a picture of their distinct genetic landscapes.

Genetic predictors of outcome

Despite significant improvements in the management of children with JIA, for many the likelihood of long-term disease activity remains high [57]. It is hoped that by studying the genetics of JIA outcomes, not only will we increase our understanding of the pathology of the disease, but it will enable us to identify those children likely to go on to experience more severe long-term outcomes earlier in their disease course, allowing paediatricians to provide more targeted care prior to the development of long-term disability.

In children with JIA it is clinically important to understand the genetic determinants of disease severity and other long-term outcomes such as pain. However, the vast majority of genetic research in JIA to date has aimed to identify variants that affect the risk of developing JIA or pathways modulating drug response in JIA. Some examples of genetic research into JIA outcomes include a study that found SNPs in the IL6 gene associated with pain [58], a protective effect against joint space narrowing on radiographs in the TGF-β1 gene [58] and in the ERA subtype, the presence of HLA-DRB1*08 predicts failure to attain disease remission [59]. These studies were performed with small sample sizes and have yet to be replicated independently. A GWAS in a large cohort would be an ideal approach to continue looking for genetic associations with various key long-term outcomes such as pain and disability.

This work requires the availability of both DNA samples and detailed long-term information on multiple key outcomes in children with JIA. With these factors in mind, a number of prospective studies are being established, including the Childhood Arthritis Prospective Study (CAPS), which is collecting samples from children in five centres across the UK and co-ordinated by the Arthritis Research UK Epidemiology Unit at the University of Manchester [60]. CAPS has been under way for a decade, with comprehensive data available on ~1000 children who are being followed up for 5 years. The Research in Arthritis in Canadian Children emphasizing Outcomes (ReAccH-Out) study is an ongoing multicentre prospective inception cohort conducted at 16 paediatric rheumatology centres in Canada and is primarily aimed at assessing JIA outcomes using epidemiological methods [61]. Similarly, another team of researchers has been collecting a long-term cohort of Nordic JIA samples [62–64]. In Australia the Childhood Arthritis Risk Factor Identification Study (CLARITY) is collecting biospecimens and extensive information about environmental factors for future outcome and genetic studies [65]. These international JIA outcome studies will be vital resources enabling us to answer some

very interesting questions, not only from a genetic perspective, but also incorporating general epidemiological data on disease presentation and course, treatment patterns and psychological aspects of disease.

Pharmacogenetics in JIA

The first-line treatment for JIA is MTX, which has been the focus of the majority of the pharmacogenetics work conducted to date. MTX is a folic acid antagonist, and although it has been used extensively in low doses to treat JIA in the last few decades, there still remain questions as to its mechanism of action and why a diverse range of patient responses is observed [66]. There are no well-validated predictors of MTX efficacy in JIA, leading to clinicians treating patients by trial and error rather than scientific evidence [67]. With ~35% of children failing to respond to MTX [68], and the notable time lag between initiating treatment and the first indications of patient response level, it would undoubtedly be beneficial for these children to be identified early and switched to alternative treatment options faster, thus minimizing their time to remission. It is hoped that this is an area where genetics research will be of benefit and ultimately lead to the development of simple diagnostic tools to assist clinician’s treatment decision making.

There has been more research into the association of the MTX metabolic pathway genes and toxicity or response in adult RA than in JIA, although with variable results and commonly without replication. This work has played a large role in informing similar work in JIA, and although limited, the pharmacogenetics research in JIA has begun to provide some important insights. Recent work in a Dutch cohort of 287 children has identified associations with MTX response over the first year of treatment in two proteins from within the ATP-binding cassette transporter superfamily (ABCB1 and ABCC3), and to a lesser degree the SLC19A1 gene [69]. Other studies have associated the MTHFR gene with MTX response and tolerability [70, 71], but have not found associations in the ATIC, AMPD1, ITPA or MTHFD1 genes [71]. However, SNPs in ATIC and ITPA were associated with increased risk for non-response in a more recent yet similar size study (~200 patients) [72]. It is apparent that the results from these studies have been inconclusive, likely due to the lack of power afforded by the small sample sizes available and the lack of appropriate validation in independent cohorts.

Researchers have recently published a prediction model incorporating both clinical and genetic data to predict whether a child will respond to MTX in their first year of treatment [73]. A derivation cohort of 183 patients was used to build the model via a backwards variable selection approach, with the dichotomized ESR and four SNPs (one from each of the MTRR, MDR-1/ABCB1, MRPI/ABCC1 and PCFT genes in the MTX pathway) combining to give the best predictive power. Using calculated predicted probabilities for MTX non-response, they found their model had 72% predictive power [73]. Although limited by sample size and only moderate predictive power in
the validation cohort (65%), this study shows that the use of genetic data provides a means to predict drug response in JIA.

To expand this work, a GWAS of children receiving MTX is underway using samples collected from multiple European countries via the Childhood Arthritis Response to Medication Study (CHARMS)–JIA GWAS International Consortium. The larger sample size available will help alleviate the power concerns of previous studies, and the global nature of the GWAS will enable the identification of genomic loci outside the MTX metabolism and folate pathways.

Future of genetics research in JIA

Although many genetic loci have been reproducibly associated with AIDs, identification of a locus is not the end of the story. To date, the common SNPs associated with disease explain relatively little of the genetic predisposition. Recently this phenomenon has been labelled the missing heritability problem, and has caused debate within the genetics community as to its cause [74]. There are several possible explanations for this missing heritability in common diseases. First, the causal SNP in most regions has not yet been identified and thus effect sizes may be currently underestimated. Second, some regions contain as yet unidentified multiple independent susceptibility variants, including rare variants, which together may increase the genetic proportion explained by an individual locus [75]. The ImmuNochip study has already begun to clarify regions containing multiple independent effects in JIA (STAT4, IL2RA, PTPN2), however, further work is required to confirm these [28]. Finally, other types of variation within the region may also be important in disease causation as well as other heritable changes such as DNA methylation, histone acetylation and microRNAs. Detailed discussion of these elements is beyond the scope of this review, however, all require thorough investigation in order to fully understand the underlying genetic architecture of a locus before embarking on the crucial phenotype–genotype characterization.

The GWAS era has provided us with an unprecedented increase in genetic knowledge, and for many complex diseases the results of multiple GWASs have been thoroughly mined, thus attention is now turning to the development and utilization of technologies available for resequencing the entire genome. This will have the advantage of identifying rare variants previously not well captured on GWAS chips. Open debate continues as to whether most disease-causing alleles are common or of low frequency [76]. Those supporting the large-scale resequencing efforts suggest that common complex diseases are in fact the result of many different rare variants with large effect sizes that were previously undetectable at a population level. Many researchers are hoping this will provide insights into disease pathogenesis and explain some of the missing heritability described above.

Refining JIA subtypes

With recent debate over the utility of the ILAR classification system for JIA [77, 78], an emerging goal for JIA research is to develop definitions of subgroups of clinically homogeneous JIA based on the genetic, clinical and biochemical presentation of the child [78]. Defining subgroups of JIA as such should allow for clearer treatment pathways. In an approach known as reverse phenotyping, key factors such as demographic, clinical and genetic variables can be used to understand a child’s probability of particular outcomes and success on treatments such as MTX. This will require the continued cooperation of paediatricians and researchers to increase the collection of JIA cohorts with a multitude of long-term data and samples.

Conclusions

It is a very exciting time in JIA research, where recent improvements across the broad fields of genetics, immunology and imaging research are enabling us to better understand JIA. There is still much work to be done until we have a comprehensive understanding of the genetic architecture of JIA. However, this review has shown that international collaborations enabling the collection of larger sample cohorts and the increasing availability of high-throughput genotyping technologies have revolutionized our knowledge of JIA susceptibility. The coming years will prove to be illuminating for JIA; important pathways involved in disease will be identified, genes will be implicated in outcomes such as disability and pain, and genetic predictors of response to treatments such as MTX will be found. This research will help us towards our ultimate goals of predicting outcomes and better targeting treatments, thus improving JIA remission rates.

Rheumatology key messages

- Current JIA genetics knowledge is primarily from studies of regions associated with AIDs.
- Larger genome-wide studies are required to understand JIA susceptibility and long-term outcomes.
- Future research will combine genetics, clinical and biochemical data predicting JIA outcomes and refining subtypes.

Acknowledgements

This report includes independent research supported by or funded by the National Institute for Health Research Biomedical Research Unit Funding Scheme. The views expressed in this publication are those of the authors and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health.

Funding: This work was supported by Arthritis Research UK (grant number 17552) and the National Institute for Health Research Biomedical Research Unit Funding Scheme.
Disclosure statement: The authors have declared no conflicts of interest.

References


60 Hyrich KL, Lair SD, Foster HE et al. Disease activity and disability in children with juvenile idiopathic arthritis one year following presentation to paediatric rheumatology. Results from the Childhood Arthritis Prospective Study. Rheumatology 2010;49:116–22.


