Introduction

The much-cited landmark publication by the WTCCC (Wellcome Trust Case Control Consortium [1]) in 2007 utilized genome-wide association studies (GWAS) to investigate the genetic basis of seven common diseases, including rheumatoid arthritis (RA). These studies were performed in a British population with approximately 2,000 individuals with each disease compared to a shared set of 3,000 healthy controls. For RA, three association signals reaching $P<5\times10^{-7}$ were revealed. While paediatric diseases were not investigated in that seminal study, the methodology has since been embraced by the paediatric rheumatology research community. This powerful tool has recently resulted in new insights into disease susceptibility and response to therapy. Details of GWAS experimental design and statistics are available in an excellent review [2].
Like the diseases in the WTCCC GWAS, many conditions in paediatric rheumatology are of unknown etiology, with evidence for a complex interaction of both genetic and environmental factors. One challenge of performing GWAS has been the relatively low incidence of paediatric rheumatic diseases, which reduces the statistical power to detect population differences between case-control cohorts. The most common childhood rheumatic disease is juvenile idiopathic arthritis (JIA), with a prevalence in high-income countries of 16 to 150 cases per 100,000 population [3] compared to approximately 1% of adults with RA [4]. JIA and Kawasaki Disease (KD) are among two paediatric rheumatic diseases where GWAS results have been validated in a replication cohort, as summarized in table 1. Figure 1 presents an overview of the diversity of genetic regions associated with JIA, KD, and the rare condition juvenile dermatomyositis (JDM), which will be discussed to exemplify the successful approaches that have been made in understanding the genetic basis of paediatric rheumatic diseases. In addition, alternative approaches to GWAS and future directions for the clinical utility in understanding the genetic basis of disease are outlined.

**Juvenile Idiopathic Arthritis**

*Defining JIA – A Heterogeneous Disease*

JIA is defined as joint inflammation of unknown origin that lasts for more than 6 weeks and starts before the age of 16 [3]. Strong evidence indicates JIA to be a complex genetic trait [5]. Using a large, well-documented Utah population, one estimate of the relative risk of JIA in siblings of patients is 11.6 [6]. Elucidating the genetics of JIA is challenging due to the collective heterogeneity of diseases that meet the above definition [7]. Different approaches have been used when investigating the genetics of JIA, each with advantages and limitations that are expertly reviewed in [8]. Analyzing all JIA subtypes as one phenotype assumes common disease pathways across the subtypes and increases sample size; however, another successful approach has been to stratify patients into homogeneous populations based on disease subtype. This latter approach minimizes cohort variability and facilitates detection of case-control differences. The classification used by the International League of Associations for Rheumatology (ILAR) [9] remains the gold-standard for defining patient subtype in both clinical practice and scientific research. Further classifications based on molecular phenotypes are gaining ground [10] and may provide a valuable approach when sufficient numbers of cases can be studied. For example, high-resolution HLA typing revealed that the HLA-DRB1*0801 haplotype is associated with disease in young onset patients with polyarticular JIA [11].

**JIA Genetics**

The first JIA GWAS included all subtypes and confirmed association with HLA and a gene that encodes for the B7-H4 costimulatory protein, VTCN1 [12] with samples sizes shown in table 1. Combining specific disease subgroups has been beneficial in a recent GWAS, where analysis of oligoarticular and polyarticular JIA cases revealed novel candidate loci, including the histone demethylase gene JMJD1C and a region including CD80 [13]. These two JIA GWAS are discussed in more detail in [14].

The largest collaborative study to date identifying JIA genetic risk factors comprised JIA cases and controls from the United States, United Kingdom and Germany [15]. Genotyping was done on the Immunochip platform, which provided dense SNP coverage of regions previously implicated by GWAS (P<5x10^{-8}) in autoimmune diseases, including RA and multiple sclerosis [16]. This knowledge allowed content to be focused on regions of importance to autoimmunity, providing advantages in both cost and SNP density. Nearly 200,000 single nucleotide polymorphisms (SNPs) representing only about 200 regions were assayed at 1/10 the cost of a genome-wide SNP array, which typically covers 1 million SNPs representing approximately 20,000 regions. This landmark study was limited to the phenotypically-related oligoarticular and polyarticular rheumatoid factor (RF) negative JIA subtypes to confirm disease susceptibility loci for HLA, PTPN22 and PTPN2 and reveal 14 new loci reaching P<5x10^{-8}, moving the field closer to understanding the mechanisms of disease susceptibility. The loci that reached genome-wide significance are shown in Figure 1 and listed in order of statistical significance: HLA-DQB1/HLA-DQA2, PTPN22, STAT4, PTPN2, ANKRD55, IL2/IL21, TYK2, IL2RA, SH2B3/ATXN2, ERAPE/LNPEP, UBE2L3, C5orf56/IRF1, RUNXI, IL2RB, ATP8B2/IL6R, FAS and ZFP36L1. While each of these loci are named...
Table 1. GWAS cohort size and genetic associations for Juvenile Idiopathic Arthritis (JIA), Kawasaki Disease (KD) and Juvenile Dermatomyositis (JDM).

<table>
<thead>
<tr>
<th>Disease</th>
<th>GWAS reference</th>
<th>GWAS Discovery Cohorts</th>
<th>GWAS Replication Cohorts</th>
<th>Genetic Associations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Cases (n)</td>
<td>Controls (n)</td>
<td>Cases (n)</td>
</tr>
</tbody>
</table>

Cohort size in discovery and replication cohorts are presented for JIA, KD and JDM. For JDM a replication cohort was not available (N/A).

Figure 1. Human karyotype showing diversity of genetic regions of association in JIA, KD and JDM. Genetic associations for the recent JIA Immunochip study [15], and GWAS studies in KD [17,18, 19] and JDM [24]. The MHC region is associated with susceptibility to all three diseases.
by the most likely candidate gene in the region, further studies are necessary to identify the causal SNPs and the mechanisms driving these associations. The relationships between these genes, which demonstrate variety in cellular location and molecular function, are shown in Figure 2. The overlap between lead SNPs in JIA and other autoimmune diseases assessed with the Immunochip are depicted in Figure 3.

**Kawasaki Disease**

Kawasaki Disease (KD) is an acute systemic vasculitis that occurs mainly in children. Japan, Korea, and Taiwan have the highest annual incidence rate for KD with several GWAS completed in these populations, as listed in table 1. A Japanese study revealed BLK, HLA and CD40 as susceptibility loci [17], and BLK and CD40 were also associated in a cohort of Han Chinese descendants in Taiwan [18]. A large GWAS combining five independent sample collections from patients with Asian and European descent revealed two loci, including a functional polymorphism in the IgG receptor FCGR2A [19].

Complications of KD are a major cause of acquired heart disease in children in developed countries. A recent paper by Kim [20] compared patients of Korean ancestry with and without coronary artery lesions (CALS) and showed that a SNP in the potassium channel KCNN2 was associated with CALs in KD patients. These findings were replicated in a second cohort. This approach demonstrates the utility of subdividing patients into clinically homogeneous subtypes.

**Environmental Influence in KD**

The interplay between genes and environment in patients with KD has been recently highlighted by an intriguing hypothesis by Rodó and colleagues published in 2014 that builds on their previous publication in 2011 [21,22]. The studies presented computer simulation data of wind air flow that indicated a source region in northeastern China that correlates with patterns of children in Japan emerging with KD symptoms between 1970 and 2010. Candida species were found in air samples between Japan and northeastern China.
Figure 3. Immunochip SNPs significantly associated with JIA: a comparison with other autoimmune diseases. Immunochip SNPs that have met genome-wide significance ($P < 5 \times 10^{-8}$) with association to JIA [15] are listed. Strength of association of these SNPs with susceptibility to other autoimmune diseases assayed by the Immunochip are indicated. Only the lead SNP for each annotated gene is presented, and the annotation for each SNP is named based on the most likely candidate gene in the region. Where multiple SNP functions exist, the main function is listed in order of importance: Non Synonymous (NS) > Synonymous (S) > Intronic (I) > Gene Region (GR). Data derived from the online database Immunobase are current as of May 2014.
that reproduced KD-like symptoms in susceptible mice, including vasculitis and elevated IL-6, TNFa and IFNγ. The authors propose that an airborne etiologic agent, such as a *Candida* fungal toxin, is an environmental KD trigger that is influenced by host genetics, leading to disease in susceptible individuals. This work implicates a novel trigger for KD. Further research will be needed to find the causative agent of KD globally and how host genetics leads to disease susceptibility upon exposure to this putative trigger.

**Juvenile Dermatomyositis**

The disease incidence for juvenile dermatomyositis (JDM) is an estimated 2 to 3 children per million per year [23]. The Myositis Genetics Consortium was formed by investigators with collections of myositis DNA samples and recently published the first GWAS. The disease cohort included both adult and juvenile DM patients of European ancestry with cohort size in table 1 [24]. 141 SNPs outside the MHC region and previously associated with autoimmune diseases were investigated, but none reached genome-wide levels of significance using the rigorous Bonferroni multiple testing correction. However, SNPs representing *PLCL1*, *BLK* and *CCL21* gene regions were associated when a false discovery rate (FDR) of <0.05 was used and overlap with association signals in systemic lupus erythematosus (SLE) and RA.

Unlike JIA, where disease is clinically different from adult RA, and KD which is mainly a disease of childhood, JDM has clinical overlaps with the adult disease. Including all age groups increased sample size in this rare disease, and the study was expedited by the involvement of international research consortia to coordinate collection of large numbers of patient DNA samples and clinical data.

**Monogenic Autoinflammatory Diseases**

The genetic basis has been more clearly defined for many of the autoinflammatory diseases, including one of the more commonly inherited fever syndromes, tumor necrosis factor receptor-associated periodic syndrome (TRAPS). Mutations in *TNFRSF1A* have been shown to cause TRAPS in an autosomal dominant pattern of inheritance, where only one aberrant copy of the gene is necessary for disease. The gene products of *TNFRSF1A* are tumor necrosis factor receptor 1 (TNFR1) proteins. Mutations often result in misfolded TNFR1 proteins that are thought to cause over-stimulation of inflammatory pathways, bypassing normal TNF signaling and resulting in recurrent fevers [25]. The Eurofever/EUROTRAPS international registry recently published data from TRAPS cases on disease heterogeneity and phenotype, with the R92Q variant being the most common, as found in 34% of cases [26]. Genes causing monogenic autoinflammatory diseases including, TRAPS, Familial Mediterranean Fever (FMF), and Blau syndrome are listed in table 2. The clinical symptoms of these diseases are reviewed in [27, 28].

**Beyond GWAS: Alternative Approaches of Investigating Genetics**

In paediatric rheumatology, patient sample size and experimental costs can be significant barriers to conducting standard GWAS. However, alternative approaches can yield great results as seen with the Immunochip in JIA. The most promising future strategies for the paediatric rheumatology field are discussed below.

**Exome Sequencing**

The advent of high-throughput sequencing has dramatically reduced the cost of individual whole-genome or -exome sequencing (WGS and WES, respectively) [29]. WGS and WES are technologies employed by both researchers and clinicians searching for genetic underpinnings of disease. WGS involves determining the nucleotide sequence at every position in the genome, whereas WES only reveals sequences within the protein-coding regions (roughly 1%) of the genome, enabling rapid detection of nonsynonymous mutations [30]. Decreased turnaround time and cost burden makes WES an attractive choice over WGS, despite the fact that WES often fails to detect sites regulating transcription or splicing events [31]. Whereas GWAS discern common variants that predispose individuals to disease, WES offers an avenue to detect new and rare variants that could contribute to disease pathogenesis. For example, WES on individuals with cardiovascular disorders detected rare variants with functional consequences on disease pathogenesis [32]. For autoinflammatory syndromes, WES promises to be of increasing clinical utility for identifying the molecular basis of these challenging diseases, enabling enhanced clinical care of the patient in this new era of precision medicine.
**Table 2. Summary of monogeneic autoinflammatory diseases, the affected gene and protein product.**

<table>
<thead>
<tr>
<th>Disease</th>
<th>Gene</th>
<th>Affected Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumor Necrosis Factor Receptor-Associated Periodic Syndrome (TRAPS)</td>
<td>TNFRSF1A</td>
<td>Tumor Necrosis Factor Receptor 1 proteins</td>
</tr>
<tr>
<td>Familial Mediterranean Fever (FMF)</td>
<td>MEFV</td>
<td>Pyrin</td>
</tr>
<tr>
<td>Familial Cold Autoinflammatory Syndrome</td>
<td>NLRP3</td>
<td>NLRP3</td>
</tr>
<tr>
<td>Muckle-Wells Syndrome</td>
<td>NLRP3</td>
<td>NLRP3</td>
</tr>
<tr>
<td>Neonatal Onset Multisystem Disease</td>
<td>NLRP3</td>
<td>NLRP3</td>
</tr>
<tr>
<td>Blau Syndrome</td>
<td>NOD2</td>
<td>NOD2</td>
</tr>
<tr>
<td>Mevalonate Kinase Deficiency (MKD)</td>
<td>MVK</td>
<td>Mevalonate kinase</td>
</tr>
<tr>
<td>Familial Cold Autoinflammatory Syndrome 2</td>
<td>NLRP12</td>
<td>NLRP</td>
</tr>
<tr>
<td>Deficiency of IL-1 Receptor Antagonist</td>
<td>IL1RN</td>
<td>IL-1 receptor antagonist</td>
</tr>
<tr>
<td>Pyogenic Arthritis, Pyoderma Gangrenosum and Acne (PAPA) Syndrome</td>
<td>PSTPIP1</td>
<td>PSTPIP1</td>
</tr>
<tr>
<td>Majeed Syndrome</td>
<td>LPIN2</td>
<td>Phosphatidate phosphatase LIPN2</td>
</tr>
</tbody>
</table>

**Identifying Novel Pathways via de novo Mutations**

WES is often used to identify variants within an individual that could be causal or contribute to disease by sequencing family trios, which consist of the mother, father, and affected child. By sequencing each trio member, de novo mutations found within the affected child, but neither parent, may be identified. On average, an estimated 74 germline de novo mutations occur with each generation [33]. The detection of these mutations is a powerful tool to identify rare variants that by themselves, or in combination with other variants, have large effects in common pathways of disease and immune modulation.

Rare variants within candidate genes revealed by GWAS influence the risk of developing autoimmune inflammatory bowel disease and type 2 diabetes [34, 35]. Moreover, de novo mutations discovered among trios for autism spectrum disorders (ASD) have been shown to increase risk of conferral for pathogenesis; genes whose functionality was altered by de novo mutations were biologically related to one another or to genes previously associated with ASD [36]. This supports the fact that de novo mutations and rare variants contribute to the overall genetic influence on disease risk that are undetected by large-scaled GWAS. WES thereby serves as a technology with which the genetic complexity underlying diseases can be further elucidated.

**Transethnic Mapping**

To date, many genetic association studies have been performed to identify genetic variants, which increase disease risk in a European ancestral subject cohort. There remains a need for studies that investigate the disease risk for all affected ethnic groups both to ensure global health benefits and to expedite causal variant discovery. The evolutionary history of the African genome has resulted in more recombination events and significantly less linkage disequilibrium between nearby genetic variants [37]. Specifically, the average haplotype block in the European genome is 22 kilobases (kb) long, while the average African genome block is only 11 kb [38]. In order to combat this limitation, it is preferable to genotype both cases and controls of other ancestries to facilitate transancestral fine mapping. First, the association analyses using the new cohorts are performed as a true replication cohort on SNPs that reach genome-wide or suggestive significance in the European cohorts [4]. Then, loci with replicated association are used to identify the physical boundaries of the association and the variants most likely to be causal. These studies hold great promise for paediatric rheumatic diseases because including...
cohorts with African, Asian, or Native American/Hispanic ancestry may further refine the associated region and decrease the number of candidate variants for biological testing [38,39].

**Response to Medication**

WES is already used as a clinically useful diagnostic tool for autoinflammatory diseases; with more understanding of the genetic response to medication, WES can also be used to deliver targeted drug therapy based on patient genotype. For example, methotrexate (MTX) is the main course of therapy for patients with JIA, but not all children respond. A range of more expensive biological therapeutics is available, and early, aggressive therapy can yield beneficial results [40]. Therefore, rapid selection of the best treatment is advantageous.

A candidate gene approach was used to define SNPs that predict MTX response status in a Dutch cohort [41]. These candidate SNPs are now being tested in larger cohorts. The first large JIA study with genome-wide coverage revealed 14 regions (P<1x10^{-5}), including a region that contains TGIF1, associated with MTX response. [42]. TGIF1 is of particular biological interest as it mediates downregulation of the aryl hydrocarbon receptor, a transcription factor important in regulatory T cell and Th17 differentiation [43]. The authors used an innovative approach to assess the genetics of response to medication by testing for genetic association with individual core outcome variables. This demonstrated heterogeneity in the genetic contributions to MTX response of the individual parameters that comprise the ACR response criteria and juvenile arthritis disease activity score (JADAS) [42]. This reductionist methodology can be applied to investigate therapeutic response in other diseases or medications. Future studies that combine genetic data with serum protein markers of response, such as MRP8/14, a protein elevated prior to MTX treatment in a subgroup of JIA patients who subsequently responded well to MTX, may further help distinguish responders from nonresponders [44].

A number of advances in adult rheumatology may inform future directions for pediatric disease, including a GWAS for response to IL-6 treatment that implicates 8 loci [45]. While further functional analysis is required, these regions implicate genes such as CD69 and CLEC2D, which are involved in multiple pathways, including hematopoietic cell proliferation, IL-2 and IFNγ production, as well as NK, T, and B cell activation. Therapeutic blockade of TNF has improved the outcome for many children with JIA, but not all. Using systems biology approaches, similar to what has been done for the TNF signaling pathway [46], to computationally model the aberrant signaling pathways in disease may lead to novel and intriguing targets for the suppression of inflammation and guide future direction in the development of new therapies.

**GWAS to Drug Targets: An Example From Rheumatoid Arthritis (RA)**

The contrast between understanding the genetic basis of adult and paediatric disease is exemplified by arthritis, where numerous well-powered GWAS studies exist for RA but not JIA. A meta-analysis of the findings of 22 GWAS included more than 100,000 subjects of European and Asian ancestries and revealed 42 novel RA risk loci at genome-wide significance [47]. This demonstrates the necessity of vast cohort sizes in the discovery of risk genes with small effect sizes. ‘Biological RA risk genes’ were characterized using a pipeline that combines data from multiple sources, including functional annotations, peripheral blood mononuclear cell gene expression data, knockout mouse phenotypes, and molecular pathway analysis to score genes. There was a significant enrichment of drug targets within the ‘biological RA risk gene’ set, with more than a quarter being known drug targets. This systematic approach includes an innovative pipeline with the potential to expedite drug discovery and ultimately benefit patients with rheumatologic diseases.

**Concluding Remarks**

Informative genetic analysis of relatively rare paediatric rheumatic diseases has been achieved through cooperative research efforts. This genomic era provides a foundation for greater mechanistic understanding of disease etiology and treatment selection. It remains a priority to understand the genetic basis of these diseases as distinct from adult conditions and to include global populations both to expedite analysis and provide widespread population benefits.

**Acknowledgments**

We would like to thank Dr. Susan Thompson for her critical reading and helpful suggestions on this manuscript.
Competing Interests
The authors have no financial conflicts of interest.

References
1. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007; 447:661-78.
21. Rodó X, Curcoll R, Robinson M, Ballester J, Burns JC, Cayan DR, et al. Tropospheric winds from northeast-
ern China carry the etiologic agent of Kawasaki disease from its source to Japan. Proc Natl Acad Sci U S A. 2014; 111:7952-7.


44. Moncrieffe H, Ursu S, Holzinger D, Patrick F, Kassoumeri


