The Journal of Molecular Diagnostics, Vol. 26, No. 10, October 2024

ELSEVIER

the Journal of Molecular Diagnostics
jmdjournal.org

A Network-Based Framework to Discover Treatment-Response–Predicting Biomarkers for Complex Diseases

Uday S. Shanthamallu,* Casey Kilpatrick,† Alex Jones,* Jonathan Rubin,‡ Alif Saleh,* Albert-László Barabási,§¶‖ Viatcheslav R. Akmaev,* and Susan D. Ghiassian*

From the Departments of Data Science and Network Medicine and Therapeutics,† Scipher Medicine, Waltham, Massachusetts; the Maret School,‡ Washington, DC; the Center for Complex Network Research,§ Northeastern University, Boston, Massachusetts; the Channing Division of Network Medicine,¶ Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts; and the Department of Network and Data Science,‖ Central European University, Budapest, Hungary*

Accepted for publication June 26, 2024.

Address correspondence to Susan D. Ghiassian, Ph.D., Scipher Medicine, 500 Totten Pond Rd., Waltham, MA 02451
E-mail: dina.ghiassian@scipher.com.

The potential of precision medicine to transform complex autoimmune disease treatment is often challenged by limited data availability and inadequate sample size when compared with the number of molecular features found in high-throughput multi-omics data sets. To address this issue, the novel framework PRoBeNet (Predictive Response Biomarkers using Network medicine) was developed. PRoBeNet operates under the hypothesis that the therapeutic effect of a drug propagates through a protein-protein interaction network to reverse disease states. PRoBeNet prioritizes biomarkers by considering i) therapy-targeted proteins, ii) disease-specific molecular signatures, and iii) an underlying network of interactions among cellular components (the human interactome). PRoBeNet helped discover biomarkers predicting patient responses to both an established autoimmune therapy (infliximab) and an investigational compound (a mitogen-activated protein kinase 3/1 inhibitor). The predictive power of PRoBeNet biomarkers was validated with retrospective gene-expression data from patients with ulcerative colitis and rheumatoid arthritis and prospective data from tissues from patients with ulcerative colitis and Crohn disease. Machine-learning models using PRoBeNet biomarkers significantly outperformed models using either all genes or randomly selected genes, especially when data were limited. These results illustrate the value of PRoBeNet in reducing features and for constructing robust machine-learning models when data are limited. PRoBeNet may be used to develop companion and complementary diagnostic assays, which may help stratify suitable patient subgroups in clinical trials and improve patient outcomes. *(J Mol Diagn 2024, 26: 917–930; https://doi.org/10.1016/j.jmoldx.2024.06.008)*

A key promise of precision medicine is the ability to match patient subgroups with the most appropriate treatments.[1] This is achieved by discovering biomarkers that connect a patient's biological status with therapeutic outcomes for a specific therapy. In precision medicine, biomarkers can be discovered using machine-learning models that unveil complex, generalizable patterns from large molecular and clinical data sets, usually comprising data from hundreds to thousands of patients. For example, analyzing extensive molecular and clinical data sets from patients with cancer, machine-learning models found biomarkers that predict response to treatment in patients with diverse cancers.[2–8] These models have substantially improved outcomes and survival rates for many cancer subtypes and greatly reduced the financial burden on health care payers.[9,10]

Biomarker discovery and validation by using a traditional machine-learning approach requires relatively large and relevant data sets. These data sets must include both high-quality molecular data collected from patients before initiation of therapy and matched clinical outcome data

Supported by Scipher Medicine Corp.

https://doi.org/10.1016/j.jmoldx.2024.06.008

indicating which patients ultimately responded to treatment and which did not. Traditional machine-learning models are not generalizable when insufficient data are available. A main challenge in modeling insufficient data is the curse of dimensionality: the amount of data needed to generalize from machine-learning results increases exponentially as the number of features increase. In such cases, machine-learning models tend to overfit the training data: the models find biomarkers that seem to be predictive for patients in the training cohort, but are not predictive for patients in other cohorts.[11–15] Important examples are chronic autoimmune diseases [ie, rheumatoid arthritis (RA)] and ulcerative colitis [UC; an immune-mediated inflammatory disease (IMID)].[11,12,16–22] The volume of publicly available molecular data from patients with complex autoimmune disease is orders of magnitude less than that from patients with cancer. [As of June 2023, the number of cancer data sets in the Gene Expression Omnibus (*https://www.ncbi.nlm.nih.gov/geo*) was over one million; the number of autoimmune disease data sets, only 27,000.[23]] This occurs partly because autoimmune diseases cause fewer deaths each year than does cancer[24]; as a result, research pertaining to autoimmune disease garners less attention and funding than does cancer research. According to the NIH RePORTER website, the NIH invested approximately $3.4 billion for cancer research in 2022, but only $380 million for autoimmune research. Thus, using a traditional machine-learning approach to find generalizable treatment-response patterns that predict biomarkers for existing and novel autoimmune therapies can be challenging.[25]

A network-based framework, PRoBeNet (Predictive Response Biomarkers using Network medicine), that radically reduces the pool of candidate biomarkers (features) is introduced, enabling machine-learning models to be successfully trained on cohorts with relatively few samples. The premise behind PRoBeNet is that the effect of a treatment (a drug targeting a specific protein) will ripple through a cascade of protein-protein interactions; reach the disease signature (comprising a panel of proteins characteristically dysregulated in that disease); and revert signature proteins to normal (in responding patients). Therefore, PRoBeNet uses as input the treatment-targeted protein, defined by a drug's mechanism of action, and the disease-signature proteins, defined by the biological effects of the disease. For example, to find biomarkers predicting responses to infliximab of patients with RA, PRoBeNet would consider as input the infliximab-targeted protein [tumor necrosis factor (TNF)] and the appropriate disease-signature cytokine proteins, such as IL-4, IL-6, and IL-10. Abnormal expression of these cytokines is detectable in serum of patients with RA years before onset of symptoms. This abnormal expression becomes increasingly prominent as the disease progresses[26,27] and is usually targeted by RA treatments.[28]

The proposed network-based framework successfully identified response-predicting biomarkers for both an approved autoimmune disease treatment and a new investigational compound. Although PRoBeNet was developed in the context of autoimmune diseases, it can be generalized to other complex diseases if the treatment-targeted proteins and appropriate disease signature molecules are known.

Key Points

- A novel network-based system biology framework, PRoBeNet (Predictive Response Biomarkers using Network medicine), was developed to predict treatment-response biomarkers in complex autoimmune diseases without relying on extensive patient data sets.
- PRoBeNet successfully identified and validated response-predicting biomarkers retrospectively for response to infliximab and prospectively for response to mitogen-activated protein kinase 3/1 inhibitors.
- The nodes in the human interactome are ranked by PRoBeNet framework by using a dual PageRank score, which considers proteins crucial to therapy's mechanism of action and disease pathogenesis.
- PRoBeNet's ability to identify response biomarkers is valuable for developing companion diagnostic tests to amplify drug efficacies and optimize personalized treatment strategies.

Materials and Methods

Human Interactome

The human interactome (HI) was assembled by compiling experimentally validated protein-protein interactions from 21 public databases, as described.[19,29] The HI is a large network of proteins and their interactions. It represents biological processes and signaling pathways in human cells. The HI comprises mostly proteins, but also include some genes. For simplicity, only proteins (and their physical interactions) are mentioned here. For each disease, a tissue-specific subgraph was constructed by extracting the largest connected component of the HI. Each tissue-specific subgraph had only proteins and gene products expressed in the specific tissue where data were available. No restrictions were imposed on the network structure, and interactions remained without direction (as in the original HI).

Personalized PageRank

The Personalized PageRank (PPR) algorithm is a variant of the conventional PageRank algorithm. The PPR algorithm provides a user-centric view of node importance in a network, where the user's interests are represented by a specific set of personalization nodes. In the HI, the algorithm operates by simulating a random walk through the network, with a bias toward personalization nodes. Starting

from these nodes, the random walk follows the edges to adjacent nodes, thereby propagating the PageRank scores. The underlying equation governing the propagation of scores in PPR is as follows:

$$PR(v) = \frac{1-d}{N} + d \cdot \sum \frac{PR(u)}{L(u)} \cdot \delta(u,v) \quad (1)$$

Where $PR(v)$ is the PageRank of node v; N is the number of nodes in the interactome; d is a damping factor set to 0.85; $PR(u)$ is the PageRank of a node u, which is connected to node v; $L(u)$ is the number of edges connected to node u; and $\delta(u,v)$ is a function that is 1 if node u is a personalization node, and 0 otherwise.

The personalization probabilities are equally distributed among all personalization nodes (ie, 1/|T| for drug-target proteins and 1/|C| for disease-specific cytokines, where T is the set of drug targets and C is the set of cytokines).

The NetworkX library version 2.6.3 (*https://pypi.org/project/networkx/2.6.3*) was used to implement the PPR algorithm. The PPR calculation was performed iteratively either until the scores converged or after 10,000 iterations. After this process, the algorithm yields a score for each node in the graph based on its relative importance to the specified personalization nodes. This process is run twice: first with drug-target proteins as personalization nodes, and second with disease-specific cytokines as personalization nodes. The nodes are ranked on the basis of the two PageRank scores, and these two ranks are merged into a single combined rank by using rank product.

Logistic Regression and Validation

The 100 top biomarkers were selected from each ranked list and used as features to train a classifier. The scikit-learn package was used to train and validate by using an L1 regularized logistic regression model.[30] Logistic regression classifier with L1 regularization can generate sparse models by encouraging some feature weights to be 0. This effectively eliminates irrelevant or less-important features from the model. By reducing the number of features used in the model, L1 regularization simplifies the model and helps with its interpretation. L1-regularized models are useful for identifying important features, especially when the feature space is large. The only limitation of using L1-regularized models is that these models may be sensitive to outliers in the data. Nested cross-validation was performed to estimate model performance in an unbiased way. In the outer loop, stratified shuffle split was used to split the data: 80% of the data were used for training and for optimizing hyperparameters; and 20%, for evaluating model performance. In the inner loop, a fivefold stratified split with grid search was used to select the best hyperparameters. The hyperparameter optimization and fivefold inner-loop cross-validation ensures that only important and relevant features are selected. This process was repeated 50 times, and the average of these six performance metrics was reported: i) The area under the receiver operating characteristic curve (AUCROC) reflects the model's ability to distinguish between classes, with higher values indicating better discriminative power. ii) The area under the precision-recall curve and average precision are especially relevant for imbalanced data, indicating the model's accuracy in identifying true positives. The area under the precision-recall curve represents the model's ability to return relevant instances, with higher areas indicating better performance. iii) Average precision aggregates the precision-recall trade-off across different thresholds, summarizing the curve into a single value for easy comparison. iv) Accuracy provides a general success rate. v) Balanced accuracy ensures equal representation of all classes. vi) F1 score offers a harmonious balance between precision and recall by combining these two into a single metric.

Cohort with UC for Retrospective Validation

Two publicly available data sets associated with UC were retreived from Gene Expression Omnibus (*http://www.ncbi.nlm.nih.gov/geo*, last accessed June 25, 2024, accession numbers GSE14580 and GSE12251).[23] In the first cohort (GSE14580), 24 patients (14 men and 10 women) with active UC, who were refractory to corticosteroids and/or immunosuppression, underwent colonoscopy with biopsies from diseased colon tissues at least 1 week before receiving their first i.v. infusion of infliximab (5 mg/kg body weight). The median age of responders at time of first infusion was 28.4 years; and of nonresponders, 45.8 years. Response to infliximab was defined as endoscopic and histologic healing 4 to 6 weeks after the first infusion. In the second cohort (GSE12251), 22 patients (10 men and 12 women) underwent colonoscopy with biopsy before infliximab infusion (two samples were derived from the same patient). The median age of responders at time of first infusion was 39.0 years; and of nonresponders, 51.5 years. Response to infliximab was defined as endoscopic and histologic healing after 8 weeks. For both cohorts, total RNA was isolated from colonic mucosal biopsies, labeled, and hybridized to Affymetrix Human Genome U133 Plus 2.0 Arrays (ThermoFisher, Santa Clara, CA). Batch effects among studies were corrected by using ComBat version 3.44.0 (*https://rdrr.io/bioc/sva*). The combined cohort consisted of 20 responders and 26 nonresponders. Patients with UC in these cohorts shared similar demographic characteristics (ie, age, weight, and duration of disease).[31]

Cohort with RA for Retrospective Validation

Patients with a clinical diagnosis of RA were included in the molecular signature response classifier test arm.[32] Patients were at least 18 years of age and had never received TNF-inhibitor (TNFi) treatment ($N = 107$). However, methotrexate is usually the first line of treatment prescribed for RA (a common disease-modifying antirheumatic drug). Approximately one-third of patients in the cohort with RA

had taken methotrexate in the past year. A few patients had taken prednisone, which is a corticosteroid. Patient data were collected in the Study to Accelerate Information of Molecular Signatures between August 2020 and August 2022.[32] All patients in this study had moderate-to-high baseline clinical disease activity scores (>10). All patients were treated with TNFi after the study began. Patients included were those for whom clinical disease activity data were available from both baseline and 6-month follow-up visits (±4 weeks). American College of Rheumatology (ACR) scores were calculated for each patient to determine response to therapy. ACR scores are subjective scores used to evaluate how well patients with RA respond to treatment. They are derived from ACR criteria and represent percentage improvement in these symptoms: the number of swollen or tender joints; patient and physician evaluations of pain and overall health (made by using the health assessment questionnaire disability index); and serologic and blood-marker levels (ie, C-reactive protein, erythrocyte sedimentation rate).[30,33,34] For example, the ACR20 score indicates a 20% improvement in the number of tender and swollen joints; the ACR50, a 50% improvement; and the ACR70, a 70% improvement. As determined by 6-month ACR50 scores, the cohort with RA ($N = 107$) comprised more nonresponders ($N = 82$) than responders ($N = 25$), consistent with insufficient infliximab response being common among patients with RA. Most patients were women (78%, 84 of 107), as in the general population with RA. Overall, these demographics suggested the cohort well represented the general population of patients with RA.

Tissue Collection of Patients with UC and CD for Prospective Validation

Colon or ileum tissue sections were obtained from four patients with UC and eight patients with Crohn disease (CD) undergoing therapeutic resection surgery. Tissues were maintained in tissue solution until use. The final cohort GSE261205 (*http://www.ncbi.nlm.nih.gov/geo*, last accessed June 25, 2024, accession number GSE261205) consisted of seven women and five men. The patients in this cohort were previously treated with a variety of disease-modifying antirheumatic drugs, such as azathioprine, infliximab (Remicade), and albuterol. Because of the small sample size, inclusion criteria with respect to previous therapy exposure were not imposed.

Ex Vivo Cultures

Biopsy tissues (5 mm^2) were cultured in triplicate in 12-well plates containing a modified Connaught Medical Research Laboratory–based cell-culture medium. Tissues were positioned such that the apical mucosal surfaces faced the liquid-air interface. Culture plates were incubated in medium (37°C, 5% CO_2) with either test compounds or vehicle. Tissues were treated with *Staphylococcus* enterotoxin B (List Biological Laboratories, Campbell, CA) to stimulate inflammatory cytokine secretion. Cell-culture supernatants (approximately 1 mL) were collected and flash frozen 18 hours after treatment. Two samples from each condition were stabilized in RNAlater (Invitrogen, Carlsbad, CA) and frozen.

Cytokine Analysis

Disease-signature cytokines were selected for analyses if their dysregulation associates with either RA or UC, or if they are generally important for inflammation.[35–38] For patient-derived tissues of patients with UC and CD, supernatant-cytokine levels were analyzed with the Luminex 200 platform by using Miliplex MAP kits (Merck Millipore, Burlington, MA), according to manufacturer's instructions. These 15 cytokines were measured: colony stimulating factor 2 (CSF2), interferon gamma (IFNG), IL-10, IL-12A/B, IL-13, IL-17A, IL-1B, IL-2, IL-4, IL-5, IL-6, IL-7, IL-8, chemokine (C-C motif) ligand 2, and TNF. All measurements were made in duplicate.

RNA Extraction and Sequencing

RNA was extracted from homogenized tissues in lysis buffer with a bead-mill homogenizer and isolated with a spin-column chromatography kit. A method optimized for gut tissue was used. To summarize, tissue lysates were treated with proteinase K, lysates were centrifuged, and ethanol was added to supernatants. Solutions were loaded onto spin columns, and samples on columns were digested with DNase. Columns were then iteratively washed, and total RNA was eluted with nuclease-free water. RNA quality was assessed: sample concentrations were confirmed to be ≥20 ng/μL (≥0.4 μg yield in ≥20 μL), and RNA integrity numbers were confirmed to be ≥6.8 (with flat baselines).

Paired-end reads (150 nucleotides long) were mapped to the human genome (GRCh38) with STAR alignment software version 2.7.3a.[39,40] Per gene abundance in fragments per kilobase of transcript per million mapped reads was calculated with the RSEM software package version 1.3.1.[41]

Cytokine-Response Score

The cytokine-response score for a donor d was defined as the following weighted sum:

$$S(d) = \frac{1}{n_c}\sum_c p(c)\,\delta_c \tag{2}$$

Here, $p(c)$ is the average percentage change in the level of the cytokine c after treatment with respect to stimulated controls,

$$p = \frac{\bar{I}_{treatment}}{\bar{I}_{control}} - 1 \tag{3}$$

n_c is the number of cytokines for which this percentage can be computed (ie, there are no indeterminate values); and δ_c is the indicator function that is 1 if cytokine c is anti-inflammatory (expected to increase when inflammation decreases) and -1 if proinflammatory (expected to decrease when inflammation decreases). A given donor d was called a responder if that donor's cytokine score $S(d)$ met or exceeded a predefined threshold, S_{cutoff} (20%).

Statistical Analysis

A two-tailed t-test from stats library in scipy package version 1.5.4[42] was performed to assess significance of differences in model performances (for each of six performance metrics) between the PRoBeNet biomarkers, all genes, and 100 randomly selected gene sets. The α was set at 0.05. For comparing different rankings of genes obtained by using various measures of centrality, Kendall τ statistic (a nonparametric approach) was used.

Ethical Approval and Consent to Participate

All procedures involving human participants in this study were performed in accordance with the local ethical standards, as defined by the relevant state, regional, or national regulatory body, and in line with the aims and principles of the 1964 Declaration of Helsinki and its later amendments. Informed written consent was obtained from all individual participants involved in the study. Approvals were granted by the West of Scotland Research Ethics Committee (22/WS/007) and by the Institutional Review Board responsible for each tissue collection site.

Overview of the Biomarker-Discovery Framework

A novel network-based framework—PRoBeNet—for discovering treatment-response biomarkers relies on protein interactions integrated by the HI. In the HI, nodes ($N = 18{,}627$) represent proteins, and edges ($N = 439{,}260$) represent experimentally validated binding interactions among these proteins (see *Materials and Methods*). To find biomarkers for autoimmune disease (RA) and IMIDs (UC, CD), a subgraph of the HI specific to each disease was constructed; thus, each subgraph comprised protein products of genes expressed in specific tissues where data were available (whole blood for RA, colon tissue for UC and CD; see *Materials and Methods*).[43]

Next, two key protein sets in each HI subgraph were considered: protein(s) targeted by the therapeutic drug (ie, TNF, targeted by infliximab), called source node(s); and the repertoire of disease-signature cytokines, called readout nodes.[27] Readout nodes were objectively selected by either of these criteria: being generally important for inflammation or being characteristically up-regulated or down-

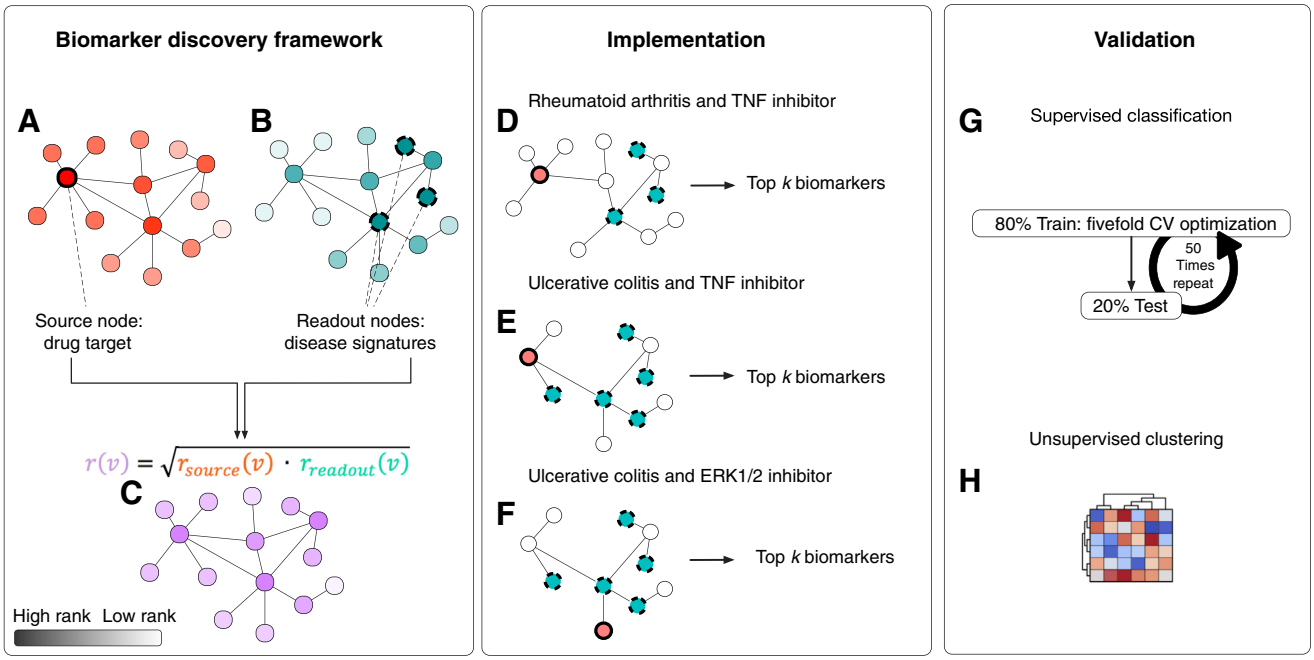

Figure 1 Overview, implementation, and validation of the biomarker-discovery framework. **A:** Schematic representation of the human interactome (HI) after running the personalized PageRank algorithm starting from the drug target (source node). **B:** Schematic representation of the HI after running the personalized PageRank algorithm starting from the diseases-signature cytokines (readout nodes). **C:** Schematic representation of the HI reflecting a final node ranking based on two rankings in **A** and **B**. **D**–**F:** Schematic representations of framework implementation: rheumatoid arthritis cytokines and tumor necrosis factor (TNF; infliximab target) mapped on whole-blood network (**D**), ulcerative colitis (UC) cytokines and TNF mapped on colon network (**E**), and UC cytokines and mitogen-activated protein kinase 3/1 (KO-947 targets) mapped on colon network (**F**). **G** and **H:** Validation of framework using patient-derived gene-expression data with supervised-classification methods for reasonable sample size and with repeated k-fold cross-validation (CV; **G**) and unsupervised clustering for limited sample size (**H**). ERK, extracellular signal-regulated kinase.

regulated in patients with RA or UC (see *Materials and Methods* and Supplemental Table S1).[22,27,28,44,45] In this framework, the expectations were that a drug would first affect its target protein(s); then, intermediary proteins; and ultimately, the disease-signature cytokines. Intermediary proteins most critical for relaying these effects were expected to comprise the most important response-predicting biomarkers. This strategy is agnostic to the amount of patient data available because it requires only three inputs: i) source node(s), ii) readout nodes, and iii) tissue-specific HI subgraph.

All nodes of the HI subgraph were ranked with the personalized PageRank algorithm to measure their random-walk–based scores relative to both source nodes and readout nodes (Figure 1, A and B).[46] Two scores were obtained for each node by running the PageRank computation twice: first, by using drug-target proteins as the starting nodes; and again, by using disease-signature cytokines as the starting nodes. The choice of the PageRank algorithm was driven by its ability to integrate the impact of the network topology, as well as by its robustness against noise and uncertainties in the network structure. Furthermore, the personalization aspect of the PageRank algorithm further allowed computations to be specific to drug targets and cytokines relevant to the disease being studied.

The two scores yielded two rankings for each node (v): $r_{source}(v)$ corresponded to node v's ranking personalized to drug-target proteins (Figure 1A); and $r_{readout}(v)$ corresponded to node v 's ranking personalized to disease-signature cytokines (Figure 1B). Source and readout rankings were combined by using a rank product[47] (Figure 1C) for node v as follows:

$$r(v) = \sqrt{r_{source}(v) \cdot r_{readout}(v)} \quad (4)$$

The rank product was chosen (as opposed to other options, like summation) because the ranking distribution follows a fat-tailed distribution, indicative of multiplicative processes. Importantly, the rank product is also robust to variations in data quality and scale, making it particularly suitable for integrating two ranks. By rank product, high-ranking nodes were more central than low-ranking ones on the paths between source and readout nodes. The top 100 ranked nodes were selected as candidates for treatment-response–predicting biomarkers and for further validation. These biomarker rankings were compared with rankings from the most common network-centrality measures, such as degree centrality, betweenness centrality, and generic PageRank. The rankings did not correlate, indicating that PRoBeNet captures metrics beyond node degree and global centrality (Supplemental Figure S1).

Results

Retrospective and Prospective Validation of Response Biomarkers

The predictive power of the PRoBeNet biomarkers was retrospectively validated in two cases and prospectively validated with an investigational drug (Table 1 and Figure 1, D–F). For UC and RA, only a third to a half of patients respond to the anti-TNF monoclonal antibody infliximab.[48] To identify biomarkers of response, models were built and validated with real-world data from patients with UC ($N = 47$) and RA ($N = 107$). Following current guidelines, the predictive models were built to assess drug response 6 to 8 weeks after patients with UC start treatment; and approximately 6 months after patients with RA start treatment (see *Discussion*). For prospective validation, biomarkers predicting response to the investigational compound KO-947, which inhibits mitogen activated protein kinase (MAPK) 3/1 (a potential target for treating patients with UC and CD), were prospectively evaluated with data from four UC and eight CD tissue-donor patients.[49] For this assessment, gene-expression levels were measured in tissue collected before treatment, and cytokine levels were assayed in tissue collected both before and after treatment (see *Materials and Methods*). For each donor, cytokine ($N = 15$) levels from samples collected before and after treatments were used to identify patients as either responders or nonresponders. Gene-expression levels from samples collected before treatment were used to predict response labels.

Repeated cross-validation was used to train and test models predicting infliximab response (Figure 1G). Because

Table 1 Three Validation Cases of Treatment-Response–Predicting Biomarkers Identified by the Network-Based Framework in Distinct Sample Types

Indication	*n*	Treatment (target)*	Sample type†	Platform	Response definition‡
UC	47	Infliximab (TNF)	Colon biopsy	Microarray	Endoscopic healing at 4 to 8 weeks
RA	107	Infliximab (TNF)	Whole blood	RNA-seq	6-Month ACR50 outcome
UC, CD	12	KO-947 (MAPK3/1)	Intestinal biopsy	RNA-seq	Change in cytokine expression§

*The first two cases are dedicated to an established drug (infliximab). The third case is dedicated to an investigational compound (KO-947).

†Sample source for data used in model development.

‡ACR50 score is calculated as described in *Materials and Methods*.

§A response was defined as ≥20% change in levels of expression.

ACR, American College of Rheumatology; CD, Crohn disease; MAPK, mitogen-activated protein kinase; RA, rheumatoid arthritis; RNA-seq, RNA sequencing; TNF, tumor necrosis factor; UC, ulcerative colitis.

of the small sample size used for the MAPK3/1-inhibitor-response model, unsupervised clustering was used to assess that model's predictive power (Figure 1H).

Discovery and Validation of Biomarkers Predicting Response to Infliximab in Patients with UC

To discover treatment-response–predicting biomarkers of infliximab in UC, TNF (the source node), the UC-signature cytokines (readout nodes) (Supplemental Table S1), and the colon-specific HI subgraph were used to rank biomarkers of response (see *Materials and Methods*). Ranked (PRoBeNet) biomarkers were then validated by using pretreatment gene-expression data from colonic mucosal biopsies from two patient cohorts with UC (see *Materials and Methods*).[50] The cohorts were combined, and batch correction was performed (see *Materials and Methods*). In this combined cohort, patient-derived samples ($N = 47$; 20 responders and 27 nonresponders) were treated with infliximab. The combined cohort and the PRoBeNet biomarkers were used to build and train an L1-regularized logistic regression model. Cross-validation was performed with 80% of the data, and model performance was evaluated with the remaining 20% (see *Materials and Methods*). Model training and validation were repeated across 50 splits to ensure the results were robust and consistent (Figure 1G). A nested cross-validation

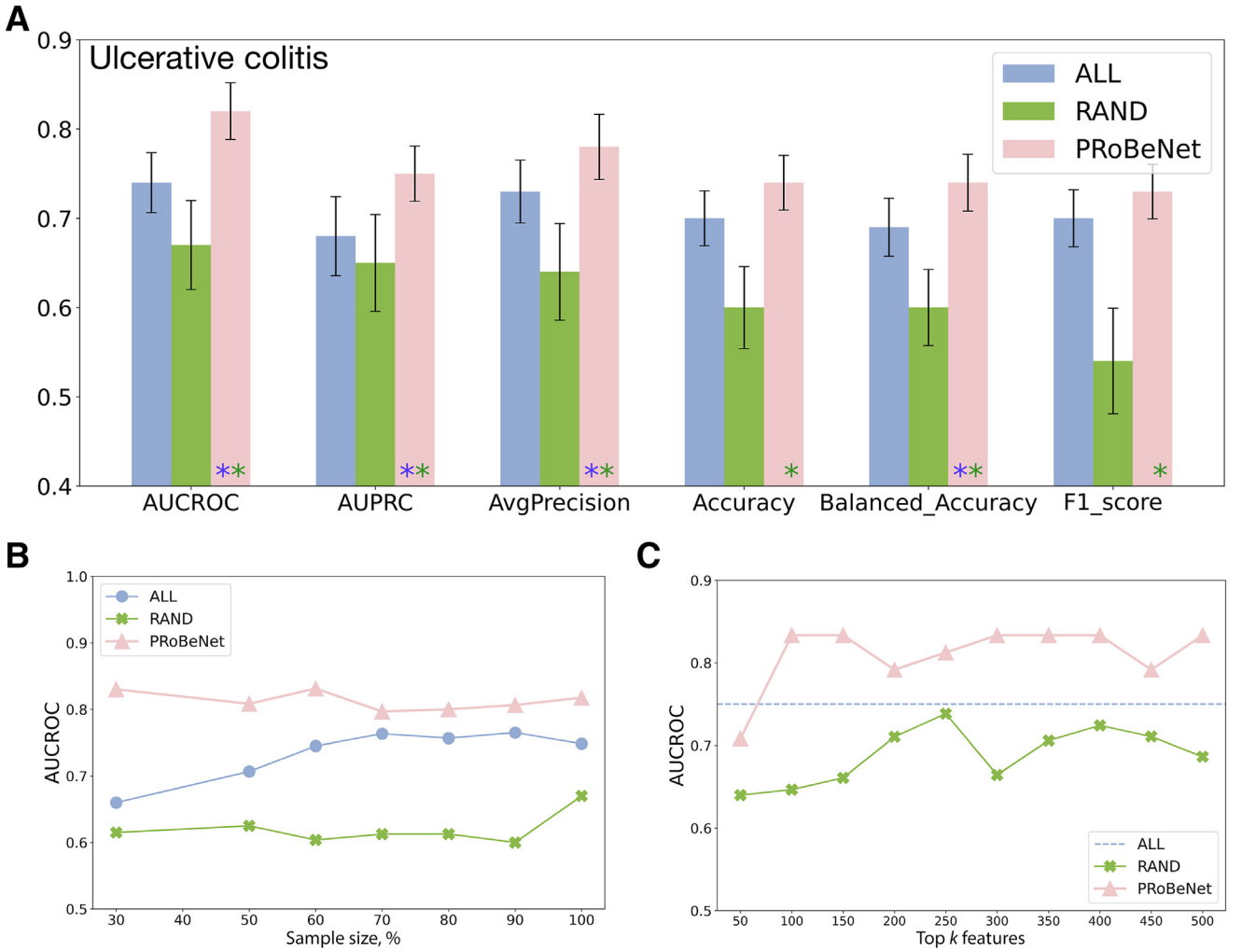

Figure 2 Retrospective validation of response predictive biomarkers in gene-expression data of patients with ulcerative colitis. **A:** Average results for six performance metrics for 50 stratified splits on three feature sets are shown. The models fed with top 100 PRoBeNet (Predictive Response Biomarkers using Network medicine) biomarkers from the proposed approach outperform models fed with all genes (ALL) and 20 sets of randomly selected 100 genes (RAND) in all performance metrics. **Blue** and **green asterisks** reflect whether the outperformance by PRoBeNet model was significant ($P < 0.05$) compared with ALL and RAND models, respectively. **B:** Sample-size analysis indicates the area under the receiver operating characteristic curve (AUCROC) obtained with the top 100 PRoBeNet biomarkers models is substantially higher than the AUCROCs obtained with either all genes or randomly selected gene sets models, even when only 30% of data are used. **C:** The effect of varying the number of biomarkers (top k) used for model training and validation. For the ALL model, there is no provision to vary the number of features; hence, all available genes are used for training and validation. ALL model's performance is indicated by the **straight dotted line**. AUPRC, area under the precision-recall curve; AvgPrecision, average precision.

scheme was used to avoid overfitting the model to the training data, and an L1 regularization was used to eliminate noninformative features.

Model performances were then compared across three feature sets used to train the model: i) RAND, 20 sets of 100 randomly selected genes (a random control); ii) ALL, a set comprising all genes (the traditional approach); and iii) PRoBeNet, a set comprising the top 100 ranked genes. The RAND model used consists of the same number of features as PRoBeNet predicted biomarkers, chosen at random. The random nature of feature selection in this model ensures inclusion of features with no known association to response and thus serving as a negative control. The RAND model should be unable to learn meaningful patterns from randomly selected genes and should thus perform poorly when using these genes. Indeed, the RAND model had the lowest prediction accuracy (AUCROC, 0.65) (Figure 2A). Noise may be introduced in the ALL model because many irrelevant or less informative features may be included that can negatively impact the model's predictive performance. Given the complexity and noise inherent in genomic data, the ALL model may not be optimal (AUCROC, 0.75). Focusing on a subset of genes (ie, PRoBeNet biomarkers) is often more productive than focusing on all genes. Indeed, the model relying on the PRoBeNet biomarkers (100 top ranked) significantly outperformed ($P < 0.05$) the other two models as assessed by several performance metrics (AUCROC, area under the precision-recall curve, average precision, classification accuracy, balanced accuracy, and F1 score) (Figure 2A). The precision-focused metrics (area under the precision-recall curve and average precision) and the harmonic F1 score highlighted the model's adeptness at correctly identifying true nonresponders of infliximab therapy in UC. Additionally, PRoBeNet classification accuracy and balanced accuracy were superior to those of other models, confirming its consistent performance even when accounting for class imbalances. Accuracy improved from 0.70 (ALL) to 0.75 (PRoBeNet). This improvement in predicting patient response becomes predominant in the context of making clinical decisions. When extrapolated to a large cohort of 10,000 patients, the improvement in accuracy afforded by PRoBeNet could better stratify patients and lead to better treatment decisions for an additional 500 patients. This improvement offers substantial real-world benefits, including optimized patient care, fewer adverse effects, and potentially improved survival rates, underlining the value of this approach in the context of treating IMIDs.

The performance of PRoBeNet biomarkers when patient data were limited was evaluated by artificially reducing the amount of patient data (through 10 random samplings) used to train and validate the model. This analysis allowed both an assessment of the power of PRoBeNet biomarkers under limited-data circumstances and an estimate of the minimum sample size required by the method. Even when using as few as 14 samples (a third of the data), the performance of the PRoBeNet model (AUCROC, >0.8) consistently surpassed the performance of the other models (AUCROCs, 0.65 to 0.75) (Figure 2B). The performance of the ALL model linearly correlated with sample size, illustrating that traditional machine-learning tools need large cohorts to accurately identify and filter biomarkers.

Next, the effect of number of top-ranked PRoBeNet biomarkers on model's performance was studied. Model performances were compared after varying the number of top PRoBeNet biomarkers used to train the model (from 50 to 500, in increments of 50). In all variations using >100 biomarkers, the PRoBeNet model outperformed the other models when the top-ranked features exceeded 100 (Figure 2C). Overall, these results confirm that the biomarkers found by PRoBeNet are more predictive than both RAND and traditionally modeled biomarkers. These results also show how PRoBeNet can take advantage of limited-size data sets for building predictive models.

Discovery and Validation of Biomarkers Predicting Response to Infliximab in Patients with RA

In theory, the PRoBeNet framework is capable of detecting and identifying response biomarkers in other disease states by modifying source nodes, readout nodes, and HI subgraphs. To assess the value of this approach in a different context, the framework was used to find infliximab-response–predicting biomarkers for RA, an autoimmune disease. Infliximab-targeted TNF protein was retained as the source node, but RA-signature cytokines (rather than UC-signature ones) were used as readout nodes (Supplemental Table S1). Also, a whole-blood–specific HI subgraph was constructed, because whole-blood RA gene-expression data were now used for validation (see *Materials and Methods*). With these three inputs (source node, readout nodes, and whole-blood subgraph), nodes in the whole-blood–specific HI were ranked, and the top 100 ranked nodes were selected as response-predicting biomarkers. Baseline gene expression, measured in whole blood from patients with RA ($N = 107$) before treatment, was used to assess the predictive power of the biomarkers (see *Materials and Methods*).[32] Patient clinical data needed to calculate ACR treatment-response scores were collected at baseline and at 6-month follow-up visits. Patients were defined as either responders or nonresponders to infliximab treatment by using 6-month ACR50 scores (see *Materials and Methods*).

Predicted biomarkers were validated for the patient cohort with RA by using the same strategy as that used to validate biomarkers from the patient cohort with UC. The patient cohort with RA and the PRoBeNet biomarkers were used to train and validate the model (Figure 1G). By multiple metrics, the model trained on PRoBeNet biomarkers significantly outperformed the models trained on either ALL or RAND genes (Figure 3A). The model trained on ALL

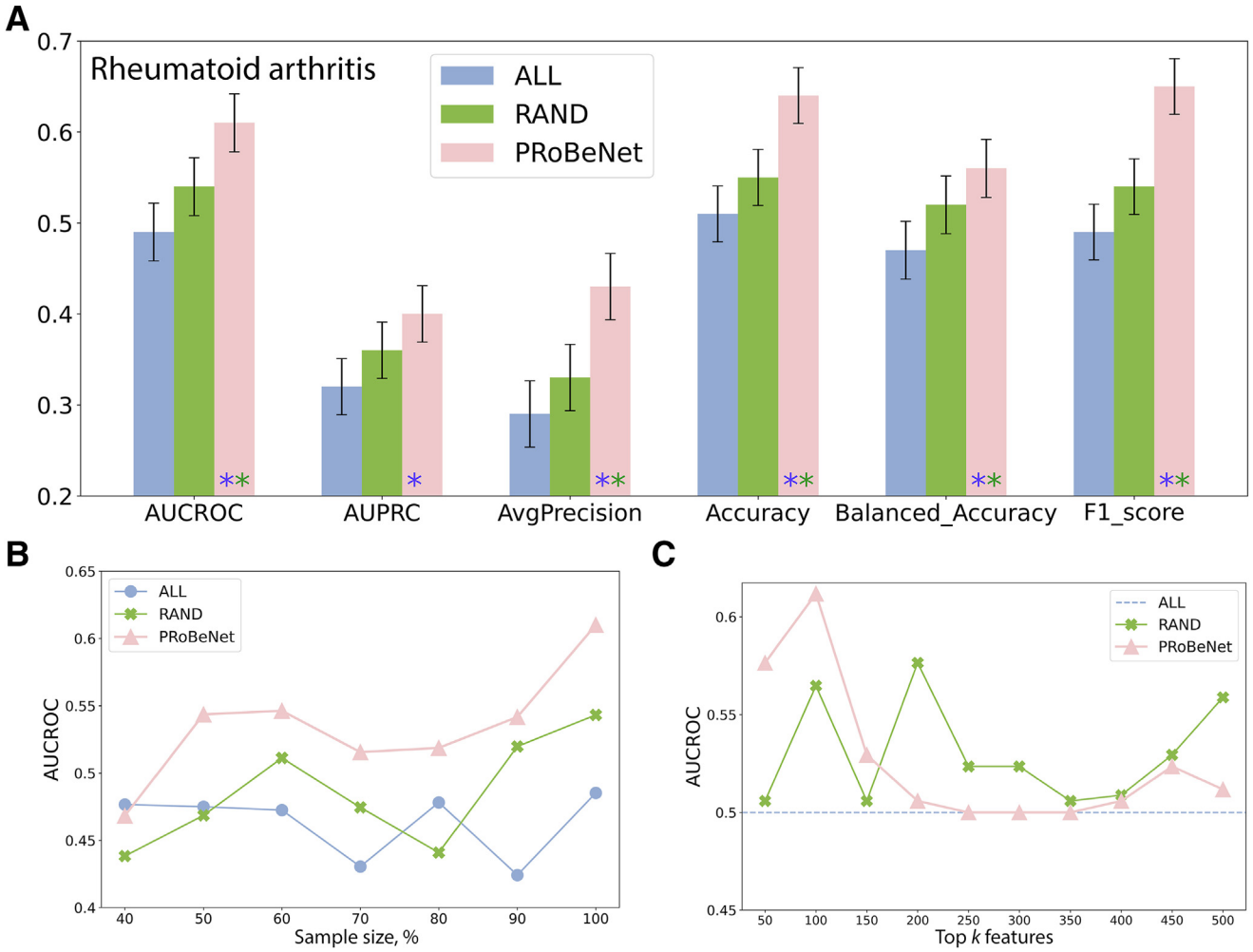

Figure 3 Retrospective validation of response predictive biomarkers in whole-blood gene-expression data of patients with rheumatoid arthritis (RA). **A:** Average results for six performance metrics for 50 stratified splits on three feature sets are shown. The model fed with top 100 PRoBeNet (Predictive Response Biomarkers using Network medicine) biomarkers from the proposed approach outperforms models fed with all genes (ALL) and 20 sets of randomly selected 100 genes (RAND) in all performance metrics. **Blue** and **green asterisks** reflect whether the outperformance by PRoBeNet model was significant ($P < 0.05$) compared with ALL and RAND models, respectively. The signal/noise ratio in RA data is low. **B:** Sample-size analysis indicates the area under the receiver operating characteristic curve (AUCROC) obtained with the top 100 PRoBeNet biomarkers is substantially higher than the AUCROCs obtained with either all genes or randomly selected gene sets, even when the available data size is small. **C:** The effect of varying the number of biomarkers (top k) used for model training and validation. For the ALL model, all available genes are used for training and validation (as in Figure 2). ALL model's performance is indicated by the **straight dotted line**. AUPRC, area under the precision-recall curve; AvgPrecision, average precision.

performed worse than the model trained on RAND genes (100 genes per set), as with many features (ALL set) machine-learning models tend to overfit to noise in training data. Although the outperformance by the PRoBeNet model compared with ALL and RAND is small, the outperformance is still significant. This can be attributed to the low signal/noise ratio in RA data.

The sample-size analysis was repeated for the patient cohort with RA to emulate the real-world scenario of data paucity. With as little as half of the data ($N = 53$), the ranked-biomarker model consistently outperformed the other models (Figure 3B). Thus, even with limited patient data, the PRoBeNet biomarkers predict responders from nonresponders better than do other models. Model performance peaked when using approximately 100 PRoBeNet biomarkers (Figure 3C), indicating that PRoBeNet captured the biomarkers with the greatest predictive value.

Physician and patient assessments of RA severity and treatment responses are subjective. This subjectivity may introduce ambiguity in ACR scores (used to build and validate response-predicting models for RA). To reduce this ambiguity, the analysis was repeated by including only extreme-responder patients with RA (who achieved $\geq$70% improvement in RA symptoms, as assessed by 6-month ACR70) and extreme-nonresponder patients (who achieved <20% improvement, as assessed by 6-month ACR20). When using these patients, the ranked-biomarker model (AUCROC, 0.64) further outperformed the other models (RAND AUCROC, 0.46; ALL, 0.5) (Supplemental Figure S2). For insight into the biology determining whether patients with RA and UC will respond to

TNFi, pathways overrepresented by PRoBeNet biomarkers (top PRoBeNet biomarkers for RA and for UC) were identified by using Metascape. Several canonical pathways related to TNF and cytokine signaling were enriched in both the RA and UC biomarkers. Comparing the enriched pathways for each disease revealed several UC-specific pathways, such as cadherin and endothelial cell adhesion, thereby yielding further disease-specific insights from the predicted biomarkers[51] (Supplemental Figure S3). Overall, these results indicate that the PRoBeNet captures response-predicting biomarkers for different diseases, even when using limited data.

Prospective Discovery of Biomarkers Predicting Response to MAPK3/MAPK1 Inhibitor

Precision medicine has the potential to improve drug-response rates by revealing biomarkers of therapeutic response, thus enabling identification of patients who carry these biomarkers. Knowledge of such biomarkers may also translate into US Food and Drug Administration approval for drugs that are administered with an accompanying diagnostic tool. To test this hypothesis, a preclinical validation experiment was designed to test the performance of PRoBeNet biomarkers in predicting responses to the investigational compound KO-947, whose mechanism of action differs from that of inflximab. KO-947 inhibits MAPK3/MAPK1, which modulate inflammatory cytokine signaling; thus, KO-947 may be efficacious for treating autoimmune diseases and IMIDs.[52] Molecular and outcome data were prospectively generated for validating this approach in a preclinical setting (see *Materials and Methods*).[53] PRoBeNet biomarkers were identified by using MAPK3 and MAPK1 as source nodes; 15 cytokines (both proinflammatory and anti-inflammatory), whose levels were measured as readout nodes (Supplemental Table S1); and the colon-specific HI subgraph. Despite the differences in UC and CD etiologies, MAPK3/1 inhibitors work in UC-derived tissues as they do in CD-derived ones (by reducing cytokines). Colon or ileum samples were collected from four patients with UC and eight patients with CD undergoing therapeutic resection (Figure 4A). Tissue samples were cultured and treated either with *Staphylococcus* enterotoxin and vehicle (baseline samples) or *Staphylococcus* enterotoxin

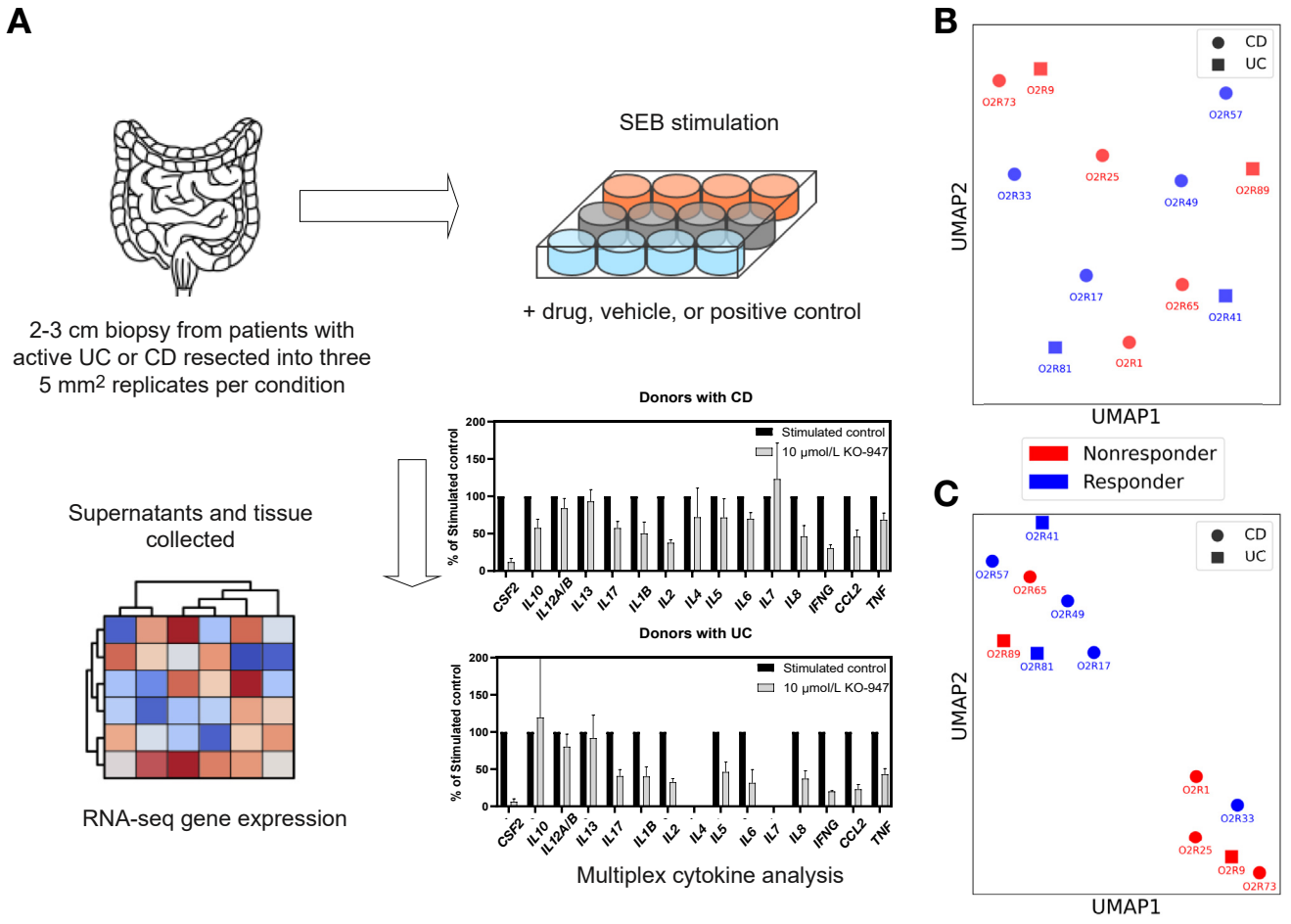

Figure 4 Preclinical validation of KO-947 (a mitogen-activated protein kinase 3/1 inhibitor) response predictive biomarkers. **A:** Study design. **B:** Uniform manifold approximation and projection (UMAP) generated using all genes as inputs. **C:** UMAP generated from the top 100 predicted biomarkers; responders are well separated from nonresponders. CCL, chemokine (C-C motif) ligand; CD, Crohn disease; CSF2, colony stimulating factor 2; IFNG, interferon gamma; RNA-seq, RNA sequencing; SEB, *Staphylococcus* enterotoxin B; UC, ulcerative colitis.

Table 2 *Ex Vivo* Responses of Tissues from Patients with UC and CD to KO-947 Treatment

Donor ID	Cytokine-response change	Binary response*	Indication
O2R9	0.06	NR	UC
O2R1	0.08	NR	CD
O2R25	0.08	NR	CD
O2R65	0.09	NR	CD
O2R89	0.17	NR	UC
O2R73	0.18	NR	CD
O2R57	0.25	R	CD
O2R17	0.25	R	CD
O2R81	0.35	R	UC
O2R49	0.40	R	CD
O2R41	0.43	R	UC
O2R33	0.46	R	CD

*Donors with >20% change in anti-inflammatory cytokine expression were considered responders.

CD, Crohn disease; ID, identifier; NR, nonresponder; R, responder; UC, ulcerative colitis.

and KO-947 (treated samples). Disease-signature cytokines secreted into media and gene expression in tissue were measured for all samples (see *Materials and Methods*). Response scores for patients were calculated by using changes in cytokine-expression levels between baseline and treated samples and across replicates (see *Materials and Methods*). The response scores for UC and CD samples differed among donor types (Table 2), confirming this *ex vivo* experimental approach was appropriate for discovering response biomarkers in a preclinical setting.

Given the few samples, the power of the identified biomarkers to segregate responders from nonresponders was validated by analyzing cytokine-derived response scores with an unsupervised clustering approach. The top 100 PRoBeNet biomarkers were used as features, and dimensionality reduction was performed by using a uniform manifold approximation and projection algorithm. Three-fourths of samples (9 of 12) were properly classified by treatment-response levels (Figure 4B). To compare the predictive power of the PRoBeNet biomarkers with the power of the traditional approach, uniform manifold approximation and projection was implemented on data from ALL; here, the samples did not clearly cluster (Figure 4C). The hierarchical clustering analysis performed with the PRoBeNet biomarkers had fewer misclassifications (3 of 12) than did hierarchical clustering analysis performed with ALL genes (5 of 12) (Supplemental Figure S4). These results were quantitatively compared by using the adjusted RAND score, a clustering similarity metric that ranges from −1 to 1, where a higher score indicates better clustering and a score of 0 suggests random labeling. The adjusted RAND score for clustering with the PRoBeNet biomarkers (0.1773) was significantly higher than that for clustering with the ALL biomarkers (−0.0433; $P = 0.0109$) (Supplemental Figure S5), suggesting the PRoBeNet biomarkers were considerably more accurate in clustering tissue samples.

These results indicate that PRoBeNet can be used to find response-predicting biomarkers for both approved drugs and investigational compounds, even with few samples.

Discussion

A novel network-based framework for discovering treatment response–predicting biomarkers, PRoBeNet, is described in this study. PRoBeNet is an important contribution to leveraging precision-medicine strategies to treat complex diseases because it does not require an extensive patient data set. Instead, the framework takes advantage of network-medicine methods, specifically drawing on the predictive power of tissue-specific HI networks. The premise behind this framework is that a therapy propagates through a cascade of protein-protein interactions to bring expression levels of disease-specific signatures (comprising cytokines abnormally expressed in each disease) closer to a normal range. PRoBeNet prioritizes intermediary nodes in the HI that play critical roles in transmitting the effects of therapies (infliximab or KO-947) from source-node target proteins (TNF or MAPK3/1) to readout-node disease-signature cytokines (for RA, UC, or CD). Predictive biomarkers were selected from prioritized intermediary nodes to determine patient responses to each treatment. The PRoBeNet biomarkers could more accurately predict treatment responses and were less contingent on training-set sizes than were biomarkers found by the traditional approach. The new approach could also reduce potential biomarkers to a more manageable number, mitigating a common drawback of traditional machine-learning approaches when dealing with small cohorts and many features (the curse of dimensionality).

A distinctive characteristic of PRoBeNet is the dual use of independent PageRank scores. The first score, recognizing proteins essential for source-node–derived effects, earmarks proteins critical to the therapy's mechanism of action. The second score recognizes proteins connected to readout nodes, highlighting proteins that strongly influence disease pathogenesis and cytokine dysregulation. By integrating both scores via the rank product (Equation 4), the framework evaluates each biomarker's relevance to both source nodes and readout nodes, ensuring that high-scoring biomarkers are not only associated with the therapy but also crucial for transmitting effects to cytokines. Importantly, this dual-score approach sheds light on the molecular mechanisms underlying a successful response. In doing so, it complements traditional, purely data-driven biomarker-discovery methods, which often rely on machine-learning or statistical models to identify biomarkers based on their correlation with response outcomes. Unlike traditional machine-learning approaches, PRoBeNet considers molecular mechanisms of response, so it may reveal valuable

biomarkers that traditional approaches fail to capture. Also, biomarkers based on differentially expressed genes (between responders and nonresponders of therapy) are not generalizable; when differentially expressed genes are mapped onto the HI, they tend to be scattered and to have sparse interactions. In contrast, PRoBeNet biomarkers consistently form a connected component in the HI. Notably, many of the differentially expressed genes are next to one or more PRoBeNet biomarkers, underlining that PRoBeNet comprehensively integrates the network (Supplemental Figure S6).

Pathway analysis of TNFi biomarkers in UC and RA suggests PRoBeNet may also explain crosstalk among pathways driving resistance to established therapies; thus, it has the potential to identify new therapies to overcome resistance. For example, genes belonging to glucocorticoid-receptor–dependent gene-regulatory networks were significantly enriched among infliximab response–predicting biomarkers (Supplemental Figure S3). Glucocorticoids (such as prednisone) rapidly reduce inflammation in patients with RA and UC.[54,55] However, TNF therapy induces glucocorticoid resistance because the TNF-signaling pathway is significantly interconnected with and coregulated with the glucocorticoid-receptor pathway.[56,57] These results suggest resistance may be overcome by using drugs conjugated to antibodies that target cell-surface TNF and glucocorticoid receptor [nuclear receptor subfamily 3 group C (NR3C1)].[58] Similar drug-antibody conjugates are used to treat patients with cancer who inadequately respond to monotherapies, but such conjugates are yet to be used to treat patients with autoimmune diseases or IMIDs.[59] Thus, this analysis correctly indicated a key mechanism of glucocorticoid resistance and supported a distinct combination therapy for overcoming this resistance.

Although biomarkers were discovered and successfully validated with PRoBeNet, there were inherent challenges. A primary challenge encountered specifically while finding RA biomarkers was a low signal/noise ratio in patients' whole-blood gene-expression data. Although RA biomarkers were found by using data from whole blood, which does not directly manifest RA disease, UC and CD biomarkers were found by using data from colon tissue, which does manifest UC and CD. Signal/noise ratios in data from tissues that do not actively manifest a disease are lower than signal/noise ratios in data from tissues that do. Indeed, the signal/noise ratio in the RA data was much lower than that in the UC data. Despite this challenge, PRoBeNet identified response-predicting biomarkers for RA, showing the successful application of the approach in data sets with poor signal/noise ratios. These results underline the effectiveness of the framework for identifying relevant biomarkers to stratify patients. Given the low response rates of patients with RA and UC to TNFi therapy, this study primarily focused on predicting initial therapy response by using baseline data collected before patients started treatment. This approach allows physicians to explore alternative therapies for patients with probable lack of response to infliximab from the outset, potentially avoiding many months of adverse effects and disease progression. Furthermore, as recommended by clinical guidelines, TNFi response is often measured at 6 to 8 weeks after treatment initiation for patients with UC, and after 6 months for patients with RA. Therefore, to increase the predictive model's clinical utility, the models were developed to predict responses at these recommended time points by using retrospective gene-expression data collected before patients started treatments.

The purpose of this study was to introduce a computational pipeline to reduce the number of features and to assess performance of these features in predicting responses at a certain time point based on current guidelines. Depending on the required application of these models, however, researchers are encouraged to overcome certain challenges, mainly introduced by data limitations. For example, if researchers seek to predict either sustainable responses after treatment or dependencies of response to earlier disease-modifying antirheumatic drugs and biologics, then researchers should consider using response measurements at multiple time points or large enough cohorts with data on earlier treatments.

In conclusion, PRoBeNet was developed for predicting a reduced pool of therapy-response biomarkers to build generalizable models for diseases with little data available (as few as 14 samples). The framework may be useful for developing companion diagnostic tests, with which appropriate patient subpopulations could be directed to specific therapies, thereby amplifying drug efficacies. It also holds promise for refining clinical development programs by indicating suitable patient populations in phase 2/3 studies. Furthermore, this approach can be generalized to other complex diseases by using different sets of molecular phenotype markers (ie, differentially expressed genes or proteins) as readout nodes. Finally, the framework holds broad value for reducing the number of features when searching for biomarkers in highly dimensional data. This work is expected to bring the scientific community closer to realizing the full potential of precision medicine for treating autoimmune diseases and other complex conditions by improving personalized treatment strategies.

Acknowledgments

We thank Ivan Voitalov, Sumit Borah, Krishna Patel, Sherry Guardiano, and Brian Cheng for support.

Author Contributions

S.D.G. and V.R.A. conceived the project; U.S.S. and S.D.G. developed the prediction framework and ran the analysis; V.R.A., C.K., and S.D.G. designed preclinical experiments; A.J. processed raw FASTQ files for patient-derived samples

from patients with ulcerative colitis; J.R. performed differential expression analysis on the publicly available patient data from patients with ulcerative colitis; U.S.S. and S.D.G. analyzed the results; and all authors discussed the content and drafted and reviewed the manuscript before submission. S.D.G. is the guarantor of this work, had access to all study data, and is responsible for data integrity and the accuracy of analysis.

Disclosure Statement

U.S.S., C.K., A.J., A.S., V.R.A., and S.D.G. are full-time employees and shareholders of Scipher Medicine. J.R. is related to Michael Rubin, a member of the Scipher board of directors and a Scipher shareholder. A.-L.B. is a founder of Scipher Medicine.

Supplemental Data

Supplemental material for this article can be found at *http://doi.org/10.1016/j.jmoldx.2024.06.008*.

References

1. Conte F, Fiscon G, Licursi V, Bizzarri D, D'Antò T, Farina L, Paci P: A paradigm shift in medicine: a comprehensive review of network-based approaches. Biochim Biophys Acta Gene Regul Mech 2020, 1863:194416
2. Vafaee F, Diakos C, Kirschner MB, Reid G, Michael MZ, Horvath LG, Alinejad-Rokny H, Cheng ZJ, Kuncic Z, Clarke S: A data-driven, knowledge-based approach to biomarker discovery: application to circulating microRNA markers of colorectal cancer prognosis. Npj Syst Biol Appl 2018, 4:1−12
3. Cheng F, Kovács IA, Barabási AL: Network-based prediction of drug combinations. Nat Commun 2019, 10:1197
4. Kyrochristos ID, Ziogas DE, Roukos DH: Dynamic genome and transcriptional network-based biomarkers and drugs: precision in breast cancer therapy. Med Res Rev 2019, 39:1205−1227
5. Tabernero J, Lenz HJ, Siena S, Sobrero A, Falcone A, Ychou M, Humblet Y, Bouché O, Mineur L, Barone C, Adenis A, Yoshino T, Goldberg RM, Sargent DJ, Wagner A, Laurent D, Teufel M, Jeffers M, Grothey A, Van Cutsem E: Analysis of circulating DNA and protein biomarkers to predict the clinical activity of regorafenib and assess prognosis in patients with metastatic colorectal cancer: a retrospective, exploratory analysis of the CORRECT trial. Lancet Oncol 2015, 16:937−948
6. Verdaguer H, Saurí T, Macarulla T: Predictive and prognostic biomarkers in personalized gastrointestinal cancer treatment. J Gastrointest Oncol 2017, 8:405−417
7. Tejpar S, Bertagnolli M, Bosman F, Lenz HJ, Garraway L, Waldman F, Warren R, Bild A, Collins-Brennan D, Hahn H, Harkin DP, Kennedy R, Ilyas M, Morreau H, Proutski V, Swanton C, Tomlinson I, Delorenzi M, Fiocca R, Van Cutsem E, Roth A: Prognostic and predictive biomarkers in resected colon cancer: current status and future perspectives for integrating genomics into biomarker discovery. Oncologist 2010, 15:390−404
8. Li J, Lei K, Wu Z, Li W, Liu G, Liu J, Cheng F, Tang Y: Network-based identification of microRNAs as potential pharmacogenomic biomarkers for anticancer drugs. Oncotarget 2016, 7:45584−45596
9. Kong J, Ha D, Lee J, Kim I, Park M, Im SH, Shin K, Kim S: Network-based machine learning approach to predict immunotherapy response in cancer patients. Nat Commun 2022, 13:3703
10. Amin S, Bathe OF: Response biomarkers: re-envisioning the approach to tailoring drug therapy for cancer. BMC Cancer 2016, 16:850
11. Glaab E: Using prior knowledge from cellular pathways and molecular networks for diagnostic specimen classification. Brief Bioinform 2016, 17:440−452
12. Kingsmore KM, Puglisi CE, Grammer AC, Lipsky PE: An introduction to machine learning and analysis of its use in rheumatic diseases. Nat Rev Rheumatol 2021, 17:710−730
13. Taylor JMG, Ankerst DP, Andridge RR: Validation of biomarker-based risk prediction models. Clin Cancer Res 2008, 14:5977−5983
14. Archer R, Hock E, Hamilton J, Stevens J, Essat M, Poku E, Clowes M, Pandor A, Stevenson M: Assessing prognosis and prediction of treatment response in early rheumatoid arthritis: systematic reviews. Health Technol Assess 2018, 22:1−294
15. Diaz-Uriarte R, Gómez de Lope E, Giugno R, Fröhlich H, Nazarov PV, Nepomuceno-Chamorro IA, Rauschenberger A, Glaab E: Ten quick tips for biomarker discovery and validation analyses using machine learning. PLoS Comput Biol 2022, 18: e1010357
16. Choobdar S, Ahsen ME, Crawford J, Tomasoni M, Fang T, Lamparter D, Lin J, Hescott B, Hu X, Mercer J, Natoli T, Narayan R; DREAM Module Identification Challenge Consortium, Subramanian A, Zhang JD, Stolovitzky G, Kutalik Z, Lage K, Slonim DK, Saez-Rodriguez J, Cowen LJ, Bergmann S, Marbach D: Assessment of network module identification across complex diseases. Nat Methods 2019, 16:843−852
17. Strunz S, Wolkenhauer O, de la Fuente A: Network-assisted disease classification and biomarker discovery. Methods Mol Biol 2016, 1386:353−374
18. Toro-Domínguez D, Alarcón-Riquelme ME: Precision medicine in autoimmune diseases: fact or fiction. Rheumatology 2021, 60: 3977−3985
19. Menche J, Sharma A, Kitsak M, Ghiassian S, Vidal M, Loscalzo J, Barabási AL: Uncovering disease-disease relationships through the incomplete human interactome. Science 2015, 347:1257601
20. Guthridge JM, Wagner CA, James JA: The promise of precision medicine in rheumatology. Nat Med 2022, 28:1363−1371
21. Lin CMA, Cooles FAH, Isaacs JD: Precision medicine: the precision gap in rheumatic disease. Nat Rev Rheumatol 2022, 18:725−733
22. Zhao SS, Moots RJ: Biomarkers for treatment response in rheumatoid arthritis: where are they? Rheumatol Immunol Res 2020, 1:1−3
23. Edgar R, Domrachev M, Lash AE: Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 2002, 30:207−210
24. Xu J, Murphy SL, Kochanek KD, Arias E: Mortality in the United States, 2021. NCHS Data Brief 2022, 456:1−8
25. Ghiassian SD, Voitalov I, Withers JB, Santolini M, Saleh A, Akmaev VR: Network-based response module comprised of gene expression biomarkers predicts response to infliximab at treatment initiation in ulcerative colitis. Transl Res 2022, 246:78−86
26. Kokkonen H, Söderström I, Rocklöv J, Hallmans G, Lejon K, Rantapää Dahlqvist S: Up-regulation of cytokines and chemokines predates the onset of rheumatoid arthritis. Arthritis Rheum 2010, 62: 383−391
27. Brink M, Lundquist A, Alexeyenko A, Lejon K, Rantapää-Dahlqvist S: Protein profiling and network enrichment analysis in individuals before and after the onset of rheumatoid arthritis. Arthritis Res Ther 2019, 21:288
28. Brzustewicz E, Bryl E: The role of cytokines in the pathogenesis of rheumatoid arthritis–practical and potential application of cytokines as biomarkers and targets of personalized therapy. Cytokine 2015, 76: 527−536
29. Mellors T, Withers JB, Ameli A, Jones A, Wang M, Zhang L, Sanchez HN, Santolini M, Do Valle I, Sebek M, Cheng F,

Pappas DA, Kremer JM, Curtis JR, Johnson KJ, Saleh A, Ghiassian SD, Akmaev VR: Clinical validation of a blood-based predictive test for stratification of response to tumor necrosis factor inhibitor therapies in rheumatoid arthritis patients. Netw Syst Med 2020, 3:91−104
30. Koh K, Kim SJ, Boyd S: A method for large-scale L1-regularized logistic regression. AAAI 2007, 22:565−571
31. Rutgeerts P, Sandborn WJ, Feagan BG, Reinisch W, Olson A, Johanns J, Travers S, Rachmilewitz D, Hanauer SB, Lichtenstein GR, de Villiers WJ, Present D, Sands BE, Colombel JF: Infliximab for induction and maintenance therapy for ulcerative colitis. N Engl J Med 2005, 353:2462−2476
32. Curtis JR, Strand V, Golombek S, Zhang L, Wong A, Zielinski MC, Akmaev VR, Saleh A, Asgarian S, Withers JB: Patient outcomes improve when a molecular signature test guides treatment decision-making in rheumatoid arthritis. Expert Rev Mol Diagn 2022, 22:1−10
33. Smolen JS, Aletaha D, McInnes IB: Rheumatoid arthritis. Lancet 2016, 388:2023−2038
34. Cohen S, Wells AF, Curtis JR, Dhar R, Mellors T, Zhang L, Withers JB, Jones A, Ghiassian SD, Wang M, Connolly-Strong E, Rapisardo S, Gatalica Z, Pappas DA, Kremer JM, Saleh A, Akmaev VR: A molecular signature response classifier to predict inadequate response to tumor necrosis factor-[alpha] inhibitors: the NETWORK-004 prospective observational study. Rheumatol Ther 2021, 8:1159−1176
35. Nakase H, Sato N, Mizuno N, Ikawa Y: The influence of cytokines on the complex pathology of ulcerative colitis. Autoimmun Rev 2022, 21:103017
36. Friedrich M, Pohin M, Powrie F: Cytokine networks in the pathophysiology of inflammatory bowel disease. Immunity 2019, 50:992−1006
37. Kondo N, Kuroda T, Kobayashi D: Cytokine networks in the pathogenesis of rheumatoid arthritis. Int J Mol Sci 2021, 22:10922
38. Mateen S, Zafar A, Moin S, Khan AQ, Zubair S: Understanding the role of cytokines in the pathogenesis of rheumatoid arthritis. Clin Chim Acta 2016, 455:161−171
39. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR: STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013, 29:15−21
40. Dobin A, Gingeras TR: Mapping RNA-seq reads with STAR. Curr Protoc Bioinformatics 2015, 51:11.14.1−11.14.19
41. Li B, Dewey CN: RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinf 2011, 12:323
42. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, van der Walt SJ, Brett M, Wilson J, Millman KJ, Mayorov N, Nelson ARJ, Jones E, Kern R, Larson E, Carey CJ, Polat İ, Feng Y, Moore EW, VanderPlas J, Laxalde D, Perktold J, Cimrman R, Henriksen I, Quintero EA, Harris CR, Archibald AM, Ribeiro AH, Pedregosa F, van Mulbregt P; SciPy 1.0 Contributors: SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods 2020, 17:261−272
43. Kitsak M, Sharma A, Menche J, Guney E, Ghiassian SD, Loscalzo J, Barabási AL: Tissue specificity of human disease module. Sci Rep 2016, 6:35241
44. Prasad B, McGeough C, Eakin A, Ahmed T, Small D, Gardiner P, Pendleton A, Wright G, Bjourson AJ, Gibson DS, Shukla P: ATRPred: a machine learning based tool for clinical decision making of anti-TNF treatment in rheumatoid arthritis patients. PLoS Comput Biol 2022, 18:e1010204
45. Lerner A, Neidhöfer S, Reuter S, Matthias T: MMP3 is a reliable marker for disease activity, radiological monitoring, disease outcome predictability, and therapeutic response in rheumatoid arthritis. Best Pract Res Clin Rheumatol 2018, 32:550−562
46. Iván G, Grolmusz V: When the Web meets the cell: using personalized PageRank for analyzing protein interaction networks. Bioinformatics 2011, 27:405−407
47. Breitling R, Armengaud P, Amtmann A, Herzyk P: Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. FEBS Lett 2004, 573:83−92
48. Roda G, Jharap B, Neeraj N, Colombel JF: Loss of response to anti-TNFs: definition, epidemiology, and management. Clin Transl Gastroenterol 2016, 7:e135
49. Voitalov I, Zhang L, Kilpatrick C, Withers JB, Saleh A, Akmaev VR, Ghiassian SD: The module triad: a novel network biology approach to utilize patients' multi-omics data for target discovery in ulcerative colitis. Sci Rep 2022, 12:21685
50. Arijs I, Li K, Toedter G, Quintens R, Van Lommel L, Van Steen K, Leemans P, De Hertogh G, Lemaire K, Ferrante M, Schnitzler F, Thorrez L, Ma K, Song XYR, Marano C, Van Assche G, Vermeire S, Geboes K, Schuit F, Baribaud F, Rutgeerts P: Mucosal gene signatures to predict response to infliximab in patients with ulcerative colitis. Gut 2009, 58:1612−1619
51. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK: Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun 2019, 10:1523
52. Lucas RM, Luo L, Stow JL: ERK1/2 in immune signalling. Biochem Soc Trans 2022, 50:1341−1352
53. Mei B, Zhu L, Guo Y, Wu T, Ren P, Deng X: Solid form selection and process development of KO-947 drug substances. Org Process Res Dev 2021, 25:1637−1647
54. Hua C, Buttgereit F, Combe B: Glucocorticoids in rheumatoid arthritis: current status and future studies. RMD Open 2020, 6:e000536
55. Bruscoli S, Febo M, Riccardi C, Migliorati G: Glucocorticoid therapy in inflammatory bowel disease: mechanisms and clinical practice. Front Immunol 2021, 12:691480
56. Van Bogaert T, De Bosscher K, Libert C: Crosstalk between TNF and glucocorticoid receptor signaling pathways. Cytokine Growth Factor Rev 2010, 21:275−286
57. Dendoncker K, Timmermans S, Vandewalle J, Eggermont M, Lempiäinen J, Paakinaho V, Van Hamme E, Dewaele S, Vandevyver S, Ballegeer M, Souffriau J, Van Wyngene L, Van Looveren K, Vanderhaeghen T, Beyaert R, De Bosscher K, Palvimo JJ, Van Montagu M, Libert C: TNF-[alpha] inhibits glucocorticoid receptor-induced gene expression by reshaping the GR nuclear cofactor profile. Proc Natl Acad Sci U S A 2019, 116:12942−12951
58. Hobson AD, McPherson MJ, Hayes ME, Goess C, Li X, Zhou J, Wang Z, Yu Y, Yang J, Sun L, Zhang Q, Qu P, Yang S, Hernandez A, Bryant SH, Mathieu SL, Bischoff AK, Fitzgibbons J, Santora LC, Wang L, Wang L, Fettis MM, Li X, Marvin CC, Wang Z, Patel MV, Schmidt DL, Li T, Randolph JT, Henry RF, Graff C, Tian Y, Aguirre AL, Shrestha A: Discovery of ABBV-3373, an anti-TNF glucocorticoid receptor modulator immunology antibody drug conjugate. J Med Chem 2022, 65:15893−15934
59. Bayat MR, Homayouni TS, Baluch N, Morgatskaya E, Kumar S, Das B, Yeger H: Combination therapy in combating cancer. Oncotarget 2017, 8:38022−38043