Investigation of genetically-regulated gene expression and response to treatment in rheumatoid arthritis highlights an association between *IL18RAP* expression and treatment response

Svetlana Cherlin[1], Myles J. Lewis[2], Darren Plant[3,4], Nisha Nair[3], Katriona Goldmann[2], Evan Tzanis[2], Michael R. Barnes[2], Paul McKeigue[5], Jennifer H. Barrett[6,7], Costantino Pitzalis[2], Anne Barton[3,4], MATURA Consortium, Heather J. Cordell[1]*

[1] Population Health Sciences Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK

[2] Centre for Experimental Medicine and Rheumatology, William Harvey Research Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK

[3] Centre of Genetics & Genomics Versus Arthritis, Manchester Academic Health Sciences Centre, The University of Manchester, Manchester, UK

[4] NIHR Manchester Biomedical Research Centre, Manchester University NHS Foundation Trust Manchester Academic Health Science Centre, Manchester, UK

[5] Centre for Population Health Sciences, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, UK

[6] NIHR Leeds Biomedical Research Centre, Leeds Teaching Hospitals NHS Trust, Leeds, UK

[7] School of Medicine, University of Leeds, Leeds, UK

Corresponding author:

Heather J. Cordell

Population Health Sciences Institute, Faculty of Medical Sciences, Newcastle University

International Centre for Life, Central Parkway, Newcastle upon Tyne, NE1 3BZ, UK

Tel: +44 (0)191 241 8669

E-mail: heather.cordell@ncl.ac.uk

Running title: Association between *IL18RAP* expression and treatment response in RA

ABSTRACT

Objectives

In this study, we sought to investigate whether there was any association between genetically-regulated gene expression (as predicted using various reference panels) and anti-TNF treatment response (change in erythrocyte sedimentation rate, ESR) using 3,158 European ancestry rheumatoid arthritis patients.

Methods

The genetically-regulated portion of gene expression was estimated in the full cohort of 3,158 subjects (as well as within a sub-cohort consisting of 1,575 UK patients) using the PrediXcan software package with three different reference panels. Estimated expression was tested for association with anti-TNF treatment response. As a replication/validation experiment, we also investigated the correlation between change in ESR with measured gene expression at the *Interleukin 18 Receptor Accessory Protein* (*IL18RAP*) gene in whole blood and synovial tissue, using an independent replication data set of patients receiving conventional synthetic disease modifying anti-rheumatic drugs, with directly measured (via RNA sequencing) gene expression.

Results

We found that predicted expression of *IL18RAP* showed a consistent signal of association with treatment response across the reference panels. In our independent replication data set, *IL18RAP* expression in whole blood showed correlation with the change in ESR between baseline and follow-up ($r = -0.35$, $p = 0.0091$). Change in ESR was also correlated with the expression of *IL18RAP* in synovial tissue ($r = -0.28$, $p = 0.02$).

Conclusion

Our results suggest that *IL18RAP* expression is worthy of further investigation as a potential predictor of treatment response in rheumatoid arthritis that is not specific to a particular drug type.

Keywords: Rheumatoid Arthritis, Treatment, Pharmacogenetics

INTRODUCTION

Tumour necrosis factor α inhibitors (anti-TNFs) are the most commonly-prescribed second-line drugs for conventional synthetic Disease Modifying Anti-Rheumatic Drug (csDMARD)-resistant rheumatoid arthritis (RA) patients. However, patients show a significant non-response rate to anti-TNF treatment[1, 2]. With recent advances in microarray and RNA sequencing (RNA-Seq) technologies, it is hypothesised that gene expression profiling might inform our understanding of the heterogeneity of responses to treatment in RA[3]. Indeed, Tanino et al.[4] identified ten genes predictive of response to the anti-TNF antibody Infliximab, based on a transcriptome analysis of white blood cells from RA patients, while Julià et al.[5] identified an eight-gene predictor model from microarray gene expression analysis on whole blood RNA samples from RA patients. Previously, using a microarray analysis of mononuclear cell RNAs, Lequerré et al.[6] was able to perfectly separate responders to Infliximab from non-responders.

However, when studying a large number of patients, measuring gene expression at a genome-wide scale might not be financially feasible, as RNA-seq remains more expensive than genome-wide genotyping approaches. The PrediXcan method/software package[7] is a cost-effective approach for estimating the genetically-regulated portion of gene expression at each gene from a genome-wide set of genes. PrediXcan estimates the component of a gene's expression determined by an individual's SNP genotypes (at SNPs in the vicinity of the gene), and then tests for association between the predicted expression and the phenotype. The estimation of gene expression is performed using whole-genome tissue-dependent prediction models trained with reference panels that have both SNP and gene expression data. Here, we applied PrediXcan to data on RA patients receiving anti-TNF treatment from the MAximising Therapeutic Utility for Rheumatoid Arthritis (MATURA) consortium[8], focussing on testing the association between the change in erythrocyte sedimentation rate (ESR) and predicted gene expression. We elected to focus on change in ESR as an objective measure of response that has been shown to have higher heritability than other measures of response[9].

METHODS

UK data set

The UK data set was comprised of imputed genome-wide SNP genotype data (9,084,265 SNPs) for up to 1583 patients receiving anti-TNF treatment from the MATURA consortium; this corresponds to the "anti-TNF, ESR data set" previously described by Cherlin et al.[10] Quality control (QC) on the imputed SNP data was performed using standard procedures outlined by Anderson et al.[11] Individuals were excluded if the reported sex did not match the sex assessed by genotype, and samples with elevated missingness rate, outlying heterozygosity rate, outlying ethnicity and relatedness were also excluded. SNPs were excluded if they had a post-imputation INFO score < 0.8. Genotype hard calls were set to missing if the posterior probability was < 0.9. The data was filtered by minor allele frequency (MAF > 0.01), Hardy-Weinberg disequilibrium (p > 0.000001) and missing genotype rate (<0.05). The SNP genotypes were encoded according to the number of copies of the minor allele possessed. The phenotype was defined as the difference between the follow-up ESR measure (measured at 6 months, or 3 months if this was not available) and the baseline ESR measure on the log scale, that is, $\log(ESR_{fu}) - \log(ESR_{bl})$. This difference was then adjusted (by taking as the final phenotype the standardised residuals from a linear regression, carried out in the statistical software package R) for baseline ESR, drug type (a 5-level categorical variable indicating Adalimumab, Etanercept, Infliximab, Certolizumab pegol and Golimumab), a separate binary indication of whether or not patients received another DMARD in addition to the anti-TNF treatment, gender and the first ten principal components (PCs) of the SNP genotypes. The final post-QC data set was comprised of 1,575 individuals and 4,542,023 SNPs.

Expanded European ancestry data set

An expanded European ancestry data set was constructed, consisting of imputed genotype data at 4,498,586 genome-wide SNPs for 3158 patients. This expanded data set consisted of a combination of the original (1,575 patient) UK data set and a separate independent data set of 1,583 US and EU patients, corresponding to a subset (to which we were granted access) of the

patients from a pre-existing international collaboration formed to study the genetics of response to TNFi agents[12]. The same QC procedures were performed separately on the 1,583 US/EU patients, and their post-QC SNP genotype data were merged with the data for the 1,575 UK patients. In this combined European ancestry data set, the phenotype, defined as the difference between the follow-up ESR measure and the baseline ESR measure on the log scale, was adjusted for the baseline ESR measure on the log scale ($\log(ESR_{bl})$) and the first three PCs (which was found sufficient to produce no inflation in the genome-wide set of test statistics for association between SNPs and phenotype) and subsequently adjusted for gender; other covariates were not available. The standardised residuals after all adjustments were then taken as the final phenotype.

Replication data set

The replication dataset consisted of 90 treatment-naive early rheumatoid arthritis patients fulfilling the 2010 ACR/EULAR RA Classification Criteria from the Pathobiology of Early Arthritis Cohort (PEAC), in whom ultrasound-guided synovial biopsies (n=87 post-QC) and whole blood samples (n=67) were subject to RNA-sequencing as previously described[13]. Notably both synovial biopsies and blood samples were taken prior to patients receiving any disease modifying treatment such as corticosteroids. 1μg of total RNA was used as an input material for library preparation using TruSeq RNA Sample Preparation Kit v2 (Illumina). Generated libraries were amplified with 10 cycles of PCR. Size of the libraries was confirmed using 2200 TapeStation and High Sensitivity D1K screen tape (Agilent Technologies) and concentration was determined by qPCR based method using Library quantification kit (KAPA). Multiplexed libraries (five per lane) were sequenced on Illumina HiSeq2500 to generate 50 million paired-end 75 base pair reads. Transcript abundance was derived using Kallisto v0.43.0 and tximport 1.4.0 using GENCODE v24/GRCh38 as reference, and transformed to regularised log expression (RLE) using DESeq2 1.14.1. RNA-Seq data have been deposited in ArrayExpress under Accession code E-MTAB-6141. Genotyping and QC was performed using the same methodology as for the UK data set as described above. Expression quantitative trait locus (eQTL) analysis on PEAC blood RNA-seq samples was performed using the matrix eQTL package in R[14] using 4 principal components to adjust for ancestry and 4 probabilistic estimation of expression residuals (PEER) calculated as per Stegle et al.[15] as model covariates.

Discovery analysis based on predicted gene expression

PrediXcan was applied to both the UK data set and the expanded European ancestry data set. In PrediXcan, an elastic net prediction model is built using a reference panel containing SNP and gene expression data. This model is then used to predict expression levels in the analysis cohort on the basis of the measured genotypes, and the resulting predicted expression levels are tested for association with the outcome of interest. We applied PrediXcan using three reference panels: (i) a MATURA reference panel comprising 210 MATURA samples (a subset of the UK samples used later for testing), for which SNP and gene expression data were available[8]); (ii) a GTEx reference panel for which PrediXcan provides pre-calculated models; this panel comprises 338 samples from the Genotype- Tissue Expression Consortium[16]; (iii) a DGN reference panel for which PrediXcan provides pre- calculated models; this panel comprises 922 samples from the Depression Genes and Networks Consortium[17].

To construct the MATURA reference panel, we used 33,170 QC-ed and batch-adjusted[18] gene expression probes from the Illumina HT-12 Gene Expression Beadchip, measured in whole-blood at baseline in 210 patients, together with QC-ed imputed genotype data at 3,978,972 genome-wide SNPs in the same patients. Probes that corresponded to different genes (according to the GENCODE version 19[19]) were removed, and probes that corresponded to the same gene were combined as specified[20]. The final gene expression data set consisted of 17,008 probes. We note that the SNPs used for constructing the MATURA reference panel corresponded to a subset of the SNPs present within the UK data set, however, they did not correspond to an exact subset of SNPs present in the expanded European ancestry data set owing to different SNPs remaining post-QC. We used default PrediXcan parameters to build the elastic net model (α = 0.4, window size = 1 Mbp, false discovery rate threshold = 0.05). The final MATURA reference panel included 1,573 genes. Models based on whole-blood GTEx and DGN reference panels were provided as part of the PrediXcan software (downloaded from http://predictdb.org/), and included 6,057 and 9,836 genes, respectively.

Replication analysis based on measured gene expression

Clinical parameters including DAS28 score and subcomponents, ESR, CRP, rheumatoid factor

(RF) and anti-citrullinated peptide antibody (ACPA) positivity/titre were collected at baseline and every 3 months. Patients were treated with methotrexate-based combination DMARDs (81%), methotrexate alone (6.8%), hydroxychloroquine alone (5.9%) or sulfasalazine alone (3.4%) or no DMARDs (2.5%). Clinical response was assessed by change in clinical parameters at 6 months and correlated with *IL18RAP* transcript levels measured by RNA-seq in baseline synovial biopsy or whole blood.

RESULTS

Discovery analysis based on predicted gene expression

The results for the UK and expanded European ancestry data sets are shown in figures 1 and 2, respectively. Although no genes pass experiment-wide significance, for the UK data set (figure 1), the most significant gene identified using the MATURA reference panel (*IL18RAP* on chromosome 2) achieved close to experiment-wide significance ($p = 4.3 \times 10^{-5}$), and this was also the top gene when using the DGN reference panel ($p = 6.7 \times 10^{-5}$). When using the GTEx reference panel, *IL18RAP* was the fourth top gene ($p = 1.5 \times 10^{-3}$).

For the expanded European ancestry data set (figure 2), *IL18RAP* was again the top gene when using the MATURA reference panel ($p = 1.4 \times 10^{-4}$). With the DGN reference panel, this gene was the fifth top gene ($p = 2.3 \times 10^{-4}$), and with the GTeX reference panel, it was the fourth top gene ($p = 4.0 \times 10^{-4}$). Online supplementary table 1 shows the significance levels and effect estimates achieved at *IL18RAP* when using either the full 3158-person expanded European ancestry cohort or when it is divided into its constituent UK or US/EU sub-cohorts. (Note that the results for the UK sub-cohort differ slightly from those obtained in the original analysis of the UK data set on account of (a) the different phenotypic adjustments made when using the full expanded European ancestry data set and (b) slightly different SNPs being available for prediction of expression, see below). The signal is seen to be predominantly driven by the results from the UK sub-cohort, with the US/EU sub-cohort showing the same direction of effect, but with the effect size considerably attenuated.

The PrediXcan models for predicting the expression of *IL18RAP* involved 77, 46 and 86 SNPs when using the MATURA, GTEx and DGN reference panels, respectively (online supplementary table 2). All 77 SNPs from the MATURA reference panel appeared in the

expanded European ancestry data set, while 37 out of 46 SNPs from the GTEx reference panel and 82 out of 86 SNPs from the DGN reference panel appeared. Density estimates for the resulting predicted expression values are shown in online supplementary figure 1; their relationship with phenotype is shown in online supplementary figure 2. The prediction R^2 statistic for *IL18RAP* (based on PrediXcan's internal tenfold cross-validation procedure using the relevant reference panel) was 0.32 ($p = 1.9 \times 10^{-19}$) with the MATURA reference panel, 0.30 ($p = 4.4 \times 10^{-247}$) with the GTEx reference panel and 0.71 ($p = 2.8 \times 10^{-28}$) with the DGN reference panel, suggesting reasonable predictive ability for expression at this gene across all panels. As expected (given the association between SNPs contributing to the prediction models and expression, and between predicted expression and response), a number of SNPs also showed direct associations with response (online supplementary table S2), although as noted previously[10] these do not meet genome-wide significance levels.

The most significant gene overall using the expanded European ancestry data set was *ARV1* on chromosome 1, which appeared both when using the DGN reference panel ($p = 9.1 \times 10^{-5}$) and the GTEx reference panel ($p = 6.4 \times 10^{-5}$). This gene was absent on the MATURA reference panel because the PrediXcan software failed to predict its expression value. In the original analysis of the UK data set, the signals for the *ARV1* gene were generally weaker than, or similar to, those seen at *IL18RAP* (figure 1). Additionally, the prediction accuracy for *ARV1* in the expanded European ancestry data set, as measured by the R^2 statistic, was very low ($R^2 = 9.6 \times 10^{-3}$; $p = 0.072$) with the GTEx reference panel, and relatively lower ($R^2 = 0.18$; $p = 2.1 \times 10^{-41}$) than that seen for *IL18RAP* with the DGN reference panel, suggesting that these results at *ARV1* should be interpreted with caution.

Replication analysis based on measured gene expression

In the replication data set, we observed a significant expression quantitative trait locus (eQTL) association ($P = 5.8 \times 10^{-11}$) between multiple SNPs across the *IL18RAP* locus and *IL18RAP* expression measured by RNA-Seq of whole blood samples in patients with early RA (figure 3a, b), thus confirming that *IL18RAP* genetic polymorphisms regulate expression of *IL18RAP* in peripheral blood in early RA patients. The expression of *IL18RAP* measured in whole blood showed correlation with the change in ESR between baseline and 6-month follow-up ($r = -0.35$; $p = 0.0091$) in RA patients treated with MTX-based combination DMARD therapy (figure 3c); specifically

each unit increase in *IL18RAP* regularised log expression resulted in a 13.4 mm/hour decrease in ESR between baseline and 6 months. Also, a correlation was observed between the expression of *IL18RAP* in synovial tissue and the change in ESR ($r = -0.28$; $p = 0.02$) (figure 3d); specifically each unit increase in regularised log expression resulted in a 11.8 mm/hour decrease in change in ESR over 6 months. Thus, our replication experiment based on actual measured gene expression (in an independent set of patients) validates the association between predicted *IL18RAP* expression and treatment response seen in the discovery cohort.

DISCUSSION

In this study, we investigated the association between the genetically-regulated portion of gene expression and change in the ESR in a large cohort of RA patients from the MATURA consortium. We found that predicted expression of *IL18RAP* showed a consistent signal across data sets analysed using different reference panels, while achieving a reasonable level of prediction accuracy as measured by the prediction R^2. Despite the consistency of the results for *IL18RAP*, some differences in the strength of the signal were observed for different data sets and reference panels. These differences require further investigation, however they can be partly explained by the different sample sizes (and SNPs available to inform prediction) in the different reference panels. In an independent replication data set of patients treated with csDMARDs with measured gene expression, the asso- ciation between expression of *IL18RAP* and change in the ESR was confirmed both in whole-blood and synovial tissue, highlighting *IL18RAP* as a gene worthy of further investigation for prediction of treatment response in RA that is not treatment-specific. No other expressed genes were consistently associated with response, providing confidence that it is the *IL18RAP* gene that is driving the association rather than serving as a proxy for another gene.

The protein encoded by *IL18RAP* enhances the IL-18 binding activity of the IL-18 receptor and plays a role in signaling by IL-18 [21]. IL-18 plays an inflammatory role in RA[22, 23] and has previously been identified as a potential therapeutic target in the treatment of RA[24, 25]. It has been suggested that IL-18 plays some part in the degradation of articular cartilage in arthritis[26]. Additionally, Rooney et al.[27] showed that synovial tissue IL-18 production measured by immunohistochemistry was correlated with serum C reactive protein in inflammatory arthritis, while

Joosten et al.[28] found a correlation between the level of IL-18 in the synovial tissue of the RA patients and ESR.

Previous studies have reported a potential association between *IL18RAP* and treatment response in RA. Analysis of CAGE sequencing data from the FANTOM5 consortium showed that *IL18RAP* is highly expressed in neutrophils, gamma delta T cells, eosinophils and NK cells[29]. Analysis of the BioGPS database (http://biogps.org/) confirms that *IL18RAP* is highly expressed in NK cells. *IL18RAP* expression is upregulated in NK and T cells in response to IFN-alpha and IL-12[30]. *IL18RAP* was found to be significantly upregulated (adjusted $P = 5.5 \times 10^{-78}$) in NK cells in single cell RNA-seq RA synovium data from Stephenson et al.[31]. Similarly in a second single cell RNA-seq study of RA synovium[32], *IL18RAP* shows increased expression in synovial tissue T cell populations. In RA synovium, the baseline expression of the S1 module (NK cell surface signature) from Li et al.[33] which includes *IL18RAP* as one of its 45 genes, correlates significantly with change in ESR. Additionally, the synovial baseline expression for another NK cell module (M7.2) which includes *IL18RAP* is also significant for the change in ESR.

Our own investigation of the relationship between measured expression of *IL18RAP* and change in the ESR in whole-blood and synovial tissue in our replication data set was motivated by our initial identification of a relationship between change in the ESR and predicted expression of *IL18RAP* in our discovery data sets, using the PrediXcan method/software. Other methods/software packages for performing transcriptome-wide association studies exist, but as shown by Fryett et al[34], they tend to perform very similarly to one another. These methods are dependent on the underlying eQTL data used to build the prediction models, and therefore would generally be expected to give very similar results. Given that the external datasets used to inform the prediction were derived from population studies and would unlikely to have been enriched for patients with RA, the risk of the association detected with ESR being spurious is low.

We elected to focus on change in ESR as an objective measure of response that has been shown to have higher heritability than other measures of response[9]. Other clinical outcomes relating to anti-TNF treatment response, such as joint destruction scores or CRP, could certainly be assessed using similar approaches. However, joint destruction scores were not available in our discovery data set, and there were many missing values for CRP, making this a less attractive

option in this instance.

Overall, our results, combined with the existing evidence, suggest that the expression of *IL18RAP* in whole blood might have utility for predicting response to treatment in RA. However, the effect observed in our replication samples (11.8-13.4 mm/hour decrease in change in ESR over 6 months) is, by itself, probably too small to be clinically useful, and the small to moderate correlations seen between *IL18RAP* expression and change in ESR (figure 3c, d) suggest that the actual predictive ability of *IL18RAP* expression alone may be limited. This approach shows the value of integrating genetic and expression data to identify factors correlated with response which could be incorporated into a multi-omic predictive model in the future. Further investigation of the relationship between *IL18RAP* expression and varying measures of treatment response in additional patient cohorts is thus warranted.

Ackowledgements

We thank all the patients who have contributed to this study as well as the research and clinical staff who supported patient recruitment. We thank the Medical Research Council and Arthritis Research UK (ARUK) for their joint funding of the MAximising Therapeutic Utility in Rheumatoid Arthritis (MATURA) study. We thank MATURA consortium members for their input and feedback during the design of this project.

Contributors

Substantial contributions to conception or design of the study: SC, PM, JHB, CP, AB, HJC. Substantial contributions to drafting the manuscript: SC, MJL, HJC. Substantial contributions to data acquisition: CP, AB. Substantial contributions to data analysis or interpretation: SC, MJL, DP, NN, KG, ET, MRB. All authors contributed to revising the manuscript critically for important intellectual content and approved the final manuscript. The funding agencies had no part in writing or reviewing the manuscript. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Funding

Support for this work was provided by the Wellcome Trust (Grant 102858/Z/13/Z) and an MRC/Arthritis Research UK award: Maximizing Therapeutic Utility in RA (MATURA) (Grant MR-K015346). The Pathobiology of Early Arthritis Cohort (PEAC) was supported by funding from the Medical Research Council (MRC) (grant number G0800648). JHB receives support from the NIHR through the Leeds Biomedical Research Centre. This project was enabled through access to the MRC eMedLab Medical Bioinformatics infrastructure, supported by the Medical Research Council (grant number MR/L016311/1). We thank Arthritis Research UK (grant number 20385) and the Musculoskeletal theme of the NIHR Manchester BRC for their support.

Competing Interests

The authors declare no competing interests.

Patient and public involvement statement

This study was conducted as part of the programme of work undertaken by the MATURA Consortium (http://www.matura.whri.qmul.ac.uk/). A patient advisory group was established in 2014 when the MATURA project commenced. The group meets regularly to: ensure MATURA strategy is maintaining relevance, accountability and direction by embedding patients and members of the public within the decision making processes; determine what level of confidence in tests, and what type of tests, would be acceptable to patients for treatment decisions; maximise patient recruitment to research studies by increasing awareness through patient groups; readily obtain patients perspective on grant applications related to stratified medicines for RA; facilitate the dissemination of the results from MATURA research, for instance by producing lay summaries of papers in conjunction with the researchers (http://www.matura.whri.qmul.ac.uk/news.php).

Patient consent for publication

Not required.

Ethics approval

This study used anonymised data for human subjects from an international collaboration of 13 studies originally published in PLOS Genetics 2013;9:e1003394. All participants provided informed consent and institutional review board and ethics approvals were in place for each of the studies and described in the original publication. Approval for additional data pre- sented in this manuscript was provided by the North West 6 Central Manchester South Ethics Committee (COREC 04/Q1403/37) and the London–Dulwich Research Ethics Committee (REC 05/Q0703/198).

Data sharing statement

Data are available upon reasonable request.

References

[1] Barrera, P., van der Maas, A., van Ede, A. E., et al. (2002). Drug survival, efficacy and toxicity of monotherapy with a fully human anti-tumour necrosis factor-α antibody compared with methotrexate in long-standing rheumatoid arthritis. Rheumatology *41*, 530–439.

[2] Hyrich, K. L., Watson, K. D., Symmons, D. P., et al. (2006). Predictors of response to anti-TNF-α therapy among patients with rheumatoid arthritis: results from the British Society for Rheumatology Biologics Register. Rheumatology *45*, 1558–1565.

[3] Häupl, T., Stuhlmüller, B., Grützkau, A., et al. (2009). Does gene expression analysis inform us in rheumatoid arthritis? Ann Rheum Dis *69 (Suppl 1)*, i37–i42.

[4] Tanino, M., Matoba, R., Nakamura, S., et al. (2009). Prediction of efficacy of anti-TNF biologic agent, infliximab, for rheumatoid arthritis patients using a comprehensive transcriptome analysis of white blood cells. Biochemical and Biophysical Research Communications *387*, 261–265.

[5] Julià, A., Erra, A., Palacio, C., et al. (2009). An eight-gene blood expression profile predicts the response to infliximab in rheumatoid arthritis. PLoS ONE *4*, e7556.

[6] Lequerré, T., Gauthier-Jauneau, A.-C., Bansard, C., et al. (2006). Gene profiling in white blood cells predicts infliximab responsiveness in rheumatoid arthritis. Arthritis Research & Therapy *8*, R105.

[7] Gamazon, E. R., Wheeler, H. E., Shah, K. P., et al. (2015). A gene-based association method for mapping traits using reference transcriptome data. Nature Genetics *47*, 1091–1098.

[8] Barton, A. and Pitzalis, C. (2017). Stratified medicine in rheumatoid arthritis – the MATURA programme: targeted treatment for patients. Rheumatology *56*, 1247–1250.

[9] Massey, J., Plant, D., Hyrich, K., et al. (2018). Genome-wide association study of response to tumour necrosis factor alpha inhibitor therapy in Rheumatoid Arthritis. The Pharmacogenetics Journal *18*, 657–664.

[10] Cherlin, S., Plant, D., Taylor, J. C., et al. (2018). Prediction of treatment response in rheumatoid arthritis patients using genomewide SNP data. Genetic Epidemiology *42*, 754–771.

[11] Anderson, C. A., Pettersson, F. H., Clarke, G. M., et al. (2010). Data quality control in genetic case-control association studies. Annals of the Rheumatic Diseases *5*, 1564–1573.

[12] Cui, J., Stahl, E. A., Saevarsdottir, S., et al. (2013). Genome-wide association study and gene expression analysis identifies *CD84* as a predictor of response to Etanercept therapy in Rheumatoid Arthritis. PLOS Genetics *3*, e1003394.

[13] Lewis, M. J., Barnes, M. R., Blighe, K., et al. (2019). Molecular portraits of early rheumatoid arthritis identify clinical and treatment response phenotypes. Cell Rep *28*, 2455–2470.

[14] Shabalin, A. A. (2012). Matrix eQTL: ultra fast eQTL analysis via large matrix operations. Bioinformatics *28*, 1353–1358.

[15] Stegle, O., Parts, L., Piipari, M., et al. (2012). Using probabilistic estimation of expression residuals (PEER) to obtain increased power and interpretability of gene expression analyses. Nat Protoc *7*, 500–507.

[16] GTEx Consortium. (2013). The genotype-tissue expression (GTEx) project. Nature Genetics *25*, 580–585.

[17] Battle, A., Mostafavi, S., Zhu, X., et al. (2014). Characterizing the genetic basis of transcriptome diversity through RNA-sequencing of 922 individuals. Genome Res *24*, 14–24.

[18] Johnson, W. E., Li, C., and Rabinovic, A. (2006). Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics *8*, 118–127.

[19] GENCODE (2019). https://www.gencodegenes.org/. Accessed: 2019-10-10.

[20] Miller, J. A., Cai, C., Langfelder, P., et al. (2011). Strategies for aggregating gene expression data: the collapseRows R function. BMC Bioinformatics *12*, 322.

[21] GENCARDS (2019). https://www.genecards.org/. Accessed: 2019-10-10.

[22] Gracie, J. A., Forsey, R. J., Chan, W. L., et al. (1999). A proinflammatory role for IL-18 in rheumatoid arthritis. J Clin Invest *104*, 1393–1401.

[23] Liew, F. Y. and McInnes, I. B. (2002). Role of interleukin 15 and interleukin 18 in inflammatory response. Ann Rheum Di *61 (Suppl 2)*, ii100–ii102.

[24] Volin, M. V. and Koch, A. E. (2011). Interleukin-18: a mediator of inflammation and angiogenesis in rheumatoid arthritis. Journal of Interferon & Cytokine Research *31*, 745–751.

[25] McInnes, I. b., Liew, F. Y., and Gracie, J. A. (2004). Interleukin-18: a therapeutic target in rheumatoid arthritis? Arthritis Res Ther *7*, 38–41.

[26] Dai, S.-M., Shan, Z.-Z., Nishioka, K., et al. (2005). Implication of interleukin 18 in production of matrix metalloproteinases in articular chondrocytes in arthritis. Ann Rheum Dis *64*, 735–742.

[27] Rooney, T., Murphy, E., Benito, M., et al. (2004). Synovial tissue interleukin-18 expression and the response to treatment in patients with inflammatory arthritis. Ann Rheum Dis *63*, 1393–1398.

[28] Joosten, L. A. B., Radstake, T. R. D., Lubberts, E., et al. (2003). Association of interleukin-18 expression with enhanced levels of both interleukin-1β and tumor necrosis factor α in knee synovial tissue of patients with rheumatoid arthritis. Arthritis & Rheumatism *48*, 339–347.

[29] FANTOM Consortium and the RIKEN PMI and CLST (DGT), Forrest, A. R., Kawaji, H., et al. (2014). A promoter-level mammalian expression atlas. Nature *507*, 462–470.

[30] Sareneva, T., Julkunen, I., and Matikainen, S. (2000). IFN-α and IL-12 induce IL-18 receptor gene expression in human NK and T cells. J Immunol *165*, 1933–1938.

[31] Stephenson, W., Donlin, L. T., Butler, A., et al. (2018). Single-cell RNA-seq of rheumatoid arthritis synovial tissue using low-cost microfluidic instrumentation. Nat Commun *9*, 791.

[32] Zhang, F., Wei, K., Slowikowski, K., et al. (2019). Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry. Nat Immunol *20*, 928–942.

[33] Li, S., Rouphael, N., Duraisingham, S., et al. (2014). Molecular signatures of antibody responses derived from a systems biology study of five human vaccines. Nat Immunol *15*, 195–204.

[34] Fryett, J.J, Inshaw, J., Morris, A.P., and Cordell, H.J. (2018). Comparison of methods for transcriptome imputation through application to two common complex diseases. Eur J Hum Genet *26*, 1658–1667.

Figure Legends

Figure 1: Manhattan plots of P -values from tests of association between genetically-regulated gene expression and the change in ESR for the 1575-person UK data set. The genetically-regulated gene expression was estimated with (a) the MATURA reference panel, (b) the GTEx reference panel, and (c) the DGN reference panel. On each panel, the red dashed line represents the experiment-wide significance level computed using a Bonferroni correction for the number of tests performed. The black diamond represents the *IL18RAP* gene. The white diamond represents the *ARV1* gene.

Figure 2: Manhattan plots of P-values from tests of association between genetically-regulated gene expression and the change in ESR for the 3158-person expanded European ancestry data set. The genetically-regulated gene expression was estimated with (a) the MATURA reference panel, (b) the GTEx reference panel, and (c) the DGN reference panel. On each panel, the red dashed line represents the experiment-wide significance level computed using a Bonferroni correction for the number of tests performed. The black diamond represents the *IL18RAP* gene. The white diamond represents the *ARV1* gene.

Figure 3: Confirmation of the *IL18RAP* expression quantitative trait locus and clinical consequences in rheumatoid arthritis. (a) Manhattan plot showing expression quantitative trait locus analysis comparing influence of SNPs at the *IL18RAP* locus on *IL18RAP* expression in blood measured by RNA-seq. (b) Scatter plot of SNP rs10439410 in the 5' upstream region of *IL18RAP* and *IL18RAP* expression in whole blood. (c & d) Correlation between the change in ESR between baseline and 6 months of combination DMARD therapy and *IL18RAP* expression measured by RNA-seq in whole blood (c) and synovial tissue (d).

Key messages

What is already known about this subject?

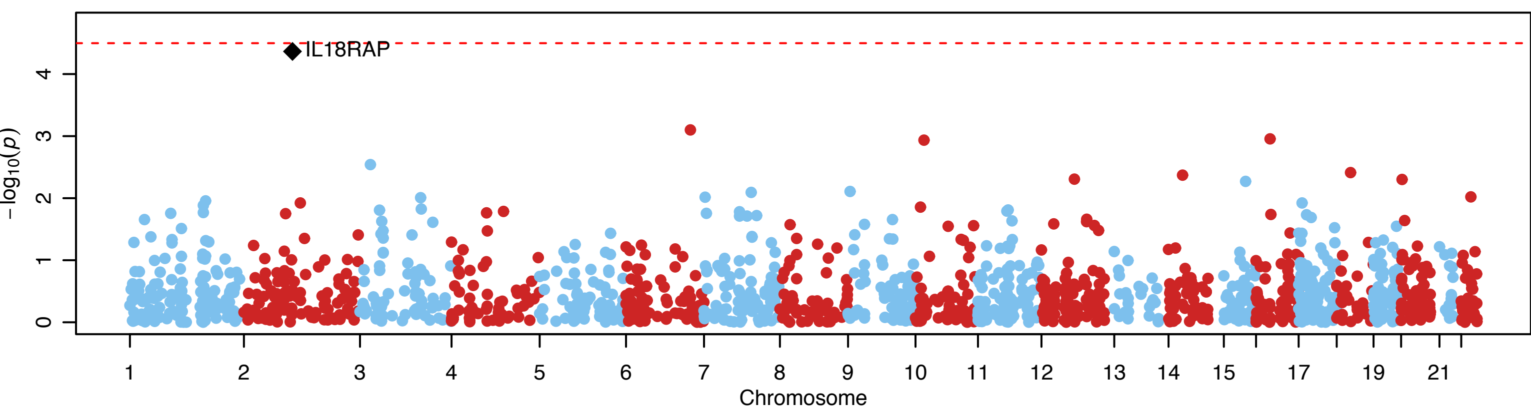
- IL-18 plays an inflammatory role in rheumatoid arthritis and has previously been identified as a potential therapeutic target.
- The protein encoded by the gene *IL18RAP* enhances the IL-18-binding activity of the IL-18 receptor and plays a role in IL-18 signaling.

What does this study add?

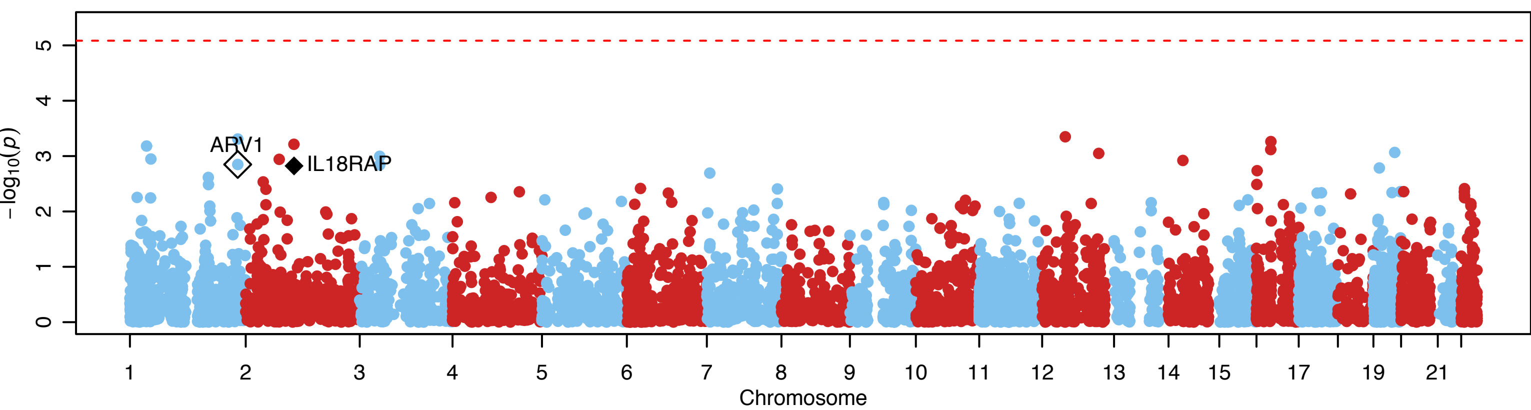
- We demonstrate a robust association between *IL18RAP* gene expression (both in whole-blood and synovial tissue) and treatment response in rheumatoid arthritis.
- The association between *IL18RAP* expression and treatment response is not specific to a particular drug type but is observed across different treatments.

How might this impact on clinical practice or future developments?

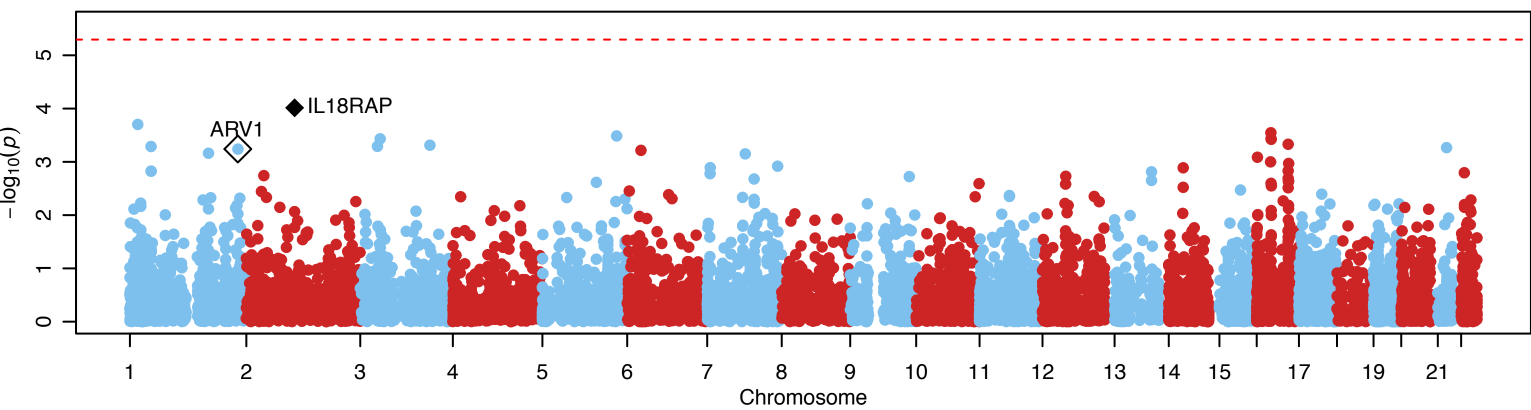
- Measurements of *IL18RAP* expression could potentially be incorporated into a multi-omic predictive model for treatment response in rheumatoid arthritis in the future.



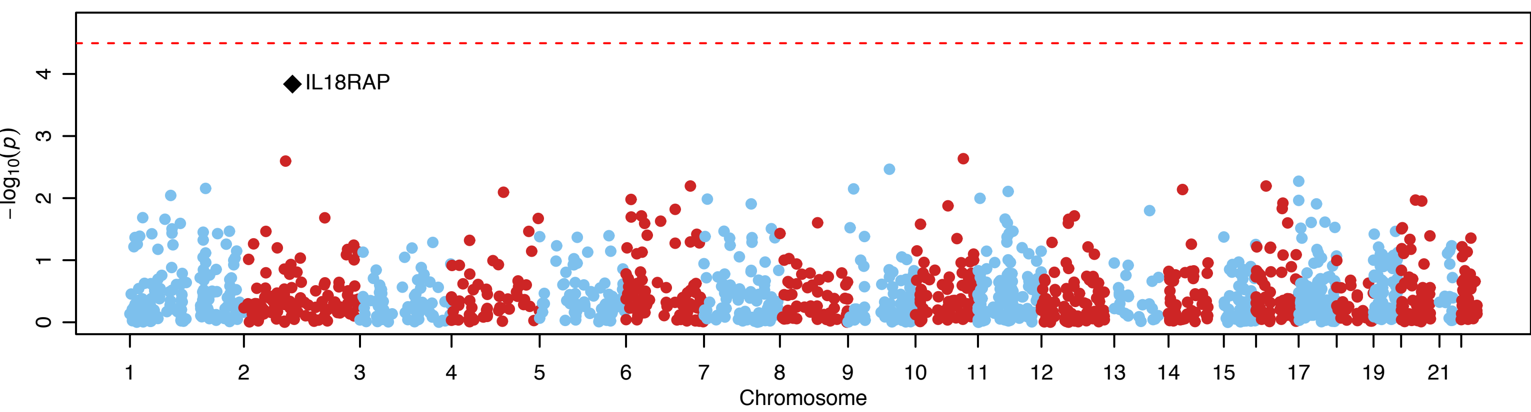
(a)



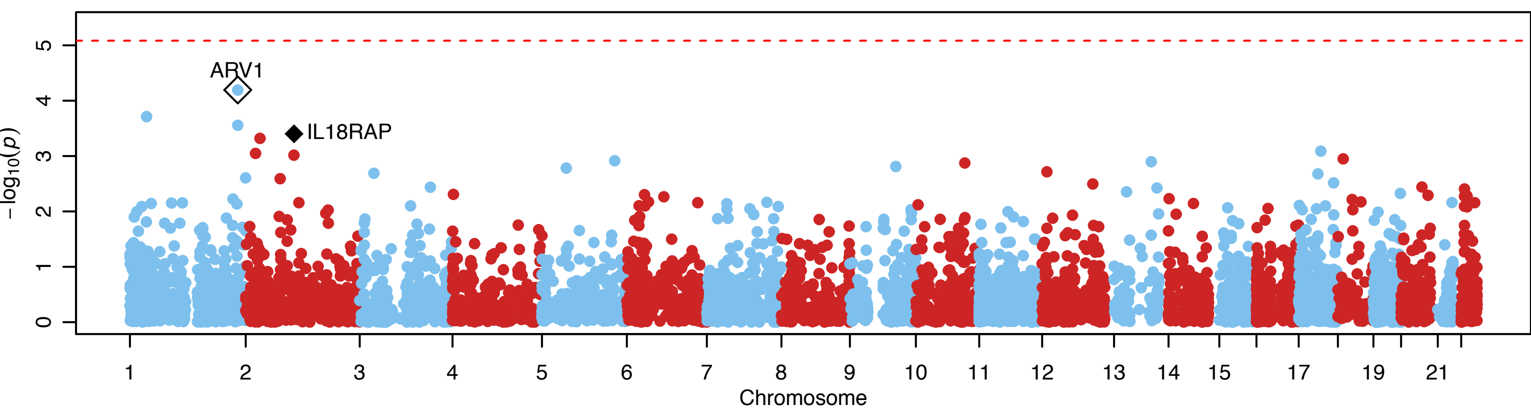
(b)



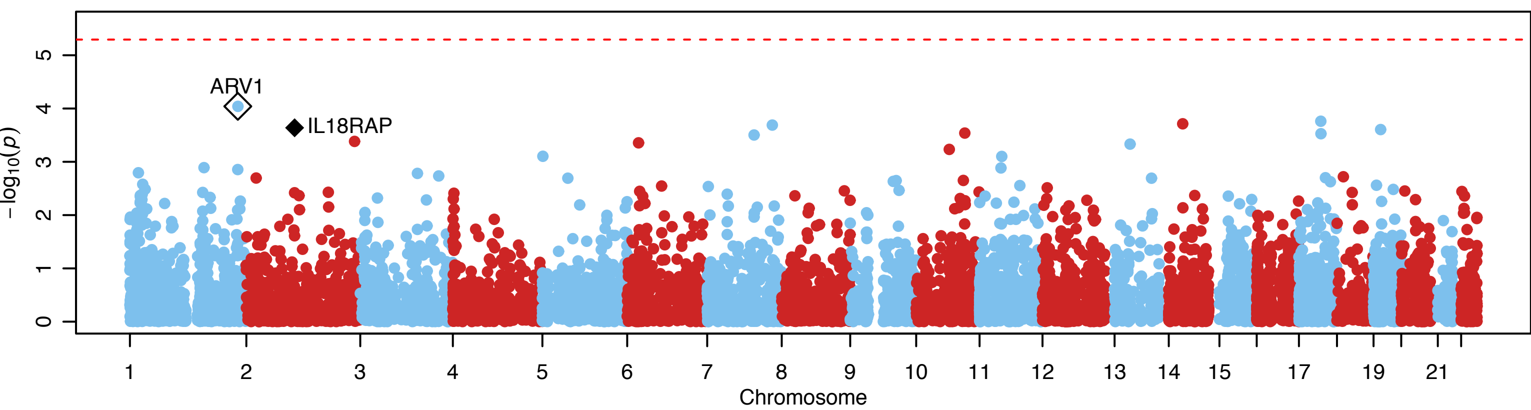
(c)



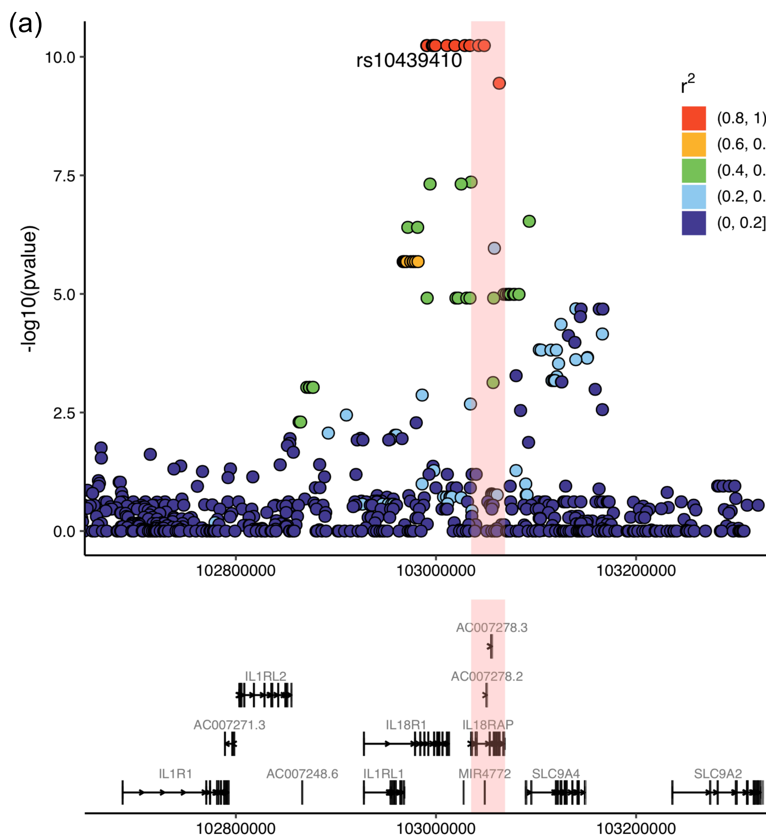
(a)

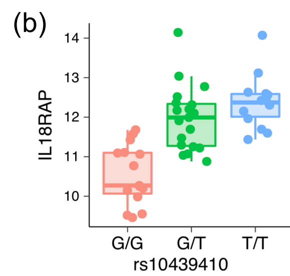


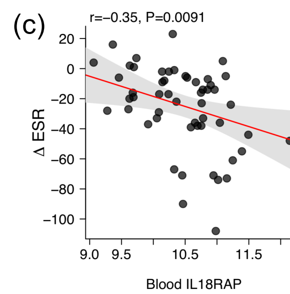
(b)

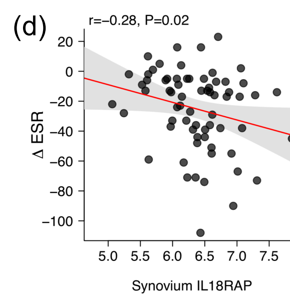


(c)


(a)
rs10439410
-log10(pvalue)
10.0
7.5
5.0
2.5
0.0
102800000
103000000
103200000
r2
(0.8, 1)
(0.6, 0.8]
(0.4, 0.6]
(0.2, 0.4]
(0, 0.2]
AC007278.3
AC007278.2
IL1RL2
AC007271.3
IL18R1
IL18RAP
IL1R1
AC007248.6
IL1RL1
MIR4772
SLC9A4
SLC9A2


(b)
IL18RAP
G/G
G/T
T/T
rs10439410


(c)
r=−0.35, P=0.0091
Δ ESR
Blood IL18RAP


(d)
r=−0.28, P=0.02
Δ ESR
Synovium IL18RAP

Investigation of genetically-regulated gene expression and response to treatment in rheumatoid arthritis highlights an association between *IL18RAP* expression and treatment response

Online Supplementary Material

Table S1. Effect sizes and significance levels achieved at *IL18RAP* when using either the full 3158-person expanded European ancestry cohort or when it is divided into its constituent UK (1575 people) and US/EU (1583 people) sub-cohorts.

Reference panel	Gene	UK		US/EU		Full	
		Effect size	*P*-value	Effect size	*P*-value	Effect size	*P*-value
MATURA	*IL18RAP*	-0.210	0.00021	-0.054	0.17344	-0.131	0.00014
GTEx	*IL18RAP*	-0.235	0.00105	-0.076	0.11999	-0.152	0.00040
DGN	*IL18RAP*	-0.128	0.00025	-0.030	0.21817	-0.079	0.00023

Table S2. The columns "Effect MATURA", "Effect GTEx" and "Effect DGN" represent the effects of the SNPs in the prediction models when using the MATURA (77 SNPs), GTEx (46 SNPs) and DGN (86 SNPs) reference panels. The columns "Coeff" and "*P*-value" represent the coefficient and the *P*-value from the association between the SNPs and the phenotype in the UK data set.

SNP	BP	Effect MATURA	Effect GTEx	Effect DGN	Coeff	*P*-value
rs13413645	-	-	-0.040	-	-	-
rs6759556	-	-	0.216	-	-	-
rs2310220	-	-	0.094	-	-	-
rs11465677	-	-	-0.003	-	-	-
rs17775170	-	-	0.009	-	-	-
rs10178585	-	-	-0.028	-	-	-
rs10185170	-	-	0.068	-	-	-
rs12328682	-	-	0.004	-	-	-
rs7599071	-	-	-0.010	-	-	-
rs1093515	-	-	-	0.015	-	-
rs13427957	-	-	-	-0.015	-	-
rs266064	-	-	-	-0.004	-	-
rs4319952	-	-	-	-0.021	-	-
rs7589943	102038936	-	0.055	-	0.006	0.96600
rs290772	102126220	-	-	0.007	-0.065	0.34800
rs75094400	102142042	0.080	-	-	-0.236	0.02390
rs17201799	102179247	-	0.054	-	-0.038	0.37600
rs12990046	102543479	0.039	-	-	-0.007	0.85600
rs11678842	102674806	-0.010	-	-	0.074	0.08400
rs13388182	102722402	-	-	-0.020	0.032	0.38700
rs13029804	102723017	-	-	0.014	-0.072	0.11000
rs1812326	102860411	-0.031	-	-	0.069	0.05930
rs10186746	102866377	-0.026	-	-	0.065	0.07800
rs62151694	102899464	0.049	-	-	-0.072	0.25500
rs13007174	102960487	0.006	-	-	-0.081	0.17300
rs13014044	102961366	0.005	-	-	-0.087	0.14800
rs1946131	102961929	0.005	-	-	-0.077	0.19600
rs1054096	102962350	0.005	-	-	-0.079	0.18600
rs12989197	102962739	0.004	-	-	-0.075	0.21100
rs12996097	102963628	0.004	-	-	-0.081	0.17400
rs13028993	102963949	0.004	-	-	-0.076	0.20100
rs12999542	102965392	-	-	0.017	-0.083	0.16300
rs13014644	102971363	0.015	-	-	-0.079	0.18500
rs13015714	102971865	-	0.018	-	0.111	0.00955
rs11465567	102978400	0.006	-	-	-0.072	0.22800

rs11465572	102980223	0.001	-	-	-0.071	0.23200
rs4134504	102982223	0.005	-	-	-0.081	0.17000
rs2058622	102985424	-	-	0.016	0.112	0.00945
rs2058623	102986170	-	-	0.015	0.113	0.00863
rs1465321	102986618	-	-	0.002	0.113	0.00863
rs11465597	102987213	-	-	0.006	-0.057	0.33600
rs10439410	102990788	0.007	-	0.031	-0.143	0.00005
rs6731157	102991191	0.006	-	-	-0.143	0.00005
rs6745614	102991213	0.007	-	-	-0.143	0.00005
rs6758936	102991369	0.007	-	0.016	-0.143	0.00005
rs2270297	102992675	-	0.011	0.050	0.115	0.00729
rs11465623	102993039	-	0.058	0.019	-0.020	0.76900
rs6753717	102993161	-	0.012	0.048	0.115	0.00729
rs2041739	102994333	0.012	-	0.018	-0.143	0.00006
rs6750020	102994714	-	0.013	0.046	0.114	0.00807
rs10208196	102996345	0.007	-	0.019	-0.143	0.00005
rs7556917	102997720	0.008	-	-	-0.143	0.00005
rs7584093	102997721	0.008	-	-	-0.143	0.00006
rs3213733	102997884	-	-	0.001	-0.102	0.03120
rs3213732	102998279	0.008	-	0.021	-0.144	0.00005
rs10204757	102998974	0.008	-	-	-0.143	0.00005
rs6760621	102999952	0.008	-	0.021	-0.143	0.00005
rs11465641	103000868	0.009	-	-	-0.143	0.00005
rs17651485	103001650	-	0.057	0.086	-0.027	0.68000
rs3771161	103003961	-	-	0.001	-0.099	0.03470
rs3771159	103004958	0.009	-	-	-0.143	0.00005
rs11903946	103005330	0.009	-	-	-0.143	0.00005
rs6706002	103006104	0.009	-	0.023	-0.143	0.00005
rs6749014	103006448	0.009	-	0.022	-0.144	0.00005
rs12712146	103008714	0.006	-	-	-0.156	0.00002
rs4851004	103009537	0.009	-	0.023	-0.144	0.00005
rs6732138	103009662	0.009	-	-	-0.146	0.00004
rs3771158	103009894	-	-	0.018	-0.095	0.04320
rs1420096	103010912	0.009	-	-	-0.145	0.00005
rs2287033	103011237	0.009	-	0.022	-0.145	0.00005
rs1568681	103014696	-0.044	-	-	0.118	0.00606
rs1420094	103015687	0.011	-	0.023	-0.143	0.00005
rs6710528	103016142	0.009	-	0.022	-0.142	0.00006
rs3732124	103018052	-	-	0.022	-0.143	0.00005
rs4851571	103019000	-	-	0.022	-0.144	0.00005
rs4851572	103019031	0.009	-	0.022	-0.144	0.00005
rs10202813	103019740	-	-	0.014	-0.099	0.03460

rs1035129	103019785	0.009	-	-	-0.145	0.00005
rs1035127	103019919	-0.045	0.027	0.049	0.114	0.00767
rs6710034	103023678	0.010	-	0.021	-0.144	0.00005
rs7589142	103024660	0.010	-	-	-0.145	0.00004
rs4851007	103024813	-0.045	0.026	0.047	0.114	0.00767
rs4851575	103025203	-0.045	0.171	0.046	0.113	0.00811
rs10181785	103025274	-	-	0.010	-0.097	0.03900
rs12712148	103025547	-	-	0.007	-0.096	0.03990
rs11687768	103025738	-	-	0.005	-0.096	0.04110
rs6543126	103027104	0.009	-	-	-0.144	0.00005
rs10203558	103027640	0.007	-	0.021	-0.144	0.00005
rs10200952	103027651	0.007	-	0.021	-0.143	0.00006
rs7586983	103028066	-	-	0.008	-0.096	0.04110
rs6543127	103028301	0.009	-	-	-0.143	0.00005
rs4851576	103028895	0.012	-	0.022	-0.143	0.00005
rs4851577	103028921	0.017	-	0.021	-0.143	0.00005
rs4851579	103028984	0.009	-	0.021	-0.143	0.00005
rs6543132	103029410	-0.044	-	-	0.114	0.00799
rs1807782	103033147	-0.017	0.023	0.045	0.115	0.00765
rs1035125	103033839	0.005	-	-	-0.143	0.00005
rs2160201	103033961	0.017	-	-	-0.143	0.00005
rs1420106	103035044	-0.043	0.022	0.045	0.113	0.00857
rs1420105	103035119	0.017	-	0.020	-0.143	0.00005
rs11465673	103035375	-	-	0.010	-0.028	0.67400
rs2293223	103035468	-	-	0.003	-0.097	0.03880
rs2293224	103035779	0.009	-	0.019	-0.143	0.00005
rs6756407	103036101	-	-	0.011	-0.026	0.68800
rs6743516	103036335	0.009	-	0.019	-0.143	0.00005
rs1420100	103037002	0.008	-	0.019	-0.143	0.00005
rs3771155	103037826	0.009	-	0.021	-0.144	0.00005
rs3755267	103038587	-0.043	0.023	0.032	0.113	0.00860
rs10206291	103038863	0.009	-	0.030	-0.144	0.00005
rs885088	103039044	0.017	-	0.019	-0.143	0.00005
rs3771154	103039360	0.017	-	0.020	-0.143	0.00005
rs2272128	103039929	-0.042	0.023	0.029	0.116	0.00728
rs6759479	103040047	0.016	-	0.020	-0.143	0.00005
rs11465689	103040167	-	-	0.036	-0.030	0.65000
rs12997015	103042401	0.002	-	-	0.011	0.91800
rs3755266	103042712	-	-	0.021	-0.143	0.00005
rs11694658	103045020	-	-	0.006	0.108	0.01160
rs7559845	103046214	-	-	0.021	-0.142	0.00006
rs2310300	103049074	-	-	0.022	-0.142	0.00006

rs10166330	103050390	-	-	0.009	-0.088	0.06110
rs3755265	103052816	-	0.052	0.018	-0.142	0.00006
rs10176820	103054420	-	-	0.009	-0.089	0.05920
rs2058659	103054556	-	0.053	0.016	-0.142	0.00006
rs13021177	103056493	-	0.042	0.029	-0.143	0.00005
rs2075186	103057251	-	-	0.002	-0.091	0.05410
rs11465736	103067930	-	-	2.79e-05	-0.096	0.04130
rs17821875	103071030	-	-	0.127	-0.014	0.82500
rs10169676	103074919	-	-	2.87e-05	-0.089	0.05970
rs10210176	103079516	-	-	0.001	-0.082	0.08120
rs10172116	103087573	-	-	2.31e-04	-0.110	0.02050
rs1403552	103088777	-	-	0.001	-0.088	0.06140
rs7597819	103092906	-	-	0.021	0.113	0.00947
rs6737668	103093081	-	0.026	0.027	0.113	0.00941
rs10469840	103093243	-	-	0.032	0.107	0.01430
rs759382	103094213	-	0.011	-	0.105	0.01420
rs9989749	103123642	-	0.033	-	0.065	0.15300
rs6751949	103125138	-	0.009	-	0.070	0.11200
rs4851607	103125632	-	0.002	-	0.079	0.07400
rs10195948	103125736	-	0.004	-	0.085	0.05510
rs2192757	103132378	-	0.004	-	0.080	0.06250
rs1916307	103134180	-0.042	-	-	0.089	0.03680
rs1523203	103135759	-	0.020	-	0.082	0.05570
rs1403551	103136446	-	0.021	-	0.082	0.05510
rs4851613	103137990	-	0.021	-	0.084	0.05080
rs6750851	103138761	-	0.021	-	0.081	0.05790
rs6750971	103138825	-	0.021	-	0.083	0.05240
rs10193407	103139298	-	0.042	-	0.083	0.05350
rs17774619	103216776	0.115	-	-	-0.393	0.02020
rs1016160	103245025	-	-	0.011	0.086	0.09020
rs17027893	103370379	-	-0.008	-	-0.022	0.65500
rs10172680	103376383	-	-0.011	-	-0.011	0.81800
rs2080294	103382425	-	-0.012	-	-0.012	0.81600
rs6705433	103386225	-	-0.011	-	-0.016	0.74600
rs6705568	103386368	-	-0.010	-	-0.014	0.77400
rs2310358	103413658	-	-0.041	-	-0.076	0.27400
rs75535834	103784955	0.154	-	-	0.157	0.29300
rs6740972	103837781	0.028	-	-	-0.042	0.29600
rs7582789	103860075	-0.003	-	-	0.126	0.20300

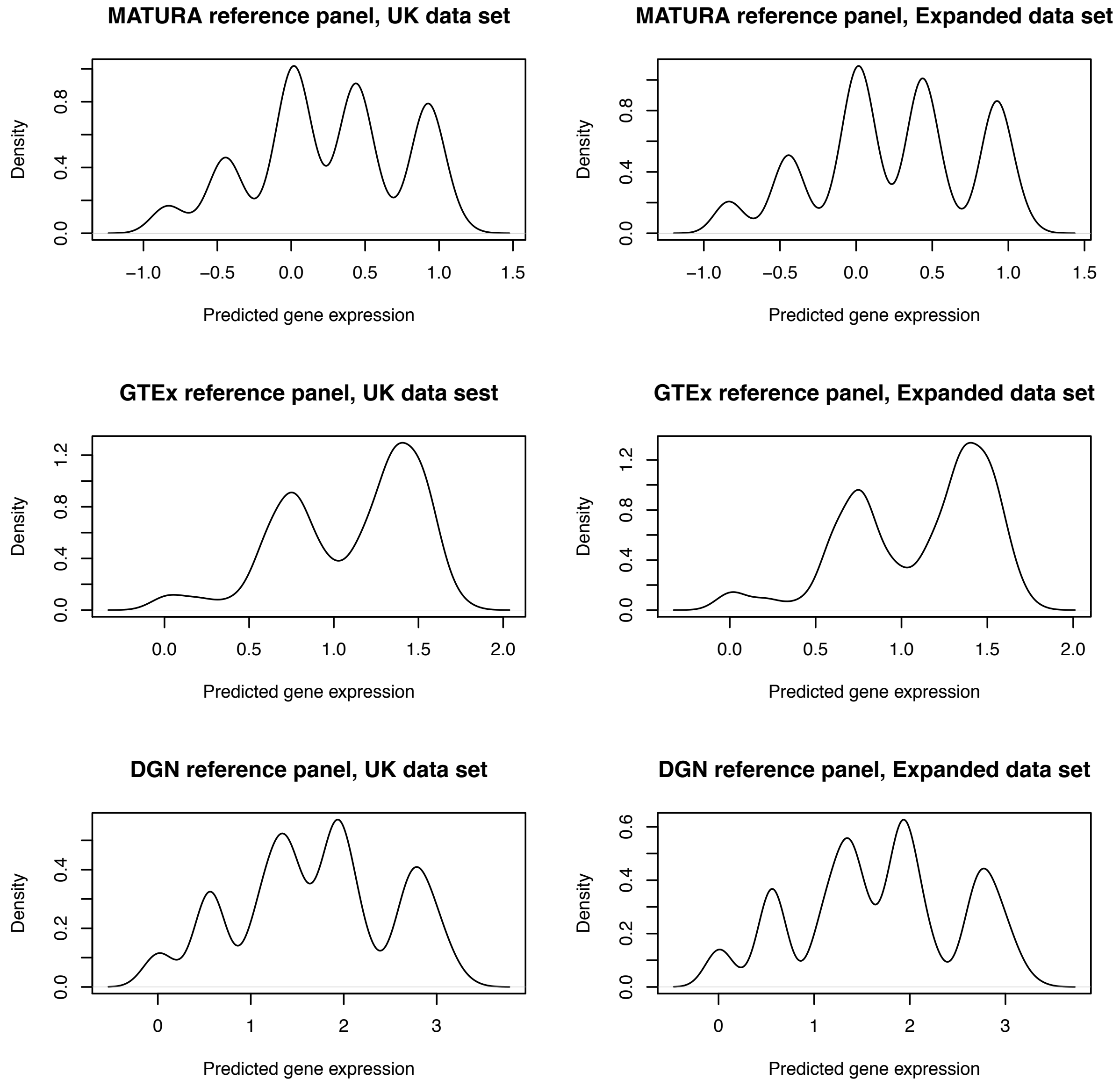


Figure S1. Density estimates (calculated using the density() function in R) of the predicted genetically-regulated gene expression at *IL18RAP* in the UK and expanded European ancestry cohorts.

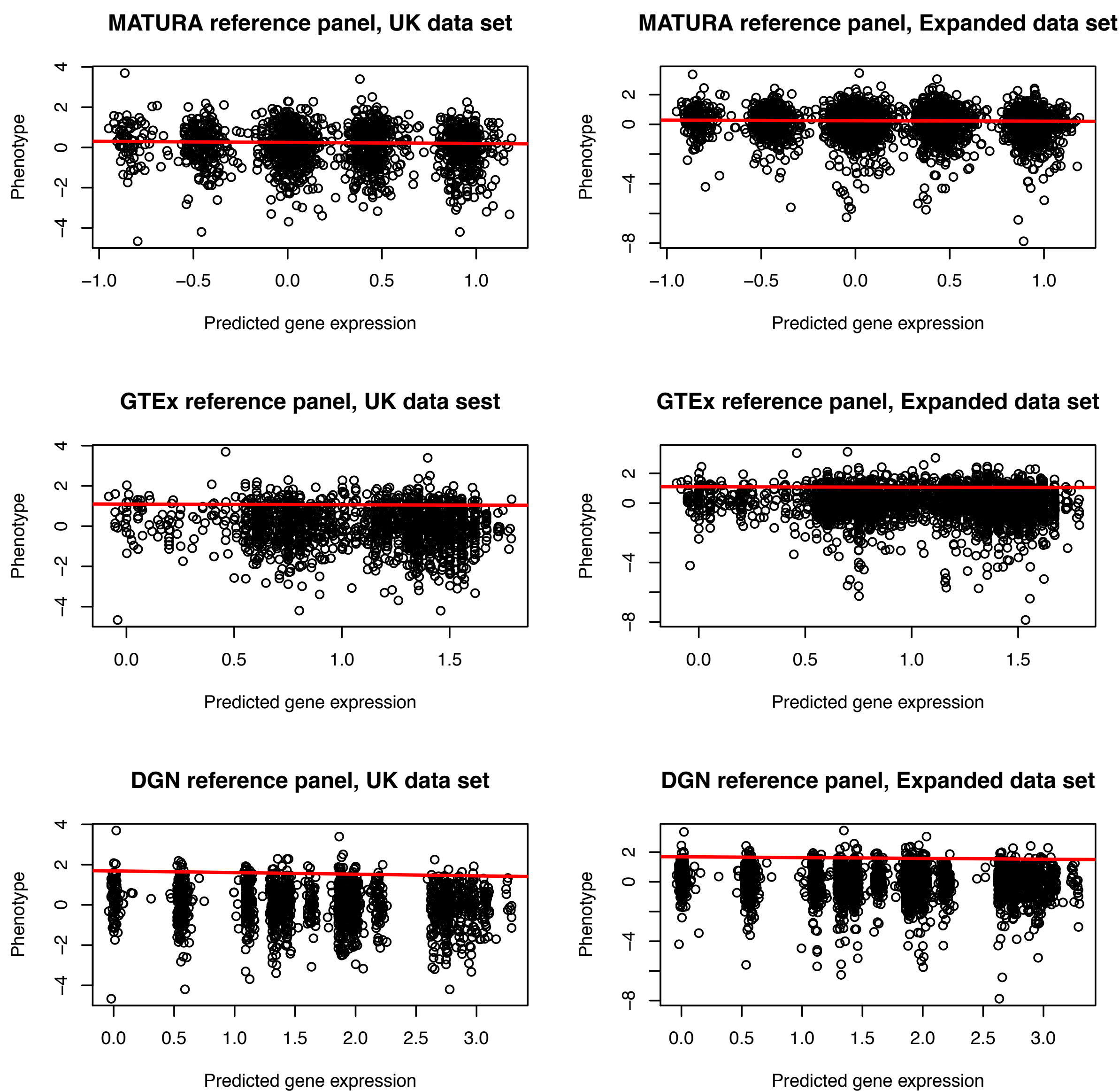


Figure S2. Plots of the linear regressions that are being fitted at *IL18RAP* in the UK and expanded European ancestry cohorts. Estimated regression lines are shown in red.