Research Article

Identifying Infliximab- (IFX-) Responsive Blood Signatures for the Treatment of Rheumatoid Arthritis

ShiJian Ding [ID],[1] **ZhanDong Li** [ID],[2] **Tao Zeng** [ID],[3] **Yu-Hang Zhang** [ID],[4] **Tao Huang** [ID],[3,5] **and Yu-Dong Cai** [ID][1]

[1]*School of Life Sciences, Shanghai University, Shanghai, China*
[2]*College of Food Engineering, Jilin Engineering Normal University, Changchun, China*
[3]*Bio-Med Big Data Center, CAS Key Laboratory of Computational Biology, Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, China*
[4]*Channing Division of Network Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA*
[5]*CAS Key Laboratory of Tissue Microenvironment and Tumor, Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, China*

Correspondence should be addressed to Tao Huang; tohuangtao@126.com and Yu-Dong Cai; cai_yud@126.com

Rheumatoid arthritis (RA) is a severe chronic pathogenic inflammatory abnormality that damages small joints. Comprehensive diagnosis and treatment procedures for RA have been established because of its severe symptoms and relatively high morbidity. Medication and surgery are the two major therapeutic approaches. Infliximab (IFX) is a novel biological agent applied for the treatment of RA. IFX improves physical functions and benefits the achievement of clinical remission even under discontinuous medication. However, not all patients react to IFX, and distinguishing IFX-sensitive and IFX-resistant patients is quite difficult. Thus, how to predict the therapeutic effects of IFX on patients with RA is one of the urgent translational medicine problems in the clinical treatment of RA. In this study, we present a novel computational method for the identification of the applicable and substantial blood gene signatures of IFX sensitivity by liquid biopsy, which may assist in the establishment of a clinical drug sensitivity test standard for RA and contribute to the revelation of unique IFX-associated pharmacological mechanisms.

1. Introduction

Rheumatoid arthritis (RA) is a severe chronic pathogenic inflammatory abnormality that damages small joints [1, 2]. Some patients with severe and progressive RA may also suffer from pathogenic lesions in various body systems, including the skin, eyes, lungs, and blood vessels [1]. According to the statistics provided by the American College of Rheumatology in 2008, more than 1.3 million people in the USA suffer from pathogenic rheumatoid arthritis with typical symptoms [3, 4]; thus, RA is the leading cause of arthritis among multiple pathogeneses.

Comprehensive diagnosis and treatment procedures for RA have been established because of its severe symptoms and relatively high morbidity [1, 5–8]. The diagnosis of RA

can be divided into two major procedures: symptom-dependent diagnosis [9] and clinical laboratory examinations [10]. A group of typical symptoms can be used to preliminary screen for RA. The typical symptoms of RA include the swelling of small joints (which may further extend to larger joints, such as the hips and shoulders), fatigue, fever, and anemia, which are quite easy to identify and recognize [11, 12]. However, the types and severity of these symptoms vary among patients. Thus, RA is difficult to diagnose based only on clinical symptoms. Apart from clinical symptoms, blood tests and imaging tests have also been applied for the accurate diagnosis of RA. Erythrocyte sedimentation rate [13] and C-reactive protein [13] are nonspecific biomarkers for the diagnosis and progression monitoring of RA that reflect the degree of inflammatory responses in the whole body. The

gold standard for RA diagnosis is clinical liquid biopsy, and rheumatoid factor [14] and anticyclic citrullinated peptide [15] antibodies are the specific biomarkers for the blood test screening of RA. Pathogenic lesions in the joints can be regarded as a diagnosis measurement for RA and can be easily identified by X-ray [16], ultrasound [17], and magnetic resonance imaging [18].

Medication and surgery are the two major therapeutic approaches for RA [19, 20]. Different from surgery-based approaches, which focus on the direct relief of joint damage, most medications focus on the prevention of abnormal inflammatory immune responses and therefore indirectly relieve systemic symptoms. Four subgroups of medications are clinically applied, namely, nonsteroidal anti-inflammatory drugs, steroids, disease-modifying antirheumatic drugs, and biological agents [21]. Among these drugs, infliximab (IFX) is a novel biological agent applied for the treatment of RA. As a chimeric monoclonal antibody, IFX has been approved by the Food and Drug Administration (FDA) for the treatment of multiple immune-associated diseases, including RA, early in 2007 [22]. IFX targets one of the pathogenic proinflammatory cytokines, tumor necrosis factor-alpha (TNF-α), and thus has been applied in classical TNF antagonist therapy against multiple chronic autoimmune inflammatory diseases, including RA [23]. The clinical application of IFX has been studied for nearly 30 years; IFX greatly improves physical functions and is beneficial for the achievement of clinical remission even under discontinuous medication [24, 25]. However, not all patients react to IFX, and its anti-inflammatory effects vary among patients with RA. IFX-sensitive and IFX-resistant patients are difficult to distinguish based on traditional clinical examination; thus, how to predict the therapeutic effects of IFX on patients with RA is one of the urgent translational medicine problems in the clinical treatment of RA.

With the development of liquid biopsy and high-throughput sequencing technologies, a recent study [26] revealed that patients with different drug susceptibility against IFX have different pretherapeutic blood expression patterns; thus, liquid biopsy and the transcriptomic profiling of patients' blood may help predict the therapeutic effects of IFX in RA and therefore assist in clinical medication. However, the application of whole transcriptome sequencing on every patient with RA is not feasible; therefore, the identification of accurate and efficient transcriptomic biomarkers and their respective pathogenic expression pattern would be the priority. In this study, we presented a novel computational method to identify the applicable and substantial blood gene biomarkers of IFX sensitivity by liquid biopsy. This study may assist in the establishment of a test standard for clinical drug sensitivity for RA treatment and contribute to the revelation of unique IFX-associated pharmacological mechanisms.

2. Materials and Methods

2.1. Data. The blood gene expression profiles of 140 patients with RA before IFX treatment were downloaded from the Gene Expression Omnibus under accession number GSE78068 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE78068) [26]. Among the patients, 42 showed

remission and 98 showed nonremission. 19595 genes were involved to constitute the blood gene expression profiles. Predicting the response of patients with RA to IFX before therapy using their blood will help in deciding previse treatments, which is the ultimate goal of precision medicine. We would like to build such IFX response prediction models for patients with RA based on their blood gene expression profiles.

2.2. Monte Carlo Feature Selection (MCFS). In this work, the MCFS method, which is a decision tree- (DT-) based feature selection method [27–29], was used to select important gene candidates. MCFS can randomly select a few feature subsets from the original features. Each feature subset consists of a small number of features from the original ones. One bootstrap sample dataset can be induced from this feature subset, and then, multiple DTs can be learned and tested on this bootstrap sample datasets. The process is repeated several times to produce many feature subsets. Each feature subset can help learn the same number of DTs.

The contribution of each feature in these DTs can be evaluated by the relative importance (RI) score, which is calculated as follows:

$$\text{RI}_f = \sum_{\tau=1}^{pt} (\text{wAcc})^u \sum_{n_f(\tau)} \text{IG}\big(n_f(\tau)\big) \left(\frac{\text{no. in } n_f(\tau)}{\text{no. in } \tau}\right)^v, \quad (1)$$

where wAcc indicates the weighted accuracy and $n_f(\tau)$ represents a node of feature f in DT τ. The information gain of $n_f(\tau)$ in the DT is measured by $\text{IG}[n_f(\tau)]$, and no.in $n_f(\tau)$ points the number of training samples in node $n_f(\tau)$. In addition, the weighting factors u and v have a value of 1 by default.

After each feature was assigned a RI value, we ranked all features in a list with the decreasing order of their RI values. In addition to the feature list, the MCFS method also outputs some most important features, called informative features, which are some top features in the list. These features are accessed by determining a threshold of RI values via a permutation test on class labels and one-sided Student's t-test.

This study used the MCFS program retrieved from http://www.ipipan.eu/staff/m.draminski/mcfs.html. Its default parameters were used.

2.3. Incremental Feature Selection (IFS). IFS is widely used to determine the optimal number of features for constructing a classification model with an integrated supervised classifier [30]. On the basis of a ranked feature list obtained from MCFS, a series of feature subsets are produced with a step interval of 10. For example, the first feature subset includes the top 10 features and the second feature subset includes the top 20 features. For each feature subset, a classifier (e.g., support vector machine (SVM)) is trained on a training data induced from this feature subset. An optimal feature subset is selected when it has the highest performance among the candidate feature subsets, where performance is evaluated by the Matthews correlation coefficients (MCC) [31] under 10-fold cross-validation [32]. The classifier with such optimal feature subset can be built and was called the optimum classifier in this study.

2.4. Support Vector Machine (SVM). SVM is a classification algorithm suitable for linear and nonlinear data [33–39]. For a dataset with two classes, SVM tries to find out an optimum hyperplane, which can divide samples in two classes with a maximum margin. However, in many cases, such hyperplane is not easy or impossible to be discovered. SVM employs a kernel trick to convert the original sample in a low-dimensional space to a new sample in some high-dimensional space, in which the optimum hyperplane can be easily constructed. For a new sample, it is also mapped into the high-dimensional space and its class is determined according to the side of the hyperplane it lies. In this study, we used the tool "SMO" in Weka [40, 41] to quickly implement SVM. Such SVM is optimized by the sequential minimum optimization algorithm [42]. For convenience, default parameters were adopted, where the kernel was a polynomial function.

2.5. Random Forest (RF). RF [43] is a metaclassifier that is widely applied in biological and biomedical researches [44–52]. RF includes many DTs as members. To construct a DT, a dataset, in which samples are randomly selected, with replacement, from the original dataset, is constructed. Such dataset has the same size of the original dataset. The DT is grown at each node by determining an optimal splitting way on some features, which were randomly selected from all features. Given a new sample, each DT gives its prediction. RF integrates these predictions with majority voting. Similar to SVM, we also employed a tool "RandomForest" in Weka [40, 41], which implements RF. Likewise, the default parameters were used, where the number of DTs was set to ten.

2.6. Rule Learning. In addition to "black-box" classification algorithms, the interpretable rules for a classification model can also be extracted to explain the feature differences between groups of patients with particular response to drug treatment. To accelerate this procedure, we directly picked up the informative features extracted by the MCFS method. These features were further filtered by the Johnson Reducer algorithm [53]. The remaining features were fed into the repeated incremental pruning to produce error reduction (RIPPER) algorithm [54] to extract classification rules. Obtained rules were represented by IF-ELSE rules. The above rule learning procedures were also integrated in the MCFS program, which was directly adopted in this study.

2.7. Measurement. The MCC [31] within 10-fold cross-validation was applied in this work to evaluate the classification performance of different classification models. The MCC for the binary problem is calculated as follows:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TN + FN) \times (TN \times FP) \times (TP + FN) \times (TP + FP)}},$$
(2)

where TP, TN, FP, and FN represent the number of true-positive, true-negative, false-positive, and false-negative samples, respectively. As a measurement for the classification model, the MCC value ranges from −1 to +1, where a value of +1

indicates that the classification model has the best performance. To date, it has wide application in bioinformatics for evaluating the performance of different classification models [55–61].

Besides, we also employed other five measurements to give a full evaluation on different classification models. They were sensitivity (SN), specificity (SP), accuracy (ACC), precision, and *F*1-measure, which can be calculated by the following equations:

$$\begin{cases} SN = \dfrac{TP}{TP + FN}, \\[2mm] SP = \dfrac{TN}{TN + FP}, \\[2mm] ACC = \dfrac{TP + TN}{TP + FN + TN + FP}, \\[2mm] Precision = \dfrac{TP}{TP + FP}, \\[2mm] F1\text{-measure} = \dfrac{2 \cdot Recall \cdot Precision}{Recall + Precision}. \end{cases}$$
(3)

3. Results

In this study, we gave a computational investigation on the blood gene expression profiles of patients with RA before IFX treatment. The entire procedures are illustrated in Figure 1. This section gave the results of all procedures.

3.1. Results of MCFS. The blood gene expression profiles were first analyzed by the MCFS method. As a result, each feature was assigned a RI value, which indicated its importance. The RI values of all features are listed in Table S1. Accordingly, all features were ranked in a list by the decreasing order of features' RI values. Such list is also provided in Table S1.

3.2. Results of IFS. Based on the feature list obtained in the above section, the IFS method is followed. It first constructed several feature subsets. Then, on each feature subset, a classifier was built using SVM or RF as the classification algorithm. Each classifier was evaluated by 10-fold cross-validation. Six measurements (see equations (2) and (3)) were obtained for each classifier, which are listed in Tables S2 and S3. For an easy observation, a curve was plotted for each classification algorithm, as shown in Figure 2, in which MCC was set as the *y*-axis and the number of features was set as the *x*-axis. It can be observed that for SVM, the highest MCC was 0.760 when top 1260 features were used. Thus, these 1260 features constituted the optimum feature subset of SVM and the SVM classifier with these features was the optimum SVM classifier. Other measurements of such classifier are listed in Table 1. As for RF, the highest MCC was 0.611 when only top ten features were adopted. An optimum RF classifier was built based on these top ten features. The MCC of the optimum RF classifier was much lower than that of the optimum SVM classifier. Other five measurements of such RF classifier are provided in Table 1. Evidently, each measurement was inferior to that of the optimum SVM
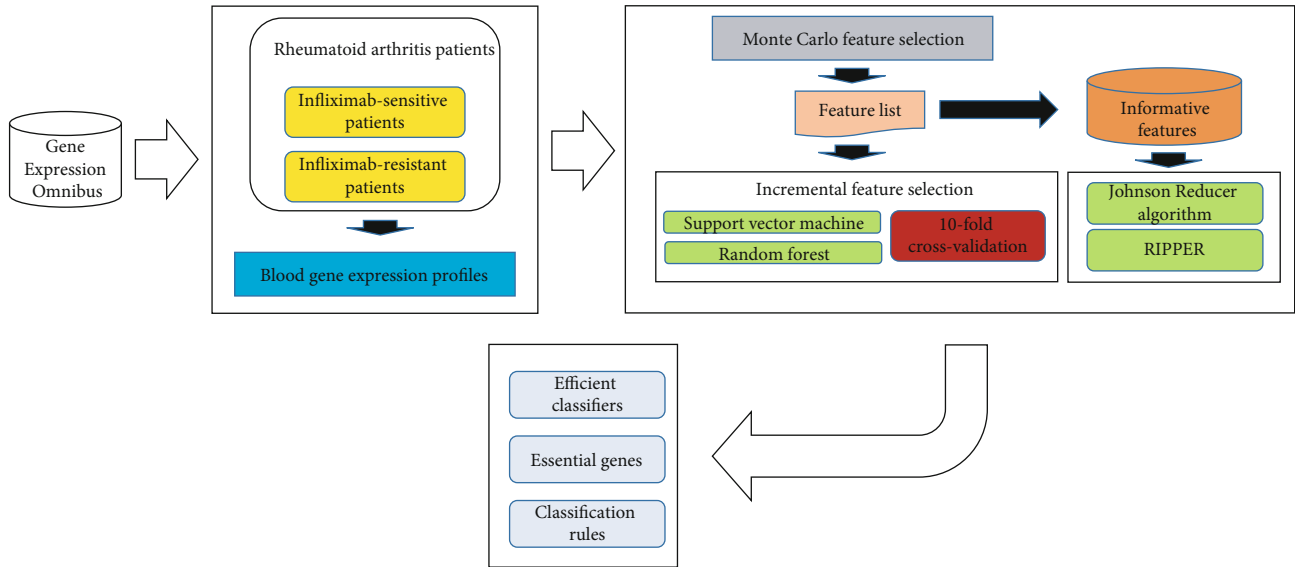
FIGURE 1: Entire procedures to investigate the blood gene expression profiles of rheumatoid arthritis patients. Profiles are retrieved from Gene Expression Omnibus, which are analyzed by the Monte Carlo feature selection method. A feature list is obtained, which is fed into the incremental feature selection method to construct efficient classifiers and extract essential genes. On the other hand, informative features, which are some top features in the list, are used to construct classification rules via Johnson Reducer and RIPPER algorithms.
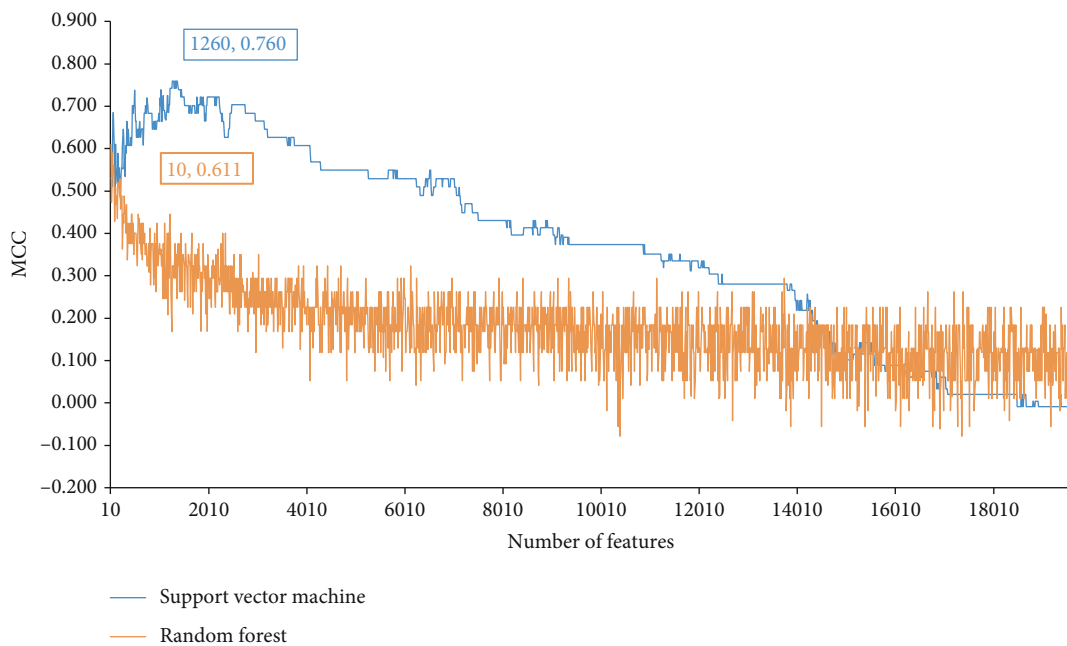

FIGURE 2: IFS curves with different classification algorithms on different numbers of features (genes). The support vector machine yields the highest MCC of 0.760 when top 1260 features are used, whereas the random forest generates the highest MCC of 0.611 when top 10 features are adopted.

TABLE 1: Performance of some key support vector machine (SVM) and random forest (RF) classifiers.

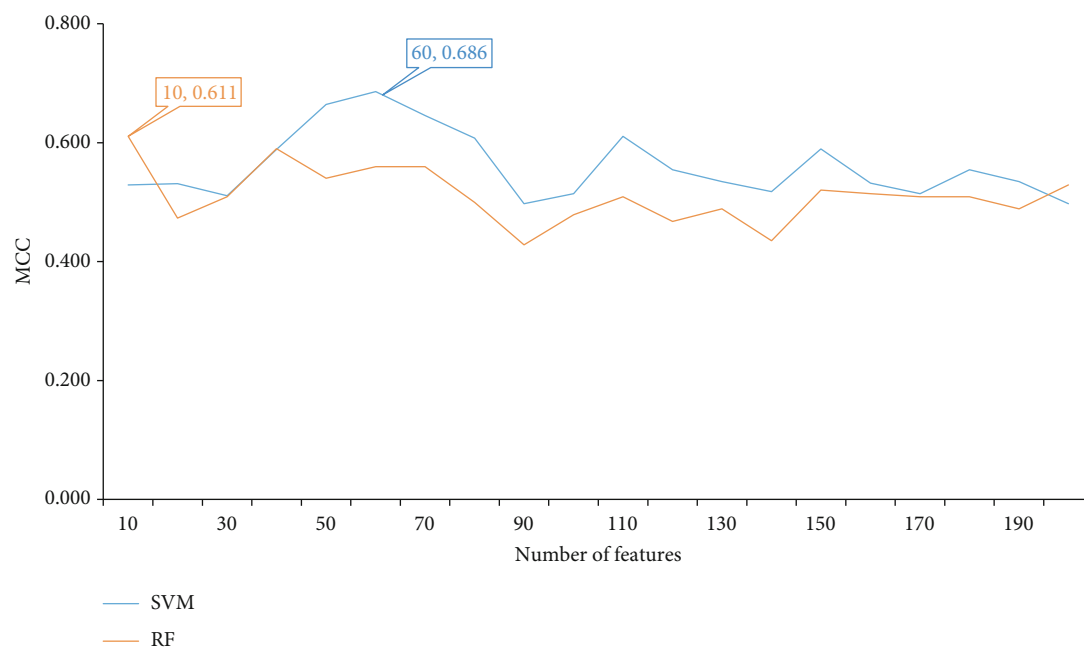| Classification algorithm	Number of features	Sensitivity	Specificity	Accuracy	Precision	$F1$-measure
SVM	1260	0.690	0.990	0.900	0.967	0.806
SVM	60	0.619	0.980	0.871	0.929	0.743
RF	10	0.643	0.929	0.843	0.794	0.711

FIGURE 3: IFS curves with different classification algorithms on top 10-200 features. The support vector machine (SVM) yields the MCC of 0.686 when only top 60 features are used. It is still higher than the highest MCC yielded by the optimum random forest (RF) classifier.

classifier. Therefore, it can be concluded that the optimum SVM classifier is better than the optimum RF classifier.

As mentioned above, the optimum SVM classifier needed much more features than the optimum RF classifier. In fact, the SVM classifier can yield good performance when much less features were used. As shown in Figure 3, when top 60 features were used, the SVM classifier can generate the MCC of 0.686, which was still higher than that of the optimum RF classifier. The detailed performance of such SVM classifier is listed in Table 1. It can be seen that all measurements, except SN, of this SVM classifier were higher than those of the optimum RF classifier. Thus, we picked up such SVM classifier as the proposed classifier because it can provide good performance and was much more efficient than the optimum SVM classifier.

3.3. Rule Learning Results. The optimum SVM and RF classifiers gave good performance. However, they were black-box algorithms. Few medical insights can be captured from these classifiers. In view of this, we further employed a rule learning procedure. The informative features yielded by MCFS were processed by the Johnson Reducer and RIPPER algorithms one by one. As a result, three rules were constructed, as shown in Table 2, where two rules were for prediction of IFX-sensitive patients and the last one was for identification of IFX-resistant patients. We counted two measurements: support and accuracy for each rule, as listed in Table 2. Each rule covered some samples, and all accuracies were quite high, implying the utility of the rules.

Furthermore, to test the effectiveness of the above rule learning procedures, we did the 10-fold cross-validation three times. Six measurements calculated by equations (2) and (3) were counted. The MCC was 0.439, and other five measurements were 0.421 (SN), 0.939 (SP), 0.783 (ACC),

0.746 (precision), and 0.538 (*F*1-measure), respectively. Clearly, the rule classifier was inferior to the optimum SVM and RF classifiers. However, it can explain the detailed gene expression pattern for distinguishing patients with RA who respond to IFX treatment or not.

4. Discussion

IFX is one of the major clinically applied drugs for RA. However, the sensitivity and effectiveness of this drug vary among patients. Recent publications confirmed that the sensitivity of this drug against RA can be predicted by obtaining the expression profiling pattern of patients' pretherapeutic blood. However, the core signatures/biomarkers for the prediction and understanding of IFX sensitivity are difficult to identify. We identified gene signatures for drug therapeutic effect evaluation and established a series of quantitative rules that explain the detailed accurate recognition of patients with different IFX sensitivity using a novel computational approach on the expression profiling of pretherapeutic blood. All the identified signatures have been confirmed by recent publications, and the detailed analysis of the representative genes and rules is discussed below.

4.1. Gene Signatures Associated with IFX Response. In this study, with some computational methods, several genes associated with IFX response were identified. Here, we selected some of them for detailed analysis, which are listed in Table 3.

DISC1, which encodes a scaffold protein, participates in the synthesis of hemoglobin in peripheral blood [62, 63]. Although no direct evidence confirmed that the blood expression of *DISC1* may directly contribute to the pathogenesis of RA, recent publications validated that *DISC1* may

TABLE 2: Classification rules yielded by RIPPER.

Index	Condition	Result	Support[#]	Accuracy[$]
1	DISC1 ≤ 5.2301	IFX-sensitive patient	7.86%	90.91%
2	SAMD11 ≥ 4.2938	IFX-sensitive patient	7.14%	90.00%
3	Others	IFX-resistant patient	85.00%	80.67%

[#]Support is defined as the proportion of samples satisfying the rule to all samples. [$]Accuracy is defined as the proportion of correctly predicted samples to the samples satisfying the rule.

TABLE 3: Some top genes associated with IFX response.

Gene symbol	Description	RI score
DISC1	DISC1 scaffold protein	0.0506
SAMD11	Sterile alpha motif domain containing 11	0.0410
EID2B	EP300 interacting inhibitor of differentiation 2B	0.0345
NTS	Neurotensin	0.0257
STAT2	Signal transducer and activator of transcription 2	0.0233
HELZ	Helicase with zinc finger	0.0198
SUMO2	Small ubiquitin-like modifier 2	0.0190

participate in TNF-α-associated biological processes [64, 65]. A further study on the biological processes of TNF-α receptors reported that our predicted gene, *DISC1*, may be functionally related to MIPT3, a microtubule-interacting protein associated with TNF receptors; thus, *DISC1* has potential regulatory effects on TNF-associated biological processes [66]. The therapeutic effects of IFX rely on the pharmacological regulation on TNF-α-associated biological processes [67, 68]; therefore, as a regulator for the TNF receptor, the different expression patterns of *DISC1* can indicate different TNF-related biological processes and affect the pharmacological effects of IFX.

SAMD11, as a transcription coactivator, contributes to the development of a photoreceptor [69]. In 2009, *SAMD11* was identified as a susceptibility gene for RA by high-throughput genotyping techniques [70]. Multiple publications have also confirmed its specific role in inflammatory regulation [71, 72] and even validated its functional relationship with IL17 [73, 74]. The pharmacological effects of IFX are mediated by TNF-α inhibition; thus, IFX is functionally related to the regulation of immune responses via the interactions between IL17 and TNF-α [75, 76]. As a specific regulator of IL17, our candidate gene, *SMAD11*, may also participate in the regulation of IFX-mediated pharmacological responses in patients with RA. Therefore, the expression level of *SMAD11* in the blood may indirectly reflect the reactiveness of IFX in patients with RA.

EID2B, as an RA-associated gene, participates in MyoD-dependent transcription, glucocorticoid receptor-dependent transcription, and muscle differentiation as a functional repressor [77, 78]. MyoD-associated biological functions induce muscle lesions by interacting with abnormal regulation on TNF-α pathways, targeted by IFX during the initiation and progression of RA [79]. Further studies on the pharmacologi-

cal effects of IFX confirmed that MyoD may directly participate in IFX-associated therapeutic metabolism in Crohn's disease, which is another autoimmune disease [80]. Our predicted gene *EID2B* may also participate in the regulation of the therapeutic effects of IFX by interacting with the core regulator, MyoD. Therefore, the expression level of *EID2B* in the peripheral blood may reflect *EID2B* expression in mesenchymal cells and immune cells in the blood and may also be a novel parameter for the prediction of the prognosis of IFX-dependent therapeutics.

NTS, which encodes a common precursor for peptides neuromedin N and neurotensin, participates in the regulation of fat metabolism in muscles [81, 82]. A recent research confirmed its specific pathogenic role in the pathophysiology of RA [83]. As for its distinctive role for the sensitivity of IFX, NTS participates in TNF-α-associated biological processes [84, 85]. Considering that IFX acts as the inhibitor of TNF-α-associated biological processes [23], the therapeutic effects of IFX are functionally associated with the abnormal activation status of TNF-α-associated biological processes. Therefore, as an effective participator of TNF-α-associated biological processes, the expression pattern of NTS may reflect the activation status of TNF-α-associated biological processes and therefore indicates the therapeutic effects of IFX.

STAT2, as an effective member of the STAT protein family, participates in the transcriptional regulation mediated by type I interferons (IFNs) [86] and the Jak kinase signaling cascade [87, 88]. The treatment effects of IFX are functionally connected to the IFN signaling cascade [89–91]; hence, IFN-associated biological processes may be crucial for the pharmacological effects of IFX. Therefore, as a transcriptional regulator at the downstream of IFN-associated biological processes, the expression level of our predicted gene *STAT2* may also be an alternative following the expression alteration of genes in the type I IFN-associated pathways. This finding indicates the potential identification ability of *STAT2* on the therapeutic effects of IFX.

The *HELZ* gene can encode a member of the RNA helicase superfamily I class. *HELZ* participates in RNA hydrolysis in multiple tissues [92]. In fact, TNF-α and its related immune regulatory functions are regulated by multiple RNA helicases, including helicase with zinc finger (HELZ) [93–95]. Therefore, the expression pattern of *HELZ* may also affect the therapeutic effects of IFX by interfering with TNF-α-associated biological processes. Similarly, *SUMO2* also affects IFX sensitivity by interfering with TNF-α-associated biological processes [96, 97].

Apart from unannotated RNA transcripts with no validated protein products, all the predicted genes have been confirmed to be functionally related to TNF-α-associated

biological processes in either physical or pathological conditions. Therefore, these genes may further affect the therapeutic effects and sensitivity of IFX in patients with RA. This finding validates the efficacy and accuracy of our prediction method and analysis.

4.2. Signature Rules Associated with IFX Response. Apart from the qualitative analysis of each top-ranked gene signatures in our prediction list, we also set up a series of quantitative recognition rules for the detailed and accurate recognition of IFX-sensitive and IFX-resistant patients. The first rule involves only the *DISC1* gene. A low *DISC1* expression (<5 FPKM) indicates that the patient may be sensitive to IFX. Based on the analysis, a low *DISC1* expression may indicate abnormal TNF-α-associated immune activation status in RA [66]. Therefore, patients with a low *DISC1* expression pattern may have an activated TNF-associated signaling pathway and are definitely sensitive to therapeutic effects. The expression level of *DISC1* is quite high in normal conditions (>10 FPKM). Therefore, our threshold may indicate low *DISC1* expression level and corresponds to our analysis above.

The second rule involves the *SAMD11* gene. Different from *DISC1*, a high *SAMD11* expression level may indicate an activated inflammatory status in the whole body and a high TNF-α expression level [75, 76]. Therefore, the therapeutic effects of IFX may also be more effective in conditions with more potential pharmacological targets. This finding corresponds to our prediction rules. According to recent publications, the expression level of *SAMD11* in normal whole blood is <1 FPKM. Therefore, under the predicted conditions, the expression level of *SAMD11* may be upregulated and result in the stronger activation of the TNF-α signaling pathway. Patients with specific *SAMD11* expression levels that follow our rules would be sensitive to IFX. By contrast, patients with expression profiling that does not follow these quantitative rules may be resistant to IFX-mediated RA therapeutics.

5. Conclusions

The identified blood gene signatures participate in IFX-sensitive pharmacological processes in patients with RA. Thus, these genes may be potential biomarkers for the distinction of IFX-sensitive and IFX-resistant patients at the transcriptomic level. Several quantitative signature rules for the distinction of patients have also been verified by other recent publications. Therefore, our newly presented method provides comprehensive qualitative and quantitative prediction standards for prognosis guidance on the clinical application of IFX on patients with RA.

Data Availability

The data used to support the findings of this study have been deposited in the Gene Expression Omnibus repository (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE78068).

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Authors' Contributions

ShiJian Ding and ZhanDong Li contributed equally to this work.

Supplementary Materials

Table S1: list of genes ranked based on the RI score from MCFS. Table S2: performance of IFS with SVM. Table S3: performance of IFS with RF. *(Supplementary materials)*

References

[1] G. R. Burmester and J. E. Pope, "Novel treatment strategies in rheumatoid arthritis," *Lancet*, vol. 389, no. 10086, pp. 2338–2348, 2017.

[2] J. S. Smolen, D. Aletaha, and I. B. McInnes, "Rheumatoid arthritis," *Lancet*, vol. 388, no. 10055, pp. 2023–2038, 2016.

[3] R. C. Lawrence, D. T. Felson, C. G. Helmick et al., "Estimates of the prevalence of arthritis and other rheumatic conditions in the United States: part II," *Arthritis Rheum*, vol. 58, no. 1, pp. 26–35, 2008.

[4] C. G. Helmick, D. T. Felson, R. C. Lawrence et al., "Estimates of the prevalence of arthritis and other rheumatic conditions in the United States: part I," *Arthritis Rheum*, vol. 58, no. 1, pp. 15–25, 2008.

[5] R. N. Maini and M. Feldmann, "Cytokine therapy in rheumatoid arthritis," *Lancet*, vol. 348, no. 9030, pp. 824-825, 1996.

[6] A. J. Hayes, "Angioneogenesis in rheumatoid arthritis," *Lancet*, vol. 354, no. 9176, pp. 423-424, 1999.

[7] I. B. McInnes and G. Schett, "Pathogenetic insights from the treatment of rheumatoid arthritis," *Lancet*, vol. 389, no. 10086, pp. 2328–2337, 2017.

[8] G. R. Burmester, J. M. Kremer, F. van den Bosch et al., "Safety and efficacy of upadacitinib in patients with rheumatoid arthritis and inadequate response to conventional synthetic disease-modifying anti-rheumatic drugs (SELECT-NEXT): a randomised, double-blind, placebo-controlled phase 3 trial," *Lancet*, vol. 391, no. 10139, pp. 2503–2512, 2018.

[9] J. M. Torpy, G. D. Perazza, and R. M. Golub, "JAMA patient page. Rheumatoid arthritis," *JAMA*, vol. 305, no. 17, p. 1824, 2011.

[10] R. Ferrari, "An examination of expectations for rheumatoid arthritis disability in Germany: comparison with Canadian data," *Seminars in Arthritis and Rheumatism*, vol. 44, no. 2, p. e4, 2014.

[11] T. Mine, K. Ihara, H. Kawamura, R. Kuriyama, and R. Date, "Knee arthritis without other joint symptoms in the elderly with seronegative elderly onset rheumatoid arthritis," *The Open Orthopaedics Journal*, vol. 10, no. 1, pp. 793–796, 2016.

[12] J. Isacson, E. Allander, and L. A. Brostrom, "17-year follow-up of symptoms and signs in the knee joint in rheumatoid arthritis," *Scandinavian Journal of Rheumatology*, vol. 17, no. 5, pp. 325–331, 1988.

[13] F. Davatchi, I. S. Abari, S. Soroosh, M. Soroosh, and B. S. Abdollahi, "Performance of the 2010 rheumatoid arthritis classification criteria," *International Journal of Rheumatic Diseases*, vol. 15, no. 5, pp. 455–461, 2012.

[14] M. Infantino, M. Manfredi, F. Meacci et al., "Anti-citrullinated peptide antibodies and rheumatoid factor isotypes in the diagnosis of rheumatoid arthritis: an assessment of combined tests," *Clinica Chimica Acta*, vol. 436, pp. 237–242, 2014.

[15] R. H. Mackey, L. H. Kuller, K. D. Deane et al., "Rheumatoid arthritis, anti-cyclic citrullinated peptide positivity, and cardiovascular disease risk in the Women's Health Initiative," *Arthritis & Rhematology*, vol. 67, no. 9, pp. 2311–2322, 2015.

[16] E. de Miguel, T. E. C. O.-D. A. I. Group, A. Pecondón-Español et al., "A reduced 12-joint ultrasound examination predicts lack of X-ray progression better than clinical remission criteria in patients with rheumatoid arthritis," *Rheumatology International*, vol. 37, no. 8, pp. 1347–1356, 2017.

[17] P. Sivakumaran, S. Hussain, L. Attipoe, and C. Ciurtin, "Diagnostic accuracy of simplified ultrasound hand examination protocols for detection of inflammation and disease burden in patients with rheumatoid arthritis," *Acta Radiologica*, vol. 60, no. 1, pp. 92–99, 2019.

[18] S. Witulski, T. J. Vogl, S. Rehart, and P. Ottl, "Evaluation of the TMJ by means of clinical TMD examination and MRI diagnostics in patients with rheumatoid arthritis," *BioMed Research International*, vol. 2014, Article ID 328560, 2014.

[19] D. Vanags, B. Williams, B. Johnson et al., "Therapeutic efficacy and safety of chaperonin 10 in patients with rheumatoid arthritis: a double-blind randomised trial," *Lancet*, vol. 368, no. 9538, pp. 855–863, 2006.

[20] L. Klareskog, D. van der Heijde, J. P. de Jager et al., "Therapeutic effect of the combination of etanercept and methotrexate compared with each treatment alone in patients with rheumatoid arthritis: double-blind randomised controlled trial," *Lancet*, vol. 363, no. 9410, pp. 675–681, 2004.

[21] M. Fechtenbaum, J. L. Nam, and P. Emery, "Biologics in rheumatoid arthritis: where are we going?," *British Journal of Hospital Medicine (London, England)*, vol. 75, no. 8, pp. 448–456, 2014, 451-6.

[22] K. E. Donahue, G. Gartlehner, D. E. Jonas et al., "Systematic review: comparative effectiveness and harms of disease-modifying medications for rheumatoid arthritis," *Annals of Internal Medicine*, vol. 148, no. 2, pp. 124–134, 2008.

[23] G. Martius, S. Cameron, M. Rave-Fränk, C. Hess, H. Wolff, and I. Malik, "The anti-TNF-α antibody infliximab inhibits the expression of FAT-transporter-protein FAT/CD36 in a selective hepatic-radiation mouse model," *International Journal of Molecular Sciences*, vol. 16, no. 3, pp. 4682–4697, 2015.

[24] J. K. Eriksson, J. K. Wallman, H. Miller et al., "Infliximab versus conventional combination treatment and seven-year work loss in early rheumatoid arthritis: results of a randomized Swedish trial," *Arthritis Care Res (Hoboken)*, vol. 68, no. 12, pp. 1758–1766, 2016.

[25] D. Dimopoulou, T. Dimitroulas, E. Akriviadis, and A. Garyfallos, "Infliximab as a treatment option for patients with rheumatoid arthritis and primary biliary cirrhosis," *Rheumatology International*, vol. 35, no. 11, pp. 1913–1916, 2015.

[26] S. Nakamura, K. Suzuki, H. Iijima et al., "Identification of baseline gene expression signatures predicting therapeutic responses to three biologic agents in rheumatoid arthritis: a retrospective observational study," *Arthritis Research & Therapy*, vol. 18, no. 1, p. 159, 2016.

[27] M. Draminski, A. Rada-Iglesias, S. Enroth, C. Wadelius, J. Koronacki, and J. Komorowski, "Monte Carlo feature selection for supervised classification," *Bioinformatics*, vol. 24, no. 1, pp. 110–117, 2008.

[28] X. Pan, T. Zeng, Y.-H. Zhang et al., "Investigation and prediction of human interactome based on quantitative features," *Frontiers in Bioengineering and Biotechnology*, vol. 8, p. 730, 2020.

[29] Y.-H. Zhang, Z. Li, T. Zeng, W. C. Lu, T. Huang, and Y.-D. Cai, "Identifying the immunological gene signatures of immune cell subtypes," *BioMed Research International*, vol. 2021, Article ID 6639698, 2021.

[30] H. A. Liu and R. Setiono, "Incremental feature selection," *Applied Intelligence*, vol. 9, no. 3, pp. 217–230, 1998.

[31] B. Matthews, "Comparison of the predicted and observed secondary structure of T4 phage lysozyme," *Biochimica et Biophysica Acta (BBA)-Protein Structure*, vol. 405, no. 2, pp. 442–451, 1975.

[32] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *In International joint Conference on artificial intelligence*, Lawrence Erlbaum Associates Ltd., 1995.

[33] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[34] J.-P. Zhou, L. Chen, and Z.-H. Guo, "iATC-NRAKEL: an efficient multi-label classifier for recognizing anatomical therapeutic chemical classes of drugs," *Bioinformatics*, vol. 36, no. 5, pp. 1391–1396, 2020.

[35] J.-P. Zhou, L. Chen, T. Wang, and M. Liu, "iATC-FRAKEL: a simple multi-label web server for recognizing anatomical therapeutic chemical classes of drugs with their fingerprints only," *Bioinformatics*, vol. 36, no. 11, pp. 3568-3569, 2020.

[36] S. P. Wang, Q. Zhang, J. Lu, and Y.-D. Cai, "Analysis and prediction of nitrated tyrosine sites with the mRMR method and support vector machine algorithm," *Current Bioinformatics*, vol. 13, no. 1, pp. 3–13, 2018.

[37] X. G. Chen, W. W. Shi, and L. Deng, "Prediction of disease comorbidity using HeteSim scores based on multiple heterogeneous networks," *Current Gene Therapy*, vol. 19, no. 4, pp. 232–241, 2019.

[38] Y. Zhu, B. Hu, L. Chen, and Q. Dai, "iMPTCE-Hnetwork: a multi-label classifier for identifying metabolic pathway types of chemicals and enzymes with a heterogeneous network," *Computational and Mathematical Methods in Medicine*, vol. 2021, Article ID 6683051, 2021.

[39] F. Yuan, Z. Li, L. Chen et al., "Identifying the signatures and rules of circulating extracellular microRNA for distinguishing

cancer subtypes," *Frontiers in Genetics*, vol. 12, p. 651610, 2021.

[40] E. Frank, M. Hall, L. Trigg, G. Holmes, and I. H. Witten, "Data mining in bioinformatics using Weka," *Bioinformatics*, vol. 20, no. 15, pp. 2479–2481, 2004.

[41] I. H. Witten and E. Frank, *Data Mining:Practical Machine Learning Tools and Techniques*, San Francisco, Morgan, Kaufmann, 2005.

[42] J. Platt, *Sequential minimal optimizaton: a fast algorithm for training support vector machines*, Technical Report MSR-TR-98-14, 1998.

[43] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[44] X. Y. Pan, Y. N. Zhang, and H. B. Shen, "Large-scale prediction of human protein-protein interactions from amino acid sequence based on latent topic features," *Journal of Proteome Research*, vol. 9, no. 10, pp. 4992–5001, 2010.

[45] Y. Jia, R. Zhao, and L. Chen, "Similarity-based machine learning model for predicting the metabolic pathways of compounds," *IEEE Access*, vol. 8, pp. 130687–130696, 2020.

[46] H. Liang, L. Chen, X. Zhao, and X. Zhang, "Prediction of drug side effects with a refined negative sample selection strategy," *Computational and Mathematical Methods in Medicine*, vol. 2020, Article ID 1573543, 2020.

[47] X. Zhao, L. Chen, and J. Lu, "A similarity-based method for prediction of drug side effects with heterogeneous information," *Mathematical Biosciences*, vol. 306, pp. 136–144, 2018.

[48] X. Zhao, L. Chen, Z. H. Guo, and T. Liu, "Predicting drug side effects with compact integration of heterogeneous networks," *Current Bioinformatics*, vol. 14, no. 8, pp. 709–720, 2019.

[49] J. R. Li, L. Lu, Y.-. H. Zhang et al., "Identification of synthetic lethality based on a functional network by using machine learning algorithms," *Journal of Cellular Biochemistry*, vol. 120, no. 1, pp. 405–416, 2018.

[50] E. Kwon, M. Cho, H. Kim, and H. S. Son, "A study on host tropism determinants of influenza virus using machine learning," *Current Bioinformatics*, vol. 15, no. 2, pp. 121–134, 2020.

[51] P. Chen, T. Shen, Y. Zhang, and B. Wang, "A sequence-segment neighbor encoding schema for protein hotspot residue prediction," *Current Bioinformatics*, vol. 15, no. 5, pp. 445–454, 2020.

[52] Y.-H. Zhang, H. Li, T. Zeng et al., "Identifying transcriptomic signatures and rules for SARS-CoV-2 infection," *Frontiers in Cell and Developmental Biology*, vol. 8, p. 627302, 2021.

[53] D. S. Johnson, "Approximation algorithms for combinatorial problems," *Journal of Computer and System Sciences*, vol. 9, no. 3, pp. 256–278, 1974.

[54] W. W. Cohen, "Fast effective rule induction," in *Twelfth International Conference on Machine Learning*, pp. 115–123, Tahoe City, California, 1995.

[55] L. Chen, S. Wang, Y. H. Zhang et al., "Identify key sequence features to improve CRISPR sgRNA efficacy," *IEEE Access*, vol. 5, pp. 26582–26590, 2017.

[56] L. Chen, C. Chu, Y.-H. Zhang et al., "Identification of drug-drug interactions using chemical interactions," *Current Bioinformatics*, vol. 12, no. 6, pp. 526–534, 2017.

[57] L. Liu, L. Chen, Y. H. Zhang et al., "Analysis and prediction of drug-drug interaction by minimum redundancy maximum relevance and incremental feature selection," *Journal of Biomolecular Structure & Dynamics*, vol. 35, no. 2, pp. 312–329, 2017.

[58] H. Liu, B. Hu, L. Chen, and L. Lu, "Identifying protein subcellular location with embedding features learned from networks," *Current Proteomics*, vol. 17, 2020.

[59] X. Pan, H. Li, T. Zeng et al., "Identification of protein subcellular localization with network and functional embeddings," *Frontiers in Genetics*, vol. 11, p. 626500, 2021.

[60] Y.-H. Zhang, Z. Li, T. Zeng et al., "Detecting the multiomics signatures of factor-specific inflammatory effects on airway smooth muscles," *Frontiers in Genetics*, vol. 11, p. 599970, 2021.

[61] Y.-H. Zhang, T. Zeng, L. Chen, T. Huang, and Y. D. Cai, "Determining protein-protein functional associations by functional rules based on gene ontology and KEGG pathway," *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, vol. 1869, no. 6, p. 140621, 2021.

[62] J. R. Liu, Q. Liu, J. Khoury et al., "Hypoxic preconditioning decreases nuclear factor κB activity via disrupted in schizophrenia-1," *The International Journal of Biochemistry & Cell Biology*, vol. 70, pp. 140–148, 2016.

[63] A. Mohammadi, E. Rashidi, and V. G. Amooeian, "Brain, blood, cerebrospinal fluid, and serum biomarkers in schizophrenia," *Psychiatry Research*, vol. 265, pp. 25–38, 2018.

[64] C. Noto, V. K. Ota, M. L. Santoro et al., "Depression, cytokine, and cytokine by treatment interactions modulate gene expression in antipsychotic naïve first episode psychosis," *Molecular Neurobiology*, vol. 53, no. 8, pp. 5701–5709, 2016.

[65] N. F. Berbari, N. W. Kin, N. Sharma, E. J. Michaud, R. A. Kesterson, and B. K. Yoder, "Mutations in Traf3ip1 reveal defects in ciliogenesis, embryonic development, and altered cell size regulation," *Developmental Biology*, vol. 360, no. 1, pp. 66–76, 2011.

[66] J. A. Morris, G. Kandpal, L. Ma, and C. P. Austin, "DISC1 (disrupted-in-schizophrenia 1) is a centrosome-associated protein that interacts with MAP1A, MIPT3, ATF4/5 and NUDEL: regulation and loss of interaction with mutation," *Human Molecular Genetics*, vol. 12, no. 13, pp. 1591–1608, 2003.

[67] I. Voloshyna, S. Seshadri, K. Anwar et al., "Infliximab Reverses Suppression of Cholesterol Efflux Proteins by TNF-α: A Possible Mechanism for Modulation of Atherogenesis," *BioMed Research International*, vol. 2014, Article ID 312647, 8 pages, 2014.

[68] S. Yoshida, T. Takeuchi, T. Kotani et al., "Infliximab, a TNF-α inhibitor, reduces 24-h ambulatory blood pressure in rheumatoid arthritis patients," *Journal of Human Hypertension*, vol. 28, no. 3, pp. 165–169, 2014.

[69] M. Corton, A. Avila-Fernández, L. Campello et al., "Identification of the photoreceptor transcriptional co-repressor SAMD11 as novel cause of autosomal recessive retinitis pigmentosa," *Scientific Reports*, vol. 6, no. 1, p. 35370, 2016.

[70] Q. Sha, R. Tang, and S. Zhang, "Detecting susceptibility genes for rheumatoid arthritis based on a novel sliding-window approach," *BMC Proc*, vol. 3, Supplement 7, p. S14, 2009.

[71] W. Liu, Q. Ma, K. Wong et al., "Brd4 and JMJD6-associated anti-pause enhancers in regulation of transcriptional pause release," *Cell*, vol. 155, no. 7, pp. 1581–1595, 2013.

[72] A. M. Zawada, J. S. Schneider, A. I. Michel et al., "DNA methylation profiling reveals differences in the 3 human monocyte subsets and identifies uremia to induce DNA methylation changes during differentiation," *Epigenetics*, vol. 11, no. 4, pp. 259–272, 2016.

[73] J. Diegelmann, T. Olszak, B. Göke, R. S. Blumberg, and S. Brand, "A novel role for interleukin-27 (IL-27) as mediator

of intestinal epithelial barrier protection mediated via differential signal transducer and ctivator of transcription (STAT) protein signaling and induction of antibacterial and anti-inflammatory proteins," *The Journal of Biological Chemistry*, vol. 287, no. 1, pp. 286–298, 2012.

[74] Y. Tian, H. Shen, L. Xia, and J. Lu, "Elevated serum and synovial fluid levels of interleukin-34 in rheumatoid arthritis: possible association with disease progression via interleukin-17 production," *Journal of Interferon & Cytokine Research*, vol. 33, no. 7, pp. 398–401, 2013.

[75] T. Hennerici, R. Pollmann, T. Schmidt et al., "Increased frequency of T follicular helper cells and elevated interleukin-27 plasma levels in patients with pemphigus," *PLoS One*, vol. 11, no. 2, article e0148919, 2016.

[76] D. A. Alves da Silva, M. V. da Silva, C. C. Oliveira Barros et al., "TNF-α blockade impairs in vitro tuberculous granuloma formation and down modulate Th1, Th17 and Treg cytokines," *PLoS One*, vol. 13, no. 3, article e0194430, 2018.

[77] E. P. D. Reis, D. M. Paixão, O. J. B. Brustolini et al., "Expression of myogenes in longissimus dorsi muscle during prenatal development in commercial and local Piau pigs," *Genetics and Molecular Biology*, vol. 39, no. 4, pp. 589–599, 2016.

[78] E. L. Huttlin, R. J. Bruckner, J. A. Paulo et al., "Architecture of the human interactome defines protein communities and disease networks," *Nature*, vol. 545, no. 7655, pp. 505–509, 2017.

[79] R. J. Boutrup, J. Farup, K. Vissing, M. Kjaer, and U. R. Mikkelsen, "Skeletal muscle stem cell characteristics and myonuclei content in patients with rheumatoid arthritis: a cross-sectional study," *Rheumatology International*, vol. 38, no. 6, pp. 1031–1041, 2018.

[80] K. Subramaniam, K. Fallon, T. Ruut et al., "Infliximab reverses inflammatory muscle wasting (sarcopenia) in Crohn's disease," *Alimentary Pharmacology & Therapeutics*, vol. 41, no. 5, pp. 419–428, 2015.

[81] J. Li, J. Song, Y. Y. Zaytseva et al., "An obligatory role for neurotensin in high-fat-diet-induced obesity," *Nature*, vol. 533, no. 7603, pp. 411–415, 2016.

[82] J. Piątek, H. Witmanowski, J. Paluszak, H. Krauss, and J. Krawczyk, "The effects of neurotensin on selected parameters of lipid metabolism in rats," *Peptides*, vol. 26, no. 5, pp. 837–843, 2005.

[83] J. P. Famaey, T. Appelboom, and J. Wybran, "The nervous system and the pathophysiology of rheumatoid arthritis: the role of neurotensin and opioid peptides," *The Journal of Rheumatology*, vol. 13, no. 3, pp. 651-652, 1986.

[84] D. Zhao, Y. Zhan, H. W. Koon et al., "Metalloproteinase-dependent transforming growth factor-α release mediates neurotensin-stimulated MAP kinase activation in hhuman colonic epithelial cells," *The Journal of Biological Chemistry*, vol. 279, no. 42, pp. 43547–43554, 2004.

[85] L. I. F. Moura, A. M. A. Dias, E. Suesca et al., "Neurotensin-loaded collagen dressings reduce inflammation and improve wound healing in diabetic mice," *Biochimica et Biophysica Acta*, vol. 1842, no. 1, pp. 32–43, 2014.

[86] C. Yue, J. Xu, M. D. Tan Estioko et al., "Host STAT2/type I interferon axis controls tumor growth," *International Journal of Cancer*, vol. 136, no. 1, pp. 117–126, 2015.

[87] M. L. Slattery, A. Lundgreen, S. A. Kadlubar, K. L. Bondurant, and R. K. Wolff, "JAK/STAT/SOCS-signaling pathway and colon and rectal cancer," *Molecular Carcinogenesis*, vol. 52, no. 2, pp. 155–166, 2013.

[88] H. Wu, Y. Wu, Z. Ai et al., "Vitamin C enhances Nanog expression via activation of the JAK/STAT signaling pathway," *Stem Cells*, vol. 32, no. 1, pp. 166–176, 2014.

[89] L. G. van Baarsen, C. A. Wijbrandts, F. Rustenburg et al., "Regulation of IFN response gene activity during infliximab treatment in rheumatoid arthritis is associated with clinical response to treatment," *Arthritis Research & Therapy*, vol. 12, no. 1, p. R11, 2010.

[90] M. Takeshita, K. Suzuki, J. Kikuchi et al., "Infliximab and etanercept have distinct actions but similar effects on cytokine profiles in rheumatoid arthritis," *Cytokine*, vol. 75, no. 2, pp. 222–227, 2015.

[91] C. P. Mavragani, D. T. la, W. Stohl, and M. K. Crow, "Association of the response to tumor necrosis factor antagonists with plasma type I interferon activity and interferon-beta/alpha ratios in rheumatoid arthritis patients: a post hoc analysis of a predominantly Hispanic cohort," *Arthritis and Rheumatism*, vol. 62, no. 2, pp. 392–401, 2010.

[92] P. A. Hasgall, D. Hoogewijs, M. B. Faza, V. G. Panse, R. H. Wenger, and G. Camenisch, "The putative RNA helicase HELZ promotes cell proliferation, translation initiation and ribosomal protein S6 phosphorylation," *PLoS One*, vol. 6, no. 7, article e22107, 2011.

[93] K. Mosallanejad, Y. Sekine, S. Ishikura-Kinoshita et al., "The DEAH-box RNA helicase DHX15 activates NF-B and MAPK signaling downstream of MAVS during antiviral responses," *Science Signaling*, vol. 7, no. 323, p. ra40, 2014.

[94] V. R. R. Mendonça, L. C. L. Souza, G. C. Garcia et al., "DDX39B (BAT1), TNF and IL6 gene polymorphisms and association with clinical outcomes of patients with Plasmodium vivax malaria," *Malaria Journal*, vol. 13, no. 1, p. 278, 2014.

[95] J. Pircher, T. Czermak, M. Merkle et al., "Hepatitis C virus induced endothelial inflammatory response depends on the functional expression of TNFα receptor subtype 2," *PLoS One*, vol. 9, no. 11, article e113351, 2014.

[96] A. M. Mabb, S. M. Wuerzberger-Davis, and S. Miyamoto, "PIASy mediates NEMO sumoylation and NF-κB activation in response to genotoxic stress," *Nature Cell Biology*, vol. 8, no. 9, pp. 986–993, 2006.

[97] J. Liu, M. Sha, Q. Wang et al., "Small ubiquitin-related modifier 2/3 interacts with p65 and stabilizes it in the cytoplasm in HBV-associated hepatocellular carcinoma," *BMC Cancer*, vol. 15, no. 1, 2015.