



# Genome-wide association study of liver enzyme elevation in an extended cohort of rheumatoid arthritis patients starting low-dose methotrexate

Marco Cavalli<sup>1,6</sup> , Niclas Eriksson<sup>1,2</sup> , Johanna Karlsson Sundbaum<sup>3,4</sup> , Matilda Wallenberg<sup>1,5</sup>, Hugo Kohnke<sup>1</sup>, Eva Baecklund<sup>3</sup>, Pär Hallberg<sup>‡,1</sup>  & Mia Wadelius<sup>\*,‡,1</sup> 

<sup>1</sup>Department of Medical Sciences, Clinical Pharmacogenomics & Science for Life Laboratory, Uppsala University, SE-751 85, Uppsala, Sweden

<sup>2</sup>Uppsala Clinical Research Center, SE-751 85, Uppsala, Sweden

<sup>3</sup>Department of Medical Sciences, Rheumatology, Uppsala University, SE-751 85, Uppsala, Sweden

<sup>4</sup>Department of Health Sciences, Luleå University of Technology, SE-971 87, Luleå, Sweden

<sup>5</sup>Svensk Dos AB, Box 2, SE-751 03, Uppsala, Sweden

<sup>6</sup>Department of Immunology, Genetics & Pathology, & Science for Life Laboratory, Uppsala University, SE-751 22, Uppsala, Sweden

\*Author for correspondence: Tel.: +46 18 611 4945; [mia.wadelius@medsci.uu.se](mailto:mia.wadelius@medsci.uu.se)

‡Authors contributed equally

**Aim:** A follow-up genome-wide association study (GWAS) in an extended cohort of rheumatoid arthritis (RA) patients starting low-dose methotrexate (MTX) treatment was performed to identify further genetic variants associated with alanine aminotransferase (ALT) elevation. **Patients & methods:** A GWAS was performed on 346 RA patients. Two outcomes within the first 6 months of MTX treatment were assessed: ALT > 1.5-times the upper level of normal (ULN) and maximum level of ALT. **Results:** *SPATA9* (rs72783407) was significantly associated with maximum level of ALT ( $p = 2.58 \times 10^{-8}$ ) and *PLCG2* (rs60427389) was tentatively associated with ALT > 1.5 × ULN. **Conclusion:** Associations with SNPs in genes related to male fertility (*SPATA9*) and inflammatory processes (*PLCG2*) were identified.

First draft submitted: 10 June 2022; Accepted for publication: 22 August 2022; Published online: 7 September 2022

**Keywords:** adverse reactions • drug-related side effects • genome-wide association study • hepatotoxicity • methotrexate • pharmacogenetics • rheumatoid arthritis

Elevation of liver transaminases is a common side effect of treatment with low-dose methotrexate (MTX), which is frequently used to treat rheumatoid arthritis (RA) [1]. Long-term therapy with MTX has been associated with the development of hepatic damage leading to fibrosis and possibly cirrhosis. Prevention of MTX-induced hepatotoxicity relies heavily on long-term follow-up of liver function tests [2]. When low-dose MTX treatment is initiated in RA patients, it is recommended to monitor the liver enzyme alanine aminotransferase (ALT) every 2–4 weeks for the first 3 months of therapy, every 8–12 weeks after 3 months of therapy and every 12 weeks beyond 6 months of therapy [3]. Risk factors for MTX-induced hepatotoxicity include excessive alcohol use, obesity, diabetes and concurrent administration of other potentially hepatotoxic drugs [4]. Pretreatment ALT elevation, female sex, statin treatment and increasing BMI were previously identified as predictors of ALT elevation during MTX therapy [5]. Genetic risk factors may also play a role. Studies have identified several single-nucleotide polymorphisms (SNPs) potentially affecting both the efficacy and toxicity of MTX by modulating the expression of key enzymes (e.g., MTHFR [6], FPGS [7], DHFR [8], ATIC [9] and SHMT [10]) in the folate biosynthesis pathway.

Identifying more genetic risk factors for MTX-induced liver toxicity would aid the development of personalized therapy for RA. The current authors previously performed a genome-wide association study (GWAS) in a cohort of 198 RA patients starting MTX treatment that revealed novel associations between ALT elevation and SNPs that could regulate the expression of *RAVER2* and *JAK1*. This was strengthened in a meta-analysis with a replication

cohort of 160 patients who underwent direct genotyping of the SNPs of interest [11]. In this follow-up study, genome-wide genotyping of the patients included in the replication cohort was performed and the results of a joint analysis of the discovery and replication cohorts are presented. Other potential associations between genetic variations and an increase in ALT during low-dose MTX therapy in this larger cohort were also assessed.

## Materials & methods

### Cohort description

Patient recruitment and study methodology have been described in detail elsewhere [11]. Briefly, a total of 198 patients starting low-dose MTX therapy for RA between 1 January 2005 and 30 April 2013, at the Rheumatology Department at Uppsala University Hospital, Sweden, were recruited and followed for a minimum of 6 months. An identical group of 160 patients starting therapy between 1 May 2013, and 30 September 2017, at the Rheumatology Department, Uppsala University Hospital, Sweden, or between 1 January 2005 and 30 September 2017, at the Rheumatology Department, Sunderby Hospital (Lulea), Sweden, were also recruited. As in the previous study, the primary outcome was an elevation of ALT >1.5-times the upper limit of normal (ULN) within the first 6 months of treatment [11], and patients were divided into cases and controls. The secondary outcome was the maximum value of ALT in all patients within the first 6 months of treatment.

### Genome-wide array data & analyses

The original GWAS cohort was genotyped with the Illumina Infinium OmniExpressExome 1 M Array (n = 198 individuals) and the second cohort (n = 160 individuals) with the Illumina Infinium Global screening array 24v3, which is optimized for imputation and has a lower cost. Genotype calls were generated with the GenomeStudio software (Illumina, CA, USA) and the Genome Reference Consortium human assembly GRCh37. Preimputation quality control included gender checks, exclusion of variants with Hardy-Weinberg equilibrium p-values  $<5 \times 10^{-8}$  or minor allele frequency (MAF)  $<0.5\%$  and exclusion of individuals with  $>2\%$  missing variants. Imputation was performed separately for each Illumina batch using the Sanger imputation server. The haplotype reference consortium (HRC r1) panel [12] was used as a reference for the pipeline with Eagle2 v2.3.3 [13] prephasing and positional Burrows–Wheeler transform (PBWT) imputation [14]. Postimputation, the datasets were filtered for quality (impute2 imputation quality  $>0.7$ ) using bcftools and converted to hard calls using PLINK v1.9 [15]. The resulting two PLINK datasets were finally merged. In this study, variants available from both Illumina batches with good imputation were used and an overall MAF cutoff of 0.01 was applied, yielding a total of 7,489,589 variants for the analysis. Due to an increasing number of spurious results at low allele frequencies, for the secondary continuous outcome (max ALT within 6 months) a MAF cutoff of 0.03 was applied, resulting in a total of 6,074,601 variants for the analysis.

PLINK v1.9 was used for logistic regression analysis on a genome-wide level for the primary outcome, and the analyses were adjusted for sex, age and the first four principal components. Analysis of the secondary outcome (max ALT within 6 months) was performed using linear regression implemented in PLINK v1.9. Prior to the analysis, max ALT was  $\log_2$  transformed due to the right-tail skewed distribution of the variable. The analysis was adjusted for age, sex,  $\log_2$  baseline ALT (ALT at the start of MTX treatment) and the first four genetic principal components. SNP effects were modeled as additive and statistical significance was set at the traditional genome-wide level of  $p < 5 \times 10^{-8}$  [16].

### Statistical analyses & power calculation

Descriptive data are expressed as mean  $\pm$  standard deviation (SD), minimum (min), maximum (max) and frequency (%). For comparative analyses between cases and controls, Student's *t*-tests or Mann–Whitney U-tests were used for continuous variables and  $\chi^2$  tests or Fisher's exact tests were used for categorical variables. Statistical power was estimated using the R package GeneticsDesign (function GPC) using the observed prevalence of 33/346 for ALT  $>1.5 \times$  ULN in the cohort. For a marker with a MAF of 0.2, there was approximately 80% power to detect an odds ratio of 6.2 at a genome-wide level, and for a marker with a MAF of 0.4, an odds ratio of 4.5 (Supplementary Figure 1).

### Functional variant analysis

The top GWAS SNPs associated with ALT elevation were intersected with different databases to obtain functional annotations, including the candidate *cis*-regulatory elements (cCREs) registry from the ENCODE project [17], a

**Table 1. Characteristics of patients with rheumatoid arthritis initiating treatment with methotrexate.**

Characteristic	Cases <sup>†</sup> (n = 33)	Controls <sup>‡</sup> (n = 313)	p-value
Female, n (%)	30 (90.9)	218 (69.6)	0.0080
Parental birthplace (%)			
Sweden/Sweden	26 (78.8)	264 (84.3)	0.27
Sweden/other	4 (13.3)	16 (7.3)	
Other/other	3 (10.0)	25 (11.5)	
Data missing	0 (0)	8 (3.7)	
Age at RA diagnosis, mean ± SD (range), years	55.7 ± 11.3 (30–80)	55.9 ± 14.3 (18–82) <sup>§</sup>	0.90
History of ALT elevation prior to MTX treatment, n (%)	14 (42.4)	52 (16.6) <sup>¶</sup>	0.0018
Rheumatoid factor positivity, n (%)	22 (66.6)	204 (65.1) <sup>#</sup>	1.00
ACPA positivity, n (%)	24 (72.7) <sup>§</sup>	200 (63.8) <sup>††</sup>	0.33
Age at MTX treatment start, mean ± SD (range), years	55.8 ± 11.3 (30–80)	56.7 ± 14.3 (19–82)	0.66
MTX maximum weekly dose, mean ± SD (min–max), mg	18.1 ± 5.0 (7.5–25)	18.2 ± 4.4 (7.5–25) <sup>#</sup>	0.83
<b>Comorbidity</b>			
Diabetes, n (%)	3 (9.0)	23 (7.3)	0.73
Hepatic disease, n (%)	0 (0)	6 (1.9)	1.00
<b>Other characteristic</b>			
BMI, mean ± SD (range), kg/m <sup>2</sup>	27.1 ± 4.5 (19.8–38.8)	26.2 ± 4.3 (17–40.9) <sup>‡‡</sup>	0.30
Ever smoker, n (%)	18 (54.5)	148 (47.3) <sup>§§</sup>	0.47
Alcohol, standard glasses/week, mean ± SD (range)	2.0 ± 1.9 (0–7)	2.3 ± 2.7 (0–17) <sup>§§</sup>	0.36
<sup>†</sup> Patients with ALT >1.5 × ULN within 6 months from start of MTX treatment. <sup>‡</sup> Patients with ALT <1.5 × ULN within 6 months from start of MTX treatment. <sup>§</sup> Missing data for 1 patient. <sup>¶</sup> Missing data for 11 patients. <sup>#</sup> Missing data for 4 patients. <sup>††</sup> Missing data for 7 patients. <sup>‡‡</sup> Missing data for 5 patients. <sup>§§</sup> Missing data for 2 patients. ACPA: Anticitrullinated protein antibodies; ALT: Alanine aminotransferase; MTX: Methotrexate; RA: Rheumatoid arthritis; SD: Standard deviation; ULN: Upper limit of normal.			

cell and tissue-specific collection of representative DNase hypersensitivity sites (rDHSs) supported by either histone modifications (H3K4me3 and H3K27ac) or CTCF-binding data; HaploReg, a tool for exploring annotations of noncoding variants on haplotype blocks [18]; the Tomtom motif comparison tool [19] and the Transcription factor Affinity Prediction (sTRAP) tool [20] to identify candidate transcription factors whose binding could be affected by the studied variants; and gene expression levels from the Genotype-Tissue Expression (GTEx) project [21] to investigate the tissue-specific expression of the associated genes.

## Results

### Cohort characteristics

A total of 346 patients of mainly Swedish ancestry with successful genotyping data were available in the combined discovery and replication dataset, their characteristics are shown in Table 1. Note that 12 of 160 patients from the replication cohort in the original study had DNA sample yields that were too low for genome-wide analyses and were excluded from the current analysis.

The study of the primary outcome was performed on 33 patients defined as cases due to at least one elevation of ALT >1.5 × ULN within the first 6 months of MTX treatment with the remaining 313 patients defined as controls (Table 1). The proportion of individuals with a previous history of ALT elevation before MTX treatment was significantly higher among cases compared with controls (42.4% vs 16.6%;  $p = 0.0018$ ). The ethnicity of the patients did not differ significantly between cases and controls ( $p = 0.27$ ; Table 1) with both parents of 79% of the cases and 84% of the controls born in Sweden.

The secondary outcome, the maximum level of ALT within the first 6 months of treatment, was investigated in a total of 341 patients (5 of 346 patients with normal ALT were excluded because exact ALT values were not recorded, although not >1.5 × ULN). All patients were treated with MTX 7.5–25 mg once a week. According to Swedish guidelines, the initial dose of 7.5 mg per week is gradually increased during 1–3 months to a maximum of

25 mg per week based on individual decisions by the treating physician. The patient is supplemented with at least 5 mg of folic acid per week. MTX dose was not associated with ALT elevation in this cohort [5].

### Genome-wide association analysis

As previously observed in the original discovery cohort ( $n = 198$ ), no statistically significant genome-wide association was found when comparing patients with  $ALT > 1.5 \times ULN$  within 6 months with controls in the extended cohort ( $n = 346$ ). Tentative associations ( $p < 1 \times 10^{-5}$ ) were observed for SNPs close to *PLCG2*, *KNOP1*, *HSD17B13* and *PRKCB* (Figure 1A & Supplementary Table 1).

The maximum level of ALT within 6 months was significantly associated at a genome-wide level ( $p < 5 \times 10^{-8}$ ) with one SNP (rs72783407) located at the *SPATA9* locus (Figure 1B & Supplementary Table 2). This SNP had a beta-value of 0.60 (95% CI: 0.40–0.81) per increase of one minor allele, meaning that the geometric mean of ALT increased 1.51-times per minor allele of rs72675408. An ALT of 1.00 would thus be predicted to increase to 1.51 in patients with one variant allele and to 2.31 in patients with two variant alleles. The possible interaction between rs72783407 and gender was investigated, however, the interaction was not significant ( $p = 0.0982$ ; Supplementary Figure 2). We also observed a near-significant association with a cluster of five SNPs (rs3920617, rs55889764, rs72675408, rs72675414 and rs72675415) at the *RAVER2* locus. A statistically significant association with rs72675408 was previously demonstrated when using data from the combined cohort ( $n = 354$ ) [11].

### Genomic landscape of the top genome-wide association study SNPs

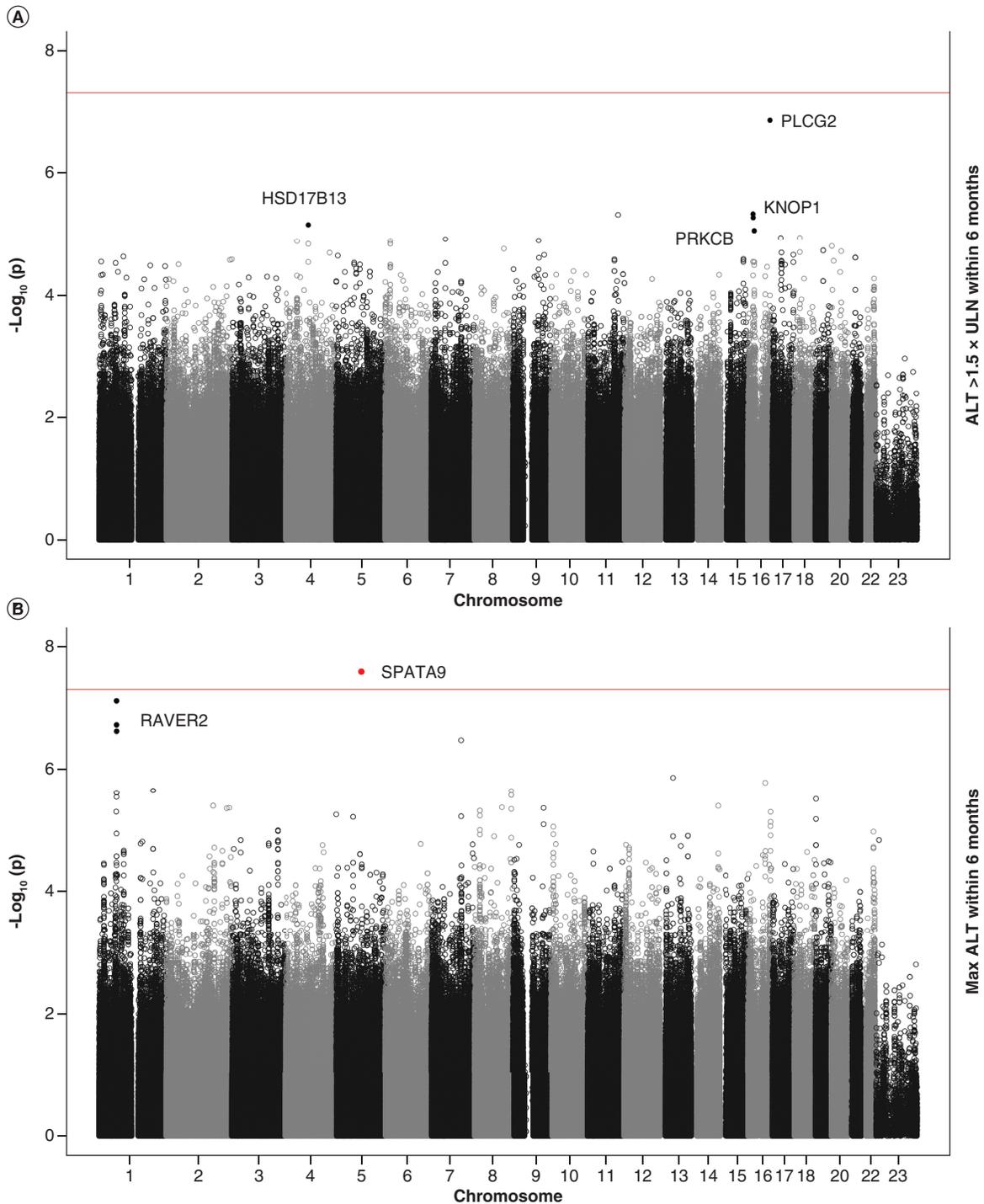
The variant rs60427389 located in the 13th intron of *PLCG2* was the SNP closest to statistical significance in association with the first outcome,  $ALT > 1.5 \times ULN$  within 6 months. *PLCG2* is expressed in several tissues including lung, skin, lymphoid tissues, blood and, in a smaller measure, the liver. The genomic region harboring rs60427389 is annotated as quiescent in blood and shows DNase signals and H3K4me1 and H3K27ac enhancer signatures in liver tissue [18]. The SNP is located <150 base pairs upstream of an ENCODE annotated cCRE (EH38E3193997) defined by DNase and CTCF signals [17]. Motif analysis showed how rs60427389 could alter the binding of several transcription factors, among others: ZNF135, ZNF524, NR2F2, NR2F1, NR2F6, RARG, RARA and MAF [18,19].

The variant rs72783407, located in the second intron of *SPATA9*, was significantly associated at a genome-wide level with the second outcome, the maximum level of ALT within 6 months. rs72783407 is an eQTL for *SPATA9* in testis [21] and is located in a genomic region annotated as quiescent in blood and liver tissue [18] where it is expressed to a much lower degree. It is predicted to alter the binding of different transcription factors: CEBPD, E2F-family, PLZF and PAX4. The transcription factor Affinity Prediction tool revealed a preferential binding of transcription factors on the reference T allele (E2F, PLZF) with respect to the alternative A allele (CEBPD and PAX4).

### Discussion

In this expanded GWAS aiming to identify associations between genetic variations and early ALT elevation in RA patients starting MTX treatment, no genome-wide signals were identified when comparing RA patients with or without  $ALT > 1.5 \times ULN$  within 6 months of the start of MTX therapy. The increased cohort size thus did not lead to a significant association, with the primary outcome mirroring the results of the original GWAS performed in the 198-patient cohort. A genome-wide near-significant association was observed for rs60427389, located in an intron of *PLCG2* that encodes for the transmembrane signaling enzyme PLCG2. Gain of function mutations in *PLCG2* have been associated with severe spontaneous inflammation and autoimmunity [22]. *PLCG2* has also emerged as a candidate gene to control the inflammation in RA and, in a meta-analytic study, it was enriched in several signaling pathways related to inflammation, antigen presentation, hypoxia and apoptosis [23]. In addition, overexpression of PLCG2 in rat hepatocytes has been shown to inhibit hepatocyte proliferation and increase apoptosis [24].

When examining genetic associations with the secondary outcome, a single SNP, rs72783407, was associated with the max value of ALT within 6 months of the start of low-dose MTX therapy. rs72783407 is located in an intron of *SPATA9* (also known as *NYD-SP16*), a gene that encodes for SPATA9, predicted to be involved in cell differentiation and spermatogenesis [25]. This membrane protein is involved in capacitation and acrosome reaction and has been investigated as a potential biomarker for male infertility. Patients with Sertoli cell-only syndrome or in maturation arrest at an early stage of spermatogenesis show no expression of *SPATA9* in the



**Figure 1. Manhattan plots for the primary and the secondary outcomes. (A) Primary outcome. (B) Secondary outcome.** ALT is  $\log_2$  transformed and adjusted by  $\log_2$  (baseline ALT), age, sex and genetic principal components 1–4. The red line denotes the genome-wide significance level  $p < 5 \times 10^{-8}$ . ALT: Alanine aminotransferase; ULN: Upper level of normal.

testis [26]. The observed association linked to *SPATA9* was unexpected considering the nature of the study cohort, composed of approximately 70% female patients. A trend was observed in which the association of rs72783407 with the outcome was more pronounced in males, although not statistically significant (Supplementary Figure 2),

a larger (male) cohort would be required to confirm this finding. The possibility that the association is a chance finding cannot be ruled out and should thus be cautiously interpreted.

A limitation of the study cohort is that both cases and controls have an overrepresentation of female patients (Table 1). However, as for many autoimmune diseases, RA displays an imbalance between the sexes, with females representing the majority of cases [27,28]. The associations reported in this study were adjusted by sex to avoid bias. Another limitation is that only around 10% experienced at least one elevation of ALT  $> 1.5 \times$  ULN within the first 6 months of MTX treatment. A significant difference was also observed between cases and controls regarding the history of ALT elevation before MTX treatment, which could lead to biased associations. We previously showed that a history of ALT elevation is a predictor of ALT elevation during MTX therapy. Correction for pretreatment ALT elevations was not employed, since it is unknown whether a genetic contribution to a predisposition for ALT elevation exists. In doing so, a genetic association with ALT elevation regardless of treatment may have been lost. Further, the patients were exposed to other potential causes of liver toxicity (e.g., alcohol use, obesity and diabetes), although no significant differences were seen in this respect between cases and controls (Table 1).

Cui *et al.* [29] recently reported two new associations at loci for *RAP1GAP2* and *ULK2*. These two SNPs were not reported in the current study due to low imputation quality (rs28661593) and the rarity of the variant, present in only 3 donors (rs76993247). Overall, the results of this study, of our current and previous GWAS [11] and of a meta-analysis [30] have not identified shared common variants associated with liver function, but rather point toward several different loci that in the future might be used to define a polygenic risk score for MTX-induced liver toxicity.

## Conclusion

No single genetic variant strongly associated with liver enzyme elevation was identified in this extended cohort of patients starting low-dose MTX. The overall results of the presented GWAS highlighted two genomic loci possibly associated with elevation of ALT levels during the first 6 months of treatment. SNPs at these loci were tentatively associated with genes whose expression could affect biological processes related to male fertility (*SPATA9*) and inflammatory processes (*PLCG2*). Further studies are required to validate these findings.

### Summary points

- Low-dose methotrexate (MTX) represents the first-line therapy for rheumatoid arthritis (RA).
- A common side effect of MTX treatment is liver toxicity with elevations in alanine aminotransferase (ALT).
- A genome-wide association study was performed in an extended cohort of RA patients starting low-dose MTX treatment to pinpoint new genetic risk factors associated with ALT elevation.
- A variant located in an intron of *SPATA9* was significantly associated with maximum ALT level within 6 months of starting MTX treatment.
- A variant located in an intron of *PLCG2* was tentatively associated with ALT  $> 1.5 \times$  the upper limit of normal within 6 months of starting MTX treatment.

### Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: [www.futuremedicine.com/doi/suppl/10.2217/pgs-2022-0074](http://www.futuremedicine.com/doi/suppl/10.2217/pgs-2022-0074)

### Author contributions

Study design: P Hallberg and M Wadelius; patient recruitment and data collection: J Karlsson Sundbaum and M Wallenberg; adjudication of cases: J Karlsson Sundbaum; DNA extraction and preparation for genetic analyses: H Kohnke; data analysis: N Eriksson; data interpretation: M Cavalli, P Hallberg, M Wadelius and N Eriksson; drafting manuscript: M Cavalli; revising manuscript content: M Cavalli, P Hallberg and M Wadelius. All authors approved the final version of the manuscript.

### Acknowledgments

We are grateful to all physicians, research nurses and supporting staff who assisted in recruiting patients and controls or administering phenotype databases. In particular, U Ramqvist, C Haglund, S Collin and E Prado (Uppsala University, Sweden). Lastly, we are grateful to the staff at the clinics of Rheumatology at Uppsala University Hospital and Sunderby Hospital. Genotyping of the patient cohort was performed by the single nucleotide polymorphism and sequencing (SNP&SEQ) Technology Platform at Uppsala

University ([www.genotyping.se](http://www.genotyping.se)). The facility is part of the National Genomics Infrastructure supported by the Swedish Research Council for Infrastructures and Science for Life Laboratory, Sweden. The SNP&SEQ Technology Platform is also supported by the Knut and Alice Wallenberg Foundation. Computations were done on resources provided by the Swedish National Infrastructure for Computing through the Uppsala Multidisciplinary Centre for Advanced Computational Science (UPPMAX).

#### Financial & competing interests disclosure

This work was supported by the Agnes and Mac Rudberg's, Brunnberg's, Selander's and Thuréus' Foundations at Uppsala University; Swedish Rheumatism Association; Swedish Research Council (Medicine 521-2011-2440, 521-2014-3370 and 2018-03307); Swedish Heart-Lung Foundation (20120557, 20140291 and 20170711); and Clinical Research Support (Avtal om Läkarutbildning och Forskning, ALF) at Uppsala University. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

#### Open access

This article is distributed under the terms of the Creative Commons Attribution License 4.0 which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited. To view a copy of the license, visit <http://creativecommons.org/licenses/by/4.0/>

#### Ethical conduct of research

This research was carried out in accordance with the latest update of the Declaration of Helsinki. The study was approved by the regional ethics committee (2010/231, Uppsala, Sweden), and written informed consent was obtained from all participants.

#### References

1. Visser K, Van Der Heijde DM. Risk and management of liver toxicity during methotrexate treatment in rheumatoid and psoriatic arthritis: a systematic review of the literature. *Clin. Exp. Rheumatol.* 27(6), 1017–1025 (2009).
2. García DS, Saturansky EI, Poncino D *et al.* "Hepatic toxicity by methotrexate with weekly single doses associated with folic acid in rheumatoid and psoriatic arthritis. What is its real frequency?". *Ann. Hepatol.* 18(5), 765–769 (2019).
3. Singh JA, Saag KG, Bridges SL Jr *et al.* 2015 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis. *Arthritis Rheumatol.* 68(1), 1–26 (2016).
4. Vuppalanchi R, Chalasani N. Risk factors for drug-induced liver disease. In: *Drug-Induced Liver Disease (3rd Edition)*. Kaplowitz N, Deleve LD (Eds). Academic Press, MA, USA, 265–274 (2013).
5. Karlsson Sundbaum J, Eriksson N, Hallberg P, Lehto N, Wadelius M, Baecklund E. Methotrexate treatment in rheumatoid arthritis and elevated liver enzymes: a long-term follow-up of predictors, surveillance, and outcome in clinical practice. *Int. J. Rheum. Dis.* 22(7), 1226–1232 (2019).
6. Karlsson Sundbaum J, Baecklund E, Eriksson N, Hallberg P, Kohnke H, Wadelius M. MTHFR, TYMS and SLCO1B1 polymorphisms and adverse liver effects of methotrexate in rheumatoid arthritis. *Pharmacogenomics* 21(5), 337–346 (2020).
7. Muller IB, Lin M, Lems WF *et al.* Association of altered folylpolyglutamate synthetase pre-mRNA splicing with methotrexate unresponsiveness in early rheumatoid arthritis. *Rheumatology (Oxford)* 60(3), 1273–1281 (2021).
8. Milic V, Jekic B, Lukovic L, Lukovic L, Bunjevacki V *et al.* Association of dihydrofolate reductase (DHFR) -317AA genotype with poor response to methotrexate in patients with rheumatoid arthritis. *Clin. Exp. Rheumatol.* 30(2), 178–83 (2012).
9. Lee YH, Bae S-C. Association of the ATIC 347 C/G polymorphism with responsiveness to and toxicity of methotrexate in rheumatoid arthritis: a meta-analysis. *Rheumatol. Int.* 36(11), 1591–1599 (2016).
10. Kremer JM. Methotrexate pharmacogenomics. *Ann. Rheum. Dis.* 65(9), 1121–1123 (2006).
11. Sundbaum JK, Baecklund E, Eriksson N *et al.* Genome-wide association study of liver enzyme elevation in rheumatoid arthritis patients starting methotrexate. *Pharmacogenomics* 22(15), 973–982 (2021).
12. Mccarthy S, Das S, Kretschmar W *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* 48(10), 1279–1283 (2016).
13. Loh P-R, Danecek P, Palamara PF *et al.* Reference-based phasing using the Haplotype Reference Consortium panel. *Nat. Genet.* 48(11), 1443–1448 (2016).
14. Durbin R. Efficient haplotype matching and storage using the positional Burrows–Wheeler transform (PBWT). *Bioinformatics* 30(9), 1266–1272 (2014).
15. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 4, 7 (2015).

16. Sham PC, Purcell SM. Statistical power and significance testing in large-scale genetic studies. *Nat. Rev. Genet.* 15(5), 335–346 (2014).
17. Moore JE, Purcaro MJ, Pratt HE *et al.* Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* 583(7818), 699–710 (2020).
18. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res.* 40(D1), D930–D934 (2012).
19. Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble WS. Quantifying similarity between motifs. *Genome Biol.* 8(2), R24 (2007).
20. Manke T, Heinig M, Vingron M. Quantifying the effect of sequence variation on regulatory interactions. *Hum. Mutat.* 31(4), 477–483 (2010).
21. Lonsdale J, Thomas J, Salvatore M *et al.* The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* 45(6), 580–585 (2013).
22. Yu P, Constien R, Dear N *et al.* Autoimmunity and inflammation due to a gain-of-function mutation in phospholipase C $\gamma$ 2 that specifically increases external Ca $^{2+}$  entry. *Immunity* 22(4), 451–465 (2005).
23. Afroz S, Giddaluru J, Vishwakarma S, Naz S, Khan AA, Khan N. A comprehensive gene expression meta-analysis identifies novel immune signatures in rheumatoid arthritis patients. *Front Immunol.* 8, 74 (2017).
24. Chen X, Lv Q, Ma J, Liu Y. PLC $\gamma$ 2 promotes apoptosis while inhibits proliferation in rat hepatocytes through PKCD/JNK MAPK and PKCD/p38 MAPK signaling. *Cell Prolif.* 51(3), e12437–e12437 (2018).
25. Cheng LJ, Li JM, Chen J *et al.* NYD-SP16, a novel gene associated with spermatogenesis of human testis. *Biol. Reprod.* 68(1), 190–198 (2003).
26. Wacławska A, Kurpisz M. Key functional genes of spermatogenesis identified by microarray analysis. *Syst. Biol. Reprod. Med.* 58(5), 229–235 (2012).
27. Sokka T, Toloza S, Cutolo M, Cutolo M, Kautiainen H *et al.* Women, men, and rheumatoid arthritis: analyses of disease activity, disease characteristics, and treatments in the QUEST-RA study. *Arthritis Res. Ther.* 11(1), R7 (2009).
28. Eriksson JK, Neovius M, Ernestam S, *et al.* Incidence of rheumatoid arthritis in Sweden: a nationwide population-based assessment of incidence, its determinants, and treatment penetration. *Arthritis Care Res. (Hoboken)* 65(6), 870–878 (2013).
29. Cui J, Chasman DI, Raychaudhuri S *et al.* Genetics are not likely to offer clinically useful predictions for elevated liver enzyme levels in patients using low dose methotrexate. *Semin. Arthritis Rheum.* 55, 152036 (2022).
30. Chen VL, Du X, Chen Y *et al.* Genome-wide association study of serum liver enzymes implicates diverse metabolic and liver pathology. *Nat. Commun.* 12(1), 816–816 (2021).