Clinical Validation of a Blood-Based Predictive Test for Stratification of Response to Tumor Necrosis Factor Inhibitor Therapies in Rheumatoid Arthritis Patients

Theodore Mellors,1 Johanna B. Withers,1 Asher Ameli,1 Alex Jones,1 Mengran Wang,1 Lixia Zhang,1 Helia N. Sanchez,1 Marc Santolini,2 Italo Do Valle,3 Michael Sebek,3 Feixiong Cheng,3 Dimitrios A. Pappas,4,5 Joel M. Kremer,5,6 Jeffery R. Curtis,7 Keith J. Johnson,1 Alif Saleh,1 Susan D. Ghiassian,1 and Viatcheslav R. Akmaev1,*

Abstract

Objectives: For rheumatoid arthritis (RA) patients failing to achieve treatment targets with conventional synthetic disease-modifying antirheumatic drugs, tumor necrosis factor (TNF)-\(\alpha\) inhibitors (anti-TNF therapies) are the primary first-line biologic therapy. In a cross-cohort, cross-platform study, we developed a molecular test that predicts inadequate response to anti-TNF therapies in biologic-naive RA patients.

Materials and Methods: To identify predictive biomarkers, we developed a comprehensive human interactome—a map of pairwise protein/protein interactions—and overlaid RA genomic information to generate a model of disease biology. Using this map of RA and machine learning, a predictive classification algorithm was developed that integrates clinical disease measures, whole-blood gene expression data, and disease-associated transcribed single-nucleotide polymorphisms to identify those individuals who will not achieve an ACR50 improvement in disease activity in response to anti-TNF therapy.

Results: Data from two patient cohorts (\(n = 58\) and \(n = 143\)) were used to build a drug response biomarker panel that predicts nonresponse to anti-TNF therapies in RA patients, before the start of treatment. In a validation cohort (\(n = 175\)), the drug response biomarker panel identified nonresponders with a positive predictive value of 89.7 and specificity of 86.8.

Conclusions: Across gene expression platforms and patient cohorts, this drug response biomarker panel stratifies biologic-naive RA patients into subgroups based on their probability to respond or not respond to anti-TNF therapies. Clinical implementation of this predictive classification algorithm could direct nonresponder patients to alternative targeted therapies, thus reducing treatment regimens involving multiple trial and error attempts of anti-TNF drugs.

Keywords: rheumatoid arthritis; anti-TNF; gene expression; precision medicine; drug response prediction
Introduction
Rheumatoid arthritis (RA) is a complex autoimmune disease with no cure. However, many potent treatment options are available to mitigate the symptoms and arrest progression of the disease. Treatment guidelines recommend early therapeutic intervention to forestall permanent functional debilitation associated with structural joint damage. Antitumor necrosis factor-α (anti-TNF) therapeutics are the first-line targeted therapy for nearly 90% of biologic-naive RA patients whose disease is not adequately controlled with conventional synthetic disease modifying antirheumatic drugs (csDMARDs), such as methotrexate. However, ~70% of these RA patients do not achieve meaningful clinical change on anti-TNF therapy. Therefore, there is a clinical need for a test that predicts which patients will not respond to anti-TNF agents before the initiation of therapy.

RA patient treatment targets are defined by the American College of Rheumatology (ACR) as disease activity scores indicating either disease remission or low disease activity. Treatment response relative to baseline is measured as ACR20, ACR50, and ACR70 where the number refers to the percent improvement in a standard set of measures. Factors used to calculate disease activity scores and ACR values include the number of swollen joints, tender joints, and both patient- and physician-reported assessments of pain and global health, as well as blood biomarker levels. ACR20 is reported as the benchmark in clinical trials for approval of new therapies; however, an ACR20 response is often insufficient to reach the clinically meaningful disease activity scores outlined in RA treatment guidelines. Rather, patients typically need to achieve an ACR50 response to reach disease activity scores corresponding to remission or low disease activity.

Complex multifactorial disorders such as RA have traditionally been difficult to study. The lack of a clear pattern of inheritance indicates that genetic information is integral to disease biology, but not sufficient for diagnosis or to guide treatment choices. Classic approaches to understand disease biology rely on either genetic analysis of mutations that drive disease-associated phenotypes or high-throughput molecular expression methodologies that are limited in their ability to sufficiently represent the complexity of the disease pathology. Advanced computational approaches that combine clinical data with integrative genetics are a necessary progression to bring modern precision medicine to autoimmune diseases. Network-based approaches to understanding human disease biology, collectively known as network medicine, have the ability to reveal underlying molecular patterns among features such as DNA markers and transcripts. Commonly, RNA sequencing (RNAseq) data have been used with the single purpose of analyzing gene expression features. However, RNAseq data also contain critical information on functionally active single-nucleotide polymorphisms (SNPs) that, when combined with transcript expression, have the potential to dramatically enhance the predictive power of a molecular signature. This study describes a multifaceted computational approach to develop a clinically useful tool to guide the treatment of RA patients.

Materials and Methods
Study populations
Discovery cohort. Patient microarray data (accession GSE15258) were obtained from the Gene Expression Omnibus. Details of sample collection and cohort information were previously reported; sample collection study protocols were approved by the local ethics committees and patients provided informed consent. Briefly, RA patients naive to anti-TNF therapy were enrolled and blood samples collected in PAXgene tubes. Therapeutic response was evaluated 14 weeks after initiation of treatment according to the DAS28-CRP EULAR response definition. Fifty-eight female patient samples were arbitrarily selected for this study.

Training cohort and validation cohort. RA patient whole-blood samples and clinical measurements were prospectively collected in the CERTAIN trial by the Consortium of Rheumatology Researchers of North America (CORRONA). The CERTAIN study was designed as a prospective comparative effectiveness study involving 43 sites and 117 rheumatologists. Institutional review board or ethics committee approvals were obtained before sample collection and study participation, and patients provided informed consent. Samples selected for the present study were from patients who were biologic-naive at the time of sample collection. Patients were treated with adalimumab, certolizumab pegol, etanercept, golimumab, or infliximab at the discretion of the treating physician and followed longitudinally for at least 6 months. In addition to a medical history, clinical assessments collected at 0 and 6 months post-therapy initiation included tender and swollen joint counts, physician and patient global disease activity scores, csDMARD dose, patient pain
evaluation, and quality-of-life surveys. Laboratory studies performed at a central laboratory included a complete blood count, C-reactive protein (CRP) levels, rheumatoid factor titer, and anti-cyclic citrullinated protein (anti-CCP) titer. Training \( n = 143 \) and validation trial \( n = 175 \) patient cohorts were balanced for response rate, age, and gender. Patients were included in the independent validation trial if they had a Visual Analog Scale pain score of \( > 15 \) out of a maximum score of 100. Consistent with the inclusion criteria of the CERTAIN study, all patients in the validation trial had a Clinical Disease Activity Index greater than 10.

**Evaluation of clinical response to anti-TNF therapy**

Among the CERTAIN study samples, response at 6 months after anti-TNF therapy initiation was defined by ACR50. An ACR50 responder was defined as an individual exhibiting \( \geq 50\% \) improvement in 28 tender joint count, \( \geq 50\% \) improvement in 28 swollen joint count, and \( \geq 50\% \) improvement in at least three out of five clinical variables (Health Assessment Questionnaire disability index, patient pain, patient global assessment, physician global assessment, and CRP level). **RNA isolation and quality control**

Total RNA was isolated from blood collected in PAXgene Blood RNA Tubes using the PAXgene Blood miRNA Kit (PreAnalytiX) according to the manufacturer’s instructions. RNA quality was assessed using the Agilent Bioanalyzer, and samples were quantitated using a NanoDrop ND-8000 spectrophotometer.

**RNAseq analysis and gene expression preprocessing**

RNA was processed using the GLOBINclear (Thermo Fisher), Ribo-Zero Magnetic Gold (Epidemiology), and TruSeq Stranded Total RNA (Illumina) kits according to the manufacturer’s instructions. Libraries were processed on a NextSeq 550 DX or a NovaSeq 6000 sequencer for 75 cycles. An average of 42.4 million reads were captured per patient, with a range of 33.7–58.6 million. Fifty-nucleotide reads were mapped to the GRCh37 human genome with STAR alignment software.38,39 Per gene abundance in fragments per kilobase of transcript per million mapped reads was calculated with the RSEM software package.40 Samples with an RNA integrity score of \( > 4 \), and \( > 7 \) million protein-coding reads were analyzed. Samples were processed in seven sequencing batches over a 1-year period. No detectable batch effect was observed between the NextSeq and NovaSeq processed libraries based on a principal component analysis.41 Hierarchical cluster analysis of RNA expression data for 37 genes is as previously described.

**SNP analysis**

Samples were aligned to the GRCh38 human genome with STAR alignment software.38 SNPs were called using a modified version of the Genome Analysis Toolkit Best Practices workflow for SNP and indel calling on RNAseq data.43–45 Thirty-nine RA-associated SNPs were evaluated (Supplementary Table S2).46

**Selection of gene expression biomarkers of nonresponse**

Gene expression biomarkers of nonresponse to anti-TNF therapy were selected from the 58-patient discovery cohort. Throughout 100 repeats, 80% of samples were randomly selected. Mann–Whitney U test was used to eliminate any gene expression features not significantly different in expression between responders and nonresponders \( ( p > 0.05 ) \). Random Forest from Scikit-learn was then used to rank the remaining features based on mean decrease impurity, a metric of feature importance. Features that ranked in the top 100 in at least 30 out of the 100 repeats were selected and considered for further evaluation in the RNAseq training cohort. The anti-TNF therapy response signal was observed in the female subpopulation, but was cross-validated in the entire training data set that included males and females. The all-patient cross-validation performance of the classifier based on the feature set selected from the female data was consistently higher.

**Model building, validation, and statistical analyses**

For model training, 143-patient samples were evaluated. Before final model building, a second round of feature selection took place to further evaluate the biomarker panel among the RNAseq training cohort. A total of 70 features were evaluated, consisting of gene expression biomarkers selected from the discovery cohort, SNPs, and clinical features (Supplementary Tables S1–S3). Each of the 70 features was ranked by assessing the decrease in cross-validated model performance (96 repeats of 20% withheld cross-validation) when a given feature was removed. Features that caused the largest drops in cross-validated model performance upon removal were considered to be the most important. The top 25 ranked features were used to build a finalized Random Forest model, which was subsequently
applied to the withheld 175-patient validation set samples. Model performance was evaluated using area under the receiver operating curve,\textsuperscript{49} positive predictive value, and specificity.\textsuperscript{50} All statistical analyses were performed using Python 3.7.6. Odds ratios and confidence intervals (CI) were calculated as previously described.\textsuperscript{51,52} Chi-squared test was used to determine the significance of differences in model performance stratified by ethnicity.

Building the human interactome and RA disease module, and performing network medicine analyses of molecular features

The human interactome was assembled as previously described\textsuperscript{30} from 21 public databases (Supplementary Table S4) containing different types of experimentally derived protein/protein interaction (PPI) data: (1) binary PPIs, derived from high-throughput yeast-two hybrid experiments (HI-Union),\textsuperscript{53} three-dimensional protein structures (Interactome3D,\textsuperscript{54} Instruct,\textsuperscript{55} Insider\textsuperscript{56}), or literature curation (PINA,\textsuperscript{57} MINT,\textsuperscript{58} LitBM17,\textsuperscript{53} Interactome3D, Instruct, Insider, BioGrid,\textsuperscript{59} HINT,\textsuperscript{60} HIPPIE,\textsuperscript{61} APID,\textsuperscript{62} InWeb\textsuperscript{63}); (2) PPIs identified by affinity purification followed by mass spectrometry present in BioPlex2,\textsuperscript{64} QUBIC,\textsuperscript{55} CoFrac,\textsuperscript{66} HINT, HIPPIE, APID, LitBM17, and InWeb; (3) kinase/substrate interactions from KinoNetworkX\textsuperscript{67} and PhosphoSitePlus\textsuperscript{68}; (4) signaling interactions from SignaLink\textsuperscript{69} and InnateDB\textsuperscript{70}; and (5) regulatory interactions derived by the ENCODE consortium. We used the curated list of PSI-MI IDs provided by Alonso-López et al.\textsuperscript{62} for differentiating binary interactions among the several experimental methods present in the literature-curation databases. All proteins were mapped in their corresponding NCBI Entrez ID and the proteins that could not be mapped were removed. The resulting human interactome includes 18,505 proteins and 327,924 interactions.

The DIAMOnD approach\textsuperscript{31} was used to generate an RA disease module. Proteins used to seed the disease module were linked to RA by at least two of five databases: GWAS Catalog,\textsuperscript{71} HuGE Navigator Phenopedia,\textsuperscript{72} ClinVar,\textsuperscript{73} OMIM,\textsuperscript{74} and MalaCards.\textsuperscript{75} DIAMOnD identified proteins that were significantly enriched in the same Gene Ontology biological process terms as the disease-associated proteins.

Proximity of the molecular features to each other on the human interactome map was calculated as previously described.\textsuperscript{76} Briefly, the closest distance was defined as the minimum path length between each protein and the other proteins in the set. Significance of the observed closest distance, reported as a z-score, was evaluated in comparison with the expected closest distance determined from 10,000 random protein sets of the same size. Randomizations were performed as previously described.\textsuperscript{76}

Pathway enrichment analysis

KEGG, BioCarta, Reactome, and Signal Transduction pathway annotations were obtained from the Molecular Signatures Database (MSigDB), Version 6.2.\textsuperscript{77} Fisher’s exact test was used to identify biological pathways. Pathways with a Bonferroni corrected \(p\)-value of < 0.05 were considered enriched. IL10,\textsuperscript{78} POMC,\textsuperscript{79} JAK1,\textsuperscript{80} ICOSLG,\textsuperscript{81} TNF, TNFSF11,\textsuperscript{82} NR3C1,\textsuperscript{83} P2RY12\textsuperscript{84} (NCT02874092), PTGER4,\textsuperscript{85} GGPS1,\textsuperscript{86} FDPS,\textsuperscript{86} TNFRSF13B (NCT03016013), IL6,\textsuperscript{87} ESR1,\textsuperscript{88} ESR2,\textsuperscript{88} ITK,\textsuperscript{89} BTK,\textsuperscript{90} TLR4\textsuperscript{91} (NCT03241108), IRAK4,\textsuperscript{92} JAK2,\textsuperscript{80} JAK3,\textsuperscript{80} HDAC1 (NCT02965599), PSMB5,\textsuperscript{93,94} ADORA3,\textsuperscript{95} ITGAM\textsuperscript{96} (NCT02698657, NCT03257852), IFNB1\textsuperscript{97} (NCT02777764; NCT03445715), and CX3CL1\textsuperscript{98} were the approved drug targets in RA.

Results

Building the human interactome and a map of RA disease biology

To begin developing the network medicine tools necessary to evaluate human disease biology, we created a map of cellular components and their physical interactions. By amalgamating publicly available data (Supplementary Table S4) of 327,924 pairwise PPIs between a total of 18,505 proteins, a comprehensive map of biology called the human interactome was created (see the Materials and Methods section).\textsuperscript{30} There are \(\sim 20,000\) proteins encoded by the human genome, and therefore, this human interactome incorporates interaction data from more than 90% of human proteins.

Disease-associated proteins tend to interact with each other in a subnetwork on the human interactome called a disease module.\textsuperscript{30} Using the DIAMOnD approach\textsuperscript{31} that aggregates potential disease-associated proteins based on their network proximity to known disease-associated proteins, an RA disease module was generated that contains \(\sim 200\) proteins. Of these, 66% were linked to RA in genome-wide association study databases and DIAMOnD identified the remaining proteins that are significantly enriched in the same Gene Ontology biological process terms as the known disease-associated proteins.
Using information from the human interactome and the RA disease module, we sought to identify a blood-based molecular signature that integrates clinical variables and molecular features to predict which RA patients will not respond to anti-TNF therapies (Fig. 1). Briefly, molecular features that discriminate between responders and nonresponders to anti-TNF therapies were selected from a publicly available microarray data set. In a cross-platform analysis, these features were combined with RA disease module-associated SNPs and clinical factors. Then, a machine-learning algorithm was trained using these features and RNAseq data. Finally, performance of the predictive drug response algorithm was validated in an independent validation trial.

Cross-platform identification of discriminatory gene expression features in whole blood predictive of inadequate response to anti-TNF therapy in RA patients

To maximize the clinical utility of a test that predicts nonresponse to therapy, a routine noninvasive or minimally invasive sample source that does not require specialized specimen collection procedures is ideal. For this reason, we analyzed gene expression data derived from whole blood. Gene expression that is discriminatory between patients considered responders and nonresponders to anti-TNF therapies was selected from a publicly available microarray discovery cohort data set of 58 biologic-naive RA patients using the Random Forest machine-learning algorithm (see the Materials and Methods section). Of the 21,818 genes in the discovery data set for which gene expression was assessed, 37 were selected as discriminatory biomarkers (Supplementary Table S1). Hierarchical cluster analysis illustrates that the gene expression profiles of these 37 biomarker transcripts can distinguish responders (n = 17) and nonresponders (n = 41) to anti-TNF therapies (Fig. 2). Two main clusters were observed, one predominantly nonresponders and the other responders, thereby substantiating the discriminatory nature of this multivariate molecular signature for response prediction. Transcriptional profiling by

---

**Fig. 1.** Flowchart describing development of anti-TNF drug response algorithm in RA. Gene expression that discriminates between responders and nonresponders to anti-TNF therapies was selected from a publicly available microarray data set. In a cross-platform analysis, these features were combined with network disease module-associated SNPs and clinical factors, and then used to train a machine-learning algorithm using RNAseq data. Finally, performance of the predictive drug response algorithm was validated in an independent validation trial. CI, confidence interval; CV, cross-validation; PPV, positive predictive value; RA, rheumatoid arthritis; SNPs, single-nucleotide polymorphisms; TNF, tumor necrosis factor.
microarray and RNAseq varies in dynamic range and exhibits some discordance in the number and extent of differential gene expression observed. Nonetheless, the majority of the transcripts (19/37; 51.4%) identified as discriminatory of anti-TNF drug response in microarray data also differentiated between responders and nonresponders in RNAseq data (Supplementary Fig. S1).

Evaluation of disease-associated SNPs from RNAseq data
RNAseq provides information on nucleotide sequence that is lacking from microarray analyses. In addition to gene expression, variations in RNA sequence may be predictive of nonresponse to anti-TNF therapy in RA patients. To this end, a training data set was generated from clinical data and whole-blood RNAseq data obtained from 143 RA patients in the CORRONA CERTAIN study. Characteristics and demographics of the patient populations are summarized in Table 1. Although SNP analysis is traditionally performed on whole-genome sequencing data, the majority of the genome is transcribed. Thus, most SNP variants can be detected in ribosomal RNA-depleted RNAseq data. Because these RNAseq data were derived from blood samples, only SNPs that are associated with RA, which have been functionally linked to gene expression changes in peripheral blood mononuclear cells through expression quantitative trait loci (eQTL) analysis, were examined (Supplementary Table S2). The genetic loci associated with the selected SNPs have a significant overlap with the RA disease module (Fig. 3B). Twenty-two such disease module-associated SNPs were above the limit of detection in the patient RNAseq data and thus included in further analyses (Supplementary Fig. S2).

Integration of SNPs, gene expression data, and clinical variables to develop a multifactorial predictive drug response algorithm
Gene expression indicative of drug response (Supplementary Table S1), RA-associated SNPs (Supplementary Table S2), and clinical factors (Supplementary Table S3) represent a set of 70 biomarker features.

FIG. 2. Discriminatory genes indicative of nonresponse to anti-TNF therapy in RA patient blood samples. Hierarchical cluster analysis of RNA expression data for 37 genes illustrates two main groupings, one predominantly nonresponders and the other responders, thereby substantiating the discriminatory nature of these genes for anti-TNF response prediction. The heatmap represents the relative RNA expression level in arbitrary units.
Table 1. Patient Demographics and Disease Characteristics at Baseline

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Training cohort (n = 143)</th>
<th>Validation trial cohort (n = 175)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years (mean ± SD)</td>
<td>55.0 ± 12.9</td>
<td>53.8 ± 11.9</td>
</tr>
<tr>
<td>Female, %</td>
<td>72.7</td>
<td>73.1</td>
</tr>
<tr>
<td>Duration of disease, years (mean ± SD)</td>
<td>4.7 ± 6.7</td>
<td>5.0 ± 7.5</td>
</tr>
<tr>
<td>Positive for anti-cyclic citrullinated peptide, %</td>
<td>66.9</td>
<td>60.9</td>
</tr>
<tr>
<td>Positive for rheumatoid factor, %</td>
<td>73.9</td>
<td>70.9</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>88.1</td>
<td>84.6</td>
</tr>
<tr>
<td>Black</td>
<td>6.3</td>
<td>6.3</td>
</tr>
<tr>
<td>Other</td>
<td>5.6</td>
<td>9.1</td>
</tr>
<tr>
<td>Current csDMARD use, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methotrexate</td>
<td>65.0</td>
<td>60.0</td>
</tr>
<tr>
<td>Hydroxychloroquine</td>
<td>3.5</td>
<td>4.0</td>
</tr>
<tr>
<td>≥2 csDMARDs</td>
<td>13.3</td>
<td>13.7</td>
</tr>
<tr>
<td>None</td>
<td>15.4</td>
<td>16.0</td>
</tr>
<tr>
<td>Concomitant prednisone, %</td>
<td>34.3</td>
<td>22.9</td>
</tr>
<tr>
<td>Baseline prednisone dose, mg (mean ± SD)</td>
<td>7.4 ± 4.7</td>
<td>8.5 ± 5.3</td>
</tr>
<tr>
<td>Anti-TNF use, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adalimumab</td>
<td>38.5</td>
<td>38.9</td>
</tr>
<tr>
<td>Etanercept</td>
<td>33.6</td>
<td>30.9</td>
</tr>
<tr>
<td>Infliximab</td>
<td>16.8</td>
<td>20.0</td>
</tr>
<tr>
<td>Certolizumab pegol</td>
<td>8.4</td>
<td>7.4</td>
</tr>
<tr>
<td>Golimumab</td>
<td>2.8</td>
<td>2.9</td>
</tr>
<tr>
<td>CDAI (mean ± SD)</td>
<td>28.5 ± 13.5</td>
<td>31.0 ± 12.6</td>
</tr>
<tr>
<td>DAS28-CRP (mean ± SD)</td>
<td>4.9 ± 1.1</td>
<td>5.0 ± 1.0</td>
</tr>
<tr>
<td>Swollen joint count (mean ± SD)</td>
<td>7.2 ± 6.0</td>
<td>8.1 ± 5.5</td>
</tr>
<tr>
<td>Tender joint count (mean ± SD)</td>
<td>10.8 ± 7.3</td>
<td>12.0 ± 7.3</td>
</tr>
<tr>
<td>ACR50 responders, %</td>
<td>30.8</td>
<td>30.3</td>
</tr>
</tbody>
</table>

CDAI, Clinical Disease Activity Index; csDMARDs, conventional synthetic disease modifying antirheumatic drugs; SD, standard deviation; TNF, tumor necrosis factor.

Independent validation trial of a drug response biomarker panel predictive of nonresponse to anti-TNF therapy in RA patients

To confirm that the drug response biomarker panel is generalizable, an independent group of prospectively collected samples (n = 175) were used to conduct a validation trial. The samples included in the validation cohort were not used for any stage of the biomarker panel development, and the algorithm has no information derived from the gene expression data or clinical outcomes from these patients.

Independent validation of the drug response biomarker panel stratified the validation cohort into predicted nonresponders and responders, with a highly statistically significant odds ratio of 6.57 (95% CI 2.75–15.70) of being a nonresponder in the respective subgroup. The odds ratio is a statistic that quantifies the strength of association between two events, in this case the signal from the biomarker panel and inadequate response to anti-TNF therapies. This means that a patient identified by the drug response biomarker panel to be a nonresponder is 6.57 times more likely to inadequately respond to an anti-TNF therapy than if that patient was a responder. The drug response biomarker panel identified patients who are unlikely to have an adequate response to anti-TNF therapies with a positive predictive value (PPV) of 89.7% (95% CI 79.0–95.7%), specificity of 86.8% (95% CI 72.4–94.1%), and sensitivity of 50.0% (95% CI 40.8–58.7%) (Table 2). Patients predicted to be nonresponders have an observed ACR50 response rate of 10.3% (7/68) with anti-TNF therapies, significantly lower than the overall response rate of 30.3% (53/175). Conversely, the predicted responders had an observed ACR50 response rate of 43.0% (46/107), which is a 41.9% improvement from that of the unstratified patient population.

Alternatively, using the more stringent threshold of response, ACR70, which requires patients to exhibit a 70% improvement in the ACR response criteria, 81.7% (143/175) of patients in the validation cohort were nonresponders. The drug response biomarker panel identified patients who are unlikely to achieve an ACR70 response to anti-TNF therapies with a PPV of 93.5% (95% CI 84.0–97.9%) and specificity of 84.4% (95% CI 62.7–96.3%). The drug response biomarker panel predicted 50.3% of these nonresponders (72/143) in addition to misclassifying five ACR70 responders.

That were used to train and develop a drug response algorithm that is predictive of nonresponse to anti-TNF therapies. Using ACR50 at 6 months as a benchmark, the training cohort population had a response rate to anti-TNF therapies of 30.8% (44/143). This is representative of the general population and reflects the real-world prospective collection approach of the CORRONA CERTAIN study. Random Forest was used to generate predictive models with 80% of the RNAseq training data set using features from the discriminatory gene expression set, SNPs, and clinical factors. The remaining 20% of the data set was withheld for performance testing during the model building process. Although 70 candidate biomarkers were considered in this analysis, not all were required to predict the anti-TNF therapy response. The biomarker panel used to develop the final model predictive of inadequate response to anti-TNF therapies included 10 SNPs, 8 transcripts, 2 laboratory tests (CRP and anti-CCP), and 3 clinical metrics (sex, body mass index [BMI], patient disease assessment).
FIG. 3. Molecular features are closely associated with each other and disease-associated proteins on the HI. (A) Visualization of a subset of the HI protein/protein network. Proteins included on the network are indicated as gray circles. Those outlined in red represent proteins encoded by SNP eQTL transcripts and those outlined in blue represent proteins encoded by DG. (B) Quantitative analysis of the proximity of molecular drug targets, proteins associated with RA (disease module proteins), and molecular features included in the drug response biomarker panel (proteins encoded by SNP-associated transcripts and proteins encoded by DGs). Proximity scores (z-scores) reflect whether two sets of proteins are significantly proximal to each other (z-score < −1.6). DG, discriminatory genes; eQTL, expression quantitative trait loci; HI, human interactome.
The drug response biomarker panel includes SNPs, and SNP allele frequencies can vary between ethnic groups. Although the training and validation set patient populations were predominantly Caucasian (88.1% and 84.6%, respectively), no significant differences (p > 0.23) were observed in the ability of the biomarker panel to predict an inadequate response to anti-TNF therapies among patients of other ethnicities. Future work using population-specific studies will address tranethnic prediction performance.

Biological interpretation of gene products that discriminate between responders and nonresponders to anti-TNF therapy

To characterize the applicability of the predictive drug response biomarker panel to RA disease biology, the protein products of the discriminatory genes and SNP eQTLs were analyzed using the human interactome and pathway enrichment analyses. The proteins encoded by the discriminatory genes and SNP eQTLs included in the drug response biomarker panel were mapped onto the human interactome map (Fig. 3A and Supplementary Fig. S3). In total, 42 proteins mapped onto the human interactome: 24 are contributed by the discriminatory genes and 18 by the SNP eQTLs. These molecular features are significantly connected in the same network vicinity of the human interactome highlighting a small, yet cohesive biological network that unifies the molecular features that predict inadequate response to anti-TNF therapies. Quantification of this proximity (see the Materials and Methods section) indicates that these different molecular features are significantly close to each other in comparison with random expectation (z-score = −3.83). Furthermore, the RA disease module (z-score = −4.92) and RA drug targets such as JAK and TNF-α (z-score = −3.97) are proximal to the SNP eQTLs and discriminatory genes collectively (Fig. 3B).

Pathway enrichment analysis was performed to gain insight into the molecular pathways involved in the anti-TNF therapy response. T cell signaling was identified as the most enriched pathway in the pathway analysis databases queried. The relevance and importance of T cell signaling to both anti-TNF therapy response and the disease biology of RA are well established.

Discussion

By incorporating microarray gene expression data, RNAseq, biological network analyses, and machine learning to large patient cohorts, this study describes the development of a drug response biomarker panel that uses whole-blood gene expression data to identify a molecular signature that predicts nonresponse to anti-TNF therapies in biologic-naive patients with RA. Validation with a prospectively collected RNAseq data set demonstrated that the drug response biomarker panel could predict nonresponse to anti-TNF therapies in an independent cohort of biologic-naive RA patients. Anti-TNF therapies failed to help nearly 70% of the unstratified patient population reach an ACR50 response. The drug response biomarker panel could have prevented half of these individuals from taking a drug that did not ameliorate the signs and symptoms of their disease. The biomarker panel included 10 SNPs, 8 transcripts, 2 laboratory tests (CRP and anti-CCP), and 3 clinical metrics (sex, BMI, patient disease assessment).

A single, large-scale, high-throughput analysis approach is yet to capture the molecular signature of RA disease biology. Many studies have hypothesized that the biology of nonresponse to anti-TNF therapies is reflected in the transcriptome of whole blood. However, none has been translated into the clinic, which is likely a reflection of both the complexity of RA disease biology and the varying methodologies used for algorithm development. Furthermore, limited sample sizes and the complexity of gene expression data analyses may have thus far prevented the development of an algorithm that is generalizable across patient cohorts and to the wider patient population. The drug response biomarker panel described in this work differs from previous attempts to predict response to anti-TNF therapies in that it integrated RNA transcript levels and RNA sequence information. SNPs can affect many aspects of cellular biology, including the propensity for regulatory elements to interact with their cognate protein partners, the ratio or identity of alternative splice variants produced from a gene locus, transcript levels, and protein sequence. Therefore, the functional readout of disease-associated SNPs contributes to the propensity of an individual to develop disease as well.

Table 2. Predictive Drug Response Biomarker Panel Validation Performance

<table>
<thead>
<tr>
<th></th>
<th>ACR50 (R)</th>
<th>ACR50 (NR)</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drug response biomarker panel (R)</td>
<td>46</td>
<td>61</td>
<td>107</td>
</tr>
<tr>
<td>Drug response biomarker panel (NR)</td>
<td>7</td>
<td>61</td>
<td>68</td>
</tr>
<tr>
<td>Sum</td>
<td>53</td>
<td>122</td>
<td>175</td>
</tr>
</tbody>
</table>

NR, non-responder; R, responder.
as the inclination for environmental factors to influence pathobiology.\textsuperscript{127,128} Many regulatory elements and genomic regions that do not encode protein are transcribed, such as in the form of enhancer RNAs\textsuperscript{129} and promoter-associated transcripts.\textsuperscript{130} Thus, many SNPs that influence spatial- and temporal-specific changes in transcription can be evaluated from RNAseq data. Analyzed together into a single-molecular signature, SNP and gene expression analyses can capture phenotypic variation and pathway associations that may otherwise be missed.

The microarray and RNAseq gene expression analysis platforms differ in RNA detection methodologies and statistical tools to determine normalized gene expression values.\textsuperscript{131,132} Despite these differences in technology, the cross-platform and cross-cohort universality of the molecular features identified in this study highlights the presence of a robust molecular signature underlying the biology of anti-TNF drug response. Examination of the molecular pathways that identify patients who will not respond to anti-TNF therapies demonstrated a connection between T cell signaling and RA disease biology.\textsuperscript{133–138} Synovial inflammation results from leukocyte infiltration into and retention in the synovial compartment, as well as from insufficient apoptosis of chronic inflammatory cells.\textsuperscript{139,140} This synovial infiltrate includes natural killer cells, CD4\textsuperscript{+}, and CD8\textsuperscript{+} T cells.\textsuperscript{141–145} Furthermore, large numbers of activated T regulatory cells can be detected in the joints of RA patients.\textsuperscript{146} The remaining discriminatory genes that are not associated with T cell signaling likely represent different aspects of RA disease that differ between those patients who will or will not respond to anti-TNF therapies. The connection to RA disease biology speaks to the reliability and applicability of the drug response biomarker panel to be a powerful clinical tool for identification of anti-TNF nonresponders.

For patients predicted to inadequately respond to anti-TNF therapies, many alternative biologic and targeted synthetic therapies are available. Because predicted nonresponders had a 10% observed response rate to anti-TNF therapies, they may see a greater benefit from being prescribed an alternative treatment with a higher reported response rate. Clinical trials assessing the efficacy these alternative therapies report ACR50 response rates of 30–40% at 6 months following alternative treatment initiation for patients who had inadequately responded to an anti-TNF therapy.\textsuperscript{147–150} The remaining patients—those lacking a molecular signature of nonresponse—would still have been prescribed an anti-TNF therapy and would have a response rate of 43% (Table 2). Inadequate responders to anti-TNF therapies who are not identified by the test likely would be directed to anti-TNF therapies, which is overwhelmingly the default first-line biologic for biologic-naïve RA patients, and thus, their treatment path would be unaffected by the results of this molecular signature test.

**Conclusion**

Customization of treatment regimens to match the individualized disease biology of each patient is a goal of modern medicine. This personalized approach to medicine is used in oncology, where particular therapies are prescribed to patients with specific genomic markers.\textsuperscript{151,152} Development and validation of a drug response algorithm that predicts nonresponse to a targeted therapy using this machine-learning and network medicine approach show great promise for advancing precision medicine in the treatment of RA and other complex autoimmune diseases where costly therapeutic interventions are met with inadequate patient response.

**Authors’ Contributions**


**Acknowledgments**

The authors thank Michael Weinblatt for clinical input and review, and Kjell Johnson for machine-learning guidance. They thank Carol Etzel, Amy Schraedr, and Austin Read from CORRONA, as well as Pauline Cronin, Jeran Stratford, Victor Weigman, and Wendell Jones from Q2. The authors also thank all of the patients and physicians who have participated in the CORRONA CERTAIN substudy.

**Author Disclosure Statement**

References


152. Bode AM, Dong Z. Recent advances in precision oncology research. NPJ Precis Oncol. 2018;2:11.