EXTENDED REPORT

Genome-wide association analysis of anti-TNF drug response in patients with rheumatoid arthritis


ABSTRACT

Background Treatment strategies blocking tumour necrosis factor (anti-TNF) have proven very successful in patients with rheumatoid arthritis (RA). However, a significant subset of patients does not respond for unknown reasons. Currently, there are no means of identifying these patients before treatment. This study was aimed at identifying genetic factors predicting anti-TNF treatment outcome in patients with RA using a genome-wide association approach.

Methods We conducted a multistage, genome-wide association study with a primary analysis of 2,557,253 single-nucleotide polymorphisms (SNPs) in 882 patients with RA receiving anti-TNF therapy included through the Dutch Rheumatoid Arthritis Monitoring (DREAM) registry and the database of Apotheekzorg. Linear regression analysis of changes in the Disease Activity Score in 28 joints after 14 weeks of treatment was performed using a multi-protein association approach.

Results 772 markers showed evidence of association with treatment outcome in the initial stage. Eight genetic loci showed improved p value in the overall meta-analysis compared with the first stage, three of which (rs1568885, rs1813443 and rs4411591) showed directional consistency across all four cohorts studied. We were unable to replicate markers previously reported to be associated with anti-TNF outcome. Network analysis indicated strong involvement of biological processes underlying inflammatory response and cell morphology.

Conclusions Using a multistage strategy, we have identified eight genetic loci associated with response to anti-TNF treatment. Further studies are required to validate these findings in additional patient collections.

INTRODUCTION

Rheumatoid arthritis (RA) is a chronic, systemic inflammatory disease characterised by polyarthritis, joint damage and functional disability. It cannot be cured, and treatment is directed towards reducing the symptoms associated with the disease.

Tumour necrosis factor α (TNFα) is a pleiotropic, proinflammatory and immunoregulatory cytokine that plays a crucial role in RA. The introduction of TNF-blocking agents, such as infliximab, etanercept and adalimumab, revolutionised the treatment of RA, most notably because of the excellent clinical efficacy and ability of these agents to prevent further structural damage in patients who failed to respond to treatment with conventional disease-modifying antirheumatic drugs (DMARDs). Despite this success, a substantial proportion of patients with RA (~50%) treated with TNF inhibitors do not display any significant clinical improvement. Given the expensive treatment regimen and the potential side effects associated with the treatment, the idea of a priori prediction of response to anti-TNF agents in patients with RA is a highly relevant topic.

Studies of clinical parameters and biomarkers have identified several factors that influence anti-TNF treatment outcome, including concurrent use of DMARDs, lower baseline Health Assessment Questionnaire score, gender, smoking, serological status, TNF levels at the site of inflammation, and the synovial microarchitect. However, these factors explain only a relatively small proportion of the observed variance in response (R²=17–29%) and are therefore not suitable to be used as predictors in a clinical setting. In addition, effort has been put into the identification of genetic markers predicting anti-TNF treatment outcome. Most of these studies are candidate gene based, focusing on polymorphisms in genes known to be involved in RA pathogenesis and genes implicated in TNFα signalling pathways. The most thoroughly investigated gene is TNFA, encoding TNFα, the target of anti-TNF treatment. Initial studies suggested a role for a variant in the promoter of the gene (−308G>A) in anti-TNF response, although recent meta-analyses do not support this association.
So far, using the candidate gene approach, the most convincing evidence of association with response to anti-TNF therapy in patients with RA is found for an RA risk allele at the PTPRC gene locus.16 17

A number of additional potential candidate loci have been suggested on the basis of results from three genome-wide association studies (GWAS).16–20 In a GWAS of 566 patients with RA, Plant et al.19 found evidence of association at seven genetic loci with response to TNF blockade, two of which mapped within genes: PDZ domain-containing protein 2 (PDZD2) and eyes absent homologue 4 (EYA4). In a small study (n=89) by Liu et al.,18 association was reported for markers in the MAFB and PON1 gene regions as well as in a region of chromosome 9 that contains the interferon kappa (IFN-κ), MOBKL2B and C9orf72 loci. The most compelling candidate for involvement in anti-TNF treatment response in this study is IFN-κ, since type 1 IFNs play a definite role in inflammatory disease and autoimmunity.20 However, these results could not be replicated by others.20 21 Krentel et al.20 reported associations of single-nucleotide polymorphisms (SNPs) within a non-coding region surrounded by the TLR4 gene and the DBC1 gene and a marker within the FOXP1 gene with treatment outcome in a cohort of 196 Danish patients. To determine whether the reported loci reflect true associations, and to search for novel loci that influence differential response to anti-TNF therapy, we performed a GWAS in a cohort of 882 Dutch patients with RA receiving anti-TNF therapy.

MATERIALS AND METHODS

Patients and study design

A multistage GWAS was performed including 984 patients with RA receiving anti-TNF medication (stage 1) with subsequent follow-up of the most significant signals in two replication cohorts (stage 2 (n=954) and 3 (n=867)).

For the initial genome-wide association analysis, patients were recruited through a collaborative effort in which 669 patients were included as part of the Dutch Rheumatoid Arthritis Monitoring (DREAM) registry (http://www.dreamregistry.nl) and 315 patients were enrolled through the database of ApotheekZorg, which facilitates the Dutch distribution of adalimumab. All patients were diagnosed with RA according to the 1987 revised American College of Rheumatology (ACR) criteria and were treated with anti-TNF according to the indications in the Netherlands: Disease Activity Score 28 (DAS28) ≥3.2; previous failure on at least two DMARDs, one of which has to be methotrexate; biological naïve.23 We used the DAS28 change at 3 months as outcome for our analysis. Patients who stopped treatment within the first 3 months were not included in the study. All patients gave written informed consent, and the study was approved by the ethics committees of the participating hospitals.

For stage 2, data from 954 RA cases treated with anti-TNF were selected from nine different cohorts as part of the ACR Research and Education Foundation (REF) ‘Within Our Reach’ project: Autoimmune Biomarkers Collaborative Network (ABCoN), Academic Medical Center Amsterdam (AMC), Behandelstrategieen voor Rheumaotide Arthritis (BeSt), Biological in Rheumatoid arthritis Genetics and Genomics Study Syndicate (BRAGCSS), Brigham Rheumatoid Arthritis Sequential Study (BRASS), Epidemiological Investigation of Rheumatoid Arthritis (EIRA), Immunex Early Rheumatoid Arthritis (ERA) study, Karolinska Institutet (KI) study, READE, formerly Jan van Breemen Institute (READE) study, Treatment of Early Aggressive RA (TEAR)—this collection has been reported on previously.16 24

Finally, stage 3 included two previously described cohorts: (1) Wellcome Trust Case Control Consortium (WTCCC) comprising 595 patients with RA from the UK19; (2) 272 patients with RA from France ascertained through ReAct.25

Genotyping and preimputation quality control (QC)

For stage 1, genotyping was performed using the Illumina HumanHap550-Duo Bead Chip or the Human660W-Quad BeadChips, according to the instructions of the manufacturer (Illumina, San Diego, California, USA).

Preimputation QC procedures were applied using PLINK software.26 SNPs that had minor allele frequency (MAF) <0.05 and call rates <95% were excluded, as well as SNPs with extreme departures from Hardy–Weinberg equilibrium (p<1×10−5). Subsequently, QC filtering was performed at the sample level. Four samples were excluded because of gender mismatch with phenotypic data, and 21 samples because of a genotyping rate <95%. Cryptic relatedness between study participants was examined by estimating identity by descent (IBD). Seven DNA samples were excluded based on a proportion of IBD (PI-HAT)>0.125. Lastly, principal components were computed to adjust for population stratification using the EIGENSTRAT package27; 59 individuals were removed as outliers, based on the EIGENSTRAT default filter. After QC, 882 individuals were left for analysis. For the replication cohorts, the same QC criteria were used.

Imputation

To obtain a marker set common to all studies and to increase overall coverage of the genome, imputation was performed using HapMap2 release 21 (downloaded from http://www.sph.umich.edu/csg/abecasis/MACH/download/HapMap-r21.html). Haplotype phasing using MaCH software (http://www.sph.umich.edu/csg/abecasis/MACH/index.html)28 was followed by genotype imputation by Minimac (http://genome.sph.umich.edu/wiki/Minimac).

Post-imputation QC criteria were MAF ≥1% and good imputation quality, which was defined as RSQR≥0.3. In total, 2,557 253 SNPs were included in the analysis.

Stage 1 GWAS

The additive genetic effect of each SNP allele on change in DAS28 at 5 months of treatment was estimated using linear regression analysis with adjustment for baseline DAS28 and the first three principal components derived using EIGENSTRAT. These analyses were performed using the Mach2qtl software package29 (downloaded from http://www.sph.umich.edu/csg/abecasis/MACH/download/HapMap-r21.html).

The Dutch samples were not genotyped in one run and therefore the results were analysed using a meta-analysis approach that combines study-specific β estimates based on the fixed-effects model and using the inverse of the variance of the study-specific β estimates to weight the contribution of each study. Calculations were performed in the METAL package (http://www.sph.umich.edu/csg/abecasis/MACH/download/HapMap-r21.html).

The Dutch samples were not genotyped in one run and therefore the results were analysed using a meta-analysis approach that combines study-specific β estimates based on the fixed-effects model and using the inverse of the variance of the study-specific β estimates to weight the contribution of each study. Calculations were performed in the METAL package (http://www.sph.umich.edu/csg/abecasis/MACH/download/HapMap-r21.html).

SNP selection for replication in stages 2 and 3

Markers demonstrating association with DAS28 change (p<10−5) in stage 1 were selected for replication. Pruning of hits based on linkage disequilibrium (LD) was performed before replication: all SNPs with a HapMap Utah residents with
Figure 1 Study design of a multistage genome-wide association studies (GWAS) of response to anti-tumour necrosis factor (TNF) medication in patients with rheumatoid arthritis (RA). We started out with meta-analysis of GWAS data from a Dutch cohort comprising 882 patients with RA treated with anti-TNF medication. We selected 772 single-nucleotide polymorphisms (SNPs) that reached \( p < 0.001 \), which were further followed-up in stage 2 samples (\( N = 954 \) individuals). Thirty-eight SNPs out of 768 investigated in stage 2 passed \( p < 0.05 \) and were further investigated in two separate cohorts in stage 3 (\( N = 595 \) and \( N = 272 \) individuals). MAF, minor allele frequency; QC, quality control.

**Explorative analysis for functional relation between genes identified in stage 1**

All markers showing association with DAS change (\( p < 10^{-3} \)) in stage 1 were investigated for functional interactions by Ingenuity Pathway Analysis (IPA) software (Ingenuity Systems, www.ingenuity.com) using an unsupervised analysis. IPA computes a score for each network accordingly to the fit of the user’s set of input genes. The score, representing the \( -\log(p\text{ value}) \), indicates the likelihood of focus genes (genes harbouring associated SNPs) in a network being found together due to random chance.

**RESULTS**

Figure 1 presents an overview of our study approach. The baseline characteristics of the patients included in the study are summarised in table 1.

**Genome-wide association analysis**

In stage 1, 2,448,996 SNPs were tested for association with anti-TNF outcome in the Dutch population. Of these SNPs, 2,359 showed evidence of association with treatment response \( (p < 10^{-3} \), online supplementary table 1). LD pruning reduced the number of SNPs prioritised for replication analysis to 772.

We aimed to replicate the findings in 954 patients from the REF collection for stage 2 of our study. A total of 768 SNPs passed the QC in the second stage replication cohort. Thirty-nine markers showed nominally significant \( (p < 0.05) \) association with treatment outcome under an additive model, 20 of which demonstrated directionally consistent association and a resulting improvement in the association signal in a stage 1 and 2 combined meta-analysis (online supplementary table 2).

In stage 3, the 59 SNP stage 2 markers were further inspected for replication in two independent GWAS, comprising 595 patients from the UK and 272 patients from France (ReAct), separately. One SNP (rs11642056) failed QC criteria in both replication cohorts, leaving 58 SNPs for analysis. None of the tested SNPs showed nominal association with treatment outcome in these cohorts (table 2). However, the meta-analysis including all cohorts showed improved association signals for eight SNPs compared with our stage 1 results, three of which, rs1568885, rs1815443 and rs4411591, demonstrated directional consistency over all four cohorts studied (table 2).

None of the SNPs previously reported to be associated with treatment outcome showed evidence of association in the Dutch stage 1 cohort (table 3).16–20

**Explorative pathway analysis of stage 1**

We explored the stage 1 dataset for potential functional relationship between genes that showed evidence of association at \( p < 10^{-3} \) with treatment outcome using the IPA. This resulted

---

Table 1 Study population characteristics

<table>
<thead>
<tr>
<th>Stage 1</th>
<th>Combined</th>
<th>Dream</th>
<th>ApotheekZorg</th>
<th>Stage 2</th>
<th>Stage 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>REF Col*</td>
<td>WTCCC</td>
</tr>
<tr>
<td>Number</td>
<td>984</td>
<td>669</td>
<td>(69.7)</td>
<td>315</td>
<td>954</td>
</tr>
<tr>
<td>Gender female (%)</td>
<td>68.6</td>
<td>67.8</td>
<td>70.4</td>
<td>75.6</td>
<td>77.3</td>
</tr>
<tr>
<td>Anti-TNF agent</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infliximab</td>
<td>225 (22.9)</td>
<td>225 (33.7)</td>
<td>315 (100)</td>
<td>415 (43.5)</td>
<td>268 (45.0)</td>
</tr>
<tr>
<td>Adalimumab</td>
<td>638 (64.8)</td>
<td>323 (48.2)</td>
<td>70.4</td>
<td>174 (18.2)</td>
<td>68 (11.4)</td>
</tr>
<tr>
<td>Etanercept</td>
<td>121 (12.3)</td>
<td>121 (18.1)</td>
<td></td>
<td>365 (38.3)</td>
<td>259 (43.6)</td>
</tr>
<tr>
<td>MTX co-medication (%)</td>
<td>73.4</td>
<td>69.2</td>
<td>82.2</td>
<td>65.6</td>
<td>85.6</td>
</tr>
<tr>
<td>DAS28 Baseline</td>
<td>5.5±1.2</td>
<td>5.3±1.3</td>
<td>5.8±1.0</td>
<td>5.5±1.2</td>
<td>6.7±0.9</td>
</tr>
<tr>
<td>DAS28 DeltaDAS (14 weeks)</td>
<td>3.6±1.3</td>
<td>3.9±1.3</td>
<td>3.1±1.1</td>
<td>3.6±1.3</td>
<td>3.9±1.3</td>
</tr>
</tbody>
</table>

*The American College of Rheumatology Research and Education Foundation (REF) collection used for stage 2 is composed of nine different cohorts: Autoimmune Biomarkers Collaborative Network (ABCoN), Academic Medical Center Amsterdam (AMC), Behandelstrategieen voor Rheumatoid Arthritis (BeSt), Biological in Rheumatoid arthritis Genetics and Genomics Study Syndicate (BRAGGSS), Brigham Rheumatoid Arthritis Sequential Study (BRASS), Epidemiological Investigation of Rheumatoid Arthritis (EIRA), Immunex Early Rheumatoid Arthritis (ERA) study, Karolinska Institutet (KI) study, Jan van Breemen Institute (READE) study, Treatment of Early Aggressive RA (TEAR).

Numbers are depicted as n (%) or mean±SD.

DAS, Disease Activity Score; MTX, methotrexate; TNF, tumour necrosis factor.
The highest scoring network (p=10^{-5}) included 26 genes identified in the genome-wide association analysis stage 1, and nine additional interacting genes (figure 2). Importantly, this network is predicted to be involved in metabolic disease and biological processes underlying inflammatory response and cell morphology and contains genes implicated in TNF signalling (NFκB) and antibody formation (IgG).

**DISCUSSION**

In this report, we described the results of the largest GWAS of response to anti-TNF treatment in patients with RA conducted to date.

Using a multistage study design, we identified eight genetic loci showing suggestive evidence of association (improved p values) with treatment outcome in our overall meta-analysis, with three markers (rs4411591, rs1813443 and rs1568885) showing directional consistency over all four cohorts studied. In the combined cohort, eight identified loci together explain 3.8% of the variance in the treatment response. Although no single SNP reached a genome-wide level of significance (p<5×10^{-8}), these variants represent excellent candidates for further investigation.

Of the eight markers with suggestive evidence of association, two map to an intergenic region in which the nearest gene is interesting in terms of its biological function. The SNPs, rs12142623 and rs4651570, are located ~400 kb downstream from the phospholipase A2, group IVA (PLA2G4A) gene. The protein encoded by PLA2G4A is a phospholipase enzyme involved in generation of eicosanoids, molecules with regulatory function in inflammatory responses. TNFα is one of the first known stimuli for PLA2G4A activation, through the action of both TNF receptor subtypes. It might be possible that the identified SNPs influence long-range regulatory elements.

In addition, three of the eight identified SNPs map within genes. The marker rs4411591 maps to the Lox100130480, encoding a hypothetical protein, while rs2378945 is located in the nucleotide-binding protein-like (NUBPL) gene. NUBPL encodes a protein required for the assembly of the respiratory chain NADH dehydrogenase (complex I). Finally, rs1813443 is located in the intronic region of contactin 5 (CNTN5), a member of the immunoglobulin superfamily, which is thought to have a role in the formation of axon connections in the developing nervous system. Little is known about the possible
inflammation, the involvement of these genes in inflammatory disease, and there are no apparent functional links to anti-TNF treatment outcome, yet. If association with anti-TNF response can be confirmed in additional replication studies, future functional studies are needed to prove the biological link with anti-TNF response.

We did not find any evidence for association for the loci identified from previously published GWAS on anti-TNF response18–20 or for the PTPRC gene (table 3)16 17 in the Dutch patients included in our stage 1 GWAS, suggesting that these genes do not play a major role in anti-TNF treatment outcome in our population. However, the power of our study to detect an association with PTPRC loci at p<0.0005 in a set size of 882 patients was 26%.

However, the top network constructed with IPA indicated involvement of the genes showing suggestive association in stage 1 GWAS in the following processes: cell morphology, metabolic disease and inflammatory response. This is in line with the expected biological function of anti-TNF; TNFα is an important mediator of insulin resistance and it also impedes insulin-mediated glucose uptake.32 More importantly, other studies showed a positive long-term effect of TNF antagonists on insulin resistance, which correlated with improvement in disease activity.33 34 The identified network also harbours two interesting interacting molecules: NFκB and IgG. NFκB is a transcription regulator that is preferentially activated by TNF and is the main downstream target of the TNF signalling pathway.35 The finding that the genes identified through the GWAS interact with IgG is particularly interesting, since response failure and side effects of anti-TNF due to immunogenicity are not rare and it has been found that infliximab antibodies are exclusively of IgG isotype.36 Both molecules might have a central role in determining the outcome of treatment with anti-TNF agents. Besides this network, we could map the most prominent associations from stage 1 to the VAV1 and SPRED2 genes. VAV1 has been found to protect T cells from Fas-mediated apoptosis in Jurkat leukaemia T cells,37 and it has been confirmed that patients with RA show differential sensitivity to apoptosis of peripheral blood lymphocytes induced by anti-TNF therapy38 SPRED2 is a known RA risk locus.39 This network might represent new leads and new additional candidates for future research.

Our study, which included 2703 patients with RA treated with anti-TNF agents, is the largest GWAS of treatment outcome to date. However, our sample size still remains modest compared with genetic studies of risk of RA40 41 and...
other complex traits.42,43 Also, there are important aspects that may affect our results. There is considerable disease heterogeneity in RA. In our dataset, there is a difference in, for example, the number of women included in the studies, but also in type of anti-TNF treatment and co-medication use (table 1). Furthermore, the REF collection used for replication in stage 2 consists of nine smaller cohorts from several populations. Hence, combining results across four different cohorts of patients that are rather diverse in subject ascertainment and assessment and previous treatment can lead to different effect estimates among studies and false negative results. In addition, the DAS28 score, used as the measure of treatment outcome in our study, is a composite score including four measures: swollen and tender joint counts, erythrocyte sedimentation rate, and self-reported general health. This score is a powerful tool for measuring treatment response in a clinical setting. However, it is likely that this complex measure is influenced by genetic effects that are individually modest and would require large sample sizes to be detected.

In the present study, we have identified eight genetic loci that show evidence of influencing anti-TNF treatment response based on a multistage approach in a population of 2703 Caucasian patients with RA. Our findings require further validation in independent cohorts and/or at a functional level.

**REFERENCES**


Genome-wide association analysis of anti-TNF drug response in patients with rheumatoid arthritis

Masa Umicevic Mirkov, Jing Cui, Sita H Vermeulen, et al.

Ann Rheum Dis published online December 11, 2012
doi: 10.1136/annrheumdis-2012-202405

Updated information and services can be found at:
http://ard.bmj.com/content/early/2012/12/10/annrheumdis-2012-202405.full.html

These include:

Data Supplement
"Supplementary Data"
http://ard.bmj.com/content/suppl/2012/12/10/annrheumdis-2012-202405.DC1.html

References
This article cites 43 articles, 12 of which can be accessed free at:
http://ard.bmj.com/content/early/2012/12/10/annrheumdis-2012-202405.full.html#ref-list-1

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Advance online articles have been peer reviewed, accepted for publication, edited and typeset, but have not yet appeared in the paper journal. Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/
Articles on similar topics can be found in the following collections

- Connective tissue disease (2914 articles)
- Degenerative joint disease (3174 articles)
- Immunology (including allergy) (3422 articles)
- Musculoskeletal syndromes (3411 articles)
- Rheumatoid arthritis (2196 articles)
- Genetics (660 articles)

Advance online articles have been peer reviewed, accepted for publication, edited and typeset, but have not yet appeared in the paper journal. Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/