



OPEN ACCESS

EDITED BY

Yiqiang Zhan,
Sun Yat-sen University, China

REVIEWED BY

Jianan Zhao,
Shanghai University of Traditional Chinese
Medicine, China
Xiaojian Shao,
National Research Council Canada (NRC),
Canada

*CORRESPONDENCE

Lingyu Fu
✉ fulingyucmu@sina.com

RECEIVED 24 February 2023

ACCEPTED 14 August 2023

PUBLISHED 07 September 2023

CITATION

Chen Y, Wang Q, Liu H, Jin L, Feng X,
Dai B, Chen M, Xin F, Wei T, Bai B, Fan Z,
Li J, Yao Y, Liao R, Zhang J, Jin X and Fu L
(2023) The prognostic value of whole-
genome DNA methylation in response to
Leflunomide in patients with
Rheumatoid Arthritis.
Front. Immunol. 14:1173187.
doi: 10.3389/fimmu.2023.1173187

COPYRIGHT

© 2023 Chen, Wang, Liu, Jin, Feng, Dai,
Chen, Xin, Wei, Bai, Fan, Li, Yao, Liao, Zhang,
Jin and Fu. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

The prognostic value of whole-genome DNA methylation in response to Leflunomide in patients with Rheumatoid Arthritis

Yulan Chen¹, Qiao Wang¹, Haina Liu², Lei Jin³, Xin Feng⁴,
Bingbing Dai⁵, Meng Chen¹, Fangran Xin¹, Tingting Wei¹,
Bingqing Bai¹, Zhijun Fan¹, Jiahui Li¹, Yuxin Yao¹, Ruobing Liao¹,
Jintao Zhang⁵, Xiangnan Jin⁴ and Lingyu Fu^{1,6*}

¹Department of Clinical Epidemiology and Evidence-Based Medicine, the First Affiliated Hospital, China Medical University, Shenyang, China, ²Department of Rheumatology, The First Affiliated Hospital, China Medical University, Shenyang, China, ³Department of Rheumatology, ShengJing Hospital Affiliated of China Medical University, Shenyang, China, ⁴Department of Rheumatology, First Affiliated Hospital of Jinzhou Medical University, Jinzhou, China, ⁵Department of Rheumatology and Immunology, Dalian Municipal Central Hospital, Dalian, China, ⁶Department of Medical Record Management Center, the First Affiliated Hospital, China Medical University, Shenyang, China

Objective: Although Leflunomide (LEF) is effective in treating rheumatoid arthritis (RA), there are still a considerable number of patients who respond poorly to LEF treatment. Till date, few LEF efficacy-predicting biomarkers have been identified. Herein, we explored and developed a DNA methylation-based predictive model for LEF-treated RA patient prognosis.

Methods: Two hundred forty-five RA patients were prospectively enrolled from four participating study centers. A whole-genome DNA methylation profiling was conducted to identify LEF-related response signatures via comparison of 40 samples using Illumina 850k methylation arrays. Furthermore, differentially methylated positions (DMPs) were validated in the 245 RA patients using a targeted bisulfite sequencing assay. Lastly, prognostic models were developed, which included clinical characteristics and DMPs scores, for the prediction of LEF treatment response using machine learning algorithms.

Results: We recognized a seven-DMP signature consisting of cg17330251, cg19814518, cg20124410, cg21109666, cg22572476, cg23403192, and cg24432675, which was effective in predicting RA patient's LEF response status. In the five machine learning algorithms, the support vector machine (SVM) algorithm provided the best predictive model, with the largest discriminative ability, accuracy, and stability. Lastly, the AUC of the complex model (the 7-DMP scores with the lymphocyte and the diagnostic age) was higher than the simple model (the seven-DMP signature, AUC:0.74 vs 0.73 in the test set).

Conclusion: In conclusion, we constructed a prognostic model integrating a 7-DMP scores with the clinical patient profile to predict responses to LEF treatment. Our model will be able to effectively guide clinicians in determining whether a patient is LEF treatment sensitive or not.

KEYWORDS

rheumatoid arthritis, Leflunomide, methylation, machine learning, prognostic

1 Introduction

Rheumatoid arthritis (RA) is an antibody-mediated persistent inflammatory autoimmune disease that is characterized by inflammation and destruction of the synovial and bone tissues (1, 2). RA is estimated to affect ~ 1% of the global population, and can develop at any age (3). The prevalence of RA in China ranges from 0.32% to 0.36%, while the prevalence of RA in Liaoning Province, located in Northeast China, is 0.5%, ranking first in China (4). The prognosis of RA is poor, with many comorbidities, requiring lifelong treatment, causing significant economic burden to patients, family members, and society (5). Currently, there is no cure for RA internationally, among the optimal RA treatment strategies recommended by the European League Against Rheumatism (EULAR) are disease-modifying anti-rheumatic drugs (DMARDs), such as, Methotrexate (MTX) or Leflunomide (LEF) (6). If these drugs fail to satisfactorily treat disease when employed as a single-agent or in combination therapies, then other biologic DMARDs, including, TNF and interleukin-6 receptor (TNFi and IL-6Ri) inhibitors may be implemented (6).

In our country, the utilization rate of the LEF, second only to MTX, is 45.9% (7). Compared to MTX, LEF treatment is typically accompanied by relatively low incidences of serious side effects and compared to biological agents, LEF has a better cost-effectiveness. As such, LEF would usually be attempted prior to biologic-based intervention (8), particularly in a developing country like China. LEF is the first DMARD that demonstrated a marked improvement in RA symptoms in the past decade (9). LEF activity is mediated by metabolite leflunomide, which interacts with nonspecific cytochrome P450 (CYP) and certain drugs inducers that undergo CYP2C9-mediated metabolic processing (10). Unfortunately, LEF treatment can also promote adverse events, namely, hepatotoxicity, gastrointestinal discomfort, headache, hypertension, peripheral neuropathy, and increased susceptibility to infection (11). As such, identifying biomarkers predictive of LEF responses is of significant value as a means of ensuring that patients can be administered personalized therapeutic regimens that are both safe and effective, thus, minimizing the risk of adverse treatment-related toxicity. In the past, the clinical predictors the erythrocyte sedimentation rate (ESR), C-reactive protein (CRP), Rheumatoid factor (RF) and anticyclic citrullinated peptide antibody (ACPA) are representative serological markers for RA prognosis (12).

As one of the main regulation modes in epigenetics, DNA methylation is reversible, that is, demethylation. It can be regulated by diet, drugs and other environmental factors, so it can be used as a good therapeutic candidate. Several prior investigations revealed that DNA methylation patterns serve as biomarkers for estimation of RA patient therapeutic responses. Glossop et al. (13) quantified DNA methylation patterns in T cells using HumanMethylation450 BeadChips, which lead to the identification of 21 cytosine-phosphate-guanines (CpGs) that, at initial diagnosis, were associated with early RA patient responses to DMARD treatment. Moreover, Nair et al. (14) conducted an epigenome-wide association analysis (EWAS) whereby they compared the differentially methylated positions (DMPs) of individuals who responded well to MTX treatment and those who did not (n=34 each). They revealed that four CpG residues were associated with marked improvements in disease activity, thus facilitating the development of personalized treatment for patients at an earlier stage of disease, based on their epigenetic profile. Whole blood DNA methylation signatures and overall increases in DNA methylation are linked to moderate reductions in 28-joint disease activity score (DAS28) values among MTX non-responders (15). In addition, Plantet et al. (16) evaluated the DNA methylation patterns among RA patients undergoing TNFi treatment. They identified two prominent DMPs belonging to the *LRPAP1* gene. Till date, however, no studies specifically assessed the value of DNA methylation patterns as biomarkers of LEF therapeutic response.

Herein, to explore DNA methylation predictor and develop a DNA methylation-based predictive model for LEF-treated RA patient prognosis, we utilized the 850k Illumina HumanMethylationEPIC Bead Chip as a tool for comparing LEF nonresponders versus responders after 6 months of LEF treatment in order to identify DMPs. Key DMPs were selected via Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses, followed by the subsequent validation of these hub DMPs using a large-sample LEF-treated RA patient population. Subsequently, we developed a prognostic model capable of predicting RA patient responses to LEF treatment. Our prognostic model can potentially assist clinicians in making the best decision for personalized patient therapy and pharmaceutical intervention. It can also aid in determining RA patient prognosis.

2 Materials and methods

2.1 Study design and patient population

To conduct this prospective investigation with an ongoing 6-month follow-up period, 272 samples were collected from patients diagnosed with RA undergoing LEF therapy between June 2018 and June 2020 from four hospitals (the First Hospital of China Medical University, Shengjing Hospital of China Medical University, First Affiliated Hospital of Jinzhou Medical University, and Dalian Municipal Central Hospital) located in the central, southern, and western regions of the Liaoning Province. Of these, 21 patients were lost to follow-up, 4 patients dropped out of treatment due to adverse reactions, 2 patients had poor blood quality, and 245 patients were included in the final analysis.

In the study, patients who met the following inclusion criteria were enrolled: 1) patients were older than 18; 2) They met the RA diagnostic criteria proposed by the American Rheumatology Society (ACR) and EULAR in 2010 (17). The exclusion criteria are below: 1) was pregnant or lactating; 2) was diagnosed with serious liver, heart, or kidney diseases or mental disorders or cancer; 3) was diagnosed with other rheumatoid disease (e.g. gout, osteoarthritis); 4) was diagnosed with other autoimmune diseases (e.g. systemic lupus erythematosus (SLE), inflammatory bowel disease, ankylosing spondylitis, and multiple sclerosis); 5) use other DMARDs or biologics instead of LEF alone. Finally, Prior to the LEF treatment, baseline blood samples and clinical data were acquired from all participants. All participants were provided with a recommended stable LEF dosage of 20 mg/day for 6 months (18, 19). This investigation followed the Declaration of Helsinki, and received ethical approval from the Medical Science Research Institute of the First Affiliated Hospital, China Medical University, with all patients providing written informed consent (approval number: AF-SOP-07-1.0-01).

All participant demographics and clinical information were acquired from the aforementioned four hospitals' medical and laboratory records, including (1) Basic information: patient's gender, age, and so on; (2) Behavioral factors: smoking behavior and alcohol usage, and so on; (3) Immune and biochemical indicators: rheumatoid factor (RF), cyclic citrulline peptide antibody (anti-CCP), erythrocyte sedimentation rate (ESR), and so on. Indicators were categorized, according to the standards of the inspection equipment used by the four medical centers, and normalized to ensure a standardized information entry. All serum indicator values were acquired from the electronic hospital information system of the four hospitals upon the patients' first blood evaluation at admission prior to LEF consumption. Following 6 months of follow-up, indicators associated with the quantity of tender joints, quantity of swollen joints, visual analog scale (VAS), and ESR were acquired again for the calculation of DAS28 of RA patients.

DAS28 values were used to assess RA patient disease activity. Following LEF treatment for 6 months, patient response status was determined according to the current DAS28 values and improvements in these values ($\Delta\text{DAS28} = \text{DAS28}_{\text{baseline}} - \text{DAS28}_{\text{current}}$) as per EULAR criteria (20, 21). A good response

was represented by a DAS28 improvement of > 1.2 , with a final DAS28 score ≤ 3.2 , while nonresponse was defined as an improvement in $\text{DAS28} \leq 0.6$ or $0.6 \leq \text{DAS28} \leq 1.2$, with a $\text{DAS28} > 5.1$ score after LEF treatment for 6 months. Patients who were between the aforementioned criteria were classified as moderate responders. To enhance the statistical power, patients with a good or moderate response were combined into a group referred to as responders for differential methylation analyses. Among the 245 patients who underwent LEF treatment for 6 months, 96 were nonresponders, and 149 were responders, as per the EULAR criteria.

2.2 Illumina 850k beadChip analysis

Overall, 20 responders were randomly selected from 149 responders identified above, following which 20 age- and sex-matched nonresponders were selected for comparison purposes. A QIAamp DNA Blood Mini Kit (QIAGEN) was employed for DNA extraction from the collected blood samples, as per kit directions. Genome-wide methylation profiles were then defined using the Illumina HumanMethylationEPIC BeadChip (Illumina, Inc., CA, USA) (850K array), which provided a coverage spanning $>850,000$ CpG sites per sample throughout the genome.

2.3 Functional enrichment analysis

GO and KEGG enrichment analyses of the screened DMGs were conducted using the R clusterProfiler package (22). $P < 0.05$ was the significance threshold. And, a network of protein-protein interactions (PPI) was created using the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) to illustrate the correlation between notable DMG.

2.4 Replication using targeted bisulfite sequencing assay

Pyrosequencing was carried out for the verification of essential methylation variant variable methylation statuses to establish that the 850K array analysis were, in fact, true methylation differences, and not mere artifacts. The aforementioned DMPs were chosen for validation using the Targeted Bisulfite Sequencing (MethylTarget) conducted at the Genesky Biotechnologies Inc. (Shanghai, China). DNA samples were quantified using the fluorometric method, and DNA integrity was assessed via gel electrophoresis using a 1.3% agarose gel. Briefly, DMPs were analyzed, sequenced, then validated. Genomic DNA conversion employed bisulfite, which converted unmethylated cytosine residues to uracil residues, prior to Polymerase Chain Reaction (PCR) reaction, which, in turn, amplified the target DNA sequences. Product sequencing was then performed on an Illumina HiSeq seq (Illumina, CA, United States). Differential methylation status was compared at each position between responders ($n=149$) and nonresponders ($n=96$)

using the t-test and logistic analysis, following which, methylated CpG sites ($P < 0.05$) were selected for additional analyses.

2.5 Data analysis

2.5.1 850k BeadChip data analysis

Data preprocessing was done with the Chip Analysis Methylation Pipeline ChAMP package (version 2.14.0) from the R Studio (version 3.6.3) (23). Raw read data from IDAT files were converted to β values, defined as the ratio of methylated to unmethylated intensity for a given CpG residue (24). Probes were removed if they did not pass the detection cut-off as follows: ($P < 0.01$), associated with known Single Nucleotide Polymorphism (SNPs), associated with the X and Y chromosomes, or non-CpG probes. The calculated β values ranged from $\beta = 0$ (unmethylated) to $\beta = 1$ (fully methylated), and the beta-mixture quantile normalization (BMIQ) was then employed to decrease bias between the Type I and Type II probes (25). Subsequently, DMPs were identified by comparing the responders and nonresponders using the ChAMP package. Following normalization, a single value decomposition (SVD) analysis was used to identify batch effects, which were then eliminated with ComBat using the gene-wise linear models from the SVA package. DMPs with a $P < 0.05$ and a $|\Delta\beta| > 0.1$ were selected for subsequent analyses. LEF-related differentially methylated genes (DMGs) were next identified, including, genes associated with both hypo- and hypermethylation. A DMP-based heatmap was constructed using the ComplexHeatmap (v1.6.0) and ggplot2 R packages (26).

2.5.2 Prognostic CpG methylation-based model construction and validation

To assess the clinical basic characteristics and potential risk factors, we employed descriptive statistics. Continuous data, analyzed via the t-test, are presented as mean (SD), whereas, categorical data, analyzed via the χ^2 test, are provided as frequencies and percentages. The main drug response confounders, such as the age, gender, smoking, Hypertension, Diabetes, alcohol-drinking and RF, Anti-CCP etc. were analyzed by t-test and χ^2 test between responder group and non-responder group. Two-tailed probabilities were employed, and correlations were deemed when P values were < 0.1 . After that, the factors like the age at diagnosis, LY, MONO, and anti-CCP which were significantly different between two group would be put in the complex model.

CpG unites with missing values in over 20% samples were eliminated, as well as samples with missing values in over 20% of CpG unites. Subsequently, the entire dataset was arbitrarily separated into a training (156/245, 70%) and a test dataset (68/245, 30%). To achieve an enhanced prediction of prognostability of the selected CpGs bioindicators, five machine learning methods, namely, logistic regression, random forest (RF), support vector machine (SVM), adaboost, and naïve Bayes (NB) from the caret R package (version 6.0-86) were employed for the prognostic model generation for LEF response in RA patients, followed by the application of a 10-fold cross-validation (27). Based on an

optimized cutoff value, the sensitivity, specificity, accuracy, F1-score, recall, and precision were obtained from the machine learning model to assess the methylation scores' ability to differentiate between LEF responders and nonresponders. In the end, we conducted a predictive model incorporating these DMPs scores along with other clinical variables.

The Hosmer–Lemeshow test were employed for model calibration evaluation. The model predictive precision for individual endpoints (discriminating ability) was assessed via AUC. The net reclassification improvement (NRI) and integrated discrimination improvement (IDI) are two complementary validation methods that assess the enhanced predictive ability of a complex model compared to a simple model (28, 29). The NRI represented the patient population who were accurately restratified by the newly developed model, relative to the existing or standard model. The IDI referred to the alteration in difference between the mean estimated possibilities between the new and existing models. By calculating the NRI and IDI values of the simple model (the seven-DMP signature) and the complex model (the 7-DMP scores with other clinical variables), the differentiation ability of the models before and after the adjustment of clinical features was compared, and the optimal model was further figured out.

CpGs (Pearson's correlation coefficient $r > 0.5$) with strong associations were eliminated, and a 10-fold cross-validation was conducted across all models. To maintain reproducibility, in all random sampling analyses, we adjusted the random number seed to 100 with the set.seed function prior to any random sampling.

3 Results

3.1 Cohort characteristics

The clinical criteria used for each indicator and patient characteristics analyzed in this investigation are summarized in [Supplementary Tables S1](#) and [S2](#). At baseline, responder and nonresponder patients were between 57.54 and 58.10 years old, respectively, with no significant differences in baseline clinical parameters between the two groups. However, we did observe a significant difference in the age at diagnosis, LY, MONO, and anti-CCP between the responder and nonresponder groups ($P < 0.10$).

3.2 Discovery epigenome-wide association assessment

A total of 3,478 probes were excluded from the analysis during the quality control and normalization process due to having a detection P -value greater than 0.01. In addition, 5,703 probes with a beadcount of less than 3 in at least 5% of the samples were removed. A further 2,970 probes were excluded from the analysis as they were not CpGs, and an additional 10 multi-hit probes were filtered out. Consequently, 16,590 probes located on chromosomes X and Y, as well as 92,002 SNP probes, were also excluded from further analyses. Overall, 740,429 CpG sites were screened for potential differences in methylation status between the responder and non-

responder samples. Based on the SVD analyses, the marked differences between the two groups were attributable to the samples. According to the Slide or Array analyses, DMPs did not exhibit any significant variance (Supplementary Figure S1A).

These comparisons ultimately led to the identification of 81 DMPs ($P < 0.05$; $|\Delta\beta| > 0.1$) between the responders to nonresponders, with 58 and 23 of these DMPs being hypermethylated and hypomethylated, respectively (Figure 1A; Supplementary Figure S1B).

Next, we mapped DMP distributions on human chromosomes using Manhattan plots, which revealed wide distribution across all non-sex chromosomes (Supplementary Figure S1C). Global methylation level comparison between these two patient subsets revealed that the

mean β values of the CpG sites in the first Exon, body, Intergenic regions (IGR), and Transcription start sites(TSS1500) regions, as well as the overall methylation level, was significantly enhanced in responders versus nonresponders. Conversely, the methylation levels in the three prime untranslated region (3'-UTR), five prime untranslated region(5'-UTR), and TSS200 regions were augmented in the nonresponders versus responders, although the difference did not reach significance (Figure 1B). Consistently, the majority of CpG sites located in the first Exon, body, IGR, and TSS1500 are hypermethylated, but in contrast, the other regions showed the opposite trend. (Figure 1C). Moreover, the CpG sites within the Island, N_Shelf, N_Shore, opensea, S_Shelf, S Shore were more

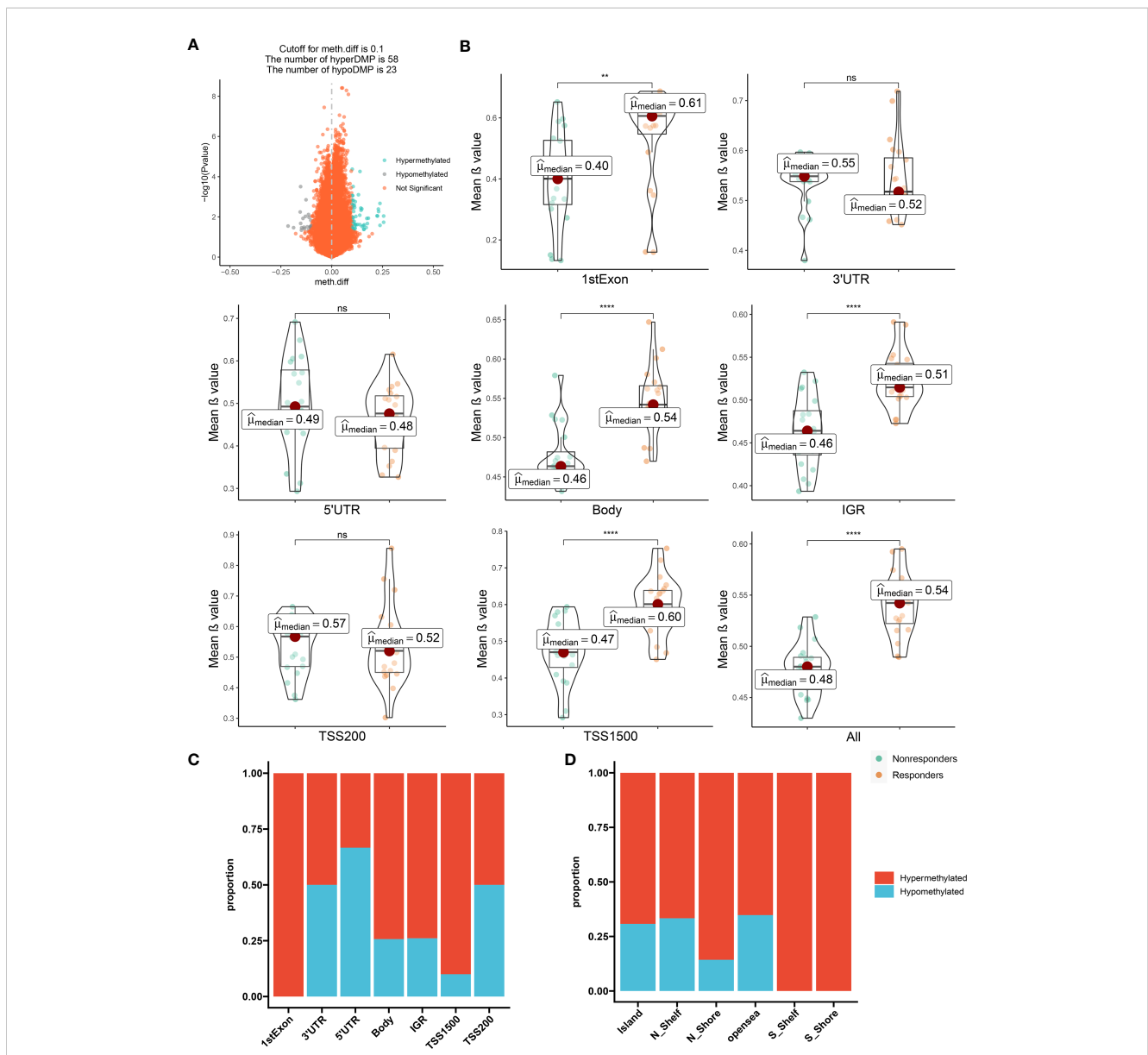


FIGURE 1
 The distribution feature of DMCpGs. **(A)** Volcano plot of DMCpGs. **(B)** Comparisons of mean β value of DMCpGs by locations of CpG sites relative to gene structure. ns=not significant; * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.001$; **** = $P < 0.0001$, by Wilcoxon tests. **(C)** Distribution of DMCpGs in relation to the gene region features. 1st Exon, the first exon of gene; 3'UTR = 3'-untranslated region; 5'UTR = 5'-untranslated region; IGR = intergenic region; TSS1500 = 1500nt upstream of TSS; TSS200 = 200nt upstream of TSS. **(D)** Distribution of DMCpGs in relation to the area related to CpG island. Island = CpG island; N_Shelf=2 kb upstream of corresponding N Shore; N_Shore=2 kb upstream of corresponding CpG island; opensea = the rest of the area; S_Shelf=2 kb downstream of corresponding S Shore; S Shore=2 kb downstream of IGR of corresponding CpG island.

susceptible to hypermethylation (Figure 1D). Hence, in terms of epigenetics, we demonstrated a distinguishable methylation profile between responders and nonresponders, indicating that epigenetics potentially modulates LEF response.

3.3 Functional clustering analyses of significant DMGs

To elucidate the potential physiological roles and signaling networks associated with the aforementioned DMGs, we conducted GO and KEGG pathway enrichment analyses (Supplement Table S3). Based on our analyses, the aforementioned DMGs were significantly enriched in GO terms associated with hydrolase activity, actin filament, muscle cell proliferation, and antigen receptor-mediated signaling pathway, as well as KEGG pathways, such as, viral carcinogenesis, rheumatoid arthritis, PI3k-Akt signaling, and MAPK signaling (Figures 2A–D). These DMGs are thus likely to be associated with the RA-related metabolic processes. In order to investigate the interactions among the DMGs, a PPI network was constructed using STRING, a useful tool for evaluating the molecular functions of proteins. We identified 30 prominent proteins and selected six significant modules (Figure 2E).

3.4 Replication of epigenome-wide findings using Targeted Bisulfite Sequencing

We then performed bisulfite pyrosequencing (MethylTarget), another approach to evaluating DNA methylation, to confirm our earlier findings regarding the 81 DMPs. We validated seven DMPs between the 96 nonresponders and 149 responders. As expected, all seven DMPs exhibited consistent regulation within the samples, and the methylation values from MethylTarget were also consistent with the EPIC array (Supplementary Figure S2). The gene location and biological process (identified by the GO pathway enrichment analyses) of the seven aforementioned DMPs are summarized in Table 1. Based on our results, two hypomethylated DMPs (cg17330251, cg23403192) were located at the *PON1* and *USP16* promoters; another hypomethylated DMP (cg24432675) resided in the gene body of *ADARB2*; one hypermethylated DMP (cg19814518) was located at the *UHMK1* promoter; and one hypermethylated DMP (cg21109666) resided in the gene body of *DISC1*. Unfortunately, no corresponding genes were annotated to the other two hypermethylated DMPs (cg22572476 and cg20124410).

3.5 Prognostic models development and validation

In light of the observed differences in DNA methylation between the LEF responders and non-responders at 6 months after baseline assessment, we sought to explore the predictability of DNA methylation profiles to estimate LEF therapeutic responses. As such, cg17330251, cg24432675, cg23403192, cg19814518,

cg21109666, cg22572476, and cg20124410 were selected for use in machine learning models (a binomial model, 10-fold cross-validation) to generate a prognostic score.

We revealed that, in the training data, all model sensitivities were between 0.79–1.00, whereas, specificities were between 0.40–1.00. The SVM model performed well in terms of accuracy, compared to the other machine learning models. In the Testing data, all model sensitivities and specificities were between 0.72–0.86 and 0.36–0.60, respectively. Thus, the SVM model again showed high precision, suggesting great strength and accuracy of the models. Moreover, we revealed that the prognostability was comparable between the training, testing, and the entire dataset in all models, indicating that our results were both robust and reliable (Figure 3; Supplementary Table S4).

3.6 Development of a prognostic classifier based on the methylation biomarkers and clinical factors

To better elucidate the prognostic value of the selected bioindicators, we employed both methylation and patient clinical profiles to construct prognostic models for LEF response stratification. The age at diagnosis, LY, MONO, anti-CCP, were identified as independent predictors of patient response status ($P < 0.1$; Supplementary Table S1). The predicted values from the DNA methylation models, which serve as independent variables in the subsequent model, can be referred to as “7-DMP scores”. Since the 7-DMP scores were the strongest univariate predictors, we next generated the SVM models using the 7-DMP scores as well as each of the remaining significant clinical factors as complex models. All models based on the 7-DMPs (simple model) and any of the clinical signatures (complex model) exhibited similar predictive performance. However, the complex model based on 7-DMP scores, the age at diagnosis, and LY provided the most accurate predictions, confirming that these were independent predictive factors. The Hosmer–Lemeshow test exhibited no significant deviation from perfect fit in simple or complex models ($P > 0.05$), suggesting good agreements between the estimated and actual outcomes. In the test set, we computed the AUC of the simple model (the seven-DMP signature) is 0.73, and the AUC of the complex model (the 7-DMP scores with clinical factors) is 0.74. The AUC revealed that the complex model (AUC=0.94) exhibited enhanced diagnostability, relative to the simple model (AUC=0.91) in the total set (Figure 4).

Based on our results of the NRI and IDI, both simple and complex models accurately restratified the RA patient’s response to LEF (Table 2). In summary, RA patients with reduced 7-DMPs scores, younger age at diagnosis, and negative LY experienced worse LEF responses, compared to their counterparts.

4 Discussion

This study employed Illumina 850k methylation arrays to assess DNA methylation profiles in whole blood samples from RA patients

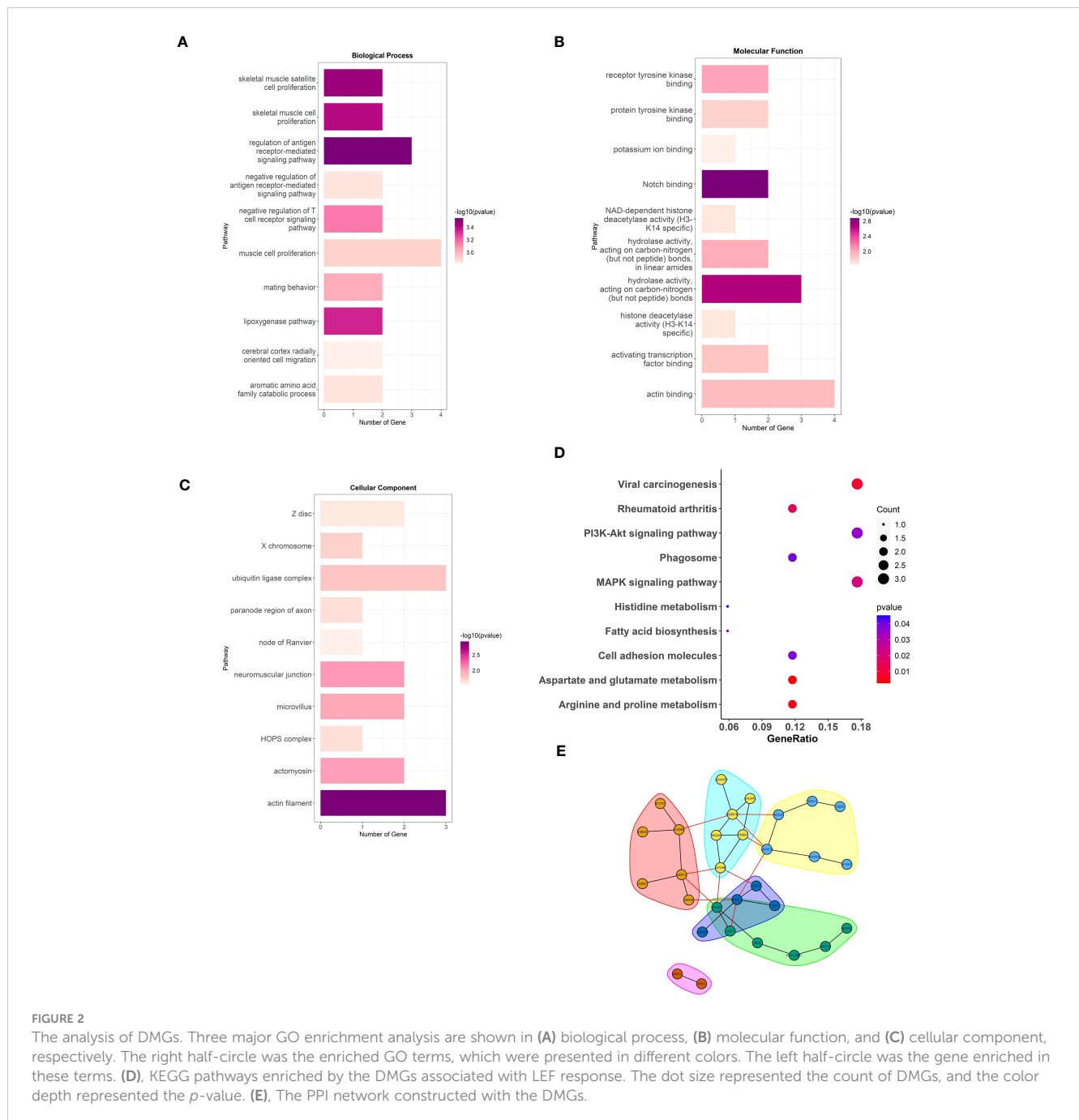


FIGURE 2

The analysis of DMGs. Three major GO enrichment analysis are shown in (A) biological process, (B) molecular function, and (C) cellular component, respectively. The right half-circle was the enriched GO terms, which were presented in different colors. The left half-circle was the gene enriched in these terms. (D), KEGG pathways enriched by the DMGs associated with LEF response. The dot size represented the count of DMGs, and the color depth represented the p -value. (E), The PPI network constructed with the DMGs.

undergoing LEF treatment. We established a DMP-based prognostic signature that incorporated cg17330251, cg19814518, cg20124410, cg21109666, cg22572476, cg23403192, and cg24432675 that were sufficient to predict RA patient response to LEF following a 6-month treatment period. To our knowledge, this is the first study that identified LEF response-related DNA methylation biomarkers using Illumina 850k methylation arrays. This investigation provides a valuable foundation for efforts to support future personalized RA patient treatment.

Recently, LEF was recommended as a second-line DMARD following MTX by EULAR owing to its marked enhancement in functional disability and health-related quality of life (6). Clinical trials involving LEF demonstrated clearly that LEF treatment

effectively reduced the signs and symptoms of disease, as well as radiographic progression (30, 31). However, LEF treatment withdrawal is a common challenge due to patient intolerance of adverse reactions, particularly upon the administration of loading doses (32). Prevention of LEF-induced undesirable effects still remains a challenging issue for the improvement of the clinical outcome of LEF-treated RA patients. As we mentioned before, in our study, DNA methylation was selected as biomarkers for predict the drug reaction because it is reversible and can be regulated by diet, drugs and other environmental factors, different from SNP and other gene mutations. Furthermore, another benefit of DNA methylation analysis is its ability to be conducted on non-invasive samples, such as blood, making it a more convenient and less

TABLE 1 the location and biological process of the validated methylated CpG sites.

CpG ID	Chromosome	Position	Gene	CpG Location	Type	Biological process	P.Value
cg17330251	7	94953956	PON1	TSS200-Island	Hypo	lipoxygenase pathway	0.030
cg24432675	10	1505472	ADARB2	Body-N_shore	Hypo	base conversion or substitution editing	0.006
cg23403192	21	30396327	USP16	TSS1500-N_Shore	Hypo	positive regulation of cellular amide metabolic process	0.039
cg19814518	1	162467080	UHMK1	TSS1500-N_Shore	Hyper	positive regulation of cellular amide metabolic process	0.002
cg21109666	1	232007841	DISC1	Body-opensea	Hyper	cerebral cortex radially oriented cell migration	0.005
cg22572476	6	28601324		IGR-N_Shore	Hyper	No found	0.002
cg20124410	13	107333224		IGR-opensea	Hyper	No found	0.008

invasive prognostic method than those requiring tissue biopsies (33).

Through analysis using an the 850k Illumina Human MethylationEPIC Bead Chip, we identified 81 DMPs. According to Manhattan maps, DMPs sites were widely distributed on non-sex chromosomes. it suggests that DNA methylation alterations are not limited to specific genomic regions or chromosomes, but rather occur globally throughout the genome in RA patients. This implies that DNA methylation may play a widespread role in the development and progression of RA, and highlights the need for a comprehensive investigation of DNA methylation patterns in this disease. We observed that methylation levels of responders were higher in the promoter region such as the first Exon and TSS1500 than the non-responders. DNA methylation in promoter regions generally suppresses transcription or serves as a marker of a silenced gene (34). This is due to the strong binding of transcription factors or the recruitment of transcriptional repressors (35) and its dysregulation plays a crucial role in oncogenesis, tumor progression, and autoimmune diseases. Therefore, CpG islands located in promoter

regions have always been of great interest to us, which is non-dynamic and less variant (36). Besides, methylation is more variable along the CpG shores, CpG shelves, and open sea (37). Therefore, it is more challenging and complicated to study methylation changes in CpG shores, CpG shelves and open sea in disease.

Based on our functional enrichment analyses of identified DMGs, we observed marked enrichments in GO terms, namely, hydrolase activity, actin filament, muscle cell proliferation, and the antigen receptor-mediated axis, all of which are closely related to RA progression (38–40). Based on the KEGG pathway analyses of these DMGs, the PI3K-Akt network was associated with key cellular functions, namely, survival, autophagy, differentiation, proliferation, and angiogenesis, and is a strong modulator of RA development and severity (41). Multiple reports suggested that MAPK signaling proteins control cellular responses to stressors and mitogenic stimuli, and that proinflammatory cytokines activate MAPK signaling in human fibroblast-like synoviocytes (FLS) in RA patients (42, 43). MAPK is therefore a critical modulator of RA progression, as well as a promising target for therapeutic intervention.

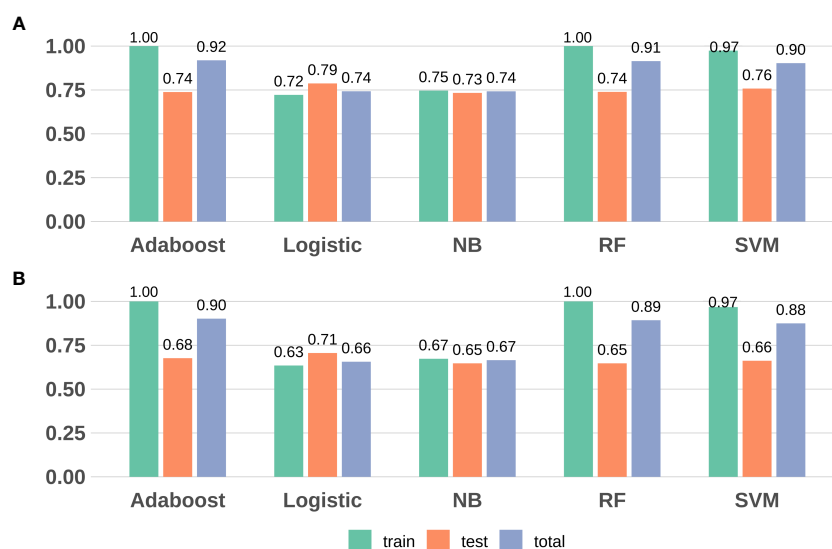


FIGURE 3

Model evaluation (F1-score and accuracy) results based on the seven DMPs across five machine learning. (A) F1-score; (B) Accuracy. RF, random forest; SVM, support vector machine; NB, naïve Bayes.

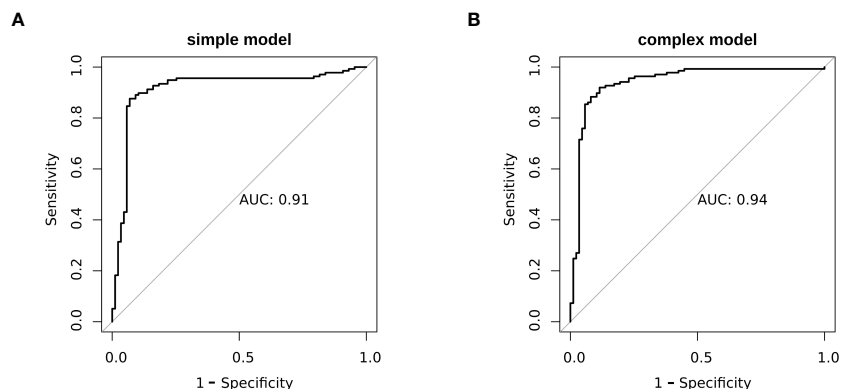


FIGURE 4
The ROC plots of models. (A) simple model (B) complex model.

RA is generally marked with enhanced heterogeneity in patient prognosis, joint damage, and therapeutic response. For a majority of these factors, the mechanistic causality is yet undetermined (44). Therefore, currently, employing highthroughput technologies to develop computational methods to process patient -omic information to identify new and precise conclusions is ever more important (45). For example, implementing machine learning algorithms in high-dimensional data analysis is a well-established approach to enhancing patient classification (46–48) or in predicting disease activity (49) in RA. Herein, we employed five machine learning approaches to analyze DNA methylation patterns, along with patient clinical information, to fine-tune the estimation of previously available stratifiers in an independent testing dataset.

The 81 DMPs were chosen for validation using the pyrosequencing, which accurately detected differential CpG sites compared to the 850k methylation array. Our study explored suitable CpG biomarkers for RA response to LEF by both 850k DNA methylation microarray and a targeted bisulfite sequencing assay. After that, the 7 DMPs were identified to be predictors for RA

response to LEF by developing prognostic models through several machine learning algorithms. Intriguingly, five of the seven DMPs were located at promoter sites or gene bodies of the annotated genes. Altered DNA methylation of genes, such as, *PON1*, *ADARB2*, *USP16*, *UHMK1*, and *DISC1*, was previously reported to be related to cancer or autoimmune diseases (Figure 5).

We identified hypomethylated cg17330251 located at the *PON1* promoter, and using GO analysis, we revealed that *PON1* was involved in the biological process of lipoxygenase pathway, which is a significant differentially modulated pathway in the GO categories. The lipoxygenase pathway may modulate eicosanoids generation, which, in turn, may impact cancer development, progression, and immune response (50). Additionally, the activation of the cyclooxygenase and lipoxygenase networks of arachidonic acid is speculated to contribute to rheumatic disease development (51). Therefore, we assumed that hypomethylated *PON1* promoters resulted in the high *PON1* mRNA expression, which, in turn enhanced the RA patient response to LEF treatment. A recent study suggested that *ADARB2* harbors the genetic MTX response variants, and demonstrated enhanced DNA methylation following MTX treatment in RA patients, compared to baseline (52). Similarly, the *ADARB2* gene body is extremely hypomethylated in HCC tumor, relative to adjacent tissues (53). This investigation suggested that the hypomethylated cg24432675 sites located at the *ADARB2* body may result in worse prognosis of LEF-treated RA patients. GO categories in both *USP16* and *UHMK1* were primarily related to the positive regulation of the cellular amide metabolic process, which may facilitate the metabolic regulation of the T cell response in RA (54). *USP16* is a critical deubiquitinase (DUB) for chromosomal segregation during mitosis (55). *USP16* also functions as a modulator of mature T cell activation, and may, therefore, serve as a new therapeutic target for T cell-mediated autoimmune disease treatment (56). In addition, *UHMK1* upregulation significantly promotes gastric cancer growth and metastasis possibly via metabolic networks (57). In contrast, impaired metabolism heavily modulates the origin and progression of autoimmune disease (58). We boldly speculated that the *USP16* and *UHMK1* may influence LEF response in RA patients via the metabolic network. *DISC1*, which

TABLE 2 Performance and internal validation of LEF nomogram.

	Simple model	Complex model
Hosmer-Lemeshow test		
χ^2	1.97	1.99
<i>P</i>	0.85	0.86
Accuracy	0.875	0.9062
F1-score	0.9028	0.9231
Recall	0.9489	0.9197
Precision	0.8609	0.9265
AUC	0.91	0.94
NRI	ref.	0.09
IDI	ref.	0.09

AUC, area under the receiver operating characteristic curve; NRI, net reclassification index; IDI, integrated discrimination improvement; ref., reference level.

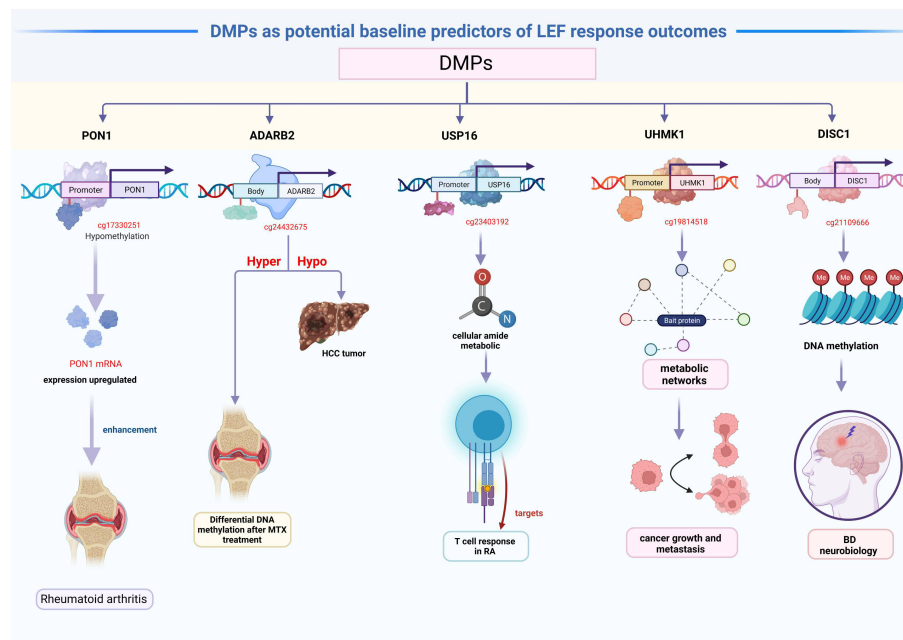


FIGURE 5
The mechanism figure of the main findings.

encodes a scaffold protein, promotes hemoglobin synthesis in the peripheral blood (59). Several studies assessed the *DISC1* contents in PBMCs of BD patients. Based on their results, the *DISC1* expression was correlated with an enhanced bipolar disorder (BD) risk, which may be due to DNA methylation, and it may be associated with BD neurobiology (60). In terms of the relevant DMPs in CD4 memory T cells, no studies till date, examined associations between their methylation status and RA patient responses to LEF treatment. Thus, further functional studies are warranted to validate these 7 DMPs role in modulating RA patients' response to LEF.

In light of the above findings, we specifically constructed a biomarker signature, based on these seven candidates CpG sites, identified by the SVM model. We demonstrated that this signature was able to readily stratify patients into low- and high-risk response groups, with a great degree of accuracy (AUC=0.91). Furthermore, when the clinical data, such as, LY and age at diagnosis were included in this model, its predictive accuracy was further enhanced (AUC=0.94). Hence, we demonstrated that the simplified models with 7-DMPs was preferable since there was no significant improvement in NRI, relative to the complex model. Moreover, there was benefit owing to the use of less independent variables, which can reduce the burden on both doctors and patients, and minimize the overwhelming cost of medical diagnosis. However, in other cases of clinical usage, a medical doctor can opt for the complex model to minimize false positives, and gain accurate prognosis.

Our findings highlighted the convenience of synergistically employing both clinical and basic research information to obtain a complete and robust patient prognosis and therapeutic evaluation. In this study, the use of a prospective design in this study may have

enabled the researchers to carefully screen and recruit participants, as well as gather data using standardized methods, thereby minimizing the likelihood of bias and confounding factors. No any other study had explored the DNA methylation as predictors for LEF-taking of RA patients, our conclusion highlights the potential of early treatment bioindicator monitoring in RA, and it raises the essential questions revolving prognosis of LEF responses. Assessing the efficiency of suitable stratification methods in the relevant context is crucial, as it can aid in predicting patient outcomes and improving therapeutic decision-making.

This study also had certain limitations. First, although our results from the 850k BeadChip array and targeted Bisulfite Sequencing Assay may be more representative and meaningful for future validation studies, further experiments *in vitro* and *in vivo* should be performed to confirm and expand upon the aforementioned findings. Second, the short time of cohort analysis and the limited number of patients with DNA methylation data. Adequate follow-up time and multi-omics data in future studies would be valuable in further elucidating the prognostic significance of LEF response in RA patients. Third, although our samples were selected from four different hospital, all the four participating study centers were Third Class A hospital which may limit the generalizability of the findings to other populations or settings. The applicability of the identified DNA methylation profiles may be limited to the specific sample of patients enrolled, and replication of the findings in an independent sample is necessary to confirm their generalizability.

In the whole, the future work includes, on the one hand, conducting prospective studies in multiple clinical centers at different levels hospitals to perform prospective validation of our predictive model. On the other hand, conducting functional

identification *in vivo* experiments at the cellular and animal models to determine the molecular mechanism underlying the results of this study

5 Conclusion

In summary, herein, we established a predictive prognosis model (the 7-DMP scores with the lymphocyte and the diagnostic age) to estimate RA patient LEF responses, and help to benefit RA patients' LEF drug choice and nonresponders' recognition. Our model showed much promise in guiding personalized patient treatment.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: PRJNA946946 (SRA) and GSE228104 (GEO).

Ethics statement

This investigation followed the Declaration of Helsinki, and received ethical approval from the Medical Science Research Institute of the First Affiliated Hospital, China Medical University, with all patients providing written informed consent (approval number: AF-SOP-07-1.0-01). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

LF and YC conceived of the study. The data analysis was completed by YC, QW, HL, LJ, XF, BD, MC, FX, TW, BB and ZF, JL, YY, RL, JZ, XJ and LF helped to interpret the findings. The manuscript was written by LF and YC and content expertise was provided by all authors. All authors contributed to the article and approved the submitted version.

References

- Smolen JS, Aletaha D, Barton A, Burmester GR, Emery P, Firestein GS, et al. Rheumatoid arthritis. *Nat Rev Dis Primers* (2018) 4:18001. doi: 10.1038/nrdp.2018.1
- McInnes IB, Schett G. The pathogenesis of rheumatoid arthritis. *N Engl J Med* (2011) 365:2205–19. doi: 10.1056/NEJMra1004965
- McInnes IB, Schett G. Pathogenetic insights from the treatment of rheumatoid arthritis. *Lancet* (2017) 389:2328–37. doi: 10.1016/S0140-6736(17)31472-1
- Bai B, Chen M, Fu L, Liu H, Jin L, Wei T, et al. Quality of life and influencing factors of patients with rheumatoid arthritis in Northeast China. *Health Qual Life Outcomes* (2020) 18(1):119. doi: 10.1186/s12955-020-01355-7
- He J, Ding Y, Feng M, Guo J, Sun X, Zhao J, et al. Characteristics of Sjögren's syndrome in rheumatoid arthritis. *Rheumatol (Oxford)* (2013) 52:1084–9. doi: 10.1093/rheumatology/kes374
- Smolen JS, Landewé RBM, Bijlsma JWJ, Burmester GR, Dougados M, Kerschbaumer A, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2019 update. *Ann Rheum Dis* (2020) 79:685–99. doi: 10.1136/annrheumdis-2019-216655
- Tian X, Wang Q, Li M, Zhao L, Zhang Z, Huang C, et al. 2018 Chinese guidelines for the diagnosis and treatment of rheumatoid arthritis. *Rheumatol Immunol Res* (2021) 2:1–14. doi: 10.2478/rir-2021-0002

Funding

This work was supported by the Program of the National Natural Science Foundation of China (82173604).

Acknowledgments

We thank Dr. Jianming Zeng (University of Macau) and all the members of his bioinformatics team, biotrainee, for generously sharing their experience and codes. We are grateful to all of the patients who participated in this study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2023.1173187/full#supplementary-material>

SUPPLEMENTARY FIGURE 1

(A) SVD plot of batch effects. (B) Heatmaps showing the top 81 CpGs from responders versus nonresponders to LEF. (C) Manhattan plots showing the $-\log_{10}$ -transformed P values for CpG sites in the DNA methylation profiling for response to LEF.

SUPPLEMENTARY FIGURE 2

The relationship of the methylation values between from MethylTarget and the EPIC array.

8. Schultz M, Keeling SO, Katz SJ, Maksymowych WP, Eurich DT, Hall JJ. Clinical effectiveness and safety of leflunomide in inflammatory arthritis: A report from the RAPPOR database with supporting patient survey. *Clin Rheumatol* (2017) 36:1471–78. doi: 10.1007/s10067-017-3687-5
9. Alldred A, Emery P. Leflunomide: a novel DMARD for the treatment of rheumatoid arthritis. *Expert Opin Pharmacother* (2001) 2:125–37. doi: 10.1517/14656566.2.1.125
10. Rozman B. Clinical pharmacokinetics of leflunomide. *Clin Pharmacokinet* (2002) 41:421–30. doi: 10.2165/00003088-200241060-00003
11. Behrens F, Koehm M, Burkhardt H. Update 2011: Leflunomide in rheumatoid arthritis - strengths and weaknesses. *Curr Opin Rheumatol* (2011) 23:282–7. doi: 10.1097/BOR.0b013e328344fddb
12. Liu S, Guo Y, Lu L, Lu J, Ke M, Xu T, et al. Fibrinogen-like protein 1 is a novel biomarker for predicting disease activity and prognosis of rheumatoid arthritis. *Front Immunol* (2020) 11:579228. doi: 10.3389/fimmu.2020.579228
13. Glossop JR, Nixon NB, Emes RD, Sim J, Packham JC, Matthey DL, et al. DNA methylation at diagnosis is associated with response to disease-modifying drugs in early rheumatoid arthritis. *Epigenomics* (2017) 9:419–28. doi: 10.2217/epi-2016-0042
14. Nair N, Plant D, Verstappen SM, Isaacs JD, Morgan AW, Hyrich KL, et al. Differential DNA methylation correlates with response to methotrexate in rheumatoid arthritis. *Rheumatol (Oxford)* (2020) 59:1364–71. doi: 10.1093/rheumatology/kez411
15. Gosselt HR, van Zelst BD, de Rotte MCFJ, Hazes JMW, de Jonge R, Heil SG. Higher baseline global leukocyte DNA methylation is associated with MTX non-response in early RA patients. *Arthritis Res Ther* (2019) 21:157. doi: 10.1186/s13075-019-1936-5
16. Plant D, Webster A, Nair N, Oliver J, Smith SL, Eyre S, et al. Differential methylation as a biomarker of response to etanercept in patients with rheumatoid arthritis. *Arthritis Rheumatol* (2016) 68:1353–60. doi: 10.1002/art.39590
17. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum* (2010) 62:2569–81. doi: 10.1002/art.27584
18. Maddison P, Kiely P, Kirkham B, Lawson T, Moots R, Prodfoot D, et al. Leflunomide in rheumatoid arthritis: recommendations through a process of consensus. *Rheumatol (Oxford)* (2005) 44:280–6. doi: 10.1093/rheumatology/keh500
19. Kalden JR, Schattenkirchner M, Sörensen H, Emery P, Deighton C, Rozman B, et al. The efficacy and safety of leflunomide in patients with active rheumatoid arthritis: A five-year followup study. *Arthritis Rheum* (2003) 48:1513–20. doi: 10.1002/art.11015
20. Gestel AM, Prevoo ML, Hof MA, van Rijswijk MH, van de Putte B, van Riel PL. Development and validation of the european league against rheumatism response criteria for rheumatoid arthritis. Comparison with the preliminary american college of rheumatology and the world health organization/international league against rheumatism criteria. *Arthritis Rheum* (1996) 39:34–40. doi: 10.1002/art.1780390105
21. Gestel AM, Haagsma CJ, Riel PL. Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. *Arthritis Rheum* (1998) 41:1845–50. doi: 10.1002/1529-0131(199810)41:10<1845::AID-ART17>3.0.CO;2-K
22. Yu G, Wang LG, Han Y, He QY. clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* (2012) 16:284–7. doi: 10.1089/omi.2011.0118
23. Morris TJ, Butcher LM, Feber A, Techendorff AE, Chakravathy AR, Wojdacz TK, et al. ChAMP: 450k chip analysis methylation pipeline. *Bioinformatics* (2014) 30:428–30. doi: 10.1093/bioinformatics/btt684
24. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: A flexible and comprehensive Bioconductor package for the analysis of Illumina DNA methylation microarrays. *Bioinformatics* (2014) 30:1363–9. doi: 10.1093/bioinformatics/btu049
25. Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics* (2013) 29:189–96. doi: 10.1093/bioinformatics/bts680
26. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* (2016) 32:2847–9. doi: 10.1093/bioinformatics/btw313
27. Kuhn M, Wing J, Weston S, Williams S, Keefer C, Engelhardt A, et al. Caret: Classification and Regression Training. *R Package Version 6.0-41* (2015). Available at: <https://CRAN.R-project.org/package=caret>.
28. Pencina MJ, D'Agostino RBSr, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med* (2011) 30(1):11–21. doi: 10.1002/sim.4085
29. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* (2008) 27(2):157–212. doi: 10.1002/sim.2929
30. Alfaro-Lara R, Espinosa-Ortega HF, Arce-Salinas CA. Systematic review and meta-analysis of the efficacy and safety of leflunomide and methotrexate in the treatment of rheumatoid arthritis. *Rheumatol Clinica* (2019) 15:133–9. doi: 10.1016/j.reuma.2017.07.020
31. Wang W, Zhou H, Liu L. The role of Chinese herbal medicine in the management of adverse drug reactions of leflunomide in treating rheumatoid arthritis. *Phytomedicine* (2020) 68:153136. doi: 10.1016/j.phymed.2019.153136
32. Cutolo M, Boloşiu H, Perdriset G. Efficacy and safety of leflunomide in DMARD-naive patients with early rheumatoid arthritis: comparison of a loading and a fixed-dose regimen. *Rheumatology* (2013) 52:1132–40. doi: 10.1093/rheumatology/kes321
33. Das PM, Singal R. DNA methylation and cancer. *J Clin Oncol* (2004) 22:4632–42. doi: 10.1200/JCO.2004.07.151
34. Jones PA, Baylin SB. The epigenomics of cancer. *Cell* (2007) 128:683–92. doi: 10.1016/j.cell.2007.01.029
35. Marchal C, Benoit M. Emerging concept in DNA methylation: role of transcription factors in shaping DNA methylation patterns. *J Cell Physiol* (2015) 230:743–51. doi: 10.1002/jcp.24836
36. Ziller MJ, Gu H, Müller F, Donaghey J, Tsai LT, Kohlbacher O, et al. Charting a dynamic DNA methylation landscape of the human genome. *Nature* (2013) 500:477–81. doi: 10.1038/nature12433
37. Trino S, Zoppioli P, Carella AM, Laurenzana I, Weisz A, Memoli D, et al. DNA methylation dynamic of bone marrow hematopoietic stem cells after allogeneic transplantation. *Stem Cell Res Ther* (2019) 10:138. doi: 10.1186/s13287-019-1245-6
38. Ling Y, Xu H, Ren N, Cheng C, Zeng P, Lu D, et al. Prediction and verification of the major ingredients and molecular targets of *Tripterygii radix* against rheumatoid arthritis. *Front Pharmacol* (2021) 12:639382. doi: 10.3389/fphar.2021.639382
39. Osborn TM, Verdrengh M, Stossel TP, Tarkowski A, Bokarewa M. Decreased levels of the gelsolin plasma isoform in patients with rheumatoid arthritis. *Arthritis Res Ther* (2008) 10:R117. doi: 10.1186/ar2520
40. Mun S, Lee J, Lim MK, Lee YR, Ihm C, Lee SH, et al. Development of a novel diagnostic biomarker set for rheumatoid arthritis using a proteomics approach. *BioMed Res Int* (2018) 2018:7490723. doi: 10.1155/2018/7490723
41. Malemud CJ. The PI3K/Akt/PTEN/mTOR pathway: A fruitful target for inducing cell death in rheumatoid arthritis? *Future Med Chem* (2015) 7:1137–47. doi: 10.4155/fmc.15.55
42. Ralph JA, Morand EF. MAPK phosphatases as novel targets for rheumatoid arthritis. *Expert Opin Ther Targets* (2008) 12:795–808. doi: 10.1517/14728222.12.7.795
43. Pakozdi A, Amin MA, Haas CS, Martinez RJ, Haines KG, Santos LL, et al. Macrophage migration inhibitory factor: A mediator of matrix metalloproteinase-2 production in rheumatoid arthritis. *Arthritis Res Ther* (2006) 8:R132. doi: 10.1186/ar2021
44. Pitzalis C, Kelly S, Humby F. New learnings on the pathophysiology of RA from synovial biopsies. *Curr Opin Rheumatol* (2013) 25:334–44. doi: 10.1097/BOR.0b013e32835fd8eb
45. Stafford IS, Kellermann M, Mossotto E, Beattie RM, McArthur BD, Ennis S. A systematic review of the applications of artificial intelligence and machine learning in autoimmune diseases. *NPJ Digit Med* (2020) 3:30. doi: 10.1038/s41746-020-0229-3
46. Guan Y, Zhang H, Quang D, Wang Z, Parker SC, Pappas DA, et al. Machine learning to predict anti-tumor necrosis factor drug responses of rheumatoid arthritis patients by integrating clinical and genetic markers. *Arthritis Rheumatol* (2019) 71:1987–96. doi: 10.1002/art.41056
47. Vodencarevic A, Tascilar K, Hartmann F, Reiser M, Hueber AJ, Haschka J, et al. Advanced machine learning for predicting individual risk of flares in rheumatoid arthritis patients tapering biologic drugs. *Arthritis Res Ther* (2021) 23:1–8. doi: 10.1186/s13075-021-02439-5
48. Mo X, Chen X, Li H, Li J, Zeng F, Chen Y, et al. Early and accurate prediction of clinical response to methotrexate treatment in juvenile idiopathic arthritis using machine learning. *Front Pharmacol* (2019) 10:1155. doi: 10.3389/fphar.2019.01155
49. Lin C, Karlson EW, Canhao H, Miller TA, Dligach D, Chen PJ, et al. Automatic prediction of rheumatoid arthritis disease activity from the electronic medical records. *PLoS One* (2013) 8:e69932. doi: 10.1371/journal.pone.0069932
50. Pidgeon GP, Lysaght J, Krishnamoorthy S, Reynolds JV, O'Byrne K, Nie D, et al. Lipoxigenase metabolism: Roles in tumor progression and survival. *Cancer Metastasis Rev* (2007) 26:503–24. doi: 10.1007/s10555-007-9098-3
51. Mishima S, Kashiwakura J, Toyoshima S, Sasaki-Sakamoto T, Sano Y, Nakanishi K, et al. Higher PGD₂ production by synovial mast cells from rheumatoid arthritis patients compared with osteoarthritis patients via miR-199a-3p/prostaglandin synthetase 2 axis. *Sci Rep* (2021) 11:5738. doi: 10.1038/s41598-021-84963-7
52. Guderud K, Sunde LH, Flåm ST, Meahlen MT, Mjaavatten MD, Norli ES, et al. Methotrexate treatment of newly diagnosed RA patients is associated with DNA methylation differences at genes relevant for disease pathogenesis and pharmacological action. *Front Immunol* (2021) 12:713611. doi: 10.3389/fimmu.2021.713611
53. Gentilini D, Scala S, Gaudenzi G, Garagnani P, Capri M, Cescon M, et al. Epigenome-wide association study in hepatocellular carcinoma: Identification of

- stochastic epigenetic mutations through an innovative statistical approach. *Oncotarget* (2017) 8:41890. doi: 10.18632/oncotarget.17462
54. Colamatteo A, Micillo T, Bruzzaniti S, Fusco C, Garavelli S, De Rosa V, et al. Metabolism and autoimmune responses: The microRNA connection. *Front Immunol* (2019) 10:1969. doi: 10.3389/fimmu.2019.01969
55. Zhuo X, Guo X, Zhang X, Jing G, Wang Y, Chen Q, et al. Usp16 regulates kinetochore localization of Plk1 to promote proper chromosome alignment in mitosis. *J Cell Biol* (2015) 210:727–35. doi: 10.1083/jcb.201502044
56. Zhang Y, Liu R, Cao Q, Fan KQ, Huang LJ, Yu JS, et al. USP16-mediated deubiquitination of calcineurin A controls peripheral T cell maintenance. *J Clin Invest* (2019) 129:2856–71. doi: 10.1172/JCI123801
57. Feng X, Ma D, Zhao J, Song Y, Zhu Y, Zhou Q, et al. UHMK 1 promotes gastric cancer progression through reprogramming nucleotide metabolism. *EMBO J* (2020) 39: e102541. doi: 10.15252/embj.2019102541
58. Galgani M, Bruzzaniti S, Matarese G. Immunometabolism and autoimmunity. *Curr Opin Immunol* (2020) 67:10–7. doi: 10.1016/j.coi.2020.07.002
59. Ding SJ, Li ZD, Zeng T, Zhang YH, Huang T, Cai YD. Identifying infliximab-(IFX-) responsive blood signatures for the treatment of rheumatoid arthritis. *BioMed Res Int* (2021) 2021:5556784. doi: 10.1155/2021/5556784
60. Shirvani-Farsani Z, Maloum Z, Bagheri-Hosseiniabadi Z, Vilor-Tejedor N, Sadeghi I. DNA methylation signature as a biomarker of major neuropsychiatric disorders. *J Psychiatr Res* (2021) 141:34–49. doi: 10.1016/j.jpsychires.2021.06.013