
This is the unspecified version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: http://openaccess.city.ac.uk/3784/

Link to published version: http://dx.doi.org/10.1038/tpj.2014.3

Copyright and reuse: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.
ORIGINAL ARTICLE

Genome-wide data reveal novel genes for methotrexate response in a large cohort of juvenile idiopathic arthritis cases

J Cobb1,2,12, E Cule3,12, H Moncrieffe4,12, A Hinks1, S Ursu4, F Patrick4, L Kassoumeri4, E Flynn1, M Bulotovic2, N Wulfraat5, B van Zelst6, R de Jonge6, P Dolezalova7, S Hirani8, S Newman8, P Whitworth9, TR Southwood9, M De Iorio10, Childhood Arthritis Response to Medication Study (CHARMS), Childhood Arthritis Prospective Study (CAPS), BSPAR study group14, LR Wedderburn4,11,13 and W Thomson1,2,13

Clinical response to methotrexate (MTX) treatment for children with juvenile idiopathic arthritis (JIA) displays considerable heterogeneity. Currently, there are no reliable predictors to identify non-responders: earlier identification could lead to a targeted treatment. We genotyped 759 JIA cases from the UK, the Netherlands and Czech Republic. Clinical variables were measured at baseline and 6 months after start of the treatment. In Phase I analysis, samples were analysed for the association with MTX response using ordinal regression of ACR-pedi categories and linear regression of change in clinical variables, and identified 31 genetic regions (P < 0.0001). Phase II analysis increased SNP density in the most strongly associated regions, identifying 14 regions (P < 1 × 10−5): three contain genes of particular biological interest (ZMIZ1, TGF1 and CFTR). These data suggest a role for novel pathways in MTX response and further investigations within associated regions will help to reach our goal of predicting response to MTX in JIA.

The Pharmacogenomics Journal advance online publication, 8 April 2014; doi:10.1038/tpj.2014.3

Keywords: juvenile idiopathic arthritis; methotrexate; pharmacogenetics; response

INTRODUCTION

Juvenile idiopathic arthritis (JIA) is a heterogeneous condition with a variable outcome and considerable ongoing disease burden.1 Studies indicate that functional disability and complications due to JIA are still common in many teenagers and young people with JIA, and that the effects of early, uncontrolled inflammation may cause irreversible damage to joints and other tissues.2–4 Thus, improving long-term outcomes of children with JIA remains a critical challenge. Recent studies in JIA have indicated that early control of joint inflammation correlates with much improved outcomes, suggesting an early ‘window of opportunity’ when disease control can translate to profound long-term benefit.5 It is known that not all children respond equally well to any given therapy. Despite increasing availability of newer therapeutic options for treating inflammation in JIA, clinicians have no validated tools to help predict likelihood of good response to a particular drug. Therefore, the current treatment strategy is to offer disease-modifying drugs in a sequential approach, with choices typically driven more by cost or safety profile than by scientific evidence.

The first-line disease-modifying agent for JIA is methotrexate (MTX). Although MTX has proven efficacy in randomised trials and a good long-term safety record, response to MTX displays considerable heterogeneity in JIA and a significant ‘non-response’ rate of 35% or more of cases.6 In addition, consensus concerning level of ‘response’ that is considered acceptable has shifted, with a target of complete control of inflammation now being advocated.7 These developments combined with ever increasing availability of newer biologic treatments, provide further imperative for the discovery of biomarkers to aid the identification of children who require early aggressive therapy, compared with those who can achieve clinically inactive disease on MTX alone.

Response to treatment is thought to be a complex trait involving multiple genetic variants and environmental factors.8 However, to date, genetic studies of MTX response have been limited for both JIA and rheumatoid arthritis (RA) and have only utilised a candidate gene approach, focussing largely on genes affecting MTX transport and metabolism, enzymes influenced by MTX and adenosine pathways,9–16 recently reviewed in ref. 17. Given that the mechanisms of action of MTX in JIA are poorly understood, candidate gene studies may miss key pathways of mechanistic importance. Pharmacogenetic studies in other diseases have shown that genes other than those directly involved

1Arthritis Research UK Centre for Genetics and Genomics, Institute of Inflammation and Repair, Faculty of Medical and Human Sciences, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK; 2NIHR Manchester Musculoskeletal Biomedical Research Unit, Central Manchester National Health Service Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK; 3Department of Epidemiology and Biostatistics, Imperial College London, London, UK; 4Rheumatology Unit, UCL Institute of Child Health, London, UK; 5Department of Paediatric Immunology, University Medical Centre Utrecht, Wilhelmina Children’s Hospital, Utrecht, The Netherlands; 6Department of Clinical Chemistry, Erasmus University Medical Centre, Rotterdam, The Netherlands; 7First Faculty of Medicine and General Faculty Hospital, Charles University in Prague, Praha, Czech Republic; 8Centre for Health Services Research, School of Health Sciences, City University London, London, UK; 9Institute of Child Health, Birmingham Children’s Hospital, Birmingham, UK; 10Department of Statistical Sciences, University College London, London, UK and 11Arthritis Research UK Centre for Adolescent Rheumatology, UCL Institute of Child Health, London, UK. Correspondence: Dr J Cobb, Arthritis Research UK Epidemiology Unit and NIHR Manchester Musculoskeletal BRU The University of Manchester Central Manchester Foundation Trust, Manchester Academic Health Science Centre, Oxford Road, Manchester M13 9PT, UK.
E-mail: joanna.cobb@manchester.ac.uk
12These authors contributed equally to this work.
13These authors made equal contribution as senior authors.
14See Appendix for the full list of authors.
Received 10 October 2013; revised 16 December 2013; accepted 19 December 2013

OPEN

The Pharmacogenomics Journal (2014), 1–9
© 2014 Macmillan Publishers Limited All rights reserved 1470-269X/14
www.nature.com/tpj
in known-drug pathways often have key roles in variation of drug response. In order to capture the genetic component more comprehensively, we brought together an International Consortium of investigators (the CHARMS-JIA GWAS International Consortium), and employed a genome-wide approach to study response to MTX in a large cohort of children with JIA.

MATERIALS AND METHODS

Study population
A cohort of children was recruited for the SPARKS-CHARM (Childhood Arthritis Response to Medication) Study, which has the overall aim to improve understanding of the variability in response to treatment observed in children with JIA and ultimately define a multifactorial model of response outcomes, and through the CHARMS-JIA GWAS International Consortium. All cases fulfilled the International League of Associations for Rheumatology (ILAR) criteria for JIA and were about to start new MTX treatment for active arthritis. The study had full ethical approval and was fully compliant with the Declaration of Helsinki; parents provided fully informed consent and patients provided age-appropriate assent. Samples and data were collected using the same inclusion and exclusion criteria at Great Ormond Street Hospital London, Birmingham Children’s Hospital, Department of Paediatrics and Adolescent Medicine Charles University Prague, Wilhelmina Children’s Hospital and University Medical Centre Utrecht, and also as part of the Childhood Arthritis Prospective Study (CAPS), a prospective inception cohort study of JIA cases from five centres across the UK.

A total of 759 individuals were included from all subtypes of JIA, classified according to ILAR criteria. Demographic and clinical data were collected at baseline (up to 4 weeks before beginning MTX treatment) and again after 6 months (median 6.2 months, range 4–8 months) of MTX treatment. MTX was given orally or subcutaneously at 10–15 mg·m⁻² per week (median 11.3 mg·m⁻² per week). No other DMARDs were taken concurrently. Steroid treatment was recorded if administered at any point between baseline and follow-up. A total of 370 children received at least one form of steroid medication: this comprised oral prednisolone in 204 children, pulsed iv methylprednisolone was taken by 99 children, whereas one form of steroid medication: this comprised oral prednisolone in 204 children, pulsed iv methylprednisolone was taken by 99 children, whereas

Genotyping
Samples were genotyped using the Illumina HumanOmniExpress infinium array, according to Illumina’s protocols in Manchester, UK. The default Illumina clustering algorithm (GenTrain2.0) was used to cluster SNPs in the software package GenomeStudio. SNPs were excluded if they had a call rate <98% and a cluster separation score of <0.4. Samples were then excluded for call rate <98%, incompatible recorded and genotype inferred gender, duplicates and evidence of identity by descent, or those with outlying heterozygosity. Combining the samples with data from HapMap Phase III individuals, principal component (PC) analysis was performed using Eigenspace v42 to identify extreme ethnic outliers. PC analysis was performed on a subset of SNPs with minor allele frequency ≥0.05, selected by removing SNPs in known regions of high linkage disequilibrium and further pruned for linkage disequilibrium between markers. Samples failing to cluster with European HapMap individuals were visually identified and removed. SNPs were excluded from the analysis if they had a minor allele frequency <0.05 and failed the Hardy–Weinberg equilibrium test (P ≤ 0.001). To assess our dataset for potential systematic over-inflation due to stratification, the genomic control inflation factor (λGC) was calculated using the same SNP subset as used in the PC analysis. Cluster plots were visually inspected for the most associated SNPs to confirm genotyping quality.

Statistical analysis
Data were available on a number of potential confounding variables: gender, sample collection centre, presence or absence of concurrent steroid treatment, age at treatment baseline, time to treatment, duration of treatment and ILAR subtype (grouped into three categories: (1) oligoarthritis: persistent and extended, (2) polyarthritis: RF-negative and positive, (3) psoriatic, enthesitis-related, systemic and unclassifiable arthritis). Each of these potential confounders were assessed for the association with each core-set variable. Moreover, presence of population stratification was checked using the first five PCs of the genotype dataset.

Statistical analysis was performed using Plink v1.07 and R v2.15 (http://www.r-project.org), and plots were generated using R and LocusZoom. SNPs were coded by minor allele count as 0, 1, 2. MTX response was defined using the ACR-pedi criteria with four categories: non-responders (reference category), ACR-pedi30, ACR-pedi50, ACR-pedi70, and association between genotype and MTX response was analysed using ordinal regression. Similar to other genetic studies of drug response measured by composite disease scores, we hypothesise that attempting to identify the underlying genetic basis of MTX response may be usefully performed by analysing each of the core-set variables individually, as it is likely that the genetic basis of each of these is different, with varying contributions to MTX response. As the core-set variables are not entirely independent from one another or the ACR-pedi status, the multiple testing burden is not as great as if we performed multiple tests on independent outcomes. Core-set variables (ESR, CHAQ, AJC, LJC, PhysVAS and ParVAS) were recoded as change between baseline and follow-up, and linear regression was used to assess the strength of association for each variable.

We conducted our analysis in two phases (I and II) with the same individuals in each analysis phase. We utilised a low stringency of significance in Phase I of the analysis in order to maximise discovery of loci for more detailed investigation in Phase II, where a more stringent significance threshold was set. In Phase I of the analysis, results from the ordinal regression of ACR-pedi categories and six linear regressions of the core-set variables (ESR, CHAQ, AJC, LJC, PhysVAS and ParVAS), totalling seven analyses, were used to identify genomic regions of interest for further investigation. The significance threshold selected (P < 0.001) allows for greater emphasis on power than reducing type I error to enable hypothesis generation, an approach taken previously.

Regions were then selected for further analysis (Phase II) by searching for clusters of associated SNPs (P < 0.001 in at least two of the seven analyses) and extending out to include all SNPs within the annotated gene (based on the Illumina HumanOmniExpress gene annotation file, hg19). The proportion of overlap in regions of interest of varying sizes (range 0.02 kb–12.8 Mb, average 1359 kb). The aim of Phase II was to refine these regions. This was performed using SNP imputation to increase the density of SNP coverage in those regions. SHAPEIT v1 was used to prepare phase genotypes and SNPs were imputed against the 1000 Genomes Project reference panel (~37 million SNPs) using IMPUTE2. Imputed SNP genotypes reaching the probability threshold 0.9 were included in the follow-up re-analysis, which focused on only these imputed regions using the same samples as Phase I, and performed ordinal and linear regressions as described in Phase I. Regions containing at least one SNP in Phase II with association P < 1 × 10⁻⁵ are the focus of the results presented here.

Power calculation
Study power was estimated at the two significance thresholds used in Phase I (P < 0.001) and Phase II (P < 1 × 10⁻⁵) of the analyses, over the range of sample sizes available, and assuming the variance explained by the additive effect of the SNP tested ranged from 0.01–0.1 under an additive genetic model.

Functional annotation
In order to gain a better understanding of the potential biological impact of our results, the most highly associated SNPs identified in Phase II (as well as SNPs in high linkage disequilibrium (r² > 0.8)) were queried using the web tool Assimilator (http://assimilator.mhs.manchester.ac.uk/cgi-bin/assimilator.pl). This facilitates collation of functional annotations from the publicly available ENCODE and UCSC Genome Browser databases. Using the advanced search options available, output was focussed on
whether any of the Phase II-associated SNPs have shown evidence of transcription factor-binding sites, evidence for open chromatin suggesting regions of active gene expression, and epigenetic marks which may be affected by drug treatment.31

RESULTS
Following stringent quality control, 694 JIA cases were available for analysis (Supplementary Table 1A and 1B) comprising individuals from all ILAR subtypes (Table 1). In total, 31% of children were non-responders (less than ACR-pedi30 response). Among responders, categorised by their highest level of response achieved, 8.6% of children reached ACR-pedi30, 14.6% ACR-pedi50 and 45.8% ACR-pedi70 response. Samples clustered together in the PC analysis (Supplementary Figure 1) and therefore were analysed together. Following the SNP quality control steps and removal of low frequency variants (minor allele frequency < 0.05), 586,062 SNPs were included in the Phase I analyses (Supplementary Table 1). Quantile-quantile plots and inflation factors showed no systematic inflation of P-values (Supplementary Figure 2), and power was estimated to range from 10–100% across the analysed sample sizes for various effect sizes (Supplementary Figure 3).

None of the potential confounding variables tested (gender, ILAR–JIA subtype, centre, age at treatment baseline, duration of treatment, time to treatment, steroid treatment or PCs generated to identify ethnic outliers) were associated with all six individual core-set variables (data not shown); therefore to reduce loss of analysis power no adjustments were made to the linear or ordinal regressions.

In the hypothesis generating Phase I of the analysis, using both ACR-pedi and the individual core-set variables, 31 genetic regions encompassing 75 nearby genes achieved our defined level of significance (P < 0.001 in at least two of the seven analyses), Table 2 and Supplementary Table 2. This included several notable associations such as genes related to TGFbeta signalling (Table 2 and Supplementary Table 2). This included several notable SNPs with 13 SNPs in each associated at r² very high linkage disequilibrium (Supplementary Table 2). Two regions showed strong evidence of very high linkage disequilibrium (Figure 4). Overlapping associations of SNPs were revealed for AJC and LJC in several genetic regions (regions 16, 17, 23 and 28).

Phase II of analysis involved imputation of SNPs within the 31 genetic regions identified in Phase I to refine the association signals by increasing SNP density. After imputation, using the increased SNP density, a significance threshold of P < 1 × 10−5 this analysis identified 14 of the initial 31 genetic regions as the most strongly associated with MTX response (MTX Phase II, Table 3, Supplementary Figure 4). Overlapping associations of SNPs were revealed for AJC and LJC in several genetic regions (regions 16, 17, 23 and 28). In one of these regions (region 23, chromosome 11 intergenic between ANGPTL5-KIAA1377) the top-associated SNP was the same (rs11225055), and in the other three regions the most significant associated SNPs for the AJC and LJC analyses were in very high linkage disequilibrium (r² ≥ 0.97). In all four regions showing association with ACR-pedi (regions 2, 12, 20 and 24), the ParVAS and/or PhysVAS scores were also associated.

Functional annotations for the most highly associated Phase II SNPs from Table 3 (plus SNPs in high linkage disequilibrium
(r² > 0.8) with these lead SNPs were assigned using Assimilator software.\textsuperscript{30} The results presented in Supplementary Table 3 suggest the regions identified in Phase II of the analysis contain evidence of many markers of regulation and highlight many possible functional mechanisms. Certain regions were less fully covered by current databases, for example, region 14 containing the gene CSMO1. Others, including region 30 containing the gene CYTH4 (cytohezin 4), have more evidence for regulatory activity, including multiple SNPs showing evidence of acting as an expression quantitative trait loci.\textsuperscript{32}

**DISCUSSION**

Recent developments in treatments and management of childhood arthritis have lead to increased expectations from clinicians, parents and patients for complete control of disease and consequent reduction of long-term adverse health outcomes.\textsuperscript{4,5,33} The first step in JIA treatment, in parallel with joint injections, is typically administration of MTX; however, it is clear that a proportion of patients treated with MTX will fail to respond adequately. Given recent recommendations for early aggressive treatment, it is important that MTX treatment is targeted to those children most likely to respond well.\textsuperscript{34} Increasing our understanding of the influence of genetic variants in MTX response could assist clinicians to choose the best treatment options for their patients and identify patients who need more aggressive treatments. Performing large-scale genetic studies searching for variants contributing to MTX response has great appeal, but has proved challenging due to the relative rarity of JIA and lack of well co-ordinated international efforts. With this in mind, the CHARMS-JIA GWAS International Consortium facilitated the collection of co-ordinated international efforts. With this in mind, the CHARMS-JIA GWAS International Consortium facilitated the collection of

---

**Table 2.** Summary of the most highly associated regions identified in phase I of analysis, indicating the genes nearby each region and the total number of SNPs associated at P<0.001 and for which analysis

<table>
<thead>
<tr>
<th>Region</th>
<th>Chr</th>
<th>Region start (bp)</th>
<th>Region end (bp)</th>
<th>Nearby genes</th>
<th>Number of SNPs in region (across all analyses)</th>
<th>Analyses with associated SNPs (at P&lt;0.001)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>162918233</td>
<td>163039777</td>
<td>C1orf110, RG54</td>
<td>7</td>
<td>×</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>165139511</td>
<td>165309711</td>
<td>LMX1A, PBX1</td>
<td>3</td>
<td>×</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>209121677</td>
<td>209420998</td>
<td>PLXNA2</td>
<td>13</td>
<td>×</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>115447624</td>
<td>115532399</td>
<td>GAP43, LSAMP</td>
<td>8</td>
<td>×</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>126096912</td>
<td>126175401</td>
<td>CCD37, KLF15, ZFDC</td>
<td>13</td>
<td>×</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>144100011</td>
<td>145662790</td>
<td>C3orf58, PLOD2</td>
<td>15</td>
<td>×</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>30284836</td>
<td>31185714</td>
<td>CDH6</td>
<td>10</td>
<td>×</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>153846868</td>
<td>166676216</td>
<td>HAND1, SAP30L, LARP1, SGCD, KID43, HAVCR2, THG1L, ADRA1B, GABRG2, CCNG1, MAT2B, ODZ2</td>
<td>27</td>
<td>×</td>
</tr>
</tbody>
</table>

---

**Abbreviations:** AJC, active joint count; CHAQ, childhood health assessment questionnaire; ESR, erythrocyte sedimentation rate; LJC, limited joint count; PhysVAS, physician's global assessment on a visual analogue scale; ParVAS, parent/patient global assessment; SNP, single nucleotide polymorphism. This formed the basis of the regions selected for Phase II (some nearby regions were merged and others were expanded to comply with the minimum size requirements in IMPUTE2). Region coordinates based on the NCBI37 assembly. See Supplementary Table 2 for full results.
these genetic regions, additional SNPs were imputed in Phase II of
the analysis, with the results in 14 regions satisfying a more
stringent cut-off of $P < 1 \times 10^{-5}$. The most strongly associated locus
was \textit{CACNA1I}, which encodes the alpha chain of a low voltage-
activated calcium channel that has been implicated in calcium
signalling in neurons and may have other roles that have yet

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig1.png}
\caption{(a) Linear regression analysis of \textit{ZMIZ1} (region 20) for the change in ESR with MTX treatment, with the top hit in this region rs2802369 coloured purple. Similar results were seen for the ParVAS, CHAQ, and ACR-pedi analyses of this region. (b) Linear regression analysis of \textit{CFTR-CTTNBP2} (region 12) for parent’s global assessment (ParVAS), with the top hit in this region rs757278 coloured purple. Similar results were seen for the ACR-pedi and LJC analyses of this region. Coordinates are based on the NCBI36 assembly.}
\end{figure}
Table 3. Summary of most highly associated Phase II results (including imputed genotypes).

<table>
<thead>
<tr>
<th>Region</th>
<th>Total number of SNPs at ( P &lt; 1 \times 10^{-5} )</th>
<th>Best SNP Chr Position (bp) Minor allele Analysis</th>
<th>( N )</th>
<th>( \beta )</th>
<th>( SE )</th>
<th>95% CI</th>
<th>( P )-value</th>
<th>Genes within region</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>6</td>
<td>rs60624478† 1 165107979 A PhysVAS 413 1.17 0.26 (0.67, 1.67) 5.66 ( \times 10^{-6} )</td>
<td>LMX1A, PBX1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>rs75104673 5 29238587 C ESR 3520 -16.89 3.8 (-24.34, -9.44) 1.09 ( \times 10^{-4} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>31</td>
<td>rs757727† 7 117376745 T ParVAS 364 1.32 0.26 (0.81, 1.82) 5.37 ( \times 10^{-7} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>17</td>
<td>rs73185959 7 311618686 A PhysVAS 415 0.93 0.21 (0.39, 1.46) 1.25 ( \times 10^{-6} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>20</td>
<td>rs10113213 8 82888608 T AJC 604 -1.99 0.43 (-3.71, -0.27) 1.06 ( \times 10^{-6} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>4</td>
<td>rs10128264 10 80959973 A PhysVAS 415 0.97 0.23 (0.51, 1.43) 4.35 ( \times 10^{-5} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>2</td>
<td>rs4536282 12 993498282 A PhysVAS 415 0.97 0.23 (0.51, 1.43) 4.35 ( \times 10^{-5} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>rs313315 12 104965584 C PhysVAS 410 -0.745 0.18 (-1.09, -0.4) 3.21 ( \times 10^{-5} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>11</td>
<td>rs8196546 15 26804131 C PhysVAS 413 0.97 0.23 (0.51, 1.43) 4.35 ( \times 10^{-5} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>3</td>
<td>rs9954608 18 33878797 T ParVAS 364 1.2 0.33 (0.61, 1.79) 1.69 ( \times 10^{-6} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>1</td>
<td>rs229527 22 37591485 A PhysVAS 415 0.97 0.23 (0.51, 1.43) 4.35 ( \times 10^{-5} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>11</td>
<td>rs2294369 22 40075400 A PhysVAS 415 0.97 0.23 (0.51, 1.43) 4.35 ( \times 10^{-5} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: AJC, active joint count; CI, confidence interval; CHAQ, childhood health assessment questionnaire; ESR, erythrocyte sedimentation rate; LJC, limited joint count; PhysVAS, physician’s global assessment on a visual analogue scale; ParVAS, parent/patient global assessment; SNP, single nucleotide polymorphism. The total number of SNPs associated at \( P < 1 \times 10^{-5} \) and the most associated SNP for each analysis are listed (including the best result for core-set variables also associated in this region that are \( P < 1 \times 10^{-5} \)) for each of the 14 Phase II regions that contain at least one association at \( P < 1 \times 10^{-5} \). Region numbers are carried over from Phase I analysis (Table 2). Region coordinates based on the NCB37 assembly. \( N \) indicates number of successfully genotyped samples at each SNP.
to be characterised. Other notable associated genes include \textit{CFTR}, \textit{ZMIZ1} and \textit{TGIF1}. Although the genes identified from this analysis need replication in independent cohorts, they provide some plausible novel candidates for further investigations into MTX response.

One association of considerable interest is the cystic fibrosis transmembrane conductance regulator gene, \textit{CFTR}. The peak association signal for this region is within the 3 end of the downstream \textit{CTTNBP2} gene (Figure 1b); this is interesting as it is known that \textit{CFTR} expression is regulated by complex structural looping involving this region of \textit{CTTNBP2}.\textsuperscript{25,26} \textit{CFTR}, also known as \textit{ABCC7}, is a member of the ATP-binding cassette transporter superfamily, specifically the multidrug resistance subfamily. These proteins are known to be important to drug transport and elimination.\textsuperscript{27} Interestingly, a gene within the same subfamily, \textit{ABCC3}, which is known to be involved in MTX efflux has recently been shown to contain a SNP (rs4793665) associated with MTX response in a cohort of 287 Dutch JIA patients.\textsuperscript{10} This finding led us to specifically review this gene within our results despite it not fulfilling the selection criteria for Phase I or II. We found that rs4793665 was not directly genotyped in this study nor were there SNPs within \( r^2 \geq 0.8 \) on the chip; however, there were 38 SNPs within the introns/exons of \textit{ABCC3} genotyped, with 15 showing association in our cohort at \( P < 0.05 \) with the MTX response outcomes analysed (except ESR). The most associated of these, rs4148411 within an intron of \textit{ABCC3}, was found in the PhysVAS analysis (\( P = 7.55 \times 10^{-7} \)) and is in low linkage disequilibrium with the SNP identified by de Rotte and colleagues\textsuperscript{10} (\( r^2 = 0.02 \)), suggesting that comprehensive further investigation of this gene is warranted.

Interestingly, our study identifies several genes related to TGF\(	extbeta\) signalling as being associated with response to MTX. \textit{ZMIZ1} has been identified in several GWAS of autoimmune diseases.\textsuperscript{37–40} It is a member of the protein inhibitor of activated STAT family, is known to regulate several transcription factors (androgen receptor, Smad3/4, p53) and TGF\(	extbeta\)/SMAD signalling, and is induced by retinoic acid.\textsuperscript{41} It is well established that TGF\(	extbeta\)/SMAD and retinoic acid have important roles in the balance between Th17 and Treg cells,\textsuperscript{42} which are known to impact directly upon JIA severity.\textsuperscript{43} Therefore, a possible role for this gene in response to treatment in JIA is of considerable interest. Corroborating this finding, another associated region contains TGF\(	extbeta\)-induced factor homeobox 1 (\textit{TGIF1}), known to be an active transcriptional corepressor of SMAD2 and to modulate the downregulation of aryl hydrocarbon receptor.\textsuperscript{44,45} Together these results suggest TGF\(	extbeta\) signalling is a strong biological candidate for a role in reducing disease activity with MTX treatment. These data are of particular interest, as they directly parallel our gene-expression profiling studies, which identified TGF\(	extbeta\) signalling, TGF\(	extbeta\)-2 and the zinc finger protein ZEB1, which interacts with SMAD signalling proteins, as being involved in response to MTX in children with JIA.\textsuperscript{14}

Four genetic regions associated with ACR-pedi status also showed associations with either ParVAS or PhysVAS. In some research studies, using the ACR-pedi status can present difficulties due to missing data observed in long-term observational cohorts. Therefore, we chose to analyse ACR-pedi, ParVAS and PhysVAS to measure MTX response similarly, leading to the possibility that both ParVAS and PhysVAS could be used to measure response to treatment, when full clinical data are unavailable.

To our knowledge, this is the first report of a large-scale genetic association study of MTX response in inflammatory arthritis; despite several international efforts in GWAS studies of JIA and RA as a whole, no previous large-scale analysis of MTX response are available to date, perhaps in part due to the considerable challenges of collecting adequate numbers of cases with detailed response data as well as DNA. A study investigating interferon-beta treatment in MS (using 53 responders and 53 non-responders) found that of the best associations most were in glutamate and interferon receptors, a cell cycle-dependent protein, and guanosine triphosphatase-activating and zinc finger proteins, all genes not known to be directly involved in the drug metabolism pathway.\textsuperscript{46} A recent GWAS in 706 RA patients treated with tocilizumab, a biologic therapy targeting the interleukin-6 receptor, found eight putative loci associated with tocilizumab efficacy; however none were in known RA risk or interleukin-6 receptor pathways.\textsuperscript{18} Similarly, our results suggest that multiple genes determine response to MTX treatment in JIA, and not just those in known MTX pathways. In fact, none of the MTX pathway genes previously investigated in candidate gene studies in both JIA and RA met our selection criteria for Phase II. This is possibly due to the small sample size and lack of power in previous studies resulting in false positive associations, and may be additionally confounded by the power limitations of our study including the availability of ACR-pedi scores on only a subset of our cohort. Despite this, the lack of strong association in the MTX pathway genes is interesting, and suggests that novel pathways and mechanisms, hitherto not known, may be important to pursue in order to understand and fully elucidate the actions of MTX and the genes involved in success or failure of MTX treatment.\textsuperscript{47} It also suggests that the previously developed MTX efficacy prediction models for both RA\textsuperscript{19} and JIA\textsuperscript{5} could be further enhanced or further optimised by incorporating additional genetic variants outside the MTX pathway genes.

Previous investigations of the genetics of MTX response in JIA have been small, often underpowered, studies taking a candidate gene approach focusing on genes in MTX drug pathways. In contrast, this study is large and comprehensively covers the genome, the first of its kind for JIA. We have identified several regions of interest, three of which show a remarkable degree of functional overlap with genes and pathways implicated by gene-expression profiling and previous candidate gene studies. By analysing each clinical outcome variable individually, we show their genetic contributions to MTX response may differ, although with interesting overlap in novel candidates including \textit{TGIF1}, \textit{ZMIZ1} and \textit{CFTR}. Future targeted replication of the exciting novel regions identified is now required to confirm these findings. This study provides an excellent basis for the future development of genetic risk models for MTX response prediction.

CONFLICT OF INTEREST

NW received an unrestricted grant from MEDAC Germany. PD received a research grant from Novartis. The other authors declare no conflict of interest.

ACKNOWLEDGMENTS

We thank all the patients and their families for participating in this study and members of the CHARMS and CAPS study groups and the CHARMS-JIA GWAS International Consortium, for collecting samples and the data. We are grateful to Professor Gudrun Moore for scientific advice in the planning of the study and Mr N Evans and Dr Susan Thompson for insightful comments on the study. The SPARKS-CHARMS study was funded by SPARKS UK (08BCH09) and the Big Lottery Fund UK (RG/1/010135231) and is supported by the UK National Institutes for Health research (NIHR) Medicines for Children Research Network. LRW has support from the Great Ormond Street Hospital Children’s Charity, Arthritis Research UK (20164) and the GOSH/UCL Ich NIHR Biomedical Research Centre. The CAPS study is funded by Arthritis Research UK (20542). PD has support from Charles University Grant Agency (GAUK 52608/2008). NW has support from Reumafonds (DAA 07-01-19). WT has support from Arthritis Research UK (17552). This report includes independent research supported 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 National Institute for Health Research or the Department of Health.

AUTHOR CONTRIBUTIONS

WT, LRW, M de I, HM, EC and JC led the study. LRW established the CHARMS study cohort. LRW, SU, HM, FP, UK, MB, NW, MB, PD, PW and TRS collected samples and data. JC, EC, HM, LRW and WT wrote the manuscript. JC and EF performed the genotyping. JC, EC, M de I performed the statistical analysis. All other authors contributed primarily to the patient ascertainment and/or data, and sample collection and preparation. All authors reviewed the final manuscript.
Supplementary Information accompanies the paper on the The Pharmacogenomics Journal website (http://www.nature.com/tpj)
APPENDIX 1
Childhood arthritis prospective study (CAPS)

Sparks-childhood arthritis response to medication study (CHARMS)
Katrin Buerkle, Joanna Cobb, Catherine Cotter, Angela Etheridge, Paul Gilbert, Anne Hinks, Shashi Hirani, Laura Kassoumeri, Sham Lal, Laura Melville, Halima Moncrieffe, Kathleen Mulligan, Stanton Newman, Fiona Patrick, Tauny Southwood, Wendy Thomson, Simona Ursu, Lucy R Wedderburn, Pamela Whitworth, Patricia Woo.

British society of paediatric and adolescent rheumatology (BSPAR) study group
Mario Abinum, A. Bell, Alan W. Craft, Esther Crawley, Joel David, Helen Foster, Janet Gardener-Medwin, Jane Griffin, A. Hall, M. Hal, Ariane L. Herrick, P. Hollingworth, Lennox Holt, Stan Jones, Gillian Pountain, Clive Ryder, Tauny Southwood, I. Stewart, Helen Venning, Lucy R. Wedderburn, Patricia Woo, Sue Wyatt.